Regulating the environmental impacts of the electricity supply industry.

HORNE, Ralph.

Published version

HORNE, Ralph. (2001). Regulating the environmental impacts of the electricity supply industry. Doctoral, Sheffield Hallam University (United Kingdom)..

Regulating the Environmental Impacts of the Electricity Supply Industry

Ralph Horne

A thesis submitted in partial fulfilment of the requirements of

Sheffield Hallam University

for the degree of Doctor of Philosophy

January 2001

ABSTRACT

The electricity supply industry (ESI) in England and Wales does not operate efficiently, in terms of optimising the balance between benefits of electricity and costs associated with environmental impacts. The optimal situation would be one where such impacts are minimised per unit of electricity service used, notwithstanding cost considerations. However, the present regulatory regime fails to account sufficiently for environmental impacts. Indeed, it cannot do so at present, due to lack of objective, complete and sufficiently accurate information.

The main methods currently advocated for valuing environmental impacts are based on the theory of neo-classical environmental economics. These aim to place monetary values on impacts, which can then, in theory, be used to internalise environmental externalities, by applying market mechanisms to correct for the market inefficiency. However, numerous objections have been raised and weaknesses identified, including, principally, the lack of a systematic approach and the inability of the technique to accurately value impacts which are not usually considered in monetary terms.

Better regulation starts with better understanding of the issue(s) to be regulated. In this case, it requires appropriate data about values of environmental impacts. While environmental economics is not rejected outright, further improvements are required and, in any event, it must be supplemented by a systematic approach, which encompasses a means of valuing non-economic elements of value. The Environmental Analysis, Valuation and Application (EAVA) Framework proposed here has been designed and developed in order to address these requirements. It also satisfies the need for objectivity, rigour, transparency, versatility, practicality and a step-by-step, sequential procedure for dealing appropriately with environmental impacts.

The EAVA Framework encompasses four separate methods which have been developed simultaneously to work together in order to address different areas of the problem. The output analysis method allows the production of a complete inventory of released incidental outputs (RIOs) which arise from the process being studied. The pathway analysis method provides a means of tracing these RIOs through the environment and generating objective data about the resulting environmental changes. The valuation method is where the only necessary subjectivity of valuation is concentrated by accommodating the views of those whose quality of life is damaged by the impacts. The unit of valuation is the "natural" unit of quality of life outcome state (QLOS), and quantification is achieved through use of the QLOS Index. The final method is the application method, where valuation data and information about unknowns or other "gaps" in knowledge or data are utilised in mechanisms to ensure decision making and operation of the process concerned correctly reflects the environmental impacts caused. It should be noted here that procedures exist throughout the EAVA Framework for identifying and quantifying "gaps". The overall result is the EAVA Framework - a single integrated process for regulating environmental impacts, from the point of origin, to the point of applying regulation.

ACKNOWLEDGEMENTS

Sincere acknowledgements are due to my principal supervisor, Professor Nigel Mortimer, at the Resources Research Unit at Sheffield Hallam University, for unstinting, critical and seemingly endless support and advice. In addition, thanks to Ana Santos, Nick Horne and Jon Kellett for constructive comments and support, as well as proofing and checking work on thesis drafts. Numerous others have assisted in the process and it is difficult to single out individuals. However, I am particularly indebted to staff in the Resources Research Unit.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Context

Issues surrounding the environmental impacts of electricity supply have received much attention over the past two decades, such that large international organisations involved in energy production and use now acknowledge the growing need to minimise their environmental impact. Through issues such as acid rain, global climate change, and the over-exploitation of oceans and forests, the environment has become a regional and global concern. Although some commentators have laid the root of the problem at the door of specific factors such as growing population or modern industrial practices, it is clear that the single cause is human behaviour and that human ability to disrupt the environment has far outpaced human ability to foresee the consequences. Too often, the environment is still treated as a free asset, to be used and abused at will. This must change if damage is to be contained.

The problem, like all real problems, is complex and challenging. Access to electricity has become a fundamental need, certainly in the western world, yet its provision is causing damage to the environment - and it is not just the environmental issues discussed in the popular media that require attention. The processes involved in generating electricity for use as an accessible and transferable source of energy currently result in an extremely wide range of incidental and unwanted environmental impacts. Without sufficient mechanisms which make those involved in producing electricity account for them and reduce them, the harmful effects will be felt by everyone and by the earth's ecosystem as a whole.

Dealing with environmental impacts after they have arisen is generally less satisfactory than preventing them from occurring in the first place, because of the potential wide-ranging effects. The two principal regulatory means which have been considered by governments to deal with environmental impacts before they arise are "command and control", where limits and controls are set on processes to reduce their potential environmental impacts to an acceptable level, and "market mechanisms", where typically, processes attract a premium according to the environmental impacts they are expected to cause. This premium, often some form of tax, is then reflected in prices of the resultant electricity produced, thus reducing demand, production and environmental impacts accordingly through the operation of the market.

The command and control type approach is well-established and, as a result, the overall level of environmental impact from electricity production is partially reduced. Those impacts which are specifically regulated for occur at a controlled level. For example, there are controls on types and levels of pollutants which can be emitted from power station stacks, controls on the design and location of mines and power stations through the planning system, and controls on the disposal of wastes from the generation process. However, it should be noted that command and control type measures are generally directed at relatively few, specific impacts, and that they generally prescribe acceptable levels of impact - so the impact continues, but at a lower level than that which could be expected without the controls.

The market generally, and market mechanisms specifically, also currently help to reduce environmental impacts. For example, competition-sponsored technology developments increase the efficiency of electricity production and use, thereby reducing the amount of wastage, with a general (though incidental) positive effect on reducing unwanted and incidental outputs from the process such as environmental impacts. Consumer pressure on producers to act in an environmentally acceptable way could translate into more "responsible" production with lower environmental impacts, although there is little evidence of this mechanism operating effectively in the electricity industry to date. So-called green taxes and levies, where production of impacts (or their sources) is penalised financially may also help to reduce impacts.

With existing policies, regulations, and markets, there are ongoing environmental impacts as a result of electricity generation, which are neither controlled nor given due weight in the market. These are termed "externalities" by economists, since they are costs paid neither by the producers of electricity nor by those purchasing it, but by third parties; the environment and the users of the impacted environment. For example, the real cost of electricity from a power station includes the environmental costs of more global climate change, but this cost is not reflected in prices. The consumer, and others who do not use the electricity, will pay by receiving the impacts of increased climate change. It is a distinct disadvantage to have such hidden or unknown costs. One result is that it is impossible to establish whether there is any net benefit in having electricity at all; maybe the hidden costs outweigh the benefits.

The neo-classical environmental economics approach to removing environmental externalities involves valuing them and incorporating them into the production system, usually by Pigouvian taxes which increase the costs of electricity and so reduce demand. Theoretically, if those who get the benefits of electricity also pay the full costs of its production, the market will operate more correctly and efficiently. Recent governments in the UK and elsewhere have favoured (but not implemented wholesale as yet) the use of such market mechanisms to internalise externalities, to deal with the problem of environmental impacts falling on third parties.

Meanwhile, the impacts go on. The lack of solutions is undoubtedly partly because much of the scientific area is one of great controversy and debate. This is mostly healthy in itself, but it provides rather less unanimity and agreement over methods and results than is desirable for policy makers and regulators who need to justify changes in the legislative framework or in electricity costs. The recognition that some costs of electricity are not borne by producers or users implies that internalisation would lead to higher electricity costs. The fact that the costs incurred at present are unknown and fall indiscriminately has not yet tipped the balance sufficiently in favour of action. In a "balance-sheet" based society, swallowing the bitter pill of higher electricity costs in exchange for the long term cure of lower overall societal costs and a healthier environment remains elusive.

The pill would be much more appetising if it were known how much damage is being caused and exactly what would be gained from avoiding the current impacts. This means good information and agreement is needed over the value of impacts, which brings the focus of the debate back to the measurement problem. In the recent past, the dominant source of information about impact values has come from studies based on the theory and methods of neo-classical environmental economics. Invariably, the aim of these studies has been to put a monetary price on environmental impacts, as the means of establishing the appropriate level of subsequent green tax or other market-based internalisation mechanism. The general method is to quantify impacts, convert to monetary terms, and "incorporate" into the market system - by a market mechanism, which will allow environmental costs to be given due weight. It is appropriate to summarise some of the current problems experienced in this area, as a prelude to considering the options for progress.

Environmental economics has developed rapidly over the past two decades, and it continues to do so. However, one of the problems which remains is that the results of valuations differ from each other for a range of reasons, some of which are unconnected to the impact, the method used or the intended application. Also, clearly, the information about environmental impacts which is used in the valuation process can be a major factor in determining the valuations produced, and this information is not available in a standard, accurate and generally accepted form. Added to this is the fact that the environmental impacts in question are diverse in nature and complex in their interactions and effects, and their observation and prediction represents, in many cases, a relatively fledgling scientific area. Furthermore, they do not lend themselves to straightforward comparison. These latter points also apply to the few non-monetary impact assessment methods, which are invariably arbitrary and subjective in impact selection and measurement. Examples include "environmental footprints" (McLaren, 1996) and "sustainability indicators" (DETR, 1999, 2000).

While environmental economics provides theory and methods pertinent to the current debate, one of the foremost policy concepts is sustainable development. The two are intrinsically related. Notwithstanding the debates around the desirability of various shades of sustainability, there is no doubt that environmental impacts have ascended the societal agenda, and that there is a paradigm shift taking place in the way the environment is conceptualised by people, particularly in the western world. Table 1.1 illustrates some examples of the changes taking place. The integration of traditionally "environmental" concepts with social and economic ones such as "quality of life" and "valuation" is therefore timely because it fits within these current trends in environmental policy making and evolving societal attitudes.

The debate on the best way of measuring, valuing and allowing for environmental impacts in regulation is intense and will continue for the foreseeable future. However, it is well-established in the existing literature that impacts arising from electricity production and use are of a considerable order of magnitude and, therefore, the impact of ignoring environmental costs in assessing technologies and projects to provide electricity is substantial. While some of these impacts are taken into account through the existing regime of policies, controls and mechanisms, many others are not. Taking account of the hitherto ignored impacts can be expected to lead to changes in the energy resource mix, with consequent improvements in long term environmental and economic sustainability.

PAST		FUTURE
Natural environment as a free asset	→	Natural environment as valuable asset
Natural environment as outside economy	→	Natural environment as a tradable commodity
Pollution and damage of natural environment as a right	→	Pollution and damage of natural environment as a wrong
Pollution and damage of natural environment as outside economy	→	Pollution and damage of natural environment as a tradable commodity
Global awareness	→	Global and local action
Industry and society as dominating natural environment	→	Living within natural environment in a mutual relationship
Notion of pristine natural environment	→	Living with impacted and changed natural environment
Conservation and preservation of natural environment	→	Rights of individuals to access the natural environment
Conception of natural environment damage as inevitable	→	Conception of natural environment damage as requiring reduction/due weight
Government and society (other) has responsibility for the natural environment	→	Individuals and small groups must take responsibility for the natural environment

Table 1.1 Paradigm Shifts in the Environment-Human Relationship

Under the current situation, the environmental costs will be paid later, as environmental degradation worsens. However, this may turn out to be very expensive in the long run. It may be cheaper in both cash and quality of life terms to act now to minimise impacts such as global climate change, rather than wait until after the worst has already happened. The problem is that, at present, we do not know the costs involved and there is too little objective-based valuation data on which to make a decision.

1.2 Approach

The origins of this thesis topic are closely associated with the emergence and development of the concepts introduced above. Clearly, the relationship between economic activity, humans and the environment is central, whether it is "sustainable development", "environmental economics", or any other mechanism which aims to combine these elements. The approach taken here will be informed by examination of such concepts, and their strengths and weaknesses.

In terms of human activities, the main focus here is one industrial sector, the electricity supply industry in England and Wales (hereafter termed the ESI). The sectoral approach and geographical limit is a constraint of scope, without constraining the underlying topic, that of regulating for impacts. The ESI is a particularly appropriate sector to study, not only because electricity provides a fundamental service, but also because the impacts arising from it are highly debated, and potentially large in significance and range of types. As a result, there is a lot of research activity in this area. Although much debate is also taking place with regard to, for example, the water sector, transport and petrochemicals, electricity remains amongst the most prominent topics with the greatest range of problems.

The logical starting point is not the "regulating", but the "impacts". Why is the ESI causing environmental impacts? The short answer is that all activities cause impacts. In a market system, unregulated impacts are a "free asset", so they are inevitable, as competitors squeeze real costs and neglect those which they do not have to pay for directly. A number of mechanisms exist to control them, such as the planning system, legislation, and regulations. However, some impacts may not be given appropriate weight or are ignored, principally because the regulatory system only takes into account some impacts. One task is therefore to assess the extent to which the ESI currently maximises economic, social and environmental efficiency. An examination of the problems inherent in attempts to achieve this is therefore required. It is the various shortcomings of the current situation which must form the criteria to be met by any improved approach or methods. The next task is to develop methods which are underpinned by theory, but which also allow an improved approach to be achieved in practice. Methods and mechanisms for dealing with environmental impacts will be considered within the assumption that the basic privatised structure of the ESI and its market framework is to be retained.

One of the key problems with dealing appropriately with environmental impacts is timing. In particular, gaps in knowledge often exist at the point when decisions should be made about what level of impact is "acceptable". In other words, if prevention is generally better than cure, then the best time to take avoiding action is before the problem is fully understood, notwithstanding that this is counter-intuitive to knowing the best action to take. This conundrum is very real, and accepting and dealing with gaps in knowledge is important to any approach to dealing with environmental impacts. Indeed, potentially significant issues requiring the most urgent internalisation are poorly understood and thus impossible to value accurately. To calculate an effect in a complex area such as human health or the environment, modern scientific method requires that it is observed first. In potentially non-reversible situations, such as global climate change, this is undesirable and directly challenges sustainability obligations. These include the need to exercise precaution by avoiding actions which could lead to uncertain and/or potentially large or irreversible environmental impacts.

There are many and varied problems associated with measuring environmental impacts. When faced with the choice of measuring the measurable and the hitherto unmeasurable, the choice is often made to pursue the former. However, measurability is not related to importance and thus, "analysts confuse things that are countable with the things that count" (Holdren, 1982). Since many environmental impacts have long been considered "intangibles", they have not been measured properly or given sufficient weight in decision making. Such intangibles have traditionally been ignored altogether, so making it impossible to achieve any relative comparison between included and excluded elements. Although knowledge moves on, so the intangible becomes tangible, it does not keep pace with damage. By the time today's intangibles become tomorrow's tangibles, it may be too late to stop irreversible environmental damage.

Even where knowledge is substantially complete, a rational method of assessing values of environmental impacts has long been considered by many to be impossible to achieve. People have various agendas, value things differently because of their interests, and generally are subjective in applying such values. A solution of the neo-classical environmental economics approach is to apply the Contingent Valuation Method (CVM), where sample groups are asked to express various preferences via questionnaire. However, CVM studies have given varying results, for example, according to the level of knowledge of the participants and how much information they are given. This is not surprising. People are rational, but make irrational decisions in the face of poor or inappropriate information.

Despite the apparent problems with environmental economics, a review of the approach is a good starting point for the following reasons:

- The overall approach assists the process of evaluating environmental impacts by bringing them closer to the benefits. Thus, rather than treating electricity and the environment as disparate, it forces a trade-off mentality to be pursued;

- Although the methods are often problematic, they have been applied and there is a wide literature on environmental economics techniques and their application.

Human values are one of the oldest concerns of civilisation. Recorded attempts to quantify values are, however, few until comparatively recently. Nevertheless, for well over 100 years, western nations have collected aggregate data on health and social conditions. In the last few decades, measures of individual health and well-being have been developed, invariably based on empirical observations through questionnaire or interview. While the data have often become the focus for social or health care reform, there has generally remained a significant gap between enquiry into values, methods to quantify them, and action based on them.

One of the apparent paradoxes of human value systems is illustrated by the example of pro-development versus anti-development views, where the former is seen as seeking change, newness, experimentation (along the lines of "variety is the spice of life", and risks need to be taken to produce the best outcomes), and the latter is seen as seeking security, conservation and preservation of the known present, with routine, familiarity and risk minimisation being the primary purpose. In reality, humans apparently need both routine and variety, and they tend to identify with both, but to varying extents at different times and over different issues. The anti-developer may have risky pastimes. They may also be opposed to development for reasons other than (or in addition to) any judgement on the basis of values or overall quality of life benefit (either personal or societal). The only way to overcome such an apparent paradox is to establish a framework within which subjective judgement is separated and isolated from objective measurement, which itself is broken down into its constituent parts. This is the pre-requisite for minimising bias and for weighting differing impacts correctly. Only then can the objective measurement of subjective phenomena become possible.

Turning to the regulatory issues, as mentioned above, a regulatory regime already exists for dealing with environmental impacts in the ESI. It is important to summarise how this currently performs, as a starting point for considering appropriate future regulations. There are "good" and "bad" approaches to regulation, and not all the current problems in regulating for environmental impacts can be attributed to lack of good information for regulators. While legislation in the early 1950s in the UK proved effective by banning the use of fuels which produced smoke in urban areas, it was retrospective, and thus a lot of suffering was caused before it was implemented. In environmental terms, the judgement that smoke from burning fuels in these areas created much higher costs than the benefits (in terms of heat and cooking facilities) was implicit in the legislation. Alternatives had to be found. Unfortunately, many smokeless fuels have since been found to cause serious environmental damage in their own right, for example, long term pollution from coking and smokeless fuel plants. Thus a major problem with such legislation is that it is reactive; it waits for a problem to occur and be identified, and then produces rules to deal with it (with a time delay which is variable depending upon the size and power of business interests, the awareness of the problem, political will, and other factors). This strategy is weak and dangerous – particularly for large, irreversible environmental impacts.

Neither "command and control" legislation nor the planning system are sophisticated enough in their current form to deal appropriately with the myriad of widely varying environmental impacts associated with the ESI. While attempts are made by policy makers to regulate or reduce some of the most severe and obvious environmental impacts which result, or which can be anticipated, the remaining impacts continue to occur. While many of these may be apparently small in initial impact, taken together, they may be more significant than one or more of the impacts which are regulated. Also, given the nature of the environment, more complex temporal/spatial aspects of impacts may be unpredicted, or unforeseen.

In short, current problems with regulating for environmental impacts have their origin in uncertainty over information, as well as the appropriate mechanism for dealing with them. Therefore, although the concern here is ostensibly about environmental regulation, a major part of the problem is in the underlying information and approach to measurement and valuation – the preparation for regulation. Consequently, the main thrust of the research here will address this issue. The approach adopted will be to review the current situation, followed by development and testing of revisions and new approaches. The direction is towards achieving a transparent, consistent and rational approach to the data collection, identification and valuation of environmental impacts. The destination is a systematic framework, tool or set of tools for undertaking these tasks for fuel cycles in the ESI.

1.3 Aims and Objectives

The aims of this thesis are to establish where weaknesses in the current regulations for environmental impacts associated with the ESI originate, and to develop a framework for producing improved information about environmental impacts, the values attached to them, and appropriate regulation to reflect these values in the ESI.

In achieving these aims, four objectives need to be met, as follows:

- Critically review the existing structure and regulation of the ESI and the current techniques used in environmental impact measurement and valuation, including their weaknesses and potential;

- Develop a workable, practical method for providing the necessary data required to identify and measure environmental impacts;

- Develop a workable, practical method for valuing environmental impacts;

- Develop and present an appropriate regulatory option for application of the impact valuation results.

1.4 Structure

The aims and objectives dictate, to a large, extent the appropriate structure of this thesis. First, a review of the history and present structure of the ESI is necessary, and this is the subject of Chapter 2. This includes setting out definitions of terms and subject/production system boundaries, and providing an overview of the general structure of the industry and the environmental impacts to which it gives rise. A review of the current environmental regulatory framework as it relates to the ESI is also necessary, and this is undertaken in Chapter 3. In particular, this examines the extent to which the current regulatory framework successfully regulates for environmental impacts. In Chapter 4, a critical review is conducted of the theoretical basis of environmental economics. This is necessary as neo-classical environmental economics is the current dominant approach to valuing environmental impacts. A critical assessment of the practical problems is also important, to inform improvements and alternatives designed to complement current practice. This is addressed in Chapter 5.

In Chapter 6, new methods designed to overcome the problems identified in Chapters 1 to 5 are outlined. Five major requirements are addressed, in order to accomplish each of the following tasks; identifying the production outputs which cause impacts, identifying the pathways of these through the environment, setting up a single scale on which the significance of each impact can be assessed; valuing the resultant environmental impacts, and; determining the most appropriate way of applying the values in practice. These major elements are developed in detail in Chapters 7 to 11 respectively. In Chapter 12, these proposed new methods are compared with existing methods, studies and practice. Finally, Chapter 13 presents conclusions and recommendations for further work.

1.5 Definitions and Scope

It is appropriate to clarify the terms ESI, regulation and environmental impacts at this point. The ESI in general means the industry which is responsible for the generation and distribution of electricity. However, for the purposes of this thesis, the definition is further constrained. Firstly, as already mentioned, it is restricted to the industry in England and Wales and, secondly, it is focused mainly upon the generation sector of the industry. Regulation as a term is used throughout to refer to the adoption of any tools which allow the application of values for environmental impacts so that they are reflected in the practice of electricity provision. In other words, it may include any mechanism, law, statute, directive, standard or fiscal instrument which could conceivably be used to implement the policy. The policy in question is assumed throughout to be the intention to achieve the optimal situation of minimised environmental impact per unit of (economically viable) electricity provided. Principles which underpin this policy, such as polluter pays, precautionary principle, or the principle of sustainable development, are outside the scope of the research topic. Background information on broad policy goals, such as sustainable development and the precautionary principle, can be found elsewhere (for example, Baker et al, 1997, O'Riordan and Cameron, 1994).

While "environmental effect" is used as a general term to mean any change on the environment irrespective of magnitude or whether it is beneficial or detrimental, an environmental impact is specifically a disruptive influence on the physical or social environment. This is in contrast to an environmental cost which is a measure of the consequent response of the environment to an impact or impacts. Therefore, cost has a (monetary) value, and is analogous to damage. Environmental impacts are occurrences which are associated with measurable, harmful changes in the environment, and which are caused by human activity. One of the most striking features of environmental impacts is their diverse nature, in terms of such attributes as; magnitude, impact area, temporal range, uncertainty, risk, synergistic tendency, resource depletion/sensitivity aspects, reversibility, ecosystem threshold, health/emotional effects, user/amenity losses, range of recipients and public perception. By way of illustration, the environmental impacts of non-renewable energy include emissions of particulates, heavy metals, carbon monoxide, hydrocarbons, components of acidic deposition and toxic air contaminants; oil and gas supply disruptions and accidents, coal mining and transport accidents and ash disposal; uranium mining injuries, nuclear waste disposal, radiation releases, reactor accidents, decommissioning costs and nuclear weapons proliferation; thermal pollution, global climate change; human anxiety and political conflict, and others (Holdren, 1987).

The complexity of impacts also requires a well-defined approach to avoid missing out or double counting individual impacts. Political, personal and other value judgements are involved in any evaluation technique, however quantitative or rational the approach appears to be. For example, one may differentiate between the origins, character (or type), costs, and indices, criteria and methodology (by which costs are measured) of impacts (Budnitz and Holdren, 1976). The fact that environmental impacts include such a wide range of phenomena and create widely varying effects is at the centre of the problem of how to best go about measuring them and dealing with them.

No precise, universal definition of environmental impact exists, and the term is used to mean a range of different effects, including those that might be called "environmental changes" and the resulting effects on the environment, life and/or humans. There are variations in use, for example, concerning whether environmental impacts are caused *by* changes in the environment or whether they *are* changes in the environment. Although this may at first appear to be a semantic issue, the lack of precision is a potential source of lack of clarity. Therefore, a more precise definition will be developed in the course of this thesis. For the present, the working definition given above will suffice. Finally, clearly, there is an overlap between social (for example, equity), economic, political (for example, security) and environmental issues, but the definition adopted here focuses primarily on the environment. Thus, socio-economic effects are not considered as environmental impacts, whereas it does include the potential loss of some far off environment which may never be "used".

2. THE ELECTRICITY SUPPLY INDUSTRY

The electricity supply industry in the UK comprises three major elements; generation, transmission and distribution, each of which form more or less discrete sectors. This layered form of organisation dates back to the structure adopted on nationalisation in 1947 (see Section 2.1). The only exceptions are in Scotland and Northern Ireland, where the operations in a particular area have tended to be controlled entirely by one organisation responsible for generation, transmission and distribution. The industry in England and Wales constitutes the overriding majority of the UK electricity industry in terms of both generation and use. As mentioned in Chapter 1, this is the main focus of attention here, and the term ESI refers to the industry in England and Wales unless otherwise qualified.

The aim of this Chapter is to establish familiarity with the ESI, by presenting a review of its historical development, current structure and recent developments up to 1999. This is partly drawn from an earlier review elsewhere (Horne, 1994), and it should be noted that further developments are ongoing which are outside the scope of this historical-based review. This Chapter is designed to provide the general background and the context within which environmental issues have hitherto been dealt with. Hence, it is a prelude to Chapter 3, which will contain a more specific review of the environmental regulatory framework of the industry.

Section 2.1 outlines the early development of the ESI, up to and including Nationalisation in 1947. Section 2.2 charts the influence of nationalisation in the rapid developments within the industry during the 1950s and 1960s, and Section 2.3 considers developments in the fuel mix in the 1960s and 1970s. Section 2.4 examines the attempts to curb the power of the industry and its monopolistic tendencies during the 1970s and early 1980s. Section 2.5 deals with the main elements of privatisation, including the Electricity Act 1989, Section 2.6 examines the current situation, and Section 2.7 draws some comparisons between the ESI in England and Wales and elsewhere. Section 2.8 looks at current environmental issues within the context of the privatised market and its regulation, which leads then into conclusions (Section 2.9) and the next Chapter.

2.1 The ESI up to Nationalisation

The development of technology to allow the bulk transmission of electricity is relatively new. Rapid progress was made between 1878, when electricity use was confined mainly to a few small lighting installations, and 1926, when the Electricity (Supply) Act introduced the first effective national co-ordination of electricity undertakings. This 48 year period saw several Acts in Britain regarding the emerging electricity industry. These included the Electric Lighting Act (ELA) 1882, which first authorised electricity supply; the ELA 1888, which further encouraged the establishment of electricity supply undertakings; the ELA 1909, which promoted, amongst other things, the co-ordination of bulk electricity supply between two or more local authorities; and the Electricity (Supply) Act 1919, which introduced central co-ordination through the establishment of the Electricity Commissioners as the official body (Hannah, 1982). The same period also saw the development of many basic electricity generation and transmission systems and technologies for electrical appliances which are still used today.

However, it was the 1926 Act which allowed the consolidation of these first decades of rapid but essentially piecemeal development in the fledgling electricity industry. Specifically, it set up the Central Electricity Board (CEB) to bring together the main generating stations by interconnecting them with a high tension main transmission system - a national grid. The creation of the CEB represented the first attempt at centralised control in the expanding ESI. It also brought the standardisation of AC frequency and led to a major reduction in generating costs through rationalisation of the existing capacity. By 1938, spare generating capacity was reduced from 80% to 15%, saving the majority of the capital cost of grid construction in the process.

The rapid growth in load in the years immediately prior to the Second World War necessitated correspondingly rapid structural expansion of operations and this, in turn, brought wide debate over the pros and cons of further increases in government control. Following the outbreak of hostilities, the Electricity Commissioners and the CEB agreed on an emergency programme of generating plant construction to provide a further 180MW of capacity in six 30MW units. Developments continued throughout the war, latterly (after 1942) overseen by the newly formed Ministry of Fuel and Power.

The election of the new Labour administration in 1945 made the nationalisation of the ESI inevitable. The Labour Party had been committed to a policy of relatively widespread nationalisation for several years, and the ESI was a high priority, although not as high as the coal industry, which was nationalised almost immediately. The Electricity Act 1947 (effected on Vesting Day, 1st April 1948) duly nationalised nearly 600 undertakings, which comprised the ESI in England, Wales and Southern Scotland and established 14 Area Boards for distribution purposes. The British Electricity Authority (BEA) was the new central authority, set up to handle both generation and transmission, including co-ordination and policy formulation. The Act was followed by nationalisation of the gas industry the following year.

The chairman of the BEA was Lord Citrine, former General Secretary of the Trades Union Congress, and his deputy was Sir Henry Self, a senior civil servant. Bearing in mind the formidable problems facing the industry at the time, the transition to the nationalised industry was smooth, largely as a result of the retention of many structures and senior posts from the old organisations. Two fundamental priorities facing the new nationalised industry from the outset were the need to ensure right of supply and security of supply. The new nationalised ESI inherited a situation where a fundamental shortage in power generation capacity had created major power failures in the first months of 1947, during the worst winter for over 100 years. Other major problems at the outset included variations in supply tariffs and the need to connect remaining domestic properties still unconnected with the grid. The BEA immediately introduced a bulk supply tariff system designed to bring some uniformity to pricing levels and further standardisation followed to this end over the next decade.

At the time of nationalisation, a quarter of households were still without an electricity supply. These were mostly located in the urban slum areas, where housing was too dilapidated to be worth connecting or landlords were unwilling to make the necessary investment. Although large areas of outlying rural districts were also unconnected, the overall number of rural dwellings was relatively low. The rate of new connections was considerable, partly associated with the accelerated housebuilding programme; 2.7 million new houses were completed in the first decade after ESI nationalisation, the vast majority of which were fitted with standard connections to the grid.

2.2 After Nationalisation

The highest priority after nationalisation was to tackle the capacity shortfall, and the BEA implemented an accelerated construction programme utilising reliable technology, and established manufacturing and construction firms. Periodic power cuts during times of peak demand was one of the main reasons for the focus of attention on generation. The major efforts made to close the gap between supply and demand in the first decade of nationalisation left its mark on the organisational structure of the industry. The balance between the need for speed and reliability on one hand, and efficiency and economy on the other was weighted somewhat in favour of the former, as the adoption of new larger unit technology in the UK lagged behind that in other countries such as the US. A critical factor later suggested in the retention of the conservative attitude for so long was the lack of sensitivity of BEA policies to price signals (Hannah, 1982).

By the early 1950s, the substantial capital programmes for generation and grid connections led to the BEA and the Area Boards being openly attacked over pricing and aggressive sales policies respectively. As a commentator of the time wrote: "The electricity industry has completely failed to set up a pricing system which will enable it to produce amounts of energy which are in the national interest. It is only interested in selling as much as possible, consistent with covering total costs, and not in the least in selling the right amounts" (Little, 1953). Early attempts at achieving cost control led eventually to a major shift towards "efficiency", the contemporary wisdom being that this lay in ever larger generating units.

With the subsequent growth in sizes of generating sets from 30MW through 60MW, 120MW, 200MW, 500MW, 660MW and later to 1000MW units, fewer units were required and so there were fewer construction sites (Openshaw Taylor and Boal, 1969). Thus, there was a need to reduce the number of divisions responsible for generation. This was achieved through rationalising the organisation of the 14 divisions (which followed the boundaries of the Area Boards responsible for distribution) into 3 project groups, under the 1957 Electricity Act. Indeed, the 1957 Act caused a considerable shake-up in the industry, with increased decentralisation, as greater financial responsibilities were devolved to the 12 Area Boards in England and Wales. With the expansion of the nuclear programme, there was a perceived need to separate generation from the central organisation to facilitate development of this new potential generator. Thus, the CEA was split into two new organisations; the Central Electricity Generating Board (CEGB), responsible for national planning and operation of generation, grid transmission and bulk sales to Area Boards, and the Electricity Council, a forum for policy and review consisting of representatives of the Area Boards and the CEGB. The CEGB was divided into 5 regions to assist delegation of duties in further attempts to devolve power from the centre. Rationalisation of plant design and construction activities also took place as three project groups were established for the purpose, replacing the arrangement whereby areas sharing boundaries with the Area Boards performed these tasks (although some rationalisation had already taken place since this initial arrangement had been established on nationalisation).

The Electricity Council was given statutory duties to ensure efficiency of supplies. Other functions included raising capital on behalf of the Area Boards, negotiating wage settlements and controlling national research and publicity. Despite this apparently key role, the Electricity Council became largely stripped of power and resources, as the Area Boards exercised their determination to limit central influence on their activities. The CEGB was the dominant organisation in the re-organised industry and, although changes in top management were made, it set out with substantially the same senior staff as previously, in the CEA. While the clear message from the Conservative government to the restructured industry and its new leaders was for change, the latter's new approaches were soon challenged both by inertia from within the industry and the realisation that the government and other critics had had somewhat unrealistic expectations from the reorganisation.

The constant political efforts to force rationalisation had, by the early 1960s, led to an embracing of very large unit technology, which took the industry somewhat to the other extreme. For example, the two 550MW units commissioned at Thorpe Marsh near Doncaster, in 1963, were the largest in Europe at the time. The inclusion of numerous very large units in the supply mix began to cause problems over the need to maintain a considerable capacity in reserve to allow for plant breakdowns in these large units and thereby ensure security of supply. The use of smaller, older and much less efficient units to provide the required reserve capacity was costly in both maintenance and operation. Also, despite the rationalisation and the rapid expansion of large plant, capacity shortages continued up to 1972. This was largely due to delays in construction and the poor initial performance of the new large generating sets. The miners' strike of 1971-2 and a 40% rise in coal prices in the period 1969-71 brought the fear of over-dependence on home coal supplies to the fore and a resultant further impetus to the shift in favour of oil which had begun in the 1950s. A fall-off in electricity demand marked the recession which followed the 1973-4 oil crisis and, thereafter, over-capacity rather than shortage became the order of the day.

2.3 Developments in the Fuel Mix

While the 1960s and 1970s brought rapid development in the domestic and industrial energy mix, the concurrent changes in the fuel mix for electricity generation were even more substantial. A White Paper was published in 1955 (DOE, 1955) outlining this initial programme which was based on the Magnox reactor technology. Largely as a result of this background, the nuclear power industry developed separately from other fuels, despite it being nominally run from within the core of operations of the CEGB.

Technological advances in petroleum exploitation made oil a relatively cheap and plentiful fuel during the 1950s. This had limited impact on the organisational structure of the CEGB, but considerable effects on the British coal industry, as oil-fired plants were built in place of coal-fired plants. An initial indication that the cheap oil bonanza was not infinite was the Suez Crisis in 1956, and this was followed by the formation of the Organisation of Petroleum Exporting Countries (OPEC) in 1960, with the 5 initial members Iran, Iraq, Kuwait, Saudi Arabia and Venezuela. However, it was not until the OPEC oil embargoes and consequent price rises which prompted the world oil crisis of

1973-4 that the CEGB's policy to broaden the fuel mix by expanding oil was brought into serious question. Apart from the increase in transport requiring petroleum fuels, the ESI was the major user of oil in Britain by 1973, and the whole European economy was becoming increasingly dependent upon this fuel source.

From 1965 onwards, gas supplies began to be derived from oil instead of coal, although this was to be a short-lived practice, as natural gas started making major inroads into the energy market from 1968, particularly in the domestic space heating sector. This market transformation was extremely rapid, with the essential precipitating event being the discoveries of the extensive natural gas reserves in the North Sea. The technology to search for and exploit these resources was relatively new, but the speed at which supplies were established and expanded is indicative of the demand for this highly useful and transportable energy resource. Also, although the most easily exploitable North Sea reserves may be depleted in the short term, further reserves exist to supply domestic and industrial demand for several decades (Stern, 1990).

The knowledge that the large long term reserves exist post-dates considerably the establishment of the industry in the North Sea. As a consequence, the first two decades of gas supply were accompanied by a general feeling of uncertainty over long term security of supply and therefore an unwillingness to put these precious resources to use for power generation. This picture has now changed since the realisation of the size of North Sea reserves in the 1980s, and the short lead in time for gas-fired plant commissioning has combined with various economic and environmental policy factors to put gas in its current position as a major element of the fuel mix for power generation, with consequent implications for other fuels. The impact of EU Policy in this change is critical, since it has only become possible following the relaxation of EU Directives governing permitted uses of the North Sea reserves.

2.4 The Long Run-up to Privatisation

Although structural and tariff changes continued to be introduced throughout the 1960s and afterwards, the effects on overall viability of the industry were limited, and concern that the industry was a drain on public expenditure and dominated by an almost uncontrollably large institution continued, particularly amongst Conservative politicians. The world economic slump and electricity overcapacity which followed the 1973-4 oil crisis exposed more than ever the consistently over-optimistic predictions of demand in the late 1960s and throughout the 1970s (DTI, 1961). The rise of natural gas as a major and very competitive energy source in domestic heating only served to underline this (Papadopoulos, 1981). Regular government overestimates of GDP also affected CEGB forecasts. Furthermore, the contraction of the economy and later cutbacks due to oil price shocks were not the only factors reducing overall energy demand in the mid 1970s. An emphasis on the need for energy conservation was rapidly developing at the time and was further boosted following the crisis. The "Save It" campaign was the most concentrated publicity campaign by a British government at that time, and the biggest energy conservation campaign before or since (Chapman, 1974, Chapman, Leach and Slesser, 1974). Unfortunately, this was largely a panic measure. It was thus ill-conceived, and has left energy conservation with a lingering negative public image. For many, energy efficiency still conjures an image close to "freezing in the dark".

While the difficulties of demand forecasting and planning in an industry which was dealing with very large generating sets with long lead times cannot be overemphasised, the CEGB was undoubtedly slow to react to the changing situation. This was partly a result of the long held view within the industry that electricity sales were of paramount importance and that they had increased at such high rates in the past that it was difficult to contemplate long term reductions in demand. That this view was able to prevail for so long was partly due to the structure of the supply industry as a whole.

The Conservative government was elected in 1979 on a privatisation platform and although the ESI was not first on the list, it was considered important, partly due to the sheer size of the CEGB. As all the public utilities could not be privatised immediately, one of the first pieces of legislation introduced was the Competition Act 1980, which allowed nationalised industries to be investigated by the Monopolies and Mergers Commission. This was followed in 1981 by further rationalisation of the CEGB structure and an accelerated power station closure programme in response to the recession. Subsequently, the Energy Act 1983 encouraged private enterprise to become involved in electricity generation, the sector with the best prospects for introducing real competition and also the greatest capital requirements.

The coal strike of 1984-5 was a major threat to the government and was treated as such. The long and bitter dispute cost the nation over 1% of GDP in 1984 and the eventual victory for the government was followed by an extraordinary acceleration of the already substantial coal mine closure programme introduced in 1983, which had precipitated the dispute. Increasing pressure on the CEGB to turn away from British coal as the dominant fuel source for generation then followed, and the election of the Conservative government for a third term in 1987 made major structural changes to the ESI, in the form of privatisation, inevitable.

2.5 Privatisation

Ostensibly, the main aim of privatisation of the ESI was to increase efficiency by bringing competition to the industry. Upon nationalisation, it had become the largest ESI in the world. By the 1980s, it had endured decades of bad press about it being bureaucratic and inefficient, especially from Conservative administrations. The CEGB was large enough to single-handedly affect the Public Sector Borrowing Requirement. Implicit in privatisation was the idea that the bureaucracy would be shattered and costs would be driven down sharply, with inevitable effects on the estimated 550,000 people employed in the industry in 1987 (Gladstone and Dewhurst, 1988). The apparent inconsistency with the creation of several generating companies, each with its own administration, out of a single, albeit large one was passed over.

The White Paper spelling out the form which privatisation of the ESI would take was published in February 1988 (DOE, 1988). In introducing it to the House of Commons, the Secretary of State for Energy, Cecil Parkinson, outlined the six principles adopted in putting forward the proposals:

- Decisions about the supply of electricity should be driven by the needs of customers;

- Competition is the best guarantee of the customers' interests;

- Regulation should be designed to promote competition, oversee prices and protect the customers' interests in areas where natural monopoly will remain;

- Security and safety of supply must be maintained;

- Customers should be given new rights, not just safeguards, and;

- All who work in the industry should be offered a stake in their future, new career opportunities and the freedom to manage their commercial affairs without interference from the government.

It was acknowledged at the time that the distribution and transmission of electricity are largely natural monopolies and hence generation was the only sector capable of being made competitive. The number of companies to be formed to share the CEGB's generating capacity was an issue, however, with the Centre for Policy Studies arguing for 10 new companies to be formed (Henney, 1987). In the event, just 2 public limited companies were formed, PowerGen and National Power. The ending of the effective monopoly operated by the CEGB was focused on as one of three conditions which must be met if competition was to develop in the sector, along with ending its obligation to provide bulk supplies and transferring ownership and control of the national grid to the distribution side of the industry.

The Electricity Act 1989, was both complex and delayed, largely due to the nature of the industry, problems with the government's privatisation programme following the stock market crash of October 1987, and the withdrawal of the nuclear industry from the flotation at a late stage. It was originally intended that National Power would operate all nuclear plant formerly operated by the CEGB. The late decision to retain the nuclear industry in the public sector resulted directly from concerns of the new management of National Power, upon examining the financial and liabilities' position of the industry. The Area Boards flotations were more straightforward, and these took place in late 1990, forming the Regional Electricity Companies (RECs).

2.6 The Privatised ESI

As with the rest of the privatisation programme, there were two elements to the flotation of the UK ESI; political and economic. The aim was to increase the role of market forces by restructuring and the freeing of access to the industry (Beesley, 1992). The political agenda was to break the CEGB, the monopolist whose market power allowed it to pursue its own priorities. The fact is that natural monopoly elements still exist. In fact, in common with other privatisations, the initial flotation involved the creation of companies with a total or virtual monopoly on a national or regional level. The Office of Electricity Regulation (OFFER, now Office of gas and Electricity Markets, OFGEM) was therefore created as an important structural element, with responsibilities to set down and ensure adherence to the regulatory framework. It was considered that, as new players entered the generation market, the regulatory burden in this sector would diminish.

In the first few years after privatisation, the reorganised ESI in England and Wales followed a generally similar structure to that which prevailed under the former nationalised regime, with sectors for generation, transmission and distribution separated into different companies as follows:

- Generators, comprising three main companies (National Power, with 70% of non-nuclear capacity, PowerGen and Nuclear Electric) and the smaller independents;

- National Grid Company plc (NGC), which operates the transmission system, the national grid;

- Public Electricity Suppliers, consisting of the distribution companies (the Regional Electricity Companies or RECs, formed out of the old Area Boards), which have their own local distribution systems (up to 132kV).

It was in generation that the early impact of privatisation was felt, firstly by the formation of the big three companies, and secondly, by the rapid development of further competition. This sector was always considered to be the most suitable for creating a competitive market, and so it proved. The basic arrangement is that large generators require a generation licence to supply electricity to the grid (a second tier licence is required to supply direct to a consumer). The generators sell electricity to the RECs under agreed supply contracts. Each contract essentially contains two elements; a capacity charge, covering capital costs and some variable costs and payable irrespective of electricity generated, and an energy charge, covering remaining variable costs, principally fuel. Although the actual costing and pricing methods are complex, the arrangements between generators and distributors are based on a simple mechanism, called the electricity pool. This pool is hypothetical and is the market place, so that generators sell their electricity to the pool and distributors buy from it. Pool prices are set on a continuing basis and generators state their plant availability, capacity and prices from which the merit order is determined. Further detail of how the pool works is not relevant here, and for an erudite yet simple explanation, the reader is referred elsewhere (Patterson, 1999).

The first years of pool operation saw considerable volatility and unpredictability in the half-hourly prices, with occasional "spikes" where prices rose suddenly for short periods of time. This was apparently exacerbated by National Power and PowerGen planning withdrawals and subsequent reinstatements of plant availability, aimed at maximising revenue by increasing pool prices through reductions in supply availability (OFFER, 1991). Such practices subsequently led to regulatory changes, principally to give OFGEM wider powers to reduce the dominance of the two main generators (Energy Committee, 1992).

While new competition has grown in generation, it should be noted that a well-established independent production sector with its own trade association, the Association of Independent Electricity Producers (AIEP), was already supplying 7% of UK electricity in 1991 (DTI, 1992). However, the old regulatory and economic regime had favoured the large CEGB generating units and thus discriminated against smaller independent generators. Since privatisation, a rather complex set of arrangements for small generators has been implemented in an attempt to address this. Essentially, smaller generators, providing less than 10MW, are likely to be Embedded Generators, meaning they are connected directly to the 132kV distribution network rather than the NGC system. They need a Connection Agreement which allows the necessary connection to and use of the distribution network and requires compliance with the Distribution Code governing the distribution system operated by the RECs. Above the 10MW threshold, generators require a Use of System Agreement with the NGC and may opt to be subject to central despatch, so that they provide electricity to the pool. Wherever the generator is situated, the structural implications of a merit order of generating capacity, mean that large stations with lower operating costs invariably supply base load capacity, that is, the off-peak demand level which is continuous and therefore ensures highest possible plant utilisation for the generator. This is inevitable as the merit order requires that power is supplied strictly on the basis of cheapest unit cost first, followed by progressively higher cost electricity until demand is met. There are therefore problems in financing and economically operating smaller independent plants which cannot rely upon high utilisation (continuous base load utilisation).

In addition to the established independent sector, around 5% of demand is met by private generators, which are invariably large industrial organisations. The RECs are also permitted to generate up to 15% of their own distribution requirements. Although this generation source has materialised, the RECs generally face similar problems to the independent generators in developing economical generating capacity, given the relatively small size of each REC compared to the big generators. This is one of the contributory factors in the recent and ongoing active speculation, re-organisation and acquisition activity in the distribution sector.

The RECs also originally owned the NGC, responsible for operating, maintaining and developing the predominantly 275 and 400kV national grid, although this has since been sold on. The grid consists of over 7000km of overhead lines, 500km of underground cable and complex computer systems for calculation of payments due following daily trading in the electricity pool. The NGC is not permitted to participate directly in generating or supplying electricity with the exception of the pumped storage capacity at Dinorwig and Ffestiniog, which it inherited from the CEGB. The tight controls on charges and operations include a requirement to operate a similar strict merit order to that operated under the CEGB. Anyone wishing to operate in the electricity market must be given access, subject to technical constraints.

The core business of the RECs was and remains their statutory duty to supply electricity. They can make two main types of charges, namely, distribution and supply, and both are strictly regulated. A second tier supply licence is required by any utility supplying electricity under contract. This may be a generator, supplier or distributor and may be a new or ex-nationalised company. The only exception to the need for a second tier licence is in the case of RECs supplying electricity under contract to customers within their own established areas. A second tier supplier can use the distribution network operated by the REC in the area, or may erect alternative lines. In the latter case use of these lines must be allowed under similar terms to those pertaining to the REC.

2.7 Comparisons

While the UK ESI privatisation is generally considered to be a forerunner of the liberalisation of energy markets which has taken place worldwide in the 1990s, it cannot be considered to be a model which has been copied rigorously. Indeed, a remarkable variety of organisational forms now exist, including retained state monopoly in France, mixed ownership in Scandinavia and strong regional interests in Germany. Britain liberalised by privatisation, whereas Norway maintained dominant public ownership and sought to create competition through a decentralised production structure (see, for example, Midttun and Thomas, 1998). In some systems vertical integration has been retained, while in others it either was not a strong feature or has been deliberately disaggregated. The origins of such variety are split between historical factors and varying motivations in the liberalisation process (Gilbert and Kahn, 1996).

One of the only features which could be said to be relatively common amongst ESIs in different states is the presence of one or a few dominant large organisations. Even this may change in the future, particularly in the EU, where an increasingly integrated electricity market may generate more cross-border trade in electricity. However, it should be noted that this is relatively minor in the case of the UK. Of more relevance is the effect of restructuring on ownership, market share and the potential for stranded assets to multiply, as existing companies with large, expensive plant are outcompeted before construction costs are covered. The potential security and responsibility issues arising from such rapid change are outside the current research topic but they clearly exist alongside the problems of environmental responsibility. Meanwhile, the market dictates that change continues apace in the struggle for market share. Several dozen large US and European power stations have changed hands since the mid-1990s. In 1995, Eastern Electricity acquired four large coal-fired plants from National Power, in an attempt to construct a more vertical electric utility. The regulator approved the sale but argued against allowing the large generators to acquire RECs. However, in general, the UK approach has been more tolerant of monopolistic tendencies in the privatised ESI than many other countries, where liberalisation has been more designed to bring about a competitive structure in the neo-classical sense, rather than simply private ownership.

2.8 Environmental Issues

The first major environmental problem recognised and acted on in relation to the ESI was particulate pollution from power station emissions. The adoption of standards by the CEGB followed in the 1950s, which involved the fitting of filters to its coal-fired stations. The second major atmospheric pollution issue was acid precipitation, and controls on sulphur dioxide, nitrous oxides followed in the 1980s and early 1990s. The current major issue is global climate change; and in particular, the need to control and reduce carbon dioxide emissions. However, for the most part, the privatisation process did not specifically address the environment, and this was therefore left to the European Union and separate environmental legislation to deal with (see Chapter 3). Moreover, targets on atmospheric pollution have generally been met to date, but largely due to coal plant closures apparently caused by market forces, rather than any planned strategy (EC, 1994). At the core of the problem is the government's conflict between environmental responsibility and those to other, more powerful agencies. Despite a shift in general attitude towards the environmental agenda between the 1980s and the 1990s, environmental responsibility has been at best a subordinate motivation in the government's overseeing of the decline of the coal and nuclear industries (Eikeland, 1998).

OFGEM, the industry regulator, is concerned primarily with price controls, financial practices and tariff structures. The regulatory system is designed primarily to provide all companies in the electricity market with incentives to operate more efficiently and to ensure benefits are shared with consumers. "Efficiency" is largely seen in terms of prices. Hence, issues such as energy efficiency or environmental efficiency (including the level of external costs) is outside the ESI regulatory framework. Indeed, there is a political conflict between the pressure to demonstrate that the privatised ESI is delivering the promised ongoing price cuts and the logic that externalities should be internalised through higher prices. Thus, arguably, privatisation has taken the industry backwards in terms of its ability to deliver energy and environmental efficiency, if it is assumed that the neo-classical approach of externality internalisation is the best way to achieve this.

Research and development into more efficient, cleaner technologies was previously undertaken by the CEGB. There is no requirement in the current regulatory regime for any organisation to undertake research and development, as this is assumed to occur within the operation of the market. "Clean coal" research has now largely ceased, despite arguments that it should continue (Bailey, 1991). However, considerable research and development has progressed within a wide range of renewable technologies. In the nuclear power industry, the same problems surrounding safety and waste disposal which prevented it being privatised along with the remainder of the ESI persist today. Although the less unprofitable parts of the industry were eventually privatised in 1996 (as British Energy), the wider environmental issues have not been addressed to date and the government has retained the majority of the industry's liabilities.

One of the main problems with the privatised structure of the ESI is that the RECs have limited choices to expand their operations and increase profits in accordance with their duties to shareholders, and that these choices will potentially aggravate environmental problems. They have three main options; diversification, for example into generation, expansion to supply customers outside their area, and development of the core business of electricity supply. Of these, the latter was considered to be the most appropriate strategy during privatisation (anon, 1989). This involves weighting the balance between energy efficiency and energy consumption in favour of the latter. A further option is to expand unregulated businesses; appliance marketing and contracting. Both expansion of appliance marketing and encouraging increased electricity consumption have potential associated environmental impacts.

While privatisation did involve commitments on behalf of the RECs to supply increasing amounts of renewable energy, it did not specifically address energy efficiency. The most attractive options available to the RECs to increase profits may indeed lead to less efficient electricity use. The UK has consistently been the amongst the highest energy users per capita in Europe and it may be in the RECs' commercial interests that this trend continues. In the past, the ESI consistently opposed combined heat and power and similar energy efficient schemes, in keeping with the image of an industry concerned to sell as much of its product (electricity) as possible. Privatisation did not place significant requirements upon the industry to change such practices. Furthermore, important market imperfections are not addressed by competition in the supply of energy commodities generally (Eyre, 1998a). On the lack of structures in the Electricity Act to encourage energy efficiency, it has been stated that: "In a market that is strongly biased towards energy supply, energy efficiency requires interventions in the market to be on a level with supply side options. This fact put energy efficiency onto the absent agenda.." (Roberts, Elliot and Haughton, 1991). While energy efficiency *per se* is outside the scope of this thesis, which is concerned with production-related impacts per unit of electricity produced, this serves as another indication of the lack of attention to environmental issues during and after privatisation.

2.9 Future Prospects

The rapid growth in significance of the ESI in the 20th Century, and the consequent influences placed upon its operations by government and other interests, have had long-reaching consequences for the way in which environmental issues arising from the industry are tackled, or not. The result is that the ESI operating today has potential flaws in operation, in terms of its ability to achieve optimal economic, social and environmental efficiency. The fact that these flaws are influenced to a large extent by historical development factors indicates the importance of this review, and that of taking into account the present when developing the methods and practices of the future.

The structural changes brought about following privatisation have coincided with, and possibly contributed to, major changes in the generation technology mix, with consequent implications for environmental impacts. Changes can be expected to continue as research and development programmes with long lead-in times have an ongoing effect. With the built-in phasing of certain structural changes to elements of the ESI over the last few years, major ongoing shifts in organisation and dominance in the market place are taking place, involving both companies set up during privatisation and others, from overseas utilities, non-electric utilities, and companies seeking to break into the electric utilities market. The RECs have contributed substantially to the shift to relatively smaller generators and generation technologies by developing principally gas-fired power stations in order to reduce their reliance upon the main generators. However, the Combined Cycle Gas Turbine may be seen as a transitional technology, which may be instrumental in the shift to much smaller, possibly even domestic-level generating units within the next decade (Patterson, 1999).

Despite government targets on renewables and policies to achieve sustainability, major long-running environmental issues exist associated with the ESI, and restructuring of the industry has done little directly to address them. Although initiatives to encourage non-fossil fuel technologies and, more recently, to levy climate change contributors have been made, they are not linked directly to the environmental impacts concerned. Indeed, since the market does not effectively value the environment, it cannot protect it (Fells, 2000). Since the full opening up of the competitive market in 1998, there has been optimism that Energy Services Companies will assist in expanding renewables through enabling environmentally aware consumers to choose "green" energy (Stanord, 1998). However, there remains the problem of how "green" green energy is, and how objective the means of measuring environmental impacts are. Meanwhile, the problem of optimising energy efficiency in an industry driven to maximise consumption remains and, in some ways, it has got worse. What is needed as a starting point to investigating how these issues can best be addressed, is an examination of the current regulatory regime as it relates explicitly to environmental impacts. This is therefore the main task of Chapter 3.

3. CURRENT ENVIRONMENTAL REGULATORY FRAMEWORK

It is suggested in Chapter 1 that environmental impacts are not fully dealt with by the current regulatory framework of the ESI. The principal aim of this Chapter is to examine this suggestion in detail, by reviewing the current environmental regulatory framework, and establishing the extent to which it fails to regulate effectively for environmental impacts. The regulatory process is introduced in Section 3.1 and the environmental regulatory framework is reviewed in Section 3.2. Failure is discussed in terms of general weaknesses (Section 3.3) and specific inadequacies (Section 3.4). A discussion of reasons for regulatory weakness leads into the next Chapter.

Before embarking on the review, it is important to clarify what effectively "dealing with" environmental impacts means. The concept of externalities is introduced in Chapter 1 and, as a starting point, this has a clear basis in economic theory and suggests a logical means of identifying regulatory weakness. If an environmental impact is not reflected in internal costs, then it is inadequately regulated for. Moreover, in order for the mechanism of internalisation to work, the cost must be linked to the impact directly. General rates and taxes are not internalisation mechanisms, since there is no incentive to reduce impacts. However, most regulation is not based on internalisation of costs, but is "command and control" in nature. Limits are set on impacts which are allowed, but cannot be exceeded. Often, licensing, consents or permissions are needed before the facility is allowed to pollute (and therefore cause impact) up to the permitted level. In such cases, some of the *potential* impact (that which would occur in an unregulated situation) is prevented, but the *actual* impact (that which is permitted) is not "internalised" in the strict economic sense. Thus, residual impacts occur under a command and control based regulatory regime.

Another important issue for the assessment of effectiveness of regulation is the range of impacts which are regulated for. This is difficult to ascertain, since a complete list of all possible impacts is not currently available, so those which are unregulated may not all be known. In general, it is assumed that significant impacts get noticed, and so they will be apparent. However, this is not always the case. The problem of identifying a complete list of impacts will be further investigated in Chapters 5 and 6. For the purposes of this review, relatively well-known impact agents, such as acidifying pollutants, and those contributing to global climate change will be considered, along with a list of impacts derived from an extensive study into ESI impact assessment (ExternE, 1995a).

3.1 Regulatory Process Overview

An ESI project starts with a decision by the developer to propose the development and the site upon which it is to take place. The next major decision is that taken through the planning process, which allows the project to proceed or not. Linked to this is the regulatory need for consents to discharge and pollute. Once construction is complete, the regulatory process switches to policing ongoing compliance during operation, with revisions of consents, etc.

During decommissioning, more regulations may apply in the form of waste management and restoration, and following decommissioning, ongoing liabilities may be regulated through, for example, contaminated land regulations. Therefore, at each stage in the project cycle, a range of regulations come into play, each being based on policies, which themselves have varying origins and intents. The general types of origins and forms of regulatory tools are summarised in Section 3.2.

3.2 ESI Environmental Regulatory Framework

The regulatory framework within which the ESI operates consists of an amalgamation of controls from a number of disparate sources, including, principally:

- International agreements;
- European legislation;
- National legislation and other regulatory mechanisms;
- Self-regulation based on codes, standards and management systems;
- Local planning framework.

The resultant framework consists of a number of identifiable components, as summarised in Figure 3.1.

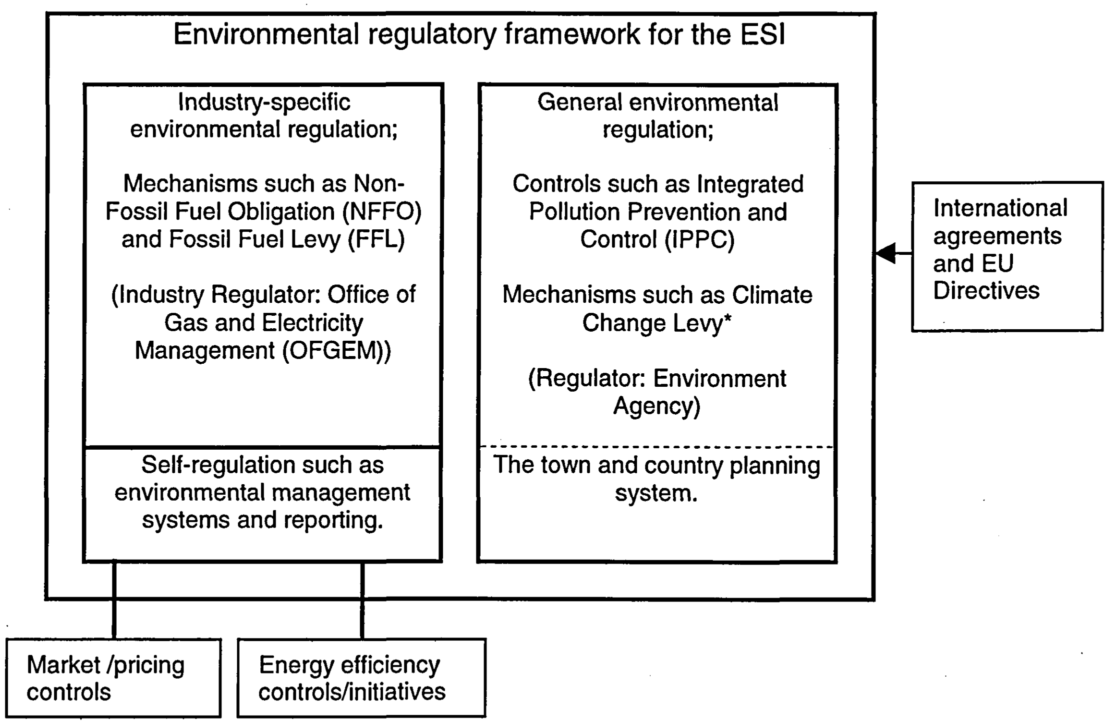


Figure 3.1 Summary of the Environmental Regulatory Framework Applying to the ESI

*Since this is directed mainly at the energy industry, it could be argued that it is placed between specific and general.

The majority of international, European and national agreements and legislation in this context have been generated as a response to pressing environmental problems, and take the form of "command and control" type limits on activities. Some agreements stop short of prescribing how pollution reductions, etc., are to be achieved, and these are of limited value here. However, other command and control legislation has been implemented and is summarised in Section 3.2.1, including both industry-specific and environmental-specific national regulations. Market mechanisms have gained more attention as regulatory instruments in recent years, and these are considered in Section 3.2.2. The local planning framework and self-regulation are considered in Sections 3.2.3 and 3.2.4, respectively.

3.2.1 Command and Control Regulations

Command and control regulations may originate at international, European or national level. International agreements include binding decisions made mutually by a number of sovereign states, for example at United Nations level. They are of particular benefit where the environmental impact is global, or where the global commons (for example, oceans or atmosphere) are threatened. They are also to be generally preferred in economic terms, because equality of regulation internationally produces a more level playing field in which the global market can operate. Equality of regulation theoretically means equality of environmental costing frameworks (although it must be noted here that there are many other factors which are not equal or level).

At European level, although the EU affects the ESI regulatory regime by influencing or changing UK policy, it also effectively affects legislation and regulation directly, since the directives and other legal instruments of the EU are often implemented in direct form in the UK. The Maastricht Treaty of 1987 formally recognised a duty on regulators to have regard to the environmental implications of their actions. In the 1980s around

20 items of environmental legislation were generated each year, rising to around 30 per year in the 1990s. This high level of activity was founded in a growing awareness that many environmental problems display a European dimension. Acid precipitation from power stations was recognised as a major transboundary phenomenon, while major pollution incidents, such as the Chernobyl nuclear accident in the Ukraine in 1986, also threatened several member states. The growing recognition that global climate change was resulting from widespread use of fossil fuels and other industrial activities, and that the polar ozone holes were a result of industrial emissions which required international action also contributed to EU action. As a result, a number of fundamental principles underlie EU environmental policy today, and these are required to be reflected in regulations:

- Prevention is better than cure;
- Polluter pays principle;
- Precautionary principle;
- Environmental effects should be taken into account at the earliest possible stage in decision making;
- Member states must not cause deterioration of the environment in other member states as a result of their actions;
- The EU should work with international organisations to promote world-wide environmental improvement.

Legislation has been developed through Directives, those in the following areas being of particular potential significance for the ESI:

- Water management, including controls on water quality and industrial discharges to water courses, and urban waste water treatment;

- Waste management, including directives on shipment of waste, incineration, landfill and the handling and disposal of substances;

- Integrated Pollution Prevention and Control, which derives from earlier directives on air pollution and is designed to incorporate cross-media pollution, i.e. pollutants which pass through various pathways, since it includes a system of basic pathway analysis on the part of the regulator;

- Eco-Management and Audit Scheme, which is designed as a voluntary means for organisations to assess their environmental performance and embark on a programme of continuous improvement (see Section 3.2.3);

- Environmental Impact Assessment, which is designed to ensure that the full environmental implications of new development proposals are examined and taken into account in deciding whether to go ahead with a project (see Section 3.2.4).

Although many environmental impact control initiatives are either initiated at international level, or have an international dimension, it is through the national regulatory regime that these are normally implemented. For a comprehensive review of the current general environmental regulatory regime at national level in the UK, the reader is referred elsewhere (for example, Spedding, 1999). However, a brief summary of its general structure is included here, which is sufficient for current requirements. The general area of modern environmental regulation in the UK has its roots in pollution control legislation, the origins of which date back to the 19th Century. However, it is the developments since the Second World War which have become increasingly extensive in controlling use and abuse of the environment. The modern era of environmental regulation began with the Control of Pollution Act (COPA), 1974, which was specifically and entirely aimed at providing a broad framework to control pollution of the environment from numerous sources. It was this which provided the extension of the more tentative Clean Air Acts of the 1950s into wider pollution control, and formed the basis for the latest generation of large, integrated, environmental-specific legislation. Today, the two most wide-ranging and significant Acts regarding environmental control and protection are the Environmental Protection Act (EPA), 1990, and the Environment Act (EA), 1995. The EPA, 1990, set out to establish a system of integrated pollution control, where "serious" pollutants are prescribed and their cross-media effects regulated through authorisations to pollute. Targets such as "Best Practicable Environmental Options" (BPEOs) were introduced, along with powers for the Environment Agency to revoke, vary and prohibit permissions and activities, and to enter and seize land and property. Powers were also granted to waste management authorities and a system of licensing of waste was introduced. Furthermore, provisions of statutory nuisance are made, to protect the public against contamination and pollution. The EA, 1995, generally strengthened the EPA, 1990, by creating the Environment Agency, which is now the combined pollution control body.

In summary, clearly, the main "day-to-day" implications of the national environmental command and control regulatory framework for the ESI are that consents are required to discharge pollutants into the atmosphere or water courses. These consents, once acquired, allow pollutants to be released up to the level stipulated, and within the time period stipulated. Records are required to be kept, and discharges and records are subject to checking by the Environment Agency, which applies BPEO and "Best

Available Technique Not Entailing Excessive Costs" (BATNEEC) principles in negotiating and stipulating licensing conditions and pollution consent levels.

Industry-specific national legislation and regulation is somewhat limited. The Electricity Act 1989 sets out the structure of the industry after privatisation, and the major structural controls governing its operation. Reviews and analysis of this wider regulatory regime are available elsewhere (for example, Kahn, 1988, Veljanovski, 1991, Beesley, 1994, Bishop et al, 1995). The Office of Gas and Electricity Markets (OFGEM) is the independent body which regulates the ESI. Its duties mainly relate to security of supply, competition and pricing regulation (rather than costing). Section 3 of the Electricity Act 1989 sets out the duties of the Director General, which include protecting the interests of electricity customers and promoting efficiency and economy on the part of licensees in supplying and transmitting electricity. Other roles do include "setting Standards of Performance for aspects of customer service and promoting the efficient use of electricity", although it is clear that it is electricity customers rather than people and the environment affected by the ESI which are the main concern.

3.2.2 Market Mechanisms

One example of a type of combined subsidy and levy in the current regulatory system of the ESI was the Non-Fossil Fuel Obligation (NFFO) and the Fossil Fuel Levy, which are directly administered by OFGEM. However, the government has more recently decided to replace NFFO with a Renewables Obligation, which will require licensed suppliers to provide a specified proportion of their electricity supplies to their customers from renewable sources (DTI, 2001). This will tie in with exemptions to the proposed Climate Change Levy (see below). The current tranche of NFFO will therefore be the last. Nevertheless, a brief summary of NFFO is relevant here, since it provides a rare example of the use of a market mechanism in the ESI.

NFFO requires public electricity suppliers to contract for renewable generating capacity. To meet the higher cost of developing renewable energy schemes, licensed suppliers are required to pay the Fossil Fuel Levy on the revenue from the sales of electricity. The additional costs can be passed on to customers and this is specifically allowed for in supply price controls. At present, this accounts for around 0.9 per cent of the customer's bill. Premium payments, funded from the Fossil Fuel Levy, are available to generators to enable them to recover the higher costs of developing renewable generation capacity. The payments are paid as an additional sum per kWh generated. To meet NFFO, suppliers enter into contracts with generators for an aggregate amount of electricity. There are different technology bands for wind power, hydro, landfill gas, municipal and industrial waste, agricultural waste and energy crops. OFGEM examines the projects put forward by the Regional Electricity Companies to ensure that they "will secure" the capacity to be specified in the Order. Contracts are awarded competitively on price and five Non-Fossil Fuel Orders have been made to date in the ESI in England and Wales, in 1990, 1991, 1994, 1997 and in 1998.

One example of environmental taxes is the planned Climate Change Levy. Following international agreements at Kyoto in 1997, and subsequent EU decisions, the UK set itself a domestic objective to reduce emissions of carbon dioxide by 20 per cent on 1990 levels by 2010. The government intends that all sectors must play their part in reducing greenhouse gas emissions, and have drawn up a programme in this light. A key element of this programme is the Climate Change Levy announced in April 1999, under which it is planned to introduce a tax on energy use by business with effect from April 2001.

3.2.3 Self-regulation

Individual ESI companies regulate their own environmental activities through, for example, codes of practice and methods of agreed standards and mechanisms for enforcement. A more formal way is through environmental management systems and certificated standards. Environmental management systems may be certified (e.g. to the European based Eco-Management and Audit Scheme (EMAS) standard or the International Standards Organisation ISO14000 series standard) or not, and for a review of their basic elements and operation, the reader is referred elsewhere (for example, Hunt and Johnson, 1995). Although there are numerous differences in the form of these two standards, the main similarity is that it is the business which compiles its own environmental effects register and targets for continuous improvement (i.e., in this case, reduction in impact). A key difference is that ISO14000 does not require public statements to be made of compliance, whereas EMAS requires a period public written environmental statement, this providing an important element of transparency.

Environmental reporting is also a means of self-regulation. One ESI company environmental report (YEG, 2000), suggests that there is uptake of self-regulation in the ESI through certified environmental management systems, with the example in question selecting waste and pollution control as major areas for continual improvement.

3.2.4 Planning Framework

Since 1947, a succession of Town and Country Planning Acts have established a regulatory regime of planning control. Although environmental protection was not originally an objective of this legislation, regulations such as Environmental Impact Assessment and the need to attach conditions to permissions have ensured that environmental criteria impinge substantially into the planning process.

Firstly, the point must be made that there is an imperfect but necessary interface between the planning system and the pollution control system, the latter being based on primary legislation and regulated by the Environment Agency (see Section 3.2.1). The Town and Country Planning system in England and Wales is best described in general terms as a hierarchical set of policy initiatives, ranging from central government-produced Planning Policy Guidance Notes (PPGs) and other guidance notes, through County Structure Plans, to District Local Plans. The Structure Plans provide a strategic long-term framework for policy but are not site specific. Local Plans are site specific, and should implement Structure Plan policies at local site level. In some areas, this two-tier plan system has been replaced with a single Unitary Development Plan, which combines both functions. All land use plans are required to include energy policies.

Development control is the area of the planning system which provides the interface between decision makers and developers. Decisions are based primarily (but not exclusively) on the provisions and policies within the development plan, as outlined in Section 54A of the Town and Country Planning Act 1990. Hence, the ultimate objective of the development control process is to decide whether the project should proceed or not. However, in the case of the ESI, most applications are determined by the Secretary of State at the Department for Trade and Industry. This process incorporates the EU-led Environmental Impact Assessment regulations, which were adopted into UK law in 1988, and have since been updated to improve, widen and strengthen them. The regulations are project-specific, and only larger projects, or those in sensitive areas, are affected. The EIA process currently involves a qualitative-based assessment of each environmental impact that is considered to be potentially significant, in turn, followed by proposals for appropriate mitigation measures to reduce impacts to “acceptable” levels, where required. Provision is also made to examine cumulative, secondary and synergistic impacts and effects. The starting point for setting assessment criteria for significance is typically the level of protection or regulation appertaining in the area to be affected by each identified impact. It should also be stressed that public participation is a requirement and wider criteria can be applied as the practitioners involved see fit or, indeed, the determining authority, which has powers to request further information or assessment.

3.3 General Weaknesses

One of the main weaknesses of the current regulatory regime is that it is influenced strongly by policies which reflect short term political agendas rather than those designed to optimise the ESI by minimising the environmental impacts per unit of electricity produced. A good example of this is at international level, where agreements and policy initiatives are often based more on concepts such as sustainable development rather than on real information about environmental impacts. The 1987 World Commission on Environment & Development Report (WCED, 1987, also known as the Bruntland report) put sustainable development on the world political agenda. A number of attempts at harmonisation of international policy followed in the 1990s, notably the Rio Earth Summit in 1992, which resulted in environmental agreements on climate change and other issues, including the Agenda 21 strategy for the 21st Century. Such international agreements have inevitably fed into European and national regulation (OECD, 1999). The case of climate change illustrates the nature of international level environmental regulation, since it is currently the "biggest" environmental issue in global policy terms, it has a truly global impact, and the level of scientific knowledge and scope for free riding is such that intense political debate has surrounded all efforts to reach agreements over regulation (for example, O'Riordan and Jager, 1996). Moreover, various regulatory options have been considered (for example, Van Ireland, ed., 1994, OECD, 1992). The problem remains, however, that unless the policy drivers and international agreements are based on hard scientific data and attempts at optimisation, then the regulations which follow will necessarily be similarly flawed. The solution to this problem is to generate policies which are based on objective information about *all* known environmental impacts.

The main weakness of market mechanisms is that, although they work in theory, in practice, they are rarely implemented successfully according to theory. Firstly, there are invariably problems in valuing many environmental impacts in monetary terms (this will be discussed in detail in Chapter 4). Secondly, even if some agreement or decision is taken on this, application is often arbitrary and does not apply equally across industries, environments and/or impacts. The result is a regime which is arguably no more optimised than the previous position, but is instead different, favouring some externalities over others. The proposed Climate Change Levy provides a timely example. Although it has been referred to as a "carbon tax", it is not. The charges to be applied to different fuels have apparently been designed with wider political motives in mind, since they bear little relation to carbon content and, therefore, to climate change effect. The fact that there are various exemptions to the proposed levy further indicates that it is not designed to internalise global climate change costs, aside from the fact that no explicit level of such costs has been expressly used in setting the level of the levy. This makes the proposed levy a particularly good example of a market mechanism, by illustrating that the theoretical basis of the internalisation process is not necessarily adhered to in practice, even when a market mechanism is chosen as the appropriate regulatory instrument.

The problem of self-regulation is that there is the potential for poacher/gamekeeper conflicts of interest. There is also no evidence that current techniques used in developing the register of environmental effects which underpins most environmental management systems, are likely to lead to an optimal situation, since, again, environmental issues are selected in an arbitrary way. Furthermore, there is no compulsion and so there remains the possibility of organisations free-riding.

The planning system inherits the weakness of the policy framework, so that some environmental issues will not be considered, particularly those for which there are no specific regulations or planning policies applying. As outlined in Section 3.2.4, the planning system in England and Wales is based around two main functions. The development control function is where decisions are made whether to allow proposed developments or not. The criteria which are used in this decision making are primarily non-financial, although clearly wider macro-economic aspects are involved. The non-economic information, such as that relating to environmental impacts and their consequences for humans is primarily supplied by developers and varies in quality, and there is currently no quality-control body to ensure or oversee this information. However, perhaps the single largest weakness within the planning system as far as minimising environmental impacts is concerned, is that the "benchmark" of what is considered acceptably close to optimal is not well defined. Since the only options for each project are to grant permission, attach conditions, or refuse permission, the only comparison which can effectively be used to set a benchmark is "do-nothing".

3.4 Specific Inadequacies

A very wide range of potential environmental effects may arise from electricity generation, distribution and use. These have the potential to have a range of significance of human consequences. The full range of environmental effects and resulting human consequences of the ESI have never been established objectively. Therefore, it follows that those which are "missing" from the current regulation framework cannot yet all be known. However, some can be identified, by comparison of a likely range of environmental impacts against the current regulatory framework summarised in Section 3.2. Such a comparison is presented in Table 3.1. It should be noted that the impacts considered represent a sample rather than a full range of impacts, but this list is derived for illustrative purposes here. Hence, it should be noted that the term "inadequacies" is not designed to indicate that the current regulatory regime is not working according to its own aim to implement current policy. Rather, inadequacy is assessed according to how well current regulation addresses a list of environmental impacts drawn from a recent study into externalities within a typical generation plant based on current coal-fired technology in the ESI (ExternE, 1995c). The aim is to illustrate inadequacy in dealing with impacts which occur in the ESI.

There are at least four potential types of specific inadequacy in the current framework. Firstly, there are "regulatory gaps" where no regulation exists to account for impacts (marked as XXX in Table 3.1). These may originate from inadequacies or omissions in policies, or in the application stages of policies through the regulatory system itself. They may occur where impacts have not yet been recognised or reacted to by policymakers or regulators. The second potential type of inadequacy is ineffectiveness, where a regulation designed to internalise does not achieve its designed aim. The third potential type of inadequacy is partial omission, where a regulation is designed to deal with only part of the impact (marked as X or XX in Table 3.1, depending on the extent of inadequacy). An example of the latter might be "command and control" type limits, which prevent critical levels of impact but allow sub-critical levels, thus only internalising part of the impact. Finally, the fourth type of failure is where regulation is introduced after the impact event, as an ameliorative or ex *post facto* measure (marked as R in Table 3.1). Although the regulation may be effective, this is contrary to the principle of "prevention is better than cure", which is enshrined within environmental protection legislation. Currently, for the many regulatory gaps which currently exist, there is little incentive for industry to implement its own controls.

Environmental Changes and impacts	*Current regulation*	*Inadequacy*
Air pollution and global climate change:		
Climate change effects on economic life	XXX R	Regulatory gap (and, if-when implemented, the Climate Change Levy (CCL) could only shorten it, not close it)
Climate change effects on future biodiversity/natural systems	XXX R	Regulatory gap (and, if-when implemented, the CCL could only shorten it, not close it)
Climate change effects on land availability/sea levels	X	The National Flood Defence Strategy is partly effective
Climate change effects on human health through unpredictable weather patterns	XXX R	Regulatory gap (and, if-when implemented, the CCL could only shorten it, not close it)
Air pollution and public health:		
Acute- particulates, ozone	XX	Control/Regional levels regularly exceeded, policing/penalties variable and often 'worth the risk'.
Chronic- particulates,	XX	Levels are set too low to guard effectively against long-term chronic outcomes, and in any case are regularly exceeded.
Chronic – unknowns such as SO_x, NO_x and ozone	XX	Lack of knowledge about outcomes, so no evidence that any levels set are effective.
Noise annoyance (industry/traffic)	X	Annoyance can occur even when compliance is occurring (-sub-critical impact)
Other public health		
Occupational – accidents in coal mines and transportation	X	Always a residual level of risk. Compensation levels may not be acceptable.
Occupational – coal mine exposure to dust and subsequent phlegm/illness/death	X	Always a residual level of exposure. Compensation levels may not be acceptable and start at a high level of illness (not, for example, for early stage disease or increased phlegm).
Possible effects of electromagnetic fields	XXX	Potential regulatory gap
Air pollution and agriculture, forests, fisheries, natural systems:		
Reduced yields from ozone and SO2	XXX	Regulatory gap
Increased liming requirements	XXX	Regulatory gap (historically there were subsidies available for lime application)
Acidification leading to increased forest death	XX R	Now substantially closed but the regulation was ex *post facto* and therefore contrary to preventative policy principles
Loss of recreation due to species and ecosystem loss/change	XX	Varies: Controls on some pollutant emissions but regulatory gaps likely, given that information is undeclared/unknown
Loss of commercial fishery	XXX	Regulatory gap
Air Pollution and building materials:		
Maintenance/replacement costs of eroded/damaged materials	XXX	Regulatory gap
Loss of heritage	X	Controls only protect prime heritage; marginal heritage is unprotected and uncompensated, hence those who damage are free riding
Costs of increased soiling and window/fascia cleaning requirements	XXX	Regulatory gap
Other		
Visual intrusion and reduced visibility	XX	Regulatory gap, since controls currently allow piecemeal and sub-critical impact, especially in unprotected zones

Table 3.1 Inadequacies in Current Regulation

Key: XXX - not dealt with; XX - ineffective/partially dealt with; X - partly dealt with/partly omitted; R - retrospectively introduced. Note: This list of environmental impacts is derived from anon, 1995. It is not based on a full range of environmental impacts. However, this is a convenient form of agreed impacts which have been already identified in the literature and is used here for illustrative purposes. This list does not take into account other stages of the fuel cycle, except in the isolated case of occupational health impacts of coal mining, which is included to illustrate the inadequacy of regulation for this specific impact type.

3.5 Why Regulation is Weak and Inadequate

Clearly, regulatory weakness and inadequacy occurs in both the extent to which current regulations control or internalise the impacts they are designed for, and in the range of impacts which are regulated. One possible reason for this is that policy makers have insufficient interest in producing more effective regulation. Despite the fact that efficiency in the way the market works is at stake, the dominant interest of the policy maker may lie elsewhere than in market efficiency. The interests of business may over-ride the interests of the environment and people affected by impacts; policy makers and regulators may lack the political will to ensure impacts are given due weight. As already mentioned in Chapter 1, while this possibility is real, the main purpose here is to establish how impacts can given due weight, assuming that this is the goal and that political will is forthcoming.

Policy making structures and motivations aside, a principle reason for the current inadequate situation is that insufficient objective-based information is available about impacts. In the absence of good information, many gaps and weaknesses in regulation will remain. In other words, a primary constraint on effective regulation is the technical data and knowledge available about the items to be regulated. In the case of the natural environment and humans affected by the ESI, the level of complexity of interactions, and therefore, of impacts, means that gaps and inadequacies in knowledge and data exist. With improvements in both knowledge and data (and data manipulation) over recent decades, many of these gaps and inadequacies are now much smaller than before, but they still exist. Inevitably, there is a lag time between new knowledge/data acquisition and application of new regulation based on it.

In summary, considerable gaps and inconsistencies exist in the current framework. The current approach to regulation is insufficient and needs strengthening. One of the main problems raised in current debates is lack of objective information about impacts. Furthermore, the *ad hoc* nature of policy and regulatory development and lack of systematic approach compounds the problems. What is required is an objective, clear and transparent method of compiling and presenting data and knowledge about environmental impacts.

The discipline which has attempted most actively and boldly to fill the data/knowledge gap for environmental regulators is environmental economics, as already introduced briefly in Chapter 1. Having established that poor data/knowledge is a major problem for environmental regulation, it is essential that a detailed examination of this discipline is now undertaken. Therefore, critical reviews of environmental economics and its applications to the ESI will be the subjects of Chapters 4 and 5, respectively. Future chapters will then concentrate on developing means of providing better data and improving knowledge for regulators.

4. THEORETICAL BASIS OF ENVIRONMENTAL ECONOMICS

In Chapter 3, it was established that sufficient information is needed about environmental impacts to enable regulators to maximise efficiency in the ESI. However, it was concluded that regulators lack such information. The technique which has been most widely recognised in attempting to address this problem is environmental economics. Therefore, a review of both this technique, generally, and work done around the ESI, specifically, is required, to evaluate its ability to address regulatory needs. Hence, the aim of this Chapter is to review critically the current approaches to environmental economics as a means of measuring environmental impacts. Section 4.1 introduces basic concepts in environmental economics, whilst Section 4.2 summarises the range of valuation methods used. Section 4.3 establishes and evaluates any weaknesses, and Section 4.4 selects from these the principal issues raised for the practical application of environmental economics methods to the ESI, which is the subject of Chapter 5.

4.1 Economics of the Environment

Since the ESI is governed by economic forces and the regulatory framework, the interaction between the economy, regulation and the environmental impacts it causes is central to dealing efficiently with such impacts. Rapid economic growth in the western world in the 20th Century, coupled with associated growth in technology and scale of exploitation of natural resources, has led to a major shift in the limiting factors in the economy. In the 19th Century, in many cases, the factor limiting growth was seen as lack of human ability or capacity to tap into seemingly unlimited resources. For example, it was lack of agricultural technology and/or labour which limited food production. Now, in many cases, the limiting factor is seen as the carrying capacity of the natural environment. Food technology, labour and production is plentiful (notwithstanding the fact that it is inequitably distributed), but the main constraint is how much land is available and how much more degradation of land, water and air can be withstood before impacts occur which cancel out (or exceed) the welfare gained from the extra food production. Hence, one of the principal debates now is sustainability. The main response of neo-classical environmental economics to the problem of degradation of the natural environment is to attribute property rights and/or economic values to it, thus bringing it into the economic sphere as capital stocks and flows on the business balance sheet. Hence, impacts are seen as costs which are external to the economy and, therefore, attributing monetary values to them (monetising), and internalising them, addresses the problem. The terms sustainability, externalities, valuation and monetisation are central to environmental economics as it relates to addressing environmental impacts from the ESI, and they are discussed further in Sections 4.1.1 to 4.1.4 below.

Prior to this, it is important to briefly outline the historical development of environmental economics and comment upon the nature of the environmental impacts for which it has been developed. Early economists acknowledged some important limitations of the environment to withstand certain types of human incursion upon it. Work by Malthus and Ricardo led to the important law of diminishing marginal returns. Although this was based largely upon the spectre of exponential population growth versus incremental growth in food production, it is now clear that, irrespective of population growth, pressure on resources increases since the process of production revolves around economic growth.

The origin of environmental economics goes back at least as far as 1920, when Pigou established the concept of social cost, meaning costs borne not by the market but by society (Pigou, 1920). Considerable development of these ideas has since taken place. Theories of public goods are also important. Indeed "the essence of environmental issues is that they involve externalities and public goods" (Winpenny, 1991). Despite such earlier theoretical work, the emergence of modern environmental economics is relatively recent and has developed since the late 1960s by a relatively small group of academics in the UK context.

Modern environmental economics arose from the need to tackle what has been described as an historical accident, namely that while some gains in human welfare are recorded by traditional accounting methods, others are not. Generally, business accounts reflect economic sectors where property rights have been well defined, whereas marginal effects of economic activity on the environment do not traditionally have such well-defined property rights, and the consequent lack of economic transfer makes them invisible in such accounts. Thus, management of public goods is problematic where access for all to the global commons may lead to tragedy through squandering, as prophesised in the 1960s (Hardin, 1968). It has been suggested that pollution arises from the common ownership of resources (Herfindahl and Kneese, 1974), and that assigning property rights to global environmental assets may solve some environmental problems (Coase, 1960). By ensuring responsibility through ownership, it is postulated that such assets would be managed in a sustainable way. While this idea is both theoretically sound and initially appealing by virtue of its simplicity, in practice, the nature of environmental systems are such that ascribing property rights to myriad of dynamic energy and material bearing phenomena quickly becomes infeasible (Daly, 1999).

As with most techniques, practical environmental economics has developed in response to topical concerns, attempting to deal with specific problems of environmental interaction with the economy as they have arisen. In the rush to develop much-needed solutions, the distinction between environmental impacts and environmental costs has often been insufficiently defined. An environmental impact is often regarded as a disruptive influence on the physical or human environment. This is in contrast to an environmental cost, which is a measure of the consequent response of the environment to an impact or impacts. Therefore, cost has a value, and is analogous to damage.

As environmental economists have concentrated on individual, pressing environmental problems, they have run the risk of not accounting for the full range of impacts appropriately. A more systematic approach is required to avoid missing out or double counting. Furthermore, the diverse nature of environmental impacts means that comparison between them is problematic and has been likened to "comparing apples and oranges" (Holdren, 1982). The use of money as the common unit of quality measurement may not be appropriate for all environmental impacts.

4.1.1 Sustainability

The current interest in sustainability is a natural extension of the realisation that economic activity often damages the environment. In simple terms, the concept of sustainability is related to the perceived need to retain environmental assets in good condition in order that they may be available in the future. It is a commonly held view amongst those concerned with sustainability and environmental degradation that economists generally have failed to accept the ultimate consequences of the changing view of the earth as a closed rather than an open system (for example, Jarret, 1966). This is an idea developed in recent times in the 1960s (Boulding, 1966). A number of important reports followed in the early 1970s, including "A Blueprint for Survival" (Goldsmith et al, 1972) and "The Limits to Growth" (Meadows et al, 1974). These contributed significantly to the change in global perception and so brought to the fore the concept of sustainability (although the term did not come into use until the late 1980s). Environmental economics is an attempt to come to grips with this change by accounting for degradation and issues around sustainability within the economic sphere. In the more generalised sense, a more environmental approach to economics has been taken in recent years by economic institutions, a notable example being the United Nations (UN) Framework Convention on Climate Change, a legally binding treaty signed by 153 member states of the UN in Rio de Janeiro in June 1992 (UN, 1993). Although not directly related to the environmental economics methods discussed below, this treaty calls for actions encompassing issues such as efficiency, global resource use and sustainability, all of which require some mechanism by which an assessment of degradation can be undertaken to allow suitable policies to be enacted.

The loss of "environmental capital" is one which is more easily related to by economic decision makers when it is couched in such economic terms. However, despite a general acceptance of related ideas, there is no consensus over what constitutes "sustainable development" or, indeed, whether the term is a contradiction in itself. Dozens of definitions have been published (for example, nearly 60 in Pezzey, 1989 and 30 in Pearce et al, 1989). An oft-quoted definition is that stated in the Bruntland Report (WCED, 1987) that sustainable development is reached when it "..meets the needs of the present without compromising the ability of future generations to meet their own needs". However, this begs for a further definition of "needs". Nevertheless, the general idea that total assets (including "environmental assets") should not be devalued over a particular (though often undefined) future period is typical of most definitions. It must be noted that there are considerable problems with assuming what values future generations will place on elements of the environment, since they will have different agendas and will exist in a technologically different world - and one with a much higher population.

4.1.2 Externalities

By addressing the issue of why the environment is not protected within the economic system, environmental economics complements the sustainability debate. Indeed, valuing the environment is a prerequisite to sustainable development, if this is to be attained within the market economic system. The central premise of environmental economics is that environmental damage is a loss of utility to society. Utility is defined as consumer satisfaction or pleasure, and the environment provides utility through supplying natural resources used in the production of goods and services, and also by providing directly consumable elements like a natural environment or clean air. The environment can therefore be considered to be a form of natural capital, with some analogy to traditional capital assets. Benefits lost through a loss of utility become externalities in situations where the producer who caused the damage does not bear the cost. This results in a loss of capital. Along with public goods and the lack of due attention to the needs of future generations, this constitutes the main reason for market failure (Winpenny, 1991). It follows from the theory of Pareto optimality that market failure causes a loss of social welfare.

A thorough review of the general theory of externalities can be found elsewhere (for example, Pirog and Stamos, 1987). For current purposes, an externality is simply a third party cost, which is not borne by producers or consumers, nor incorporated directly into the economic system. This is in contrast to internal costs, which are incorporated into the system and therefore influence prices directly. In a market economy, the price mechanism ensures that prices paid by customers for resources such as electricity include all of the internal costs associated with production of the product or service. For electricity, this includes construction costs of infrastructure, distribution network and power generating plants, operation and maintenance costs, fuel supply costs and others. External costs, in contrast, are not reflected in prices but are imposed on society through some loss of welfare. An externality is therefore equivalent to damage and, in some literature, it is defined as that damage which should be internalised, since the optimal amount of pollution is unlikely to be zero (Joskow, 1992).

Potential non-environmental externalities associated with electricity supply include security, liability, research and development, administration costs of regulation, income and employment effects, taxes and subsidies. Environmental externalities are just one type of externality associated with energy supply. In 1989 the US Department of Defence spent over $15bn to safeguard oil supplies in the Persian Gulf, with more than double this subsequently being spent on the war in 1991. A contemporary and conservative estimate of the effective subsidy is $23.50 per barrel of oil imported into the USA (Hubbard, 1991). Clearly non-environmental externalities should not be disregarded, but they are not considered here. It is clear that, wherever market prices fail to reflect substantial (external) costs, the market mechanism cannot deliver optimal allocation of resources. Thus, the presence of significant externalities is damaging for the economy, and these should be internalised if the market is to operate efficiently.

4.1.3 Valuation and Monetisation

Valuation involves measuring the worth of something quantitatively, in this case, the (negative) worth of each impact. However, if values are to be compared, they need to be expressed in similar units. Comparability between different impacts is desirable to allow decisions to be made about which are most important, and what weight should be attached to dealing with each of them. This is a basic argument for the monetisation of environmental impacts, where values are converted to or expressed in money, so that comparison is possible through the use of common (money) units. Money is usually considered to be the most appropriate common unit amongst environmental economists, since the eventual aim is to incorporate environmental costs into the existing economic system, which is based on money. Indeed, one of the key uses of the monetisation of environmental effects to economists is the potential for incorporating them into Cost Benefit Analysis (CBA). This allows for the inclusion of environmental costs in the decision making process. Indeed, results of valuation studies have recently begun to be included in integrated energy-environmental-economic models for a variety of applications, including full-cost accounting (Holub et al, 1999), to produce ready estimates of current externality burdens (for example, Levy et al, 1999), and to predict the viability of fuel cycles into the future (for example, Lee et al, 1997). Costings could also be incorporated into environmental assessments and, thus, become more integrated into project evaluation (Adger and Whitby, 1990). At whatever stage and by whatever means, accounting for environmental effects by incorporating costings into decision making could result in these costs and benefits being adequately reflected in the price of energy in the market place, which they currently are not.

As stated in Section 1.2, a key problem with valuing environmental effects is that "analysts may confuse things that are countable with the things that count" (Holdren, 1982). This can be interpreted in at least two ways. First, since environmental impacts are largely unmonetised at present, CBA treats them as intangibles and generally ignores them, giving them an effective value of zero. Second, when environmental economists begin applying values to environmental impacts, they tend to select those which lend themselves most readily to monetisation, and neglect those which do not. However, the situation is not static, and many environmental goods and services considered as intangibles two decades ago are now classified as measurable (Johansson, 1990).

Although monetisation of environmental externalities alone cannot solve all the problems associated with sustainability and resource scarcity, it is an important reminder that the environment is not a free commodity. In theory, it allows attempts to be made to redress the balance between the costs included in conventional economic analysis, such as CBA, and the so-called unquantifiables, hitherto ignored or, at best, undervalued. It is also a prerequisite to effective application of "polluter pays" type legislation, and bringing markets closer to true free markets by attributing costs appropriately. However, the key question remains; do valuation methods exist which provide accurate, objective and useful monetary values for all environmental impacts?

4.2 Valuation Methods

What is actually being measured when a value is determined for an environmental externality using an economic valuation method? To the economist, value is an expression of the satisfaction of a want or preference and is specifically related to people. Generally, value is ascribed to capital, which is a store of assets (things of value). Gains in value are “benefits”, while losses are “costs”. There are various ways to determine values for the environment and any effects upon it. At the simplest level, if qualitative environmental impacts are simply given scores, these could be converted into monetary terms. The key to the usefulness of the results are their completeness, along with the scoring process and the method of conversion.

In general, all economic values are either use values, existence values or option values, and Total Economic Value (TEV) is the sum of these. An explanation of types of value and approaches to valuation is peripheral here, and is included in Appendix A. This also contains commentary on risk and uncertainty, since this is a feature of the future status of the environment, its maximum pollution loadings, etc.

The critical problem in establishing a market value for an environmental impact is that invariably, no market currently exists for the environmental goods in question. Three methods for overcoming this have been identified (Budnitz and Holdren, 1976), each of which can be further subdivided. The first is to create a market conceptually by eliciting people's responses to questions to establish their willingness to pay or accept, the second is to look for a surrogate market, where an existing market which is indirectly affected by the environmental effect is identifiable, and the third is to combine information about physical effects and surrogate market valuation. It is important to note that, while in all cases the intention is to establish a value, this value is not always an estimate of Total Economic Value. Furthermore, the varying approaches produce different results. In terms of the development of methods, economists have turned most of their attention to assessing benefits such as clean air, rather than measuring control costs (Cropper and Oates, 1992).

Although economic valuation methods have been classified in various ways, there are relatively few rigorous approaches which are used widely. Classifications may differentiate between indirect and direct methods (for example, Pearce and Markyanda, 1989) or, similarly, between market-oriented and survey-oriented methods (Hufschmidt et al, 1983). The former is the general split used in Sections 4.2.1 and 4.2.2 below. Although there is contradictory usage of the terms "direct" and "indirect" in relation to environmental valuation in the literature, they can be used to describe two general sets of approaches. Direct valuation methods aim to establish revealed preferences, either through eliciting responses from individuals (thereby creating imaginary markets) or by studying suitable surrogate markets. Indirect costs are obtained by summing individual effects, sometimes expressed in terms of replacement values. The terms have also been used differently, differentiating between indirect methods which rely on household behaviour to reveal valuations of non-market goods, and direct methods where individuals' valuations are elicited directly through surveys (Smith et al, 1986).

Sections 4.2.1 and 4.2.2 provide a summary of direct and indirect methods respectively. Further detail is provided in Appendix A and elsewhere (Horne, 1995). The review is based on three extensive reviews (Pearce et al, 1992, Turner and Bateman, 1990, Pearce and Markyanda, 1989), all of which contain further review references. In addition, various texts give comprehensive reviews of valuation methods and wider environmental economics issues (Barde and Pearce, 1991, Johansson, 1991, Markyanda and Richardson (eds), 1992, Oates, 1992, Pearce et al, 1991, 1993, 1994, Teitenberg, 1994, Turner, 1993).

4.2.1 Direct Valuation

The most widely used and applicable direct valuation method is the Contingent Valuation Method (CVM). This is a survey-based method, where people are asked for their Willingness To Pay (WTP), and/or Willingness To Accept (WTA) a particular environmental change or range of changes through a questionnaire. Respondents are offered various options and bid-values are suggested, these being raised or lowered until a satisfactory monetary value is reached. Analysis of the questionnaire results then leads to establishment of a mean value. CVM is probably the most widely used method of valuing environmental impacts, most of the studies having been carried out in the USA. However, despite experience, wide variations occur between results, depending upon the approach taken. For example, in the Equivalent Variation approach, it is assumed that the project is carried out, and so WTP for avoiding and/or WTA deterioration in the environment is elicited. This provides widely varying results to those obtained using the Compensating Variation approach, where the project is assumed to be at the planning stage (that is, in the present rather than the future), so WTP for environmental improvements and/or WTA for environmental losses under the project are sought (see Section 4.3).

The main benefits of CVM over other valuation methods are that it can be used to measure the value of almost any aspect of the environment, and it is the only method which can comprehensively measure non-use values. The latter is particularly important, since it has been shown that such values can form a significant proportion of TEV (Madariaga and McConnell, 1987). Indeed, in at least one study, non-use values of freshwater fish in Norway were found to be an order of magnitude higher than use values. Environmental assets with low use values (for example, protected sites of natural interest) may therefore be expected to reveal even proportions of non-use values (Navrud, 1989).

As with all questionnaires, a major determinant of the results is how the questions and the exercise is conceived and framed in order to minimise researcher subjectivity and maximise respondent objectivity. Furthermore, a good CVM study must be informative; clearly understood; realistic by relying on established patterns of behaviour and legal institutions; have uniform application to all respondents; and, hopefully, leave the respondent with a feeling that the situation and his or her responses are not only credible but important.

The Household Production Function method is most often used where the output of the environmental good is marketable. The basic approach is to observe household spending behaviour to value the costs they are prepared to undergo to avert/substitute for environmental damage or to experience an environmental benefit. The former results in the calculation of Avertive Expenditures, while the latter results in calculation by the Travel Cost Method (TCM). The former is only relevant where households spend money on measures to offset environmental impacts, such as noise insulation. The latter is based on the premise that the value people place on a particular piece of the natural environment is inferred from the time and cost they incur in travelling to it.

Therefore, it can only be used to estimate environmental goods which involve travel costs, and is usually applied to sites with minimal or no entry fees, the travel cost being analogous to the entry fee. The main application in the UK has been for valuing recreational sites (Willis, 1990). Non-use values are clearly excluded from TEV calculations since only users travel to the site, and various technical problems are associated with TCM (Kealy and Bishop, 1986).

In the Hedonic Pricing method, as with the TCM, valuations are obtained through the study of surrogate markets. It assumes that people choose the amount of an extra-market good they use by altering their consumption of a marketed good. The most common use is in house price methods (sometimes called the Property Value, PV, approach), where environmental aspects of location such as noise or pollution are valued by using complex analytical methods to remove all other influences on property prices to leave a value for the aspect in question. Shortcomings exist due to the need to assume a well-functioning property market, which often is not the case, and due to its narrow range of application. Another Hedonic method is the Wage Risk (or Wage Differential) Method, which measures the employees' willingness to accept a risk, such as dangerous, unhealthy or disagreeable working conditions.

4.2.2 Indirect Valuation

Direct methods assume that people whose values are being elicited have perfect knowledge about impacts and markets. Where this is clearly not the case, indirect methods must be used, since revealed preference cannot be reliably used to elicit values from individuals. There are various methods for establishing market prices for environmental effects, or, if market prices are inappropriate, shadow (substitute) prices. The main ones applicable here are Dose-response and Alternative or Replacement

Cost. Each involves the calculation of approximate replacement, compensation and/or alternative costs.

With Dose-response, a linkage is established between the source and magnitude of the causal human activity and the resultant environmental impact. This impact is then measured and valued at market or shadow prices. With notable exceptions, the Dose-response method is not generally used, since it involves considerable work in identifying ecological interactions, many of which are poorly understood. Alternative Cost is a related method which investigates defensive expenditure necessary to remove the environmental damage, such as costs of double glazing to reduce noise in buildings (Pearce and Turner, 1992). This is also known as the Replacement Costs method; the cost of replacement or repair of the environment is calculated.

Also of peripheral interest here are the Human Capital and the Delphi methods. In the Human Capital method, people are treated as units of economic capital, and effects on health are quantified through loss of earnings and resource costs of health care. The Delphi method involves eliciting views from a panel of experts as to the value of environmental assets, and is used as a method of producing quick and cheap estimates. However, it is not a true market economic approach, since it does not reflect preferences or market forces.

4.3 Problems with Valuation Methods

In general, problems arise either with the theoretical or practical aspects of a particular method, or with the comparison of approaches taken by different methods. For the former, it could be expected that, given time and research effort, some theoretical and practical problems could be resolved, so that reliable and meaningful results could be obtained. However, the differences in theoretical approach between different methods mean that different types of value are being measured. Therefore, it can be argued that different results may accrue from using different approaches, even where each method is apparently accurate. The reliability and accuracy of valuations determine the efficiency level of the economic system into which they are incorporated. Clearly, poor valuations lead to inefficiencies in the economic system and, while improvements continue to be made, any case for ignoring externalities diminishes. However, there currently remain a number of unresolved issues in the field, not least of which is the partial nature of monetary estimates and methods by which they are derived.

One particular problem is associated with the lack of standard approaches or agreed methods for valuing environmental goods. Several major studies have compared CVM to other approaches discussed above (for a review, see Pearce and Markyanda, 1989). Different approaches invariably lead to widely differing values (Johannson, 1990), indicating that there is a lack of comparability between methods. Conversely, several studies have shown different valuation methods to be valid in comparison (Brookshire et al, 1982, Seller et al, 1985, Smith et al, 1986, Loomis et al, 1991). While it may not be possible to favour one method on theoretical grounds, it is also difficult to avoid the conclusion that at most, only one of these values can be correct. A solution used by some workers is to combine results or methods. This can be seen as a means for averaging, which assumes that both values were right, but different. It can also be seen as combining the shortcomings of the various methods, which have already produced spurious and differing results, and so cannot possibly lead to the right answer. Another problem is that workers tend to consistently select different methods to value different effects. This can lead to systematic bias.

Specific methods have also been criticised on methodological or structural grounds. CVM is arguably the most theoretically sound and extensively used valuation method, and yet various shortcomings have been identified and discussed extensively in the literature (for example, Pearce et al, 1992). It has been pointed out that respondents in CVM studies cannot address effects if they are unaware of them (Stirling, 1992). The assumption that respondents have perfect information is problematic. A good CVM study must, therefore, list all the effects that each respondent is aware of and incorporate a mechanism for topping up valuations to allow for effects missed out. Other potential errors have been summarised as being; "in responses to CV (CVM) quotations caused by purposeful respondent misstatements and differential valuation stimuli (that is, bias problems - strategic bias, information bias, instrument/vehicle bias, starting point bias, hypothetical bias, operational bias)" (Turner and Bateman, 1990). Strategic bias occurs when respondents state untrue values, often if they feel they are in a "free rider" situation, where they can gain the benefits without paying the costs. Indeed, it has been noted (Green et al, 1990) that the free rider hypothesis (Samuelson, 1954), in which people would systematically lie in response to a CVM survey, led to the rejection of CVM as a method by some economists (Feenberg and Mills, 1980). However, subsequent work has also suggested that it is not a significant problem (Marwell and Ames, 1981). Hypothetical bias is more structural in nature, arising because the transactions taking place in the questionnaire are not real; design bias includes the layout/type of information or type of bidding offered (bid vehicle), and starting point bias occurs where the starting bid offered to respondents affects their final bid due to their impatience or the suggestion of an appropriate bid size. A detailed discussion of biases is presented elsewhere (Schulze et al, 1981).

CVM is left with a further recurrent problem; the question of whether values should be based on WTP or WTA. Measures of WTP to maximise utility (have more environmental benefit or avoid environmental costs) and WTA to compensate for utility change (foregoing environmental benefit or putting up with environmental costs) should theoretically reveal equal or nearly equal values (Willig, 1976; Cropper and Oates, 1992). However, in reality, CVM results typically show WTA to be 3-5 times higher than corresponding WTP (Winpenny, 1991; Pearce and Markyanda, 1989; Cummings et al, 1986). This indicates that people value the loss of something they already have much more highly than the possible gain of something they do not yet have (Pearce and Markyanda, 1989). This has also been called loss aversion, and it can be shown that income constrains WTP (unless limitless borrowing is possible), whereas it does not constrain WTA bids (Hanley, 1990). A further explanation is that risk averse consumers, given one chance to value a good (rather than the repeated valuations operating in a normal market) will overstate WTA and understate WTP (Hoen and Randall, 1987). It has been suggested that the WTA/WTP disparity is dependent upon the substitutability of the environmental good in question, that is, how difficult it is to replace an environmental loss. Thus, WTA and WTP should converge where the environmental good can be easily substituted by an ordinary market good.

Explanations therefore exist for the WTA/WTP disparity. Indeed, only in an infinitely large market with zero transaction costs and perfectly divisible goods would the results be the same (Brookshire et al, 1980). Further explanations for WTA/WTP disparities have been developed through assessment of the problems and a review of CVM studies (Hanley, 1989), and discussion elsewhere in the literature (for example, Brookshire and Coursey, 1987; Pearce and Markyanda, 1989). The issue of WTA and WTP is significant to all valuation methods, since it shows clearly that environmental economic values are not absolute, but depend on whether the environment is to be gained or lost, and probably how substitutable it is perceived to be.

Problems with TCM include selecting appropriate applications for the method. It has been suggested that "the method should not be used unless there is evidence for the site in question that the key relationship (enjoyment increases with distance travelled) is approximately correct" (Green et al, 1990). The main shortcomings in practice are that visitors (travellers) are assumed to not enjoy travelling aspects of the trip (see also

Winpenny, 1991), some trips are multi-purpose (the traveller may visit several sites in one trip, or be on holiday and so have completed part of the trip already), and there may or may not be other similar sites nearby, thereby affecting trip length. A further difficulty is in choosing an appropriate rate for an individual's travelling time; if work time is given up to travel, then the work rate of pay is appropriate, but more often it is leisure time, and this is more problematic. The Department of Transport have produced their own figures for working and non-working time, based on attempts to establish shadow prices, although there are statistical problems associated with the method (see Hanley, 1990). A further point is that the travel cost must be regarded as the minimum cost a traveller is prepared to pay; he or she may be prepared to pay much more and, thus, the site value may be higher.

More general objections can be made to valuation methods and to environmental economics. Firstly, if traditional economic approaches have failed the environment in the past, how can advocating more intense but similar approaches provide the answer to solving environmental problems in the future? This argument has been taken further, in accusing economists of intellectual imperialism and an aggression towards other disciplines, arising from the competition in academia (Stirling, 1992). The failure of the so-called free market is another area where the fallibility of the market economic system is exposed. However, it should be noted here that, albeit within market capitalist countries, many of the worst examples of pollution have arisen in state controlled industries, while the former Eastern European states stand as an example of the way in which state capitalist planning can also fail the environment.

Valuation is also considered by many people (although to varying degrees) to be a rather inappropriate way to protect things which are considered to be priceless -as many elements of the environment. Pricelessness (clearly a barrier to valuation) is related to the irreversibility of many environmental impacts (Goodin, 1980), which itself creates numerous other problems for valuation (for example, see anon, 2000). The uneasiness with the whole approach may stem from the fact that the environment is seen as having intrinsic value, whereas in valuation an economic value is sought. The two are largely considered inconsistent and so pursuing economic value will lead to abandonment of intrinsic value, unless some method can be found to combine them. Philosophical and ethical limitations are therefore inherent in environmental economics, since "monetary valuation of environmental externalities relies on specific ethical foundations" (Söderholm and Sundqvist, 2000). Philosophical approaches other than the utilitarian are neglected by neo-classical environmental economics.

A further area of concern is that of geographical scale. Should externalities be valued on the basis of the total value attributed by people living locally, nationally or globally? It is generally argued that, since some externalities have global effects and elements of TEV are, therefore, attributable to areas distant from the source, the national or global perspective is appropriate (Buchanan, 1990). However, this raises questions of the practical size for valuation studies; of how people value externalities in different regions (especially between the developing and developed world), and what values should be attributed in each case. Is the environment cheaper in developing countries?

The undoubted subjective nature of the size of environmental effects and, therefore, what values to place on them, seems to demand a method which allows for that subjectivity. Methods from social sciences which are notionally objective but which are based on value judgements, such as CVM, should offer the possibility of measuring subjective elements of value. The finding that "judgement is an inevitable component of any empirical model of an economic process ... (does) ... not imply benefit estimation is unfeasible for practical purposes. Rather, ... (it) ... suggests that it is not a mechanical process" (Smith et al, 1986). However, the problem is not that CVM is subjective, nor that it is measuring a subjective phenomenon. Instead, it is that the subjectivity is not precisely located and minimised. The only way of achieving this is to incorporate a means of maximising objectivity in the process, thus isolating subjective elements. Problems with various biases suggest that economic valuation methods have not yet effectively achieved this.

4.4 Principal Issues for Practical Application

The concerns over methodological variations within environmental economics undoubtedly may affect results considerably, raising questions about the efficacy of monetisation *per se*. However, the key issue is not that monetisation is an inherently flawed exercise, but that it is affected by three main problems. Firstly, it suffers from the lack of a systematic approach and is generally insufficiently rigorous, transparent and objective. Secondly, it is dogged by emergent, experimental methods, and gaps in knowledge and data when it is applied in practice. Thirdly and fundamentally, no method can claim to be able to measure TEV, but can only capture part of the total value to humans of the environmental impact. Thus, even if valuers were to use a rigorous method to identify and measure all impacts, only part of the total value would be captured in money terms. Some elements of value would be left out because the valuation method effectively excludes them. Others, arguably, would be miss-valued, simply because we are not accustomed to measuring them in money terms. These may be specifically referred to as non-economic environmental impacts. In short, they are insufficiently valued using environmental economics methods and at worst, they are not considered and so receive a default value of zero.

The general objections to environmental economics are insufficient to warrant the rejection of the technique. At their heart, these objections point to the existence of non-economic environmental impacts - those elements of impact which cannot be sufficiently captured using economic approaches. These elements of impact must therefore be valued in non-economic terms. In Chapter 5, the general weaknesses identified here with environmental economics methods are considered more specifically in relation to their application to valuing environmental impacts of the ESI. One of the purposes of this is to assess to what extent the general weaknesses have affected valuations to date in the ESI. Also, it provides industry-specific findings as to where the current problems with valuation lie and, therefore, where the solutions must be found.

5. PRACTICAL APPLICATION OF ENVIRONMENTAL ECONOMICS

Having examined general issues of neo-classical economics valuation methods, it is necessary to consider their specific application to the ESI. Firstly, it is important to note that uncertainty should not be a reason for policy makers to reject impact valuation, just as they do not currently reject cost-effectiveness analysis because estimates of future load growth, construction costs, and fuel costs are uncertain. Secondly, the main attraction of monetisation is that it brings impacts into the economic sphere. Without this, utilities are prevented from accounting for any impacts outside current regulations, given their obligation to customers and the regulator to act efficiently in setting lowest prices possible. Section 5.1 highlights the issues raised in applying such methods to the ESI. Section 5.2 reviews critically numerous attempts to value the environmental impacts of the ESI using these methods, and Section 5.3 summarises the problems with the resulting valuations. Section 5.4 investigates appropriate ways forward for impact valuation, whether through strengthening environmental economics methods or by pursuing alternative ones.

5.1 Environmental Economics and the ESI

Some valuation methods are clearly inapplicable to valuing environmental impacts in the ESI because of structural limitations, validity or other constraints. For example, TCM is inappropriate for assessing health effects of pollution, whereas CVM or Dose-response may be expected to yield useful results. Table 5.1 presents a relevance assessment of the five main valuation methods against a range of environmental impacts. While the potential for using specific valuation methods *per se* may be reaffirmed, there are potential problems with specific application, for example, different methods producing different results because they are measuring different elements of value, as discussed in Section 4.3 above.

Stage in ESI	Type of Environmental Impact	Relevant Valuation Methods				
		CVM	AE	PV	CMA	HC
Fuel/source of raw materials, manufacturing, processing and transport	Physical, natural environment	✗	✗	✗	✓	✗
	Pollution, health	✓	✓	✓	✓	✓
	Pollution, economic activity	✓	✓	✗	✓	✗
	Resource depletion	✓	✓	✗	✓	✗
	Aesthetic, visual, noise	✓	✓	✓	✗	✗
Power generation, waste disposal	Physical, natural environment	✗	✗	✗	✓	✗
	Air pollution, health	✓	✓	✓	✓	✓
	Air pollution, nat. envt. ex. CO_2	✓	✓	✓	✓	✗
	Air pollution, CO_2	✓	✓	✗	✓	✗
	Air pollution, economic activity	✓	✓	✗	✓	✗
	Radiation	✓	✓	✗	✓	✗
	Aesthetic, visual, noise	✓	✓	✓	✗	✗
	Risk of major accident	✗	✗	✓	✓	✗
Transmission and distribution	Physical, natural environment	✗	✗	✗	✓	✗
	Transmission, health	✓	✓	✓	✓	✓
	End users, safety	✗	✗	✗	✓	✓

Table 5.1 Generalised Scoping List of Environmental Impacts and the Applicability of Environmental Valuation Methods (after Horne, 1995)

✗ = unlikely to be usable.

✓ = some applicability to at least one aspect of the impact.

(CVM = Contingent Valuation Method; AE = Avertive Expenditures; PV = Property Value, a Hedonic Method; CMA = Conventional Market Approach, encompassing Dose-response and Replacement Cost; HC = Human Capital. Note that Delphi is not considered as it is not a true market approach, and other methods are rejected because of their lack of suitability to most or all likely major impact areas in the ESI).

There are three ways in which external costs have been measured in ESI costing studies:

- Control costs are derived from the cost of avoiding environmental impacts by preventing their occurrence. Clearly, they are only relevant where this is possible, for example, by introducing better safety or environmental control standards, or installing flue gas desulphurisation equipment. Control costs are relatively simple to calculate and incorporate;

- Mitigation costs are derived from the cost of reducing harmful environmental effects, usually at the point of impact. The main difference from control costs is that mitigation costs assume the environmental impact occurs and addresses the cost of cleaning up or reducing its significance. Examples include treating accident victims, liming, afforestation and sea defence construction. Again they are relatively simple to calculate, although there are invariably external elements which remain unmitigated;

- Damage costs are the costs of damage caused to the environment. The calculation of damage costs extends much further than control or mitigation costs because an attempt is made to include all significant environmental externalities in the valuation process, not just those which can be avoided by controls or mitigated by treatment of affected areas. Damage costs are either market costs (for example, damage to agricultural crops, forests used for products traded) or non-market costs (human health, environmental amenities). They are more comprehensive in measuring environmental externalities and more problematic in quantification than control or mitigation costs. Research suggests they dominate in any comprehensive assessment of environmental externalities (Ferguson, 1993).

5.2 Review of ESI Impact Valuation Studies

It is apparent that problems exist with valuation and with valuation studies carried out on the ESI to date and therefore it is relevant to review the principal studies concerned. The valuation methods used vary. However, generalising across the literature on costing studies, it appears that the major externalities arise directly as a result of power generation, although those which arise elsewhere have been shown to be significant in some cases. However, the highlighting of global climate change and acid deposition as the principal external costs (both largely due to their impact on people's health, and, in particular, mortality) does present problems for the valuation approach. This is because no study has claimed confidence in calculation of these externalities and, indeed, the major reliable studies have singled out global climate change as having largely unquantifiable effects due to our present knowledge of this mechanism. In all cases, the establishment of dose-response functions is critical. Previously, critical loads (the point in the dose-response relationship where a small increase in dose would lead to a large increase in response) have been considered incompatible with monetisation; but they can be monetised through the use of an appropriate dose-response function, where the threshold corresponds to the critical load. Below this threshold, the externality is negligible, whilst above, it is large.

More detailed notes on the eight principal costing studies reviewed below is given in Appendix B and elsewhere (Horne, 1996a). This review of the development of costing studies is not intended to be exhaustive, but to illustrate the range of approaches taken to the problem. One of the first large studies of ESI external costs was undertaken in 1988 (Hohmeyer, 1988). As an initial low-cost estimation exercise, it attempts a comparison of fuel cycles on a general level. However, the approach, where existing aggregated air pollution data is apportioned to give the proportion attributable to electricity generation, has clear limitations, and could not be used as a basis for the detailed assessment of external costs. Aggregated data cannot effectively inform the decision making process where there is a possibility that site specificity and the type of technology are important factors, as is the case with many environmental impacts. The data sources are approximate and generalised, while the range of impacts considered is limited and no consideration is given to stages in the fuel cycle other than electricity generation.

One of the first large CVM-based studies of ESI impacts was sponsored by the US utility, the Bonneville Power Administration in 1990 (Hinman et al, 1990). The questionnaire was based around willingness to pay to avoid hydro, fossil fuel and nuclear technologies, and responses suggested that coal-fired air pollution, water pollution, nuclear waste storage, radioactive material transport and ozone depletion were major concerns. Of less concern were global climate change risks, new nuclear generation capacity, fish losses, radon in homes, and new dam construction for hydro schemes. In the same year, another energy externality study was published based on damage cost estimations (Hall, 1990). This incorporated more detail than in the Hohmeyer study, but was still a "first rough cut" culmination of recent estimates, and missed out numerous potential impacts.

At the time, the Pace University study of 1991 (Ottinger et al, 1991) was the most detailed damage cost-based externality study, and it has been widely used subsequently as a reference study. The costs are based entirely upon numerical estimates from earlier work. However, an attempt is made to allow for site specificity and, through the use of dispersion modelling, determine reference areas of impact. Also in 1991, a major study based on the control cost approach was published (Tellus Institute, 1991). The approach taken, where costs are based upon complying with the existing legislative framework, assumes that this accurately reflects current societal values for the environment. This assumption is widely considered to be invalid. Thus, the results are of limited value.

The UK Department of Energy commissioned a review of the available literature on monetary estimation of the social costs of energy production and the report was published in 1992 (Pearce et al, 1992). Although many sources are quoted, the report draws particularly heavily on the Pace University study, so incorporating its assumptions and weaknesses. The most striking aspect of the review is lack of data for major concerns, such as nuclear accidents and global climate change, "owing to the absence of suitable literature" (Pearce et al, 1992). The report does not suggest what should be done to internalise unknown externalities. Without such proposals, the danger is that such an approach leads to confusing things that are countable with things that count.

Ferguson (1993, 1994a, 1994b) has produced preliminary costs of electricity generating technologies in the UK, based on a rigorous approach to scoping and order-of-magnitude damage cost calculations. The conclusion is that human health costs are likely to dominate externalities, and these are highest in fossil fuel based generation (particularly global climate change-induced famine), and nuclear generation (particularly public aversion to accidents).

In 1991, the European Commission (EC) and US Department of Energy launched a joint research project to assess the external costs of fuel cycles, and this was continued by the EC under the JOULE programme, as the ExternE project. This major study first reported in 1995 (ExternE, 1995a, 1995b, 1995c, 1995d, 1995e, 1995f) and has subsequently produced updated summary, methodology and results, along with further reports, including on global climate change damages and national implementation (ExternE, 1998, 1999). The accounting framework developed has allowed damage cost estimates to be produced for case study projects for all the major fuel cycles/options, subdivided into separate studies as follows; Coal, Oil, Natural Gas, Nuclear, Wind, Photovoltaics, Biomass, Small Scale Hydroelectric, and Energy Conservation. This is the most comprehensive example to date of an external costing study for fuel cycles, and the first to be based substantially on original data sources. It is also based on a fuel cycle approach which borrows from the principles of life cycle inventory analysis. This comprises of setting system boundaries, producing an emissions inventory, classifying emissions into impact categories, describing and quantifying impacts, and valuing them in turn.

While other studies have attempted a broadly similar approach, they have generally been restricted to certain (new) technologies (for example, Keoleian and Lewis, 1997) or the level of detail is much lower than with the ExternE project (for example, Akai et al, 1997). However, at least four other major studies, all conducted in the US, share the same basic damage function approach to that used in ExternE (Thayer, 1991, Lee et al, 1994, Rowe et al, 1995 and Desvouges et al, 1995).

Other aspects of the ExternE study include the stated need for transparency in how results are calculated and to indicate uncertainty associated with the results and the extent to which the external costs have been fully quantified, as well as to show consistency with respect to boundaries of the fuel cycle system under examination. The ExternE methodology (ExternE, 1995b) also states that "no impact should be ignored for convenience. Instead, it should be retained for consideration alongside whatever analysis has been possible. An advantage of the present analysis is that such gaps have been identified". The case study approach is designed to illustrate the role of site specificity, and uses closely specified technology options.

Thus, the ExternE study is the most comprehensive external costing study to date and provides a major step forward in terms of producing data for use in valuation. However, the valuation method used is based mainly on drawing generic values from the neo-classical environmental economics literature and applying them to the quantified impacts. Thus, in terms of results, these may be more accurate than those of previous studies, but in terms of efficacy of valuation method, many of the general weaknesses of environmental economics monetisation still apply. Notwithstanding this, it should be stressed that the data collection methodology and standard modelling framework developed in the ExternE project is of use in itself and indeed, has already been adopted and applied to assessing environmental impacts of electricity (for example, Krewitt et al, 1999, Sáez et al, 1998). Furthermore, even with shortcomings and problems in valuation, ExternE results provide important provisional implications for energy policy and regulation (Eyre, 1997).

5.3 Problems with ESI Impact Valuations

There are a number of problems incumbent on the monetary valuation process, as discussed in Section 4.3 and elsewhere (for example, Horne, 1995, Stirling, 1997, 1998). Furthermore, the review of valuation studies above indicates that there has been incomplete and inappropriate identification and valuation of externalities to a greater or lesser extent in different studies. Therefore, it is now appropriate to summarise the main problems with ESI impact valuations conducted to date, given both their theoretical basis and practical approaches. An awareness of these shortcomings is clearly important in informing the requirements of future attempts at impact valuation.

5.3.1 Scoping Externalities

Power station stacks are a direct result of the need to remove local pollution problems, and serve as an indication that some environmental impacts have already been internalised, long before any detailed costing studies were published. Some of the environmental impacts of stack emissions were recognised and legislated for by the UK government in the 1950s. Generally, it may be expected that known, large (or potentially large) impacts may already be internalised, subject to the dual assumptions of a responsible government accountable to a knowledgeable electorate. This may explain why many costing studies have produced externality estimates of the same order of magnitude as current electricity prices, since lower costs are ignored and higher ones have already been internalised by some means, although other reasons for this convergence may be that such externalities are subconsciously derived by economists through massaging assumptions or are consciously derived as a "less bitter pill for the economy to swallow".

If the most pressing, large impacts have been internalised, this suggests that those which remain will be generally not catastrophic or as significant as those which have already been legislated for. However, the main problem with this concept is the assumption of knowledge and accountability. Technological advances introduce new areas of unknown potential environmental impact, and ongoing expansion of economic activity may increase the overall impact level. Furthermore, impacts come to light after the event, and these may be very significant, for example, carbon dioxide-induced climate change. Thus, at the heart of external costing is the need for a mechanism to proactively identify all possible impacts, not just to reactively consider a selection of those which are already known. All studies to date have simply assumed a list of impacts, generally drawn from those currently being debated in the literature.

5.3.2 Discounting

The valuation of environmental impacts is particularly problematic where they have a long lead-in time, and may occur at some point in the future. It is generally accepted that lower social discount rates should apply to valuing these types of future impacts, in accordance with the desire to meet sustainability targets intended to ensure that environmental stocks are available for future generations. Indeed, there is some logic in applying a zero discount rate to future impacts, although many neo-classical environmental economists argue that this would lead to "total current sacrifice" and too high a burden on current economies, so a 1-5% discount rate is more appropriate in such cases (for example, Pearce et al, 1992). However, even low rates effectively delete large future costs. For example, radioactive waste impacts over 10,000-100,000 years are negligible after 1,000 years with a 1% rate (a 10% rate would render them negligible after only 100 years). While a cure for radiation or cancer may be found, or there may be no humans left to take account of in 1000 years, undoubtedly, choice of discount rate plays a significant role in externality cost calculations, especially for long term future impacts.

Other issues to be taken into account in deciding which discount rates should be applied include the assumed increasing overall wealth of future nations and, specifically, what weight is given to natural, human and economic capital (Jones-Lee and Loomes, 1992). It has been argued that, in the face of increasing scarcity, natural resources will be valued more highly over time and that, instead of applying varying discount rates, some measure of the sustainability of the environment to be affected should be considered (Winpenny, 1991). The issue of discounting thus provides a problem for monetising environmental externality costs, simply due to the ongoing lack of consensus as to the appropriate rate to apply (for example, Portney and Weyant, 1999).

5.3.3 Resource Depletion

Related directly to the debate about discount rates, is the issue of the uniqueness of natural resources and the potential failure of recognised economic mechanisms to reflect the irreplaceability of natural capital and the irreversible nature of many environmental losses. The application of depletion premia to non-renewable commodities has been envisaged (Hohmeyer, 1988). These Reinvestment Surcharges would offset the effect of high discount rates which currently promote early exhaustion of resources. They are calculated for energy resources as follows (highest to lowest); crude oil, uranium, natural gas, hard coal, lignite. Another approach is to address the problem from the point of view of depletion levies or taxes. For example, a figure of £37 per barrel (1985 rates) and a cost escalation factor of 3% for resource depletion is envisaged in the US-ISEW study (Jackson and Marks, 1994). A similar but alternative or complementary mechanism would be one of depletion allowances, to encourage the development of non-renewable resources. Thus, the problem of non-renewable resource depletion could be dealt with by complementing externality costs with an appropriate sustainability-based regulatory mechanism. However, such input-related depletion premia are excluded from further consideration, since here the main focus is on output-related impacts per unit of electricity produced.

5.3.4 Global Valuations

While people in the same region and the same socio-economic class may apply similar values to environmental impacts (although even here, considerable variations and "spike" values appear in CVM studies), major problems occur across cultural, socio-economic and national boundaries. Cultural, social, ethical, philosophical and economic aspects of impact values vary widely both temporally and spatially, raising numerous fundamental valuation issues (Rothman, 2000). Firstly, even allowing for variations in national income and ability to pay, Bangladeshis may value differing environmental effects differently from North Americans, according to their priorities and level of information. Secondly, a dangerous conclusion from WTP/WTA data is the suggestion that human life in poor sectors of Bangladesh is valued at approximately one tenth that in the US, simply due to the lower values attached to the environment by people who have higher, more urgent priorities derived from lower life quality, and much lower incomes. Thirdly, there is the problem of transboundary impacts, where one group of people can effectively free-ride on another group, by gaining the benefits, while the impacts occur in a neighbouring/other area. The problem is greatest in the case of global impacts, especially where the receptor areas differ from the originator areas. US industry is the single largest contributor to global climate change, yet

Bangladesh is a principal receptor, in terms of loss of productive land under global sea level rise and consequent famine and death.

The solutions to such problems exist, but they have not been fully addressed by economic valuation studies. Cultural differences can be addressed by eliciting values for each cultural/socio-economic group separately and, as has been suggested elsewhere (Ferguson, 1994b), overseas impacts must be measured using home (polluting) country values. Global pollution requires global solutions - but these solutions must come from the polluters themselves. However, the thorniest problem is the implied differential value of life from neo-classical economic valuations. At its core, this arises because what is being valued is not the impact, but the element of the impact that can be stated in money terms, and these are two potentially widely varying quantities. Everyone feels lung cancer, or the sensation of drowning by flood or starving by famine approximately the same. It is this value, not willingness to pay values, which should be elicited for human health effects.

5.3.5 Information

The Bonneville Power Administration study highlights the problem of information within the valuation process. Information bias is possibly one of the greatest problems with CVM studies. If people are not given information, they rely on what they have been told by the media, which is usually inadequate. Put crudely, if global climate change risks are given a low media profile then they will incur lower external cost valuations. Information provision and type is an issue in all CVM studies and this links with the problem of how an impact or risk can possibly be valued if it is substantially unknown, whether this ignorance lies with the public at large or with the total level of scientific knowledge. In the wider context, this problem applies to all valuation methods, where the level of scientific knowledge is insufficient to allow the dose-response or cause-effect relationship to be confidently established. Furthermore, information is often interpreted differently by different disciplines where inter-disciplinary research is undertaken. Integrating economics and physical science is critical to the credibility of damage/benefit estimates, and they must work together to understand the limitations of various theories, models, databases and assumptions in each discipline (Rowe and Oterson, 1983).

5.3.6 Aggregation

During valuation, numbers are brought together from different sources, incorporating varying accuracy, precision, variability, and validity. This creates potential problems with results. Indeed: "Aggregation above the first levels of data handling is only useful, in general, to specific decision makers involved in utilising analyses made for specific purposes. In these cases, the biases of decision makers in identifying and prioritising evaluation criteria are the only biases which should be accepted" (Rowe and Oterson, 1983). Knowledge of the specific dangers of aggregation is important each time a valuation exercise is undertaken. If figures are only derived from transparent sources, where assumptions are clear and explicit, then sensitivity analysis in its various forms can be used to test significance under ranges of assumptions.

Specific problems with aggregation include combining results of valuations derived from differing economic methods, poorly scoped or inadequately valued effects (or left out due to difficulties in valuation), and lack of consistency and standardisation in the approach to valuation or its precursor stages. The Pace study (Ottinger et al, 1991) draws on studies using CVM methods for most environmental effects, but uses TCM in some aquatic effects, (incomplete) mitigation costs for air pollution costs; global climate change (afforestation) and acid rain (liming); and control costs for global climate change (retrofit carbon dioxide scrubbers) and some aquatic effects (closed cycle cooling). Hohmeyer (1988) uses mitigation and CVM approaches as well as indirect methods. Both Ottinger et al and Hohmeyer accept that visual impacts are a major, if not the principal externality for wind power, but neither quantify them satisfactorily (this requires site specificity issues to be overcome). Studies typically display both complexity and lack of standardisation; costs are expressed for individual technologies, or as avoided costs for other technologies, or damage costs can be calculated with respect to initial pollutant loadings, intermediate effects or ultimate consequences (Stirling, 1992). As well as the expected problems with using differing scientific or data sources, there are less obvious problems with using similar basic data. Such an approach may be the cause of structural problems, where errors are embedded in the data so that a range of studies gets similar wrong answers, which are misleading.

5.3.7 Valuing Risks

The comparison of risks is problematic, as they have various dimensions. The omission of one or more of these may lead to seriously deficient understanding of comparative risks. Statistical methods for evaluating risks vary, and non-normal distributions and rare events present particular difficulties in both quantification and comparison. There are choices to be made between incremental (single plant) and marginal (cumulative) analysis, and comparisons of average and marginal, net and gross risks. Illusory precision must be avoided where such calculations are concerned. Uncertainties surrounding risk calculations extend far beyond those associated with predicting the likelihood and implications of events in the future. Technology becomes less risky as it matures and various safety or environmental checks and balances are built into the system. However, knowledge and awareness of risks grow as technology matures, so that unforeseen implications of technology surface as it becomes proven in use. Different technologies are on different developmental trajectories in terms of any rates of change in their riskiness (Appendix A contains additional notes on risk and uncertainty).

5.3.8 Unquantifiables

Whether a given externality is unquantifiable or not depends upon the degree of consensus over valuations and what risks are associated with the data on which the valuations are based. The most important conclusion to be drawn for externalities which have been deemed unquantifiable in the neo-classical environmental economics literature, is that such impacts should be given extra priority in research and regulation, rather than being ignored due to valuation difficulties, as usually happens. If a lack of knowledge or the presence of a particular risk or uncertainty renders an impact unquantifiable, the activity which causes it should be avoided until the valuation problem can be overcome. Confusing things which are countable with things that count is only part of the problem. In the long run, avoidance may be cheaper than the unknown. A regulatory framework based only on quantifiable costs excludes the unknown. This must be avoided since it will lead to unknown environmental implications as well as potential gross inefficiency in the operation of the market.

Two clear candidates as unquantifiables are global climate change and nuclear accident risks. Externality studies have highlighted global climate change and generally failed to value it satisfactorily, although recent attempts are much improved (for example, ExternE, 1998). Societal aversion to the risk of a major nuclear accident is real and can be valued, although there is still debate over why or what this value is. The cost of risk aversion is potentially large and magnitude weighting functions have been applied to demonstrate sensitivity, in an attempt to overcome the problem of lack of method or empirical data for quantification (Ferguson, 1993). This leads to some meaningful figures. In cases where externalities are found to be unquantifiable, alternative approaches to incorporating external costs need to be addressed.

5.4 Problems with Current Valuation Methods

The discussions of valuation methods (Chapter 4) and their application to the ESI (Sections 5.1 to 5.3) have identified a range of potential problems with economic valuation methods. Issues range from theoretical to practical, and affect single methods or all methods. The principal issues are:

- Incorrect, insufficient or restrictive scoping and selection of environmental impacts to be valued, including general lack of rigour and partial impact selection processes;

- Use of poor, incomplete, or over-aggregated data;

- Lack of transparency in valuations;

- The discrepancies between a single value measured under different approaches (and systematic bias, where similar methods are always used for the same impacts);

- Lack of attention to difficult or unquantifiable impacts, effectively valuing them at zero (included here are valuation problems arising from the relationship between pricelessness and irreversibility);

- Starting point and design biases, where conduct of the CVM survey or bidding process leads to wrong values;

- Strategic bias, where people deliberately give wrong values because they feel that they may benefit in some way (for example, by free-riding), and hypothetical bias, where people give wrong values because they see the survey as a game and so it does not matter;

- Theoretical problems over the discrepancy which often exists between values of WTP and WTA;

- Vulnerability of valuation methods generally to public perception and, specifically, public perception and experience of money (which stems from the questionable assumption that people make rational choices when faced with complex situations and partial information);

- Varying levels of knowledge/information about impacts, which leads to varying levels of accuracy and confidence with valuations (this includes, but is not restricted to, information bias problems);

- Variations in scope and scale of impact valuation areas (spatially and temporally), with changes in values across income scales and national boundaries often not reflected in studies;

- Failure to capture TEV, since most methods cannot capture existence values or other aspects of TEV, so valuations are generally underestimates;

- Insufficient regard as to the needs of regulators and decision makers, leading to lack of confidence in valuations by those intended to use monetary values in the regulatory framework.

The environmental economics literature has hitherto concentrated on developing valuation methods, contemplating problems with emerging methods, and developing methods to include environmental costs in financial accounting, such as Environmental Cost Benefit Analysis (for example, Pearce et al, 1989, Johansson, 1991). Considerable progress has been made in both external costing generally (Smith, 2000), and in identifying some of the critical factors in producing externality values for electricity production (Rowe et al, 1996). Confidence in some externality estimates has grown as more consensus has apparently been reached over certain results, such as some air pollutant-health impacts (Krupnick and Burtraw, 1996). Since the early 1990s, it has been concluded that the current technology mix would change if (first estimate) external costs were internalised (for example, Hohmeyer, 1990, 1992). Realistic reviews of external costs are feeding into regulatory processes in the UK (for example, Eyre, 1998b). However, fundamental problems exist, since "the uncertainties and methodological problems associated with valuation prevent any reliable statement about the relative magnitudes of externalities resulting from fossil and nuclear generating options" (Eyre, 1993). Not surprisingly, even where studies compare similar fuel cycle technologies, they still come to varying conclusions about both relative and absolute external costs (for example, Sáez et al, 1998 and Faaij et al, 1998, Freeman, 1996).

It should be noted that difficulties such as comparability of diverse impacts, standardisation of impact assessment criteria, the problem of unquantifiables, and discounting and discount rate selection are likely to apply to any method of regulating environmental impacts of the ESI, not just one derived from economic valuation. Similarly, standard problems arise in formulating regulations which are unconnected to the underlying mechanism (in this case, economic valuation), including the need to address problems and concerns of current industry, and inertia to change, the problem of retrospective legislation and the need to treat new projects and existing plant equally, the desire to adopt recognised (and, as far as possible, politically expedient) regulations for the purposes of acceptability, and the need to reach international agreements for global impacts, where appropriate. However, addressing the problems which currently exist with neo-classical environmental economics methods, assessing alternatives such as more physical-based, multi-criteria analysis (for example, Mirasgedis and Diakoulaki, 1997), and supplementing an improved approach with alternative means of valuing essentially non-economic impact values makes the task of regulation based on such values both simpler and more successful.

5.5 Improvements and Alternatives

The need for better valuation is pressing. In the early 1980s, regulators in the western world initially embraced environmental economics as a potential solution to the problem of environmental regulation. Then, as the 1990s revealed problems over valuations, other, more traditional regulatory drivers re-established themselves, such as expediency. The most recent ESI-related market mechanism in the UK is the Climate Change Levy, and it is not really a market mechanism at all, but an old-fashioned tax, albeit a biased form of green tax. One possible explanation for this tax being apparently unrelated to the impact – for example, why it is unrelated to impact values or to carbon content of fuels - is that the government which drafted the regulation has not been persuaded of the accuracy or benefit of using environmental economics-derived values in framing the regulation, or has otherwise exercised political expediency. This indicates the urgency of improving environmental impact valuation to inform better ESI environmental impact regulation.

Clearly, improvements in environmental economics valuations are needed if it is to be deemed credible enough for decision makers to have confidence in basing new regulation on. Two major areas of weakness which need addressing are the completeness of valuations and the efficacy of methods. The latter can only be dealt with by a more rigorous approach to identifying fuel cycle systems and their outputs, which are, after all, the origin of all impacts. Arbitrary listing or choosing of impacts to value must be replaced by a systematic approach based on these outputs, traced through to their destinations as impacts, with associated dose-response or cause-effect relationships. This information is largely gathered through systematic observation of such relationships, and so, for impacts which may be irreversible or very large, risk-based methods may also be used. For example, awaiting closer scientific observations of global climate change occurring to obtain better valuations of effects may be less efficient than averting such risks (the decision being based on the values obtained for risk aversion).

Methods require varying practical and theoretical development, as discussed in Section 4.3. However, a major problem is that many impacts are not generally thought of in money terms, so asking people to put a monetary value on them is likely to lead to them applying "wrong" values, which are influenced by their understanding and experience of money from their economic lives. Taking essentially non-economic aspects of human existence, such as health, and applying monetary values is inherently problematic. The non-economic element of environmental impact can only be established accurately by eliciting values in non-economic terms, such as life quality, mobility, pain or distress. This is the only way in which the issue of completeness of valuations can be addressed. A means of valuing impacts in non-economic terms, to capture non-economic value, is required.

In summary, environmental economics may be suitable as a technique for valuing the economic aspects of environmental impacts arising from the ESI, provided further improvements are made. However, it must also be supplemented by a means of valuing non-economic aspects of environmental impacts. Furthermore, any and all methods must conform to the requirements for objectivity, rigour, transparency and completeness, which have not been met in valuation studies to date. In Chapter 6, a solution to the problem of non-economic aspects of value will be outlined, in the form of a new approach, specifically designed for this purpose.

6. A SYSTEMATIC FRAMEWORK

Critical reviews of the ESI current environmental regulatory framework (Chapter 3) and the theory and methods of neo-classical environmental economics (Chapters 4 and 5) have demonstrated where barriers to improved regulation lie. These barriers, in the form of weaknesses, shortcomings, omissions or inadequacies in valuation can now be turned into signposts, pointing out what needs to be done. The main issue addressed in this and Chapters 7 to 11 is the failure of current techniques for adequately valuing non-economic environmental impacts and ensuring they are adequately accounted for in the regulatory framework. This Chapter outlines the general principles and framework of a procedure for data collection, valuation and application for environmental impacts arising from the ESI. In order to achieve this, the following objectives must be met:

- Production of a list of criteria which must be complied with by the procedure if the current problems are to be avoided;

- Definition of environmental impact, and what precisely is to be valued;

- Establishment of a systematic framework approach for the procedure and the sequential Steps within it;

- Outline of a systematic means of producing data about the sources of impacts;

- Outline of a systematic means of producing data about how these sources become impacts;

- Outline of a systematic means of valuing impacts in preparation for appropriate application, for example, through regulation.

In accordance with these objectives, Section 6.1 contains a set of criteria which the systematic framework must meet, and introduces possible solutions to the current problems. Section 6.2 clarifies the object of the end-point of the valuation process, by defining the term environmental impact. Section 6.3 establishes the concept of a framework approach, and provides an overview of the systematic framework. Section 6.4 outlines how the framework approach can be applied to the ESI, and Section 6.5 assesses the framework by comparison against the criteria set in Section 6.1.

6.1 Criteria

Based on the current problems detailed in Chapters 3 to 5, and the needs of regulators, a systematic framework must meet the following set of criteria:

- Systematic and rigorous throughout;
- Transparent throughout, with maximum accessibility (ease of non-technical understanding);
- Clear system boundaries;
- Objective, in as much as subjectivity is confined to the values of receivers of impacts, not the measurer;
- Breaks the process into its component parts, with clear links between and within each Step;

- Sequential, step-wise approach, with each individual Step small enough to ensure that data are not lost;

- Fully inclusive, able to identify all possible changes in the environment arising from a given activity;

- Clear definition of what is to be valued and how;

- Incorporates tests to ensure data quality and completeness is preserved throughout;

- Based on logical and rational theory recognised in the literature, since the validity of objective measurement of subjective phenomena relies on a sound theoretical basis;

- Can meet the needs of reliability, validity and sensitivity;

- Demonstrably practical and simple in operation, with efficiency maximised through aggregation and data transfer, within boundaries set by other criteria;

- Valuation must be possible on a single scale, to enable comparisons between unlike non-economic impacts to be undertaken.

One of the main problems of the existing situation and methods is that regulation and valuation tend to be approached from the end-point, looking backwards. There are good reasons for this, the principal one being that time is short and regulators and valuers are usually working to time- and resource-constrained objectives. Therefore, existing methods and data sources are invariably used, which incorporate existing shortcomings and the problems of aggregation, etc. Very few studies have attempted to follow the impact valuation and regulation process in the same direction as the impacts originate, which means starting with human activity to exploit natural resources, progressing by tracing materials and energy through the production process and out into the environment, and ending with quantified impacts stated in comparable terms and appropriate means of reflecting values in regulations and decision-making. Such an approach could be described as "bottom-up" rather than the "top-down" approach taken by researchers who, understandably, are working to deadlines and must take short cuts in order to produce values and regulations. As Chapters 3, 4 and 5 have demonstrated, there are no short cuts. If the environmental regulatory system is to be optimised, impacts must be valued sufficiently accurately, and all elements of value must be measured. Therefore, appropriate solution to this problem is to adopt a systematic, step-by-step, bottom-up approach which starts with human activity. This is the systematic framework.

In order for the bottom-up approach to lead to the correct value, the object of value must be precisely determined, at least in as far as the data requirements of the valuation method can be set. Thus, the term environmental impact must be defined, along with a definition of what is to be valued, before the systematic framework incorporating the valuation method can be determined in detail.

6.2 Environmental Impact

An environmental impact is usually considered to be a detrimental effect on the environment arising from human activity. However, during valuation, the precise impact is not always well-defined. As discussed in Chapter 4 and Appendix A, environmental economics theory assumes environmental impact is analogous to the

Total Economic Value of the loss, which is made up of use value, existence value and/or option value. However, this is an insufficient definition, since it neglects or insufficiently represents elements of value which are not considered or experienced in money terms. In other words, it particularly misses out non-economic aspects of environmental impacts. Where attempts are made to provide an imaginary economy to non-economic aspects, to allow economic valuation of non-economic phenomena, the process is fraught with problems, from survey bias to incorrect responses stemming from the unwillingness or unfamiliarity of respondents towards the process. The solution is two-fold. Firstly, a clear definition of environmental impact is needed and, secondly, the valuation must be undertaken using a method which avoids survey bias as far as possible and seeks to establish a value in a currency to which respondents can relate. The latter is addressed within the systematic framework approach.

Having already established that an environmental impact is defined as something arising from human activity, it is also widely accepted that the term "impact" indicates a detrimental outcome. However, the point at which the detrimental outcome occurs is critical to the definition. This hinges on both the human activity-induced physical changes which take place in the environment and the subject of the detrimental outcome, so each of these must be examined in turn.

It is possible to define the (physical) environmental change element of environmental impact succinctly; it is the objectively measurable deviation from equilibrium arising from the human activity concerned. In order to be physically, objectively measurable, this must involve transfer of energy and/or matter from the human activity to the environment. Since the environment includes everything, this means that the environmental change commences at the point at which energy and/or matter leaves the immediate location of the human activity. However, while this definition of environmental change is simple, the range of potential changes is vast. This is because energy and/or matter can have a knock-on effect through numerous types of environments following a single emission from a human activity, and because of the highly variable nature of the environment and the changes which can occur to it.

The variable properties of the environmental change element of environmental impacts include dose, area of effect, synergistic effects, irreversible effects, the state of knowledge about effects, threshold considerations, risk and significance of changes caused. A fundamental variable is dose, which is the amount of energy/matter coming from the human activity. This clearly affects the size and nature of physical changes it causes. The size of the potential impact area is also important, as is the sensitivity of the receiving environment(s). In particular, how robust it is in being able to withstand change caused by the human activity (sometimes called "carrying capacity"). A synergistic impact is where two or more environmental changes act in combination to cause an effect which exceeds the sum of the individual effects. An irreversible impact is one where the environment changed by the human activity will not return to its original equilibrium state in the foreseeable future. Another variable is the level of existing knowledge about the dose-response relationship of a given human activity/environmental change combination (response refers to the change occurring as a result of the dose). For complex, non-linear dose-response relationships, there may also be varying thresholds - points in the dose-response relationship where a small increase in dose leads to a large (detrimental) effect in the receiving environment. Finally, there are variations in the risks of a potential environmental change occurring, based on the likelihood of human activities leading to various possible environmental change outcomes.

Clearly, there are numerous variables associated with environmental changes and their measurement. However, a succinct definition is possible, and this is often not explicitly stated in valuation studies. It is important since it indicates that the physical, objectively measurable element of environmental impact can be separated from the subjective element – which is associated with the second part of the environmental impact process; the subject of the detrimental outcome. The subjects or recipients of the detrimental outcome are the humans who experience a consequence from a given environmental change. The reason that impacts must always be viewed in terms of human consequence is that it is humans who are doing the valuing, and they can only value the consequences from their perspective. For some impacts, such as those on human health, the point of impact is clear, since it is the point at which the health deteriorates. Human health impacts are not always included as environmental impacts, but they generally are, and should be, since humans are part of the environment.

The fact that humans are doing the valuing does not mean that the environment or other life forms are not valued. If a human activity output, traced through a pathway, does not physically impinge upon a human, it may still have human consequences. If a released substance were to cause a change in the population of Emperor Penguins, with no discernible knock-on effect on other species or outside Antarctica, there would still be human consequences. Even for humans who will never visit Antarctica, the knowledge that this species is suffering causes human suffering, albeit psychological or emotional rather than as a result of direct physical contact with the causative agent. The value attributed to the Emperor Penguins in this example is comparable to the "existence value" element of environmental economics and it is very real. Humans effectively act as a surrogate, valuing the environment on behalf of the non-human life within it. Therefore, human consequences include what has been called "intrinsic value" of the environment, in as much as humans can appreciate intrinsic value and determine it in comparison to other human consequences.

So, in contrast to the objectively measurable physical environmental change element of environmental impact, the human consequence element is inherently subjective, and must be measured only by those who experience the consequence. For example, while the probabilistic risk of a potential impact occurring can be determined objectively, the perceived risk is subjective, and can only be determined by the potential receiver of the consequence. In summary, an environmental impact is a consequence, as experienced by humans in emotional, mental or physical terms, of an environmental change arising from energy/matter leaving a human activity. There is a direct, physical and objectively measurable link between the original outputs from a human activity and the environmental changes they cause. Consequences are the (subjective) human response to environmental change. Thus, environmental impact is not an unduly complex concept. However, clarity of definition is necessary as, without it, there is a tendency to mix inherently objective and inherently subjective elements, which is a poor basis for accurate valuation. Having defined environmental impact, it is now possible to turn to the issue of how the framework approach can be used to achieve the necessary accuracy of valuations for use by regulators.

6.3 Overview of the Systematic Framework

While it has now been established that environmental impact is not conceptually complex, it has also been indicated that there are a very large number of possible journeys between the point at which each impact originates, and the impact itself. The point of origin, in every case, is the release of matter and/or energy during a production process. The point of impact is the human body. In between, matter or energy changes may occur in a range of natural environments, both spatially and temporally. The natural environment is the total arena within which human activities are undertaken, which leads to releases of energy and/or matter, which, in turn, leads to environmental change and subsequent impacts. In approaching an understanding of this phenomenon, a comparable conceptual arena is needed, within which it is possible to trace the materials and/or energy to the impacts they cause. This is the basis of a systematic framework.

6.3.1 From Measurement to Valuation

A key concept of the systematic framework is that, whatever the complexity of human activity-induced environmental changes, they can be precisely measured by tracing energy and/or matter. For some impacts, the causal energy/matter is immediately apparent. For example, dust from a smoke stack falling on a populated area will have effects upon people, possibly damaging their health or causing annoyance or inconvenience. Heat energy in power station cooling water may affect fluvial ecology when it is discharged into a river and here, again, the causal agent of environmental change is clear. However, spatial factors, such as dispersion of pollutants, locality of processes and site-specific issues, and temporal factors, such as impacts occurring at different times and lasting for varying lengths of time can complicate the tracing of the energy/matter flow.

Some emissions of energy/matter are less tangible than others. For example, energy leaves human activities in the form of noise and light as well as heat. Light may be generated within the activity, such as by the use of floodlighting, or may be natural light reflected from buildings and other elements of the activity and into the surrounding environment. These sources of light cause both light pollution and landscape impacts, through allowing the physically measurable presence of human activity in the environment to be seen. Other less tangible energy/matter-environmental change linkages have effects on health or on particular ecosystems over a period of time. These include the human health effects of radiation, smoke and dust, and the forest and fish deaths associated with acid deposition. Even less tangible are those impacts which are predicted by scientists but remain unproved or only recently proven, such as health effects of electromagnetic fields or global climate change. The latter has only been accepted as a problem over the past two decades or so, and is now generally regarded as the single largest global environmental threat. It is reasonable to assume that other impacts which are currently held to be even less tangible or significant, or about which insufficient knowledge exists, will come to be recognised as important in the future. Electricity production cycles have always led to the atmospheric emission of carbon dioxide from fossil sinks. The traditional impact-oriented view ensures that the next discovery of an environmental harm agent after carbon dioxide will be equally surprising and unforeseen. In order to address this problem of being taken by surprise, it is necessary to look at the complete process which leads to environmental impacts, starting with the human activity.

6.3.2 The Need for a Framework

The lack of a standard and integrated approach to identifying the sources of all possible impacts has led to a range of problems with subsequent valuations, not least, that they are not demonstrably inclusive. A systematic framework does not automatically address the issue of choosing what to value, valuation methods, or policy making, but it does inform how and in what order these tasks need to be carried out. It also provides transparency and clarifies where problems and weaknesses lie. Most importantly, it provides a means of linkage between hitherto poorly linked elements of the system, in a structure which is clearly sequential and contains data and information requirements. Only by defining all the necessary Steps, and the data inputs and outputs between them, can data loss and subjectivity be minimised. To date, linkage between such elements has not been strong enough to allow valuations to be sufficiently reliable and accurate.

An important reason for drawing the various Steps of the valuation and application process together into a framework is to clarify the ongoing data needs from and to each Step. Weak links can be more easily established where a clear framework has been identified. Also, as with any emergent study area, there is a steep learning curve involved both within elements of the process and in the relationships between them. This learning curve can be incorporated into the framework through a series of feedback loops, where knowledge gained is fed back into earlier Steps of the framework. Feedback loops are a means of passing information back along the line of data flow, so that the framework can become a dynamic mechanism, with ongoing improvements being made.

The standardisation which a systematic framework brings is also beneficial (provided enough rigour is incorporated to prevent systematic bias). This allows input into the problem to be achieved from a wide range of disciplines, working on individual Steps within the framework. Transparency and clarity over system boundaries and data requirements at each Step allow disparate groups to provide the necessary specialist input at each Step. Since the form of data outputs is stipulated at each Step in the framework, each specialist can provide outputs which meet the needs of those involved in the next Step. This is the only means of ensuring that gaps, inadequacies and biases do not creep in to the data production process, the valuation process, or the application process.

6.3.3 Framework Steps

In order to meet the criterion requirement that the procedure must be capable of identifying all possible changes in the environment arising from a given activity, the starting point must be a rigorous examination of the entire production cycle being studied. The sum total of human activities which lead to the production of a good or service is called the life cycle. This includes inputs to and outputs from the point of production itself, and the implications of these inputs and outputs. This wide area of consideration immediately creates a degree of complexity and difficulty since particular supply chains of materials must be accounted for during project design and planning. However, such a breadth of area of consideration is necessary to ensure that all environmental impacts arising from the production life cycle are taken into account.

The next Step is to examine each Stage within the life cycle, and list the outputs released from it, in particular, those outputs which are incidental to the useful product or service of the life cycle. Given the complexities of most life cycles, these are numerous. Many of them can give rise, directly or through various mechanisms, or pathways, to impacts. It is here that many complexities arise due to lack of knowledge and information about the various mechanisms operating. However, all outputs are physical and so can be stated as physical quantities. Also, all environmental impacts arise from physical outputs, either directly or through intermediate pathways. Detrimental changes to the environment (of which humans are a part) can then be valued in terms in which humans understand as significant and these values can be reflected in regulations governing the production system. This relationship between the life cycle and resultant impacts and regulation can be represented in simple terms as shown in Figure 6.1.

Life Cycle
⇩
Outputs
⇩
Pathways
⇩
Receiving Environment
⇩
Impacts
⇩
Valuation
⇩
Application

Figure 6.1 Schematic Relationship Between Life Cycle, Resultant Impacts, and Measures to Deal With Them

By starting with the life cycle, examining the entire life cycle-pathway-environmental impact relationship and breaking it down into its constituent parts, understanding of the process and information about which parts of it are poorly understood can be improved. Constituent parts can be examined and further broken down for the purpose of:

- Reducing confusion and uncertainty by decreasing the number of variables under consideration at any one stage/time, thus improving overall understanding of the problem as a whole;

- Identifying areas where understanding is particularly weak and requires further attention (general conclusions like “advances required in underlying science” neither assist in the process of understanding nor have any capacity to lead to useful outputs now or in the future);

- Identifying areas where there are dislocations in the fuel cycle-impact-regulation data flow and describing them in detail, pointing the way to possible solutions.

Following this approach leads to a systematic framework which consists of nine discrete, separate but linked Steps, and these are summarised in Table 6.1. These nine Steps can be grouped into four methods. The first is the output analysis method, which comprises Steps 1 to 3. This is entirely concerned with the production system life cycle and, specifically, defining quantities of the matter and energy leaving the system which is incidental to the product or service being provided. The system boundary around the life cycle is defined by existence of potential environmental changes. Materials and energy which can cause environmental changes cross the system boundary and thus define it. Environmental impacts therefore arise exclusively as unintended or incidental outputs of the life cycle, referred to as Released Incidental Outputs (RIOs). The output analysis method culminates in Step 3; production of the

RIO Inventory, which is a quantified list of each output, checked for completeness, with data being supplied by design and process engineers, etc.

The second method, which comprises Steps 4 to 6, is called the pathway analysis method. This involves tracing quantities of RIO through the environment (including humans), and recording all changes which result. For each RIO in the Inventory, all the potential pathways must be identified, by considering all the environments which the RIO may affect. By measuring quantities of each RIO at each point on each pathway, gaps where quantities of RIO are unaccounted for can be identified (initial RIO = RIO at all subsequent pathways + RIO unaccounted for). Elements which are unaccounted for or lost/missing indicate where research is required to identify and measure potential changes as yet unknown (estimates may be used). This is a multidisciplinary undertaking involving ecologists, biologists and scientists (including human health specialists) in gathering data on quantification, mobility, transfer and pathways of RIOs and end-point environmental and human changes. The outcome of the pathway analysis method is therefore a long list of objectively quantified pathways and environmental changes, attributed individually to source RIOs. These lists form the Pathway and the Environmental Change Inventories.

The third method, comprised of Steps 7 and 8, involves the production of human consequence data, followed by valuation, and this is primarily a task for social and political scientists. The valuation method uses data from the Pathway Inventory and about human consequences of environmental changes to quantify the environmental impacts of the life cycle. These values are then applied through the fourth and final method, the application method, in Step 9. Hence, the four methods involve different disciplines, and are clearly linked, but have separate, sequential roles within the systematic framework.

OUTPUT ANALYSIS METHOD	Step 1. Life cycle definition The life cycle comprises all elements of the production system. The system boundary around the life cycle is defined by existence of potential environmental changes. Materials and energy which can cause environmental changes cross the system boundary and thus define it. Environmental impacts therefore arise exclusively as unintended or incidental outputs of the life cycle, referred to as Released Incidental Outputs (RIOs). Step 2. Stage definitions Each part of the production system is identified as a Stage, each with its own system boundary. Step 3. RIO Inventory Once Step 2 has been conducted successfully, production of a complete Inventory of RIOs, each in stated quantities, can be undertaken. The test of completeness is; total primary inputs of energy and matter = total RIO + product.
PATHWAY ANALYSIS METHOD	Step 4. Pathway Identification For each RIO in the Inventory, all the potential pathways must be identified, by considering all the environments which the RIO may affect. By measuring quantities of each RIO at each point on each pathway, gaps where quantities of RIO are unaccounted for can be identified (initial RIO = RIO at all subsequent pathways + RIO unaccounted for). Unaccounted RIOs are thus located and quantified, which may indicate potential unmeasured environmental impacts. Step 5. Pathway Inventory All pathways and destinations are recorded on a Pathway Inventory, including elements unaccounted for and where they are "lost" from the RIO-pathway process. Step 6. Environmental Change For each item in the Pathway Inventory, quantification of the resultant change to the environment in each case is required as a pre-requisite to assessing human consequences. Thus, the Environmental Change Inventory is produced. Each unaccounted for element signals where research is required to identify and measure potential changes as yet unknown (estimates may be used).
VALUATION METHOD	Step 7. Human Consequence Establishing the implications of each environmental change for humans is the immediate pre-requisite to valuation calculations, and involves gathering all the information directly required for this. Step 8. Valuation Valuation involves calculation of impact value using the human consequence data.
APPLICATION METHOD	Step 9. Application The final Step is to incorporate valuation results into the regulatory framework, so that due weight is given to these values in the decision making process.

Table 6.1 <u>Summary of the Systematic Framework</u>

6.4 Application to the ESI

The pathway analysis approach could potentially apply to any life cycle. However, the particular application here is the ESI. The ESI life cycle involves the production and distribution of electricity. There are a wide range of possible ESI life cycles, the main differences being attributable to the generation technology used. Common economically viable technologies in the UK include coal, gas and oil fired, nuclear, hydro, wind and biomass based generation, while several other technologies are either viable elsewhere or are undergoing research and development. Clearly, the framework approach applies to all ESI technologies, whether current or future. For the purposes of outlining this application, the nine Steps of the systematic framework are split into four main methods and discussed below; the output analysis method (Steps 1 to 3, Section 6.4.1), the pathway analysis method (Steps 4 to 6, Section 6.4.2), and the valuation and application methods (Steps 7 to 9, Section 6.4.3).

6.4.1 Outline of the Output Analysis Method

The life cycle of an electricity production system is called the fuel cycle. As with other production cycles, the process of describing the fuel cycle-impact relationship is not well documented nor made explicit in the majority of literature on environmental economics and the internalisation of environmental externalities. Few attempts have been made to produce major valuation studies which consider how impacts were selected for valuation in any detail; one such is a parametric assessment technique (Clarke, 1994). Where selection is mentioned, the most common method is some form of peer and/or literature review. However, this common approach misses out an important initial part of the valuation process; establishing the fuel cycle and an inventory containing all the RIOs. These are the source of all environmental impacts, and omitting this contributes to the uncertainty which surrounds the results of valuations. In particular, it leads to the possibility of omitting important impacts and can introduce unnecessary subjectivity. The magnitude of RIOs can be agreed, since, given knowledge of the process under consideration, they are identifiable, measurable, physical quantities. Thus, recording of RIOs as quantities is an essential common starting point in examining the fuel cycle-impact relationship.

The fuel cycle can be reduced to a process diagram and a simplified example, for that of a conventional coal-fired plant fitted with flue gas desulphurisation equipment, is illustrated in Figure 6.2. Drawing a system boundary around a Stage within the fuel cycle allows all inputs to and outputs from it to be identified, and then quantified. All Stages need to be identified, and this is achieved by examining all materials and energy inputs to each known process. Inputs indicate earlier Stages in the fuel cycle. The outputs from each Stage fall into three categories; intended products needed for a future process (or final product/service - electricity), incidental outputs which remain within the system (and so indicate downstream Stages in the fuel cycle) and RIOs (which leave the system). When all Stages of the fuel cycle are included, including inputs, this is referred to as the output analysis.

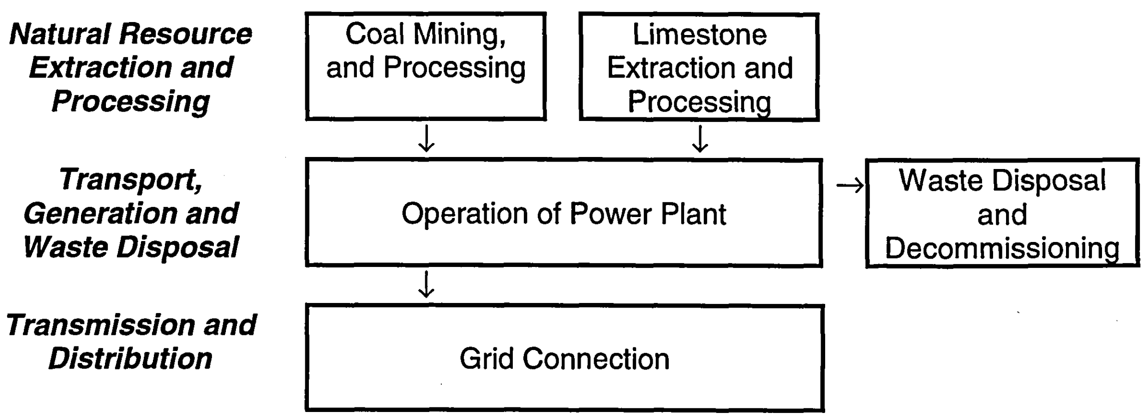


Figure 6.2 Simple Process Diagram for Conventional Coal Fired Plant with the Addition of a Flue Gas Desulphurisation Plant Designed to Remove Sulphur Dioxide

Note: Each box represents a Stage within the output analysis.

The process of listing each RIO by examining inputs and outputs for each process Stage is important, since impacts associated with elements which feed indirectly into the industry under consideration may be significant and must be considered appropriately in order to maximise industry efficiency. The consideration of each Stage of the electricity production cycle is important to ensure that all RIOs are considered, including those which occur indirectly to each Stage, for example, as a result of inputs into the generation process, or otherwise not at the point of generation in space or time. Furthermore, it is important that the RIO Inventory is complete, hence the test of completeness indicated in Table 6.1 as part of Step 3. This is the basic method of mass and energy balancing, and is based on the system principle that what goes in, must come out. The output analysis method is presented in detail in Chapter 7.

6.4.2 Outline of the Pathway Analysis Method

The pathway analysis method is at the centre of the systematic framework, and provides the critical means of bridging a gap found in many valuation studies, between the production system and the impact. It is split into three Steps; pathway identification, production of the Pathway Inventory, and production of the Environmental Change Inventory.

Once RIOs have been identified, it is necessary to begin to trace the path they take from being an incidental output to becoming an environmental impact. This is the process of pathway identification. Clearly, some pathways are relatively simple, while others are long and complex. Whichever is the case, the appropriate starting point for examining pathways is to commence with each individual RIO. Each physical or energy output from a process initially enters air, land or water. Thus, each incidental output falls into one or more of the following:

- Materials and Energy Outputs directly to air/atmosphere;
- Materials and Energy Outputs directly to land/lithosphere;
- Materials and Energy Outputs directly to water/hydrosphere;
- Materials and Energy Outputs directly to humans.

For example, stack emissions and dust from stock piles can be measured and stated in quantities, and listed under the first category above. This category also contains energy-related emissions such as wave-forms; light, sound and so on, which are transmitted through air. The second category contains lists of materials to be sent to waste disposal facilities and substances which escape from the process to site or surrounding land. The third lists water outputs and the constituents contained within waste water, such as substances in solution and energy in the form of heat. The fourth lists all outputs which directly affect humans without going through the natural (non-human) environment.

Thus, for each RIO in the RIO Inventory, all the potential pathways can be identified, by first considering all the environments which the RIO initially *directly* affects (these are called RIO-1s). For each RIO-1, a further list is required for all the environments the RIO-1 passes *directly* into after leaving its first destination. These second destination pathways are called RIO-2s. For each RIO-2, a list of RIO-3s is then produced by the same method, and the process is repeated until no more destinations are found. All pathways and destinations for each quantity of each RIO are then recorded to produce the Pathway Inventory. Every unique pathway-RIO combination forms a separate entry. In Step 6, for each item in the Pathway Inventory, quantification of the resultant change to the environment in each case is calculated to produce the Environmental

Change Inventory. This inevitably involves predicting, modelling and otherwise describing the interaction between (complex) natural systems and RIOs.

Pathway identification, inventory production and environmental change measurement (Steps 4, 5 and 6) are all major undertakings involving ecologists, biologists and scientists in gathering data on quantification, mobility, transfer and pathways of RIOs and end-point environmental changes. However, of these, it is the quantification of environmental changes for the Environmental Change Inventory (Step 6) which is most challenging, since this involves quantifying the relationship between the amount of a given incidental output and the response of the environment, the so-called dose-response relationship. Here, the information on receiving environments carried out during pathway analysis is important in defining different natural systems and parts thereof, for which relevant specialists can then become involved in measuring specific environmental changes. It is noted that many complex ecological and natural science phenomena come to the fore at this point, such as dose-response relationships, thresholds, cumulative and synergistic effects, organic and inorganic chemical conditions and reactions, to name but a few.

In summary, the pathway analysis method consists of three separate, sequential Steps, with each Step undertaking a limited range of tasks to ensure data are not lost. While it is recognised that objectivity cannot ever be entirely eliminated, this does not preclude the aim of minimising it where it is unwarranted, and the basic structure of the method prevents subjectivity from creeping into the process. Thus, pitfalls are avoided, such as discussion of human values of environmental impacts encroaching on pathway analysis, or leaving out of potentially significant RIOs by pathway analysts, both of which can occur in studies with less well defined methodologies. The pathway analysis method is presented in detail in Chapter 8.

6.4.3 Outline of the Valuation and Application Methods

The final three Steps of the systematic framework concentrate on valuation and the application of values. The valuation method commences with Step 7, which is the point at which each objective environmental change is converted into subjective human consequence(s). It is the immediate pre-requisite to valuation calculations, and involves gathering information needed for this about every environmental change on the Pathway Inventory, namely, demographic data, duration dynamics data, and probabilistic risk of occurrence of change. Demographic data includes the distribution, number and density of population affected, while duration dynamics data includes the duration of the environmental change and any variation in its characteristics over time.

The other data set required for valuation is that which comprises the main point of subjectivity within the systematic framework approach; the surveyed population response to each human consequence. As established in Section 6.2, environmental impacts culminate in emotional, mental and/or physical consequence for humans. In principle, established methods of valuation in environmental economics can be used to value those impacts which are in the former category. Nevertheless, it should be stressed that the use of such valuation methods should be subject to the wider systematic framework to ensure that the appropriate approach is taken to data collection, etc. Furthermore, the valuation method chosen should comply with the criteria laid out in Section 6.1.

However, the main focus of attention here are those impacts which cannot be expressed in money terms, for these require both the systematic framework and a currency of valuation. If money is not the common currency in which such non-economic impacts should be valued, then what is? Since it has been established that impacts are consequences in physical, mental and/or emotional terms, these units must be directly related to human life. There are two dimensions to life – quantity and quality. The former is directly related to duration of consequence, and this must be reflected in valuation. This leaves the requirement for valuation of non-economic impacts using a unit of quality of life.

Therefore, the human response dataset must consist of human responses for each human consequence, expressed in terms of the outcome for quality of life on each person affected (with any aggregated data including range of response and mean response). Thus, while other approaches may contain subjective elements distributed through the valuation method, using the systematic framework, subjectivity is isolated as far as possible to this single dataset of quality of life outcomes within Step 7. Moreover, subjectivity intentionally lies within the responses themselves, rather than in the means of obtaining them. It should be noted that the measurement of quality of life outcome for each human consequence requires development and application of a single index containing all possible quality of life outcome states on a single ratio scale, since this is the only means by which impact values in comparable units can be produced. Description of this index is presented in detail in Chapter 9.

Step 8 of the systematic framework comprises a systematic means of valuing impacts. In its simplest form, this involves performing the following function to arrive at impact value for each human consequence;

Impact value = quality of life outcome x risk x duration x number of people affected

The four datasets required for the valuation calculation are now known. One issue remains; that of checking the data for the necessary accuracy. Sensitivity analysis can be used as a data quality checking tool for this purpose. In short, this involves varying a piece of data by its possible range, allowing for possible inaccuracies. Commonly, high, low and most likely estimates are given, and the sensitivity of the valuation to this range can be assessed by computing values based on each estimate. Provided all possible values fall within an acceptable range, the data is sufficiently sensitive. The valuation method is presented in detail in Chapter 10.

The final Step in the systematic framework is the method for applying valuation results, so that due weight is given to these values in the decision making process. The logical approach is to take the current regulatory regime as the starting point. Clearly, additional regulation is required – regulation being used in its widest sense, including any mechanism, law or assessment tool which addresses the weaknesses and inadequacies identified as current and ongoing in Chapter 5. The development of such regulation enables the comparison of the impact values of competing projects by decision makers, along with other criteria. The application method is presented in detail in Chapter 11.

6.5 Criteria Comparison and Discussion

The systematic framework meets the criteria requirements for rigour, transparency, accessibility, objectivity, rationality and logic, step-wise approach, and inclusivity. It also incorporates clear system boundaries, and breaks the process into its component parts, with clear links within and between each sequential Step. Section 6.2 defines what is to be valued and the systematic framework establishes how valuation should be undertaken, using a single scale based on quality of life outcomes for non-economic impacts. Tests are incorporated to ensure data completeness and quality is preserved throughout. Completeness is addressed by means of mass and energy balancing (Step 4) and quality is addressed by use of sensitivity analysis. The framework has a sound theoretical basis, and further theory, for example, of scaling and of quality of life outcomes, is provided in detail in subsequent Chapters, as detailed description of each

Step is presented. It has not been conclusively established at this point that the last two criteria, regarding reliability, validity and sensitivity, and practicality, simplicity and efficiency in operation, are met by the systematic framework. However, none are contravened by it. Whether it meets these criteria can only be established by practical demonstration, and this is undertaken in Chapters 8 and 10, where detailed description of the pathway and valuation methods are presented.

In summary, the criteria, drawn from weaknesses of the current situation, can potentially be met by the proposed systematic framework. However, this does not mean that all problems are eliminated. While (environmental) science is engaged in developing understanding of all possible environmental pathways, it clearly has not achieved this, or reached a point yet where there is confidence that all pathways are even identified, or are likely to be in the foreseeable future. There is therefore considerable uncertainty surrounding the existence and detailed mechanisms of pathways. This uncertainty serves as both a reminder of how much more work is required to understand the environment, and an indication that human activities have impacts above and beyond those which have been discovered and are understood at present. Thus, it is essential that the systematic framework allows for the existence of as yet unidentified pathways and impacts. These "empty boxes" or gaps in knowledge, can be filled in when improved understanding of the underlying science allows, and are identified through the mass and energy balance checks within Step 4, and entered on the Pathway Inventory in Step 5. The acceptance and identification of what is not yet known is useful in itself, since it exposes gaps in knowledge which require filling. Current typical approaches such as simply generating a list of pathways and impacts from studies to date tend to ignore uncertainty and gaps by starting with pathways and impacts and only considering current knowledge.

The application of the systematic framework to the ESI is the main consideration here. However, it could be applied in a wide range of valuation methods, situations and industries. It provides a clear number of sequential Steps, all of which must be undertaken if accurate and appropriate valuation and regulation are to be achieved. Application of the framework should take place at the project planning level, for two reasons. Firstly, different life cycles, technology types and sizes of plant produce different impacts, and secondly, many impacts are site specific, so they change in significance from site to site. Although broad consideration has established in outline that the systematic framework approach can theoretically provide the data and valuations needed, much detail remains to be explained. Therefore, the next task is to examine each sequential Step within the framework and present methods for accomplishing each Step in detail. Following the "bottom-up" approach of the systematic framework, this commences in the following Chapter, with the production life cycle cycle, and the output analysis method.

7. OUTPUT ANALYSIS METHOD

An outline of a 9-Step procedure within a systematic framework is introduced in the previous Chapter. The aim of this Chapter is to present the first 3 Steps in detail, referred to collectively as the output analysis method. Although the outline of the method as already presented is criteria-led, these must now be complied with at the level of detail. Therefore, the appropriate starting point for this Chapter is a summary of the pertinent criteria it must meet and this is dealt with in Section 7.1. Sections 7.2 to 7.4 detail Steps 1 to 3 of the systematic framework respectively and, in Section 7.5, an assessment of compliance with the criteria is undertaken, including ensuring that the data outputs of Step 3 provide appropriate data inputs to Step 4.

7.1 Output Analysis Method Criteria

Several of the criteria laid out in Section 6.1 apply directly to the output analysis method. It must be systematic, with clear boundaries between system components and full transparency and accessibility of both data and process. It must also be broken down into sequential Steps, each linked forwards and backwards, and each small enough to ensure that data are not compromised. Finally, checks must be incorporated to ensure data quality and completeness. Despite these rigorous requirements, the method must also be practical and as simple as possible in operation.

7.2 Step 1: Life Cycle Definition

For the ESI, the life cycle can be defined as the full fuel cycle of electricity provision. Step 1 is the task of defining the total system boundary, that is, the dividing line between the system of processes and the environment. The simplest way of describing the fuel cycle system boundary is as a single box, as shown in Figure 7.1.

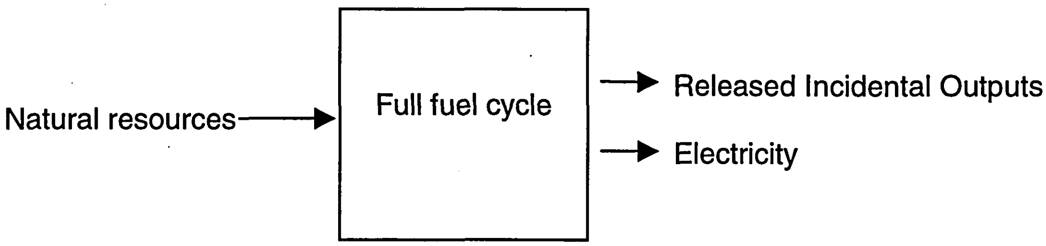


Figure 7.1 Full Fuel Cycle System

Therefore, the full fuel cycle is defined by the existence of inputs and outputs. Inputs are the natural resources required for the process which are still in their natural, undisturbed state. There are no other energy or material inputs (conventionally, human labour is not included). Any other inputs for any part of the process, such as pre-manufactured elements, packaging and pre-processed materials, simply indicate the existence of earlier parts of the full fuel cycle. The boundary must be drawn earlier in the process, and must include all the processes back to the point of natural resources. For example, coal is required to make steel to construct a power plant, and this coal in its natural state, still in the ground and unidentified, is a natural resource, meaning that its exploration, extraction, and transport is part of the full fuel cycle. The system boundary is thus defined as the point immediately prior to which the natural resource inputs required for the full fuel cycle are identified and utilised from the natural environment.

On the output side, the only intended output is electricity at the point of use. Therefore, every activity up to the delivery of electricity at the point of use is part of the full fuel cycle process. All other outputs of energy and materials from the process are incidental. These incidental outputs are either unreleased and retained within the system for further treatment or use (in which case they indicate downstream Stages in the fuel cycle) or they are released from the system and are called Released Incidental

Outputs (RIOs). RIOs have the potential to cause environmental changes; indeed, they are the exclusive cause of environmental impacts arising from the fuel cycle. However, at this point, the main function of identifying RIOs is that they cross the system boundary and thus define it. Therefore, Step 1 establishes a clear system boundary, defined by flows of materials and energy across it. However, since in reality, the full fuel cycle is very large and disparate, typically spread over a number of industrial sectors and involving numerous sites and manufacturing plants, the life cycle definition thus far is rather conceptual. There is a need to break the full fuel cycle process down further, into site or process specific parts. This is the task of defining Stages in Step 2.

7.3 Step 2: Stage Definitions

Each part of the production system within the full fuel cycle can be defined as a Stage, each with its own system boundary. It is necessary to define a number of system boundaries around discrete Stages within the fuel cycle, in order to ensure that all RIOs can be identified and that data are not lost. Usually, each Stage will occur at a different geographical location and/or involve different process equipment and activities to other Stages. A typical (though not necessary) starting point for embarking on defining all Stages within the fuel cycle of the ESI is the point of electricity generation.

Drawing a system boundary around the generation site allows all inputs to and outputs from this particular Stage to be identified, and then quantified. Inputs will all be of manufactured or primary materials, or energy. All these inputs indicate earlier Stages in the fuel cycle. For example, steel and concrete for construction indicate steel fabrication and concrete plants as earlier Stages, while coal would indicate at least one coal mine as an earlier Stage. Outputs will all be either RIOs, electricity, or incidental outputs which are subjected to further treatment. Each incidental output not released from the full fuel cycle points to a subsequent Stage. For example, pulverised fuel ash leaving the power station to a landfill indicates the receiving landfill as a subsequent Stage. Electricity is not considered (except in later normalisation of impact values, see Chapter 10), while the RIOs themselves are needed for compilation of the RIO Inventory in Step 3.

Each of the new Stages indicated by inputs and outputs to the generating site must now be subjected to a similar examination of inputs and outputs. For each Stage, inputs will all be either manufactured or natural resources, including energy. Outputs will be RIOs, or incidental outputs which are subjected to further treatment, or intended outputs (products) which are subject to a future process in the production of electricity. Again, more new Stages will be added to the fuel cycle, and subsequently examined in the same way. The process of Stage identification ends when the only inputs remaining to the full fuel cycle are natural resources, and the only outputs are electricity and RIOs. Thus, the simple box drawn around the full fuel cycle, as shown in Figure 7.1 and described in Step 1, is now split into numerous Stages, each based invariably on an individual site and/or process within the cycle. However, the principle of definition by natural resource inputs and RIO/electricity outputs established in Step 1 remains, and is central to the task of dividing the cycle into Stages. A highly simplified version of part of a flow diagram produced in Step 2 to show Stages in a fuel cycle is given in Figure 7.2, for illustrative purposes.

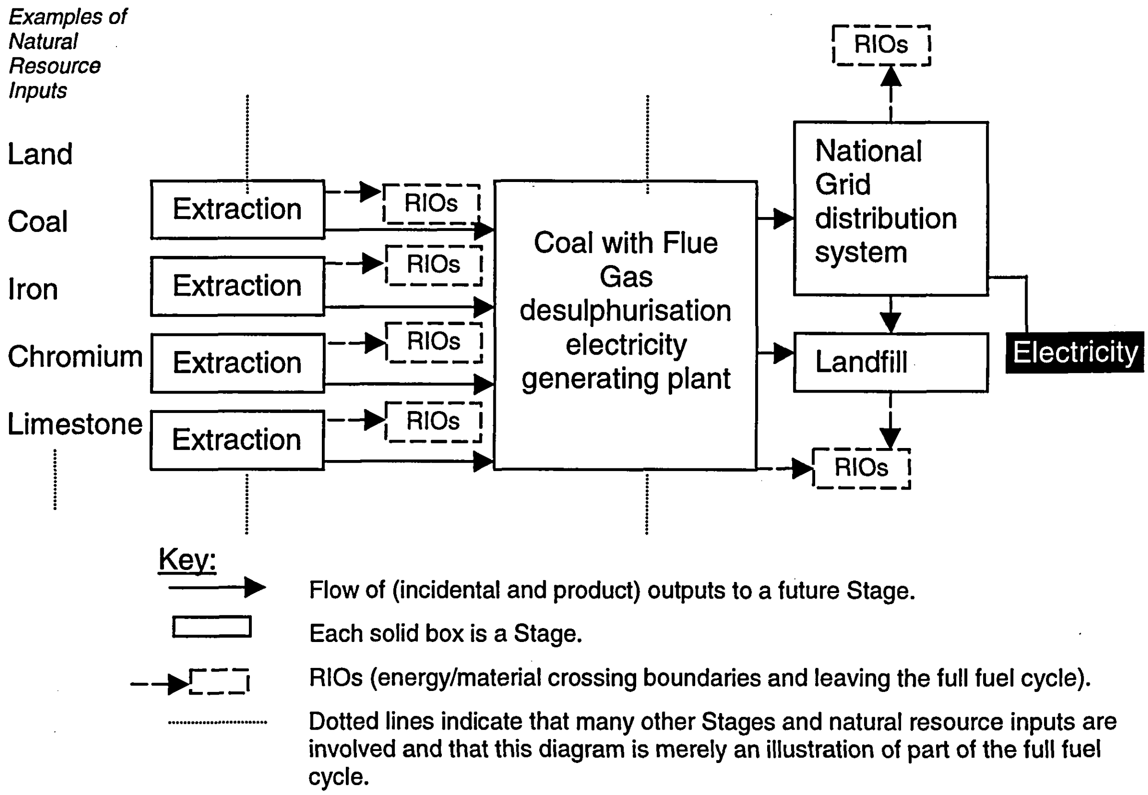


Figure 7.2 Part of a Full Fuel Cycle System Showing Stages and RIOs

7.4 Step 3: Rio Inventory

Once Step 2 has been undertaken successfully, production of a complete inventory of RIOs, each in stated quantities, can be undertaken. When compiling the inventory, it is important to list the RIOs by fuel cycle Stage. RIOs cannot be aggregated at this point, since they occur at different sites and will therefore have different pathways and impacts. The result is an exhaustive list of RIOs of materials and energy produced at each Stage of the fuel cycle, with specific quantities stated in each case. Materials include those from stacks, leakages and discharges, dumping and disposal/loss (except where it is subjected to further treatment and is therefore still in the fuel cycle), dusts, odours etc. Energy includes heat, sound, and light (including the physical presence of structures in landscape leading to visual impact) and any other physically measurable output other than the intended product or service.

Although the natural resources are not required to be listed on the RIO Inventory (they are inputs, not outputs), a complete list of quantified natural resources is required to be compiled as each Stage is defined. As the RIO Inventory is being compiled, quantities of natural resources must be balanced against quantities of RIO plus product, to check that RIO quantities as entered on the inventory are correct and no material/energy is missing or "lost". If any imbalances cannot be (or are not) corrected, quantities of natural resources/energy inputs unaccounted for must be entered as missing data or "gaps" on the RIO Inventory.

The RIO Inventory is therefore a complete, unaggregated list of outputs to the environment from each Stage of the full fuel cycle. Each output must be quantified in Système Internationale d'Unités (SI) units and normalised to electricity output. The normalisation process is assisted by the commencement of Stage identification at the point of electricity generation. As each material/energy item is traced into or out of the generation process, it can be normalised as the quantity of material per unit of electricity produced (kWh) can be recorded for entry on the RIO Inventory. The RIO Inventory itself is the culmination of the engineering/design input into the systematic framework, and comprises a clear, transparent set of data inputs for environmental specialists to use subsequently, in the pathway analysis method (see Chapter 8).

7.5 Output Analysis Method Criteria Compliance

Given that there are a myriad of different possible RIOs and paths each might take through the environment, the identification of all possible environmental impacts arising from a human activity is practically impossible, unless the method for identifying them starts with the identification of each RIO, rather than the identification of each impact. This has already been established as a reason for the systematic framework approach, and the logical starting point for it, the output analysis method. Criteria compliance has been established at outline level in Chapter 6. However, it is important that this compliance is also carried through to the level of detail, with a comparison of each task against the criteria set in Section 7.1.

Clear system boundaries are established in Steps 1 and 2. This process is broken down into two Steps for clarity, and in order to comply with the requirement that each Step is small enough to ensure that data are not lost. The approach is systematic, rigorous and sequential throughout the 3 Steps, with transparency and accessibility inherent in the presentation of the data in the RIO Inventory. The only measure of data quality and completeness which is required is the mass and energy balance check to ensure that "total primary inputs of energy and matter equals total RIO plus product". This is incorporated into Step 3 as a means of checking that the RIO Inventory is complete.

Simplicity is inherent in the output analysis method because it is broken down into its component parts in a series of linked and transparent Steps. Regarding practicality in operation, while there is no doubt that the RIO Inventory is a lengthy list, requiring considerable input from designers and engineers, the information required already exists within the design detail for each process Stage. Therefore, the tasks of Steps 1 to 3 are inherently straightforward and involve compilation of data rather than primary collection and measurement. Any shortened version of the RIO Inventory, such as the removal of small quantities or other potentially insignificant material and/or energy outputs, would render the exercise incomplete and the criteria unfulfilled. Since no information is collected about impacts at this point, it is incorrect to remove any materials or energy from consideration on the basis that they may not cause significant impacts. The fact is, they may.

Clearly, a long list of RIOs must be completed, invariably involving the collection of data which has not been collected before. The first time the method is applied to a given technology, it will be time and resource consuming. However, the information exists, and the entries are known from design details of each production Stage. Furthermore, the method continues to become increasingly practical as information is gathered, while new information technology is a further asset in providing the means to compile RIO Inventory information in accessible form.

Nevertheless, it should be noted that current barriers to gathering RIO Inventory information exist, including problems of confidentiality, unfamiliarity and resources required to keep and produce such information. However, RIOs cannot and should not be regarded as confidential (commercial) information. Impacts affect others, and in tracing these impacts (which unequivocally should be public information), RIO information is required. Once data has been gathered once, problems with unfamiliarity and gaps in existing systems are reduced. Provision of RIO information, treated as, for example, pollution licence or planning permission application information, is not technically problematic. There are resource implications for providers (developers and their designers and engineers), albeit relatively minor. Given good practice of free exchange of data in transparent and accessible form, these can be minimised. However, in the event of continuing reluctance of providers to engage in such good practice (and the bottom line is invariably costs and resources), it may be appropriate to make provision of RIO data a regulatory requirement (see Chapter 11).

While the RIO Inventory is demonstrably accurate, given the test of completeness, a note must be made regarding the potential for deviation between design in prospect and reality. For example, during construction, more or less material may be used that is explicitly required by the architects' or engineers' design. Due to human error or unforeseen circumstances, more waste may be generated than expected during construction. A particular natural resource on which the fuel cycle project is based may be unavailable in the event, and a new source used for the natural resource. Potentially most significantly, the generation plant in operation may not perform to design expectations. If the plant operates at a lower than design load factor, it will produce less power and this will affect normalised inputs and outputs for fixed components. Unforeseen design-reality deviations do not comprise an inaccuracy within the output analysis method, but it should be noted that quantities based on established practice should be used wherever possible. Where reality checks indicate that quantities should be changed from those on design drawings, or where there is uncertainty over precise quantities, this should be recorded in the RIO Inventory. This ensures that the RIO Inventory contains the most appropriate quantities of each RIO.

In summary, the output analysis method meets all the criteria set. The result of its application is a complete list of RIOs; the RIO inventory. The next task is to trace these RIOs through the environment in order to establish where and how they may cause changes which result in environmental impacts. This tracing process is at the centre of Steps 4 to 6 of the systematic framework, collectively called the pathway analysis method, which is the subject of Chapter 8.

8. PATHWAY ANALYSIS METHOD

As outlined in Chapter 6, the pathway analysis method is designed as an objective means of identifying the physical origins of each environmental impact arising from a human activity, in this particular case, a full fuel cycle within the ESI. It has already been established that this involves tracing physical quantities of energy and/or matter from the point at which they leave the production process, through the environment. However, the details of how this is undertaken, when the tracing stops, and what data are required as the output from the pathway analysis method are issues which now require detailed description. Much of the material used in this Chapter is drawn from detailed development and demonstration of the pathway analysis method undertaken elsewhere (Horne, 2000a, 2000b).

Although the outline of the pathway analysis method presented in Chapter 6 is criteria-led, as with the output analysis method, it is critical to the success of this method that the criteria set out must also be referred to and complied with at the level of detail. A summary of the pertinent criteria it must meet is presented in Section 8.1. Sections 8.2 to 8.4 detail Steps 4 to 6 of the systematic framework, respectively, and, in Section 8.5, a demonstration of the method is summarised. In Section 8.6, the detailed pathway analysis method as described is compared against the criteria set out in Section 8.1 to establish the level of compliance.

8.1 Pathway Analysis Method Criteria

Many of the criteria laid out in Section 6.1 apply directly to the pathway analysis method. Common to all the methods in the systematic framework are the requirements for a systematic approach, with clear boundaries between system components and full transparency and accessibility of both data and process. Likewise, each method must be broken down into sequential Steps and be as practical and simple as possible in operation. For the pathway analysis method, practicality and simplicity must be demonstrated, and efficiency must be maximised through aggregation and data transfer, within boundaries set by other criteria. The pathway analysis method must also be fully inclusive. In other words, it must be designed to identify all possible changes in the environment arising from a given activity, including "gaps", where RIOs are untraced or missing from the pathway analysis. It therefore follows that checks must be incorporated to ensure data quality and completeness. These criteria are demanding, but the pathway analysis method must comply with them to minimise the weaknesses and pitfalls of current practice. They have been achieved at outline level in Chapter 6 and must now be complied with at the level of detail.

8.2 Step 4: Pathway Identification

The pathway analysis method starts with the quantified energy/materials which result from application of the output analysis method, in the form of the RIO Inventory. Each RIO is mobile in two senses. It may be transferred via pathways from place to place through time, and it may undergo alteration. Energy may be transformed into various types, and materials may undergo chemical and physical changes. For each RIO, every potential pathway must be identified, by tracing all the possible physical routes and changes which the RIO may undergo after leaving the production system. Quantities of RIO are entered along each pathway, tracing the dissipation and/or concentrations of material/energy. Hence, differentiating between different receiving environments (and attributing quantities of RIO to each) is central to the pathway identification process.

There are two conflicting pressures on the pathway identification process. Firstly, there is the need to reduce the task to a practical size, which requires interpolation, extrapolation and transfer of current data and understanding between various RIO-pathway-change situations. Secondly, these RIO-pathway-change relationships are complex and natural environments are unique, which acts against such practices on the grounds that loss of accuracy will result. The resolution of these pressures is to adopt an appropriate methodological approach which reconciles them by ensuring that data losses are not unacceptable, while allowing necessary aggregation and transfer of generic data where possible. There are two elements to this approach. Firstly, a pathway Coding system must be used in pathway identification. Secondly, all pathways and destinations of RIOs are recorded on a Pathway Inventory, as introduced in Chapter 6, and caveats must be attached to individual entries which involve interaction with complex natural systems, as appropriate (see Section 8.3).

The pathway Coding system is a simple but precise means of recording the path taken by a RIO through the environment. Because of the diversity and complexity of environmental impacts, some form of characterisation or grouping of them is common to environmental assessment methods, and it is possible to categorise the environment according to its physical (inorganic and organic) attributes. At the simplest level, four subsystems can be identified at each pathway point; air, land, water and organic (living) matter. Within each, it is necessary to identify further subsystems to assist in the identification of specific environmental changes. Table 8.1 provides a generic list of pathway Codes based on these four subsystems of the natural environment. It should be noted that, in any given site-specific case, further Codes are added within each subdivision which indicate specific sites or areas, so that specific quantities of given burdens can be quantified. Each subsystem boundary is defined by the physical extent of the RIO at each destination it reaches on its way through the environment.

Air/atmosphere	
A1	Low level, local surroundings
A2	Medium level atmosphere, regional
A3	High level atmosphere, global
Land use/lithosphere	
L1	Landfill site or other disposal facility
L2	Built environment and transport
L3	Agricultural area
L4	Planted and managed forest
L5	Natural forest
L6	Natural mountainous zone
L7	Natural arid zone
L8	Natural coastal zone
L9	Natural wetland
L10	Other natural/semi-natural area
Water/hydrosphere	
W1	Potable groundwater supply
W2	Potable surface water supply
W3	Other groundwater
W4	Other fluvial, lakes
W5	Estuarine waters
W6	Marine waters
Life forms/biosphere	
B1	Plant life
B2	Animal life (excluding humans)
B3	Humans

Table 8.1 Generic Pathway Codes

Notes: Codes B1 and B2 can be further subdivided, for example, by genus and species. Code B3 applies to physical matter/energy (RIOs) ingested/in contact with humans. All Codes can be suffixed to identify particular site-specific locations.

In practice, Coding is undertaken as follows. For each RIO in the Inventory, all the potential pathways can be identified, by first considering all the environments which the RIO may initially affect *directly* (these are called RIO-1s). The quantity of the RIO which enters each environment is stated. Then, for each RIO-1, a further list is required for all the environments the RIO-1 will pass *directly* into, after leaving its first destination. These second destination pathways are called RIO-2s, and the quantities again are required. For each RIO-2, a list of RIO-3s is then produced by the same method, and the process is repeated until no more destinations are found. Hence, as the RIO passes through each subsystem in the environment, a new Code ending is added to its Code String to reflect each destination. For example, the RIO sulphur emitted from a coal power station stack first enters the atmosphere and will split into three portions as it enters the lower, middle or upper atmosphere, at which point the pathway splits into three discrete corresponding Codes, A1, A2 and A3. Among many other RIO-2s, a fraction of the A1 element may then enter human lungs, while another may enter a river. Hence, the two pathways A1/B3 and A1/W4 are identified and must be quantified. The smaller the sub-systems can be made (that is, the further each can be broken down), the more accurately the flow of sulphur through it can be predicted. Therefore, the A1/W4 pathway will invariably be broken down further. A new Code may be added to show one of the next destinations as a specific stretch of a stream, which may be given the Code WS1, making the pathway Code String A1/W4/WS1. Note that any discrete subsystem can be given a new Code - the only provisos in assigning a new Code are that it should indicate generic environment type and be different from all existing Codes. Further Codes along this pathway may be required to show quantities of RIO entering specific animal species locally. The general Code String for this would be A1/W4/WS1/B2, although other B-prefix codes could be used to identify individual species.

Points at which RIOs reach a state of prolonged rest or lack of interaction are termed "sinks". Thus, sinks are a special type of subsystem, and provide a RIO destination with a longer residence time. Sinks are therefore stores or pools, which may vary through time, and which may become more or less stable due to the various dynamic factors which create and maintain them in equilibrium. Fundamentally, sink size is determined by input/output rates. Each pathway ends when the RIO is dissipated entirely into pre-existing natural systems or otherwise reaches a natural sink, predictable cycle, state of dynamic equilibrium or long-term residence.

It is essential that the pathway analysis method produces data which are complete, so that potential impacts are not overlooked, and that the data are sufficiently accurate for the purposes of valuation. Tests to ensure that these conditions are met are therefore required. To check all pathways are considered, once all environmental changes are listed, along with all human consequences, RIO quantities at each pathway destination can be summed to provide a mass balance comparison with the known total RIO release. Where errors or losses of RIO are found, stated gaps are left in the Pathway Inventory to reflect this, or estimates may be included based on known RIO releases and assumptions (for example, using modelling) of likely points of loss. Where energy is the form which the RIO takes, a similar approach is used, but for energy instead of matter. The principles of mass and energy balance accounting are well-established, and are central to the important task of checking that inputs are equal to outputs. In this case, invariably, it is the extent of completeness of the Pathway Inventory which is established, along with the approximate locations and quantities (by RIO) of gaps.

To check whether data are sufficiently sensitive, the range of likely actual values must be established, using information and assumptions about variance and margin of error. Then, highest and lowest estimates can be established and used in addition to the calculated central point in all subsequent calculations up to and including valuation, as introduced in Section 6.4.3. This test is a basic form of sensitivity analysis, and it can be made more sophisticated if necessary, for example, by the use of confidence intervals. The outcome of this test is that, following valuation, the potential range of values is found to be significant or not. If it is, then the source of sensitivity is known and can be re-examined with a view to improving data by reducing the potential range to within non-significant limits.

8.3 Step 5: Pathway Inventory

Each pathway destination for each RIO has a discrete entry on the Pathway Inventory. Each entry also contains the quantity of the RIO, and indication of the subsystems through which the RIO has passed up to that point. The latter is presented in the form of the (unique) Code String. Thus, each RIO pathway may give rise to many Pathway Inventory entries – one for each pathway destination. The following rules must be adhered to in compiling the Pathway Inventory:

- RIOs are not mixed (unless they are exactly alike and are found in the same site specific environment Code);

- Quantities are explicitly stated for each entry (i.e. for each pathway destination point);

- Any unknowns, gaps and approximations of data are included as separate Pathway Inventory entries;

- Any data loss indicated by mass/energy balance accounting is given a separate Pathway Inventory entry;

- All Pathway Inventory entries for a given RIO are represented by a unique Code String (differentiation is undertaken where necessary by the application of site specific Codes, for example, where two natural forests are subject to the same RIO, they are assigned site specific Codes).

To assign quantities of RIO to each item on the Pathway Inventory, each part of the atmosphere, lithosphere, hydrosphere and/or biosphere into which each RIO pathway passes must be studied. Each destination provides a potentially complex physical-organic system interaction which must be described, invariably using some means of measurement, modelling and prediction. Some of the typical issues and concepts involved in this process are outlined in Sections 8.3.1 to 8.3.4 below. Where such complex systems lead inevitably to estimations or uncertainties in Pathway Inventory quantity data, caveats must be attached to the appropriate entries as indicated in the rules above.

8.3.1 Air/Atmosphere

The atmosphere is a series of layers of mixed gases, in a dynamic system governed by the principal energy sources of solar and reflected radiation, and gravity. The distribution of a RIO into this subsystem varies according to the nature of the RIO. Initially, it may enter one or more of the three parts of the atmosphere, as represented by the generic Codes presented in Table 8.1. These are arbitrary but generally conform to three classes or types of atmospheric pollutants; those which are short-lived or too dense to rise above lower levels (local), those which reach sufficient levels to get transported by regional weather systems (regional), and those which have high residence times and can reach higher atmospheric levels over long periods, allowing them to travel long distances across the global atmosphere (global).

Although reactions do take place which may alter the initial form of a given RIO, the measuring and monitoring of atmospheric systems and our understanding of how they work, even on a localised level, has rapidly improved over recent years. Relatively sophisticated computer models are now available to allow reliable prediction of the dispersal patterns of many common RIOs within the atmosphere, and even to predict if/where/when these may change form, be deposited into hydrosphere/lithosphere, etc.

8.3.2 Land Use/Lithosphere

RIOs which enter land have varying distribution patterns according to land use, and the purpose of the generic Codes is to allow an initial differentiation of RIO quantities into basic land use types. The Codes (Table 8.1) may appear to indicate ecosystems rather than lithospheric systems (rock types, etc.), but specific life forms are not included, and each Code simply refers to a specific land use type. They are used for RIOs which enter inorganic parts of the land environment, such as soils or rock. Where the RIO then enters organic parts, subsequent biosphere destination Codes will be required (see Section 8.3.4 below). Otherwise, in general, energy/matter pathways within land tend to be localised in nature, unless the RIOs themselves are gaseous or liquid. Specific parcels of land which are unique are given unique, site specific Codes, as may any lands with similar land uses affected by similar RIOs, for differentiation purposes.

8.3.3 Water/Hydrosphere

The Codes for land covered by water are designed to provide for differentiation of types of water bodies. As with atmosphere and lithosphere, unique Codes are allowed for site specific water bodies. Since water is the receiving environment, and it is often mobile and has complex properties, there is generally greater possibility of transfer of the RIO through the hydrospheric system than in land. Thus, there is a need to consider mobility and dispersion modelling of RIOs within groundwater, fluvial systems, lakes, estuaries, seas and oceans, as well as the potential for chemical changes and interference with current patterns of energy and matter equilibrium within the system. As in the case of atmospheric modelling, this invariably involves both monitoring and the use of predictive computer modelling techniques.

8.3.4 Life Forms/Biosphere

In use, the biosphere is the most complex of all the Coding groups. Many life forms may be affected by a given RIO, for example, through food chains or mutual ecosystem dependency relationships. Therefore, detailed study of the life forms and their interactions with each other and with each RIO is needed, and this will invariably result in the generation and use of a large number of life form-specific Codes to describe individual quantities of RIO in individual parts of the biosphere. There is considerable complexity involved in measuring quantities of RIO reaching each part of the system, given the interdependency of organic systems. For example, concentration of a particular RIO in a plant species near the base of the food chain may lead to entirely different concentrations of the same RIO in birds of prey or mammals near the top of the food chain, due to natural concentration and accumulation processes. Nevertheless, the task is limited to measuring RIOs physically entering the subject or subsystem. Indeed, establishing quantities (and uncertainties) of each RIO reaching each part of the biosphere is the sole aim of the exercise. Discussion of effects is precluded, as is any consideration of pollutant thresholds or dose-response relationships. Although a separate generic Code is assigned to humans, there is no difference in the pathway analysis as it relates to this species (indeed, it is envisaged that other species are also assigned unique Codes during analysis). In accordance with pathway rules governing all pathway analysis, this Code is used only when RIO energy/matter physically contacts humans. In reality, there are two main groups of possibilities here; those where the pathway continues after humans, for example, where physical matter passes through them and into waste streams; and those where the pathway ends at humans, for example, where energy in the form of noise emanating from a production process-related structure reaches humans.

8.4 Step 6: Environmental Change

Once the Pathway Inventory has been produced, the resultant changes in the environment can be quantified. As introduced in Chapter 6, this inevitably involves predicting, modelling and otherwise describing the interaction between RIOs and (complex) natural systems. The Coding of receiving environments and quantification of RIOs carried out during pathway analysis is the means of defining different natural systems and parts thereof, for which relevant specialists can then become involved in measuring specific environmental changes. The aim of Step 6 is to measure objectively the change each Pathway Inventory entry will (or may) cause in the environment. Change is any shift in system equilibrium caused by the RIO energy/matter. Therefore, the result of this Stage is an inventory of environmental changes; the Environmental Change Inventory, derived from the Pathway Inventory. A set of rules governs the process of identifying each environmental change from each Pathway Inventory entry, as follows:

- Each environmental change is defined as a quantified shift in equilibrium of an inorganic system, or a quantified change in biomass, health or population of an organic system;

- Pathway Inventory entries are not summed or aggregated for change assessment (unless they act similarly, in which case the summed environmental change should be transparently attributed to the appropriate Pathway Inventory entries);

- There are two principal types of environmental changes; those caused by direct physical contact with a Pathway Inventory item, and those which arise elsewhere as a result of a causal link – this latter requiring particular attention as it is more easily overlooked in environmental change analysis;

- Any unknowns, gaps and approximations of data must be included in all estimates (i.e. caveats must be applied and sensitivity tests undertaken, with ranges of uncertainty being provided along with a central estimate);

- All environmental changes should be presented in site/type specific groups, and each examined and compared systematically to other group entries and to other groups for potential system-level, synergistic, cumulative, or threshold effects. Where present, the combined environmental changes should be adopted;

- Every item on the Environmental Change Inventory should retain its unique Pathway Inventory Code (where Inventory items have been summed to assess cumulative/synergistic effects, the Code should indicate the range of Pathway Inventory items it represents, by use of footnotes, hyphens, etc.), and should be supplemented with a brief descriptor to provide transparency and accessibility.

There is a relationship between the amount of a given incidental output and the response of the environment, the so-called dose-response relationship. This must be described as part of the process of identifying and quantifying environmental changes. The relationship between dose and response is often non-linear. For example, it may contain thresholds; levels at which a small increase in incidental output will cause a large response in the form of a disproportionately increased environmental impact. Also, as indicated in the rules above, environmental changes can occur as a result of a causal link - where there is no direct change in RIO energy/matter, but where it causes another (indirect) change in the environment. As with the pathway identification process, the creation and use of sophisticated computer models is necessary to combine and compile input/output data and assist in predicting changes resulting from flows of materials and/or energy through the environment. Different environments raise different implications for this process, and these are illustrated in Sections 8.4.1 to 8.4.4 in terms of each environment type/site along each pathway.

8.4.1 Environmental Changes in Air/Atmosphere

The prediction of environmental changes within the atmosphere requires a greater understanding of the system than predicting atmospheric pollutant loadings, as in pathway analysis. Here, the main task is to identify the amount of change arising from interference with the flow of energy and matter through the system. As with pathway tracing, this change is measured in objective quantities, although the unit of RIO dose may be different from the unit(s) of change(s). For example, while the dispersal of fossil fuel-produced CO_2 from a given production process can be relatively simply predicted, its complete impact in terms of global climate change has only recently been agreed with any broad consensus. The effect of this on weather patterns and sea levels is, even now, only reaching any consensus. Nevertheless, the criteria requirements of transparency, stating confidence levels in results and applying ranges of data under uncertainty allow even relatively poorly understood environmental changes to be predicted.

RIOs which enter the atmosphere are often relatively benign (apart from their capacity to interfere with energy flows), until they are deposited out of the atmosphere, whereupon they enter one of the categories dealt with below. Hence, the atmosphere is often a mechanism of transport, rather than a direct physical source of environmental change. Incidentally, the exception regarding energy flows is of major importance, since it is broadly agreed that global climate change due to interference with atmospheric energy flows is currently the environmental change of single greatest concern to the human race, but this is a matter for valuation (see Chapter 10). Also, the link in this case is causal rather than involving direct contact with the physical matter of the RIO, so it is also a good example of the need to systematically identify environmental changes indirectly caused by RIOs. It therefore illustrates an important difference between pathway analysis and environmental change analysis. In the former, it is physical quantities of RIO alone which are measured, whereas in the latter, it is the results of these quantities which are studied – and these can occur in an environment which is remote in space, type and time from the location of the causal quantity of RIO.

8.4.2 Environmental Changes in Land Use/Lithosphere

Environmental changes affecting land are evidently inorganic, and are principally quantified in terms of:

- Changes in soil/ground concentrations of RIO or RIO-derived materials;
- Changes in soil/ground chemical conditions (acidity, oxidation-reduction potential, etc.);
- Changes in soil/ground physical conditions (structure, porosity, permeability, etc.).

It is possible but unlikely that significant changes in land could accrue from RIO-energy (energy directly from RIOs).

8.4.3 Environmental Changes in Water/Hydrosphere

As for land, environmental changes in the hydrosphere are inorganic. However, there are particular complexities in water-based natural systems, partly due to the physical properties of aqueous bodies, and partly due to the unique chemical properties of water. Groundwater presents problems due to the lack of ready access to measure and monitor the system physically, while marine waters present similar problems at depth. They also exhibit physical and chemical complexity, with various inputs and outputs and internal processes operating on both local and large scales. Similar problems are presented by estuarine and shoreline environments, where additional complexity occurs due to the juxtaposition (and mixing) of different systems with different chemical and physical properties. Nevertheless, the principal changes that may occur within the inorganic parts of any given water-based system can be reduced to the following categories:

- Changes in concentrations of RIO or RIO-derived materials;

- Changes in chemical conditions (acidity, oxidation-reduction potential, etc.);

- Changes in physical and dynamic conditions (currents, energy profiles and flows, suspension and solution loads, deposition processes, etc.).

8.4.4 Environmental Changes in Life Forms/Biosphere

The objective measurement of changes to organic systems, particularly at species level, is the single most challenging of the pathway analysis tasks. Given that many species have not yet been identified, and they are being affected by numerous industrial processes, it is clear that lack of data will present itself as a problem. However, taking as a starting point the total number of possible species and the total number of RIOs is neither productive nor appropriate for the pathway analysis method, which operates from the simple to the complex, rather than the opposite. Indeed, for any given RIO, a simple classification of the number of possible changes to which any given species (or range of species, if appropriate) may be subjected, provides the appropriate starting point, as follows:

- Changes in population levels;

- Changes in population health;

- Changes in habitat equilibrium and dynamics (including food/water supply, shelter/defence, parasites, diseases, predators, etc., where interaction with other related Pathway Inventory entries may be required).

The first two are relatively straightforward in concept. However, the third presents difficulties for study at species level, since the wider issues raised about population survival and health will inevitably relate to other Pathway Inventory items. For example, water supply is affected by changes in the hydrosphere, which may be affected by the RIO under consideration (or another RIO) under different Pathway Inventory entries. Hence, assessing this third type of change involves systematic comparison and combination of numerous potentially related Pathway Inventory entries. Also, although the dose is known, the effect may be governed by a number of factors. Firstly, the relationship may not be linear, so that a given level of RIO (dose) may cause different responses (amounts of change) in cases where background or previous levels of RIO or related agents differ. A threshold or thresholds may occur, at which a major change in response occurs when a particular level of dose (or level, load or residue) is reached. Moreover, establishing thresholds and dose-response relationships is not necessarily enough to allow prediction of environmental change. Other variables include residence times of RIOs, variations over time (including in inorganic systems), and reactions and mixtures of RIOs acting in combination. The latter can lead to cumulative effects, or even synergistic effects, where the combination of two (or more) RIOs have an effect greater than the sum of their individual effects. This can occur due to thresholds being reached, or by combination reactions, or by a source of stress being aggravated simultaneously.

There are a number of proven techniques for overcoming the inevitable complexity of measuring environmental changes at individual species level. Clearly, field sampling of actual cases must be undertaken where possible, and although this presupposes that

RIOs have already been released into the environment, it is also possible to transfer data into predictive situations. Indicator species are often used in surveys, to reduce the number of species studied to a manageable range. These are species which are known to be sensitive to particular stresses but, ideally, are otherwise numerous and relatively easy to locate. They are used to give an indication of the general health of an ecosystem or a range of species.

Exactly the same initial approach is taken to measuring environmental changes in humans as for other organisms; loss of physical life and function (population levels and health). However, examination of changes in habitat equilibrium and dynamics through comparison with other related Pathway Inventory items and associated changes are not included in environmental change analysis. Neither is the establishment of causal links between systems involving changes in human organisms. These tasks are excluded here because they involve subjective judgement and all subjectivity is dealt with in the valuation method. Indeed, they are at the centre of the process of producing data on human consequences in the next Step (see Chapter 10). The Environmental Change Inventory entries with the Code for "humans" contain only physical, objective, quantified descriptions of changes in humans arising from RIOs which are in direct contact with humans. Environmental changes record change in function or experience (in this case, in humans), whereas human consequences record the result or, more specifically, *outcome* in terms of change in quality/quantity of life for humans.

8.5 Pathway Analysis Method Demonstration

The aim of demonstrating the pathway analysis method is to show how it can be undertaken practically, by examining how each Step can be undertaken and to what extent data requirements can be met with current knowledge. In a full impact assessment exercise, the pathway analysis method would need to be applied to all

RIOs from the fuel cycle in question. However, for demonstration purposes, it is sufficient to select a single RIO at the outset, provided this has been relatively well researched, but not a special case or a particularly simple case, since this would raise questions over whether and how the method could be applied successfully to more complex RIOs. A RIO which satisfies these requirements is the sulphur dioxide (SO_2) stack emission from a coal fuel cycle electricity generation plant, or, more specifically, the sulphur (in whatever form). A more detailed discussion and justification for this selection is presented in Appendix C, along with substantial background information and literature review on the demonstration.

As sulphur from a stack is traced through pathways, it inevitably splits and disperses. Hence, while a complete RIO is the starting point, for demonstration purposes, only a part of this RIO needs to be traced through to the point of producing environmental change data. This approach provides an adequate demonstration of the pathway analysis method and is made possible because the pathway analysis method is in the form of separate, sequential Steps. It is important to differentiate between the natural cycle of sulphur through the environment, and the tracing of the demonstration RIO sulphur. The sulphur cycle is the collection of processes in which sulphur is stored and transferred through the global system. Although there is considerable overlap and mixing between RIO-sulphur and the sulphur cycle, the latter does not necessarily provide all the pathway possibilities. These can only be determined by following Steps 4 to 6, using Coding and rules, as specified above.

The first destination of the demonstration RIO is the atmosphere. It will be divided between the three generic atmosphere Codes in Table 8.1; A1, A2 and A3. Quantities are determined by two means; measuring of sulphur at various levels under experimentally controlled conditions and modelling of dispersion plumes. However, following this, the next (RIO-2) pathway destinations will include environments into which sulphur is deposited out of the atmosphere. These will include various subsystems, and the smaller the sub-systems can be made (that is, the further each can be broken down), the more accurately the flow of sulphur (or its causative agents) through it can be predicted or measured. Figure 8.1 shows a simplified schematic representation of how the demonstration RIO is dispersed through pathways.

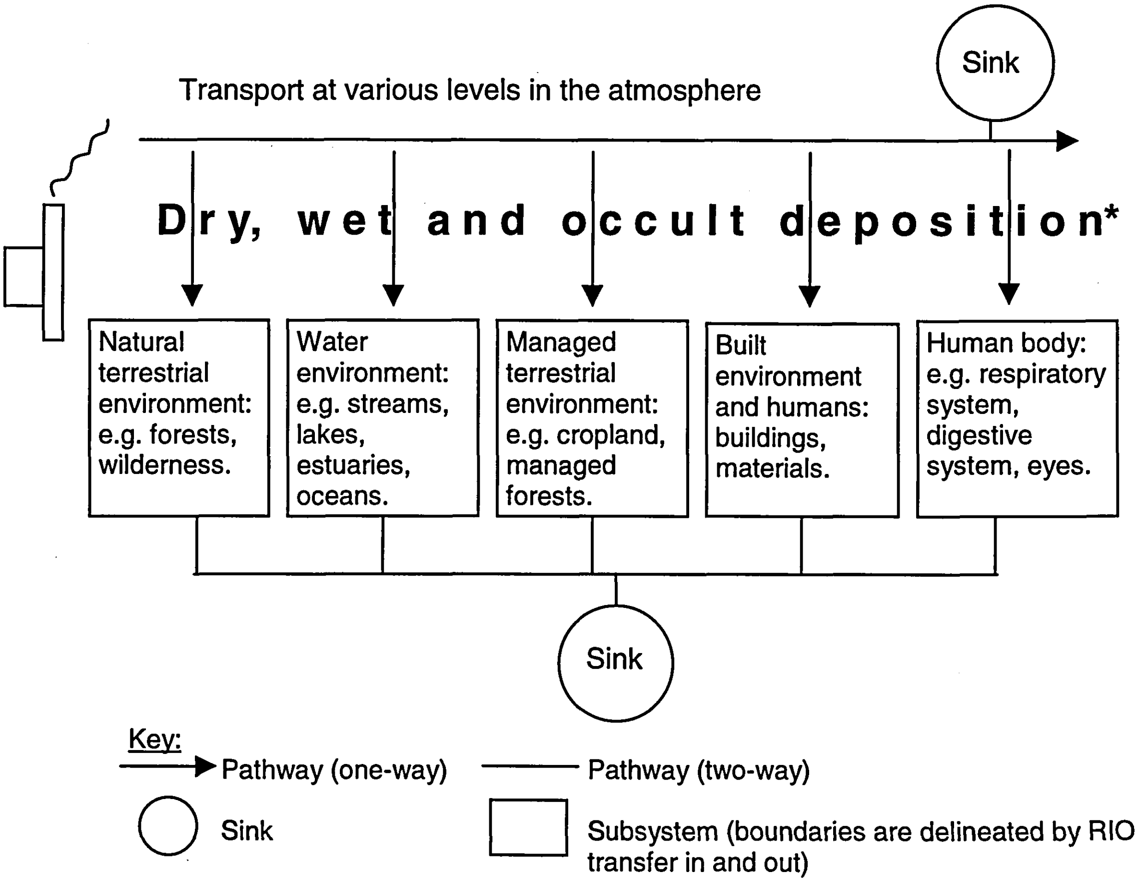


Figure 8.1 Simplified Schematic Representation of How the Demonstration RIO is Dispersed Through Pathways

*Note: Dry deposition refers to particulate/dust deposition without water droplets present; wet deposition refers to deposition through precipitation; and occult deposition refers to wet droplet deposition without precipitation, that is, through direct contact with mist/fog/smog.

However, this large scale model is too simplistic to satisfy the requirements of the pathway analysis method. Therefore, the task of pathway identification is to break each of the large subsystems into smaller ones, as described in Section 8.2 above. The "water-tight" method which is needed in order to identify all the possible pathways involves Coding each stage of physical progression of the RIO (i.e. each pathway destination). The list of generic Codes is the starting point for this exercise (Table 8.1), although the choice of Codes themselves is immaterial, provided each is unique and each RIO is traced through each pathway using a unique, sequential Code String. With Coding complete, each Code String is entered on the Pathway Inventory, with a separate, quantified entry for each destination in each String. Then, for each entry, environmental changes are identified using the rules and procedure described in Section 8.4, to produce the Environmental Change Inventory. An example of a part of a data presentation table, which contains Pathway Inventory and Environmental Change Inventory information for part of the demonstration pathway, is presented in Table 8.2.

Pathway Inventory for Pathway A1/B3/B3RS/	RIO quantity	Code	Causal links checked?	Causal change? (If no, it is a direct physical change)	Environmental Change Inventory	Quantities
Nasal Cavity		/B3NC	No	No	mucus generation	
Pharynx		/B3Ph	No	No	?? (no change identified)	
Oesophagus		/B3Oe	No	No	?? (no change identified)	
Trachea		/B3Tr	No	No	?? (no change identified)	
Bronchial and Lung activity		/B3Lu	No	No	mortality	
		/B3Lu	No	No	morbidity: long term bronchial illness/lung cancer	
		/B3Lu	No	No	episodic short term bronchial problems	
??? (gap = total A1/B3/RS/ - A1/B3/B3RS/B3NC + A1/B3/B3RS/B3Ph + A1/B3/B3RS/B3Oe + A1/B3/B3RS/B3Tr + A1/B3/B3RS/B3Lu)		/B3Lu/?	No	No	?? (gap and no change identified)	

Key:
?? Potential direct environmental change missing/unaccounted for
??? Possible pathway missing

Table 8.2. Pathway Inventory and Environmental Changes Data Presentation Form

Note: A1/B3/B3RS/ is pathway coding, denoting that the sulphur has passed into the atmosphere mixed layer, then directly to the human respiratory system. Further codes are added to this Code String as shown in column 3.

Note: Causal links have not been investigated.

Note: Quantities columns are indicated for completeness: quantities would be entered in a full analysis.

The Code String A1/B3/B3RS/, which Table 8.2 refers to, indicates that the RIO-sulphur has passed into the atmosphere mixed layer, then directly to the human respiratory system. Further codes are added to this Code String as shown in the third column, to describe the next destinations of fractions of this RIO. Descriptors of environmental changes provide clarity and help immediate understanding of the entry (in accordance with transparency and accessibility requirements). Therefore, although these are not essential to accuracy of impact change identification, they are a necessary part of the Environmental Change Inventory. Double question marks denote areas where pathway inventory production has identified that there are potential direct effects which have not been described. Therefore, this potential remains unknown, and must be coded and specified, and, if possible, quantified in terms of RIO involved. Thus, the presentation of data illustrates current gaps in data and knowledge. Quantities of RIO, environmental changes or indirect links and associated changes are not required for demonstration purposes, but this information is required in a complete exercise.

Table 8.2 must be seen as a small part of a much larger, multi-layered table containing information in separate steps and on many different subsystems and RIOs. The most appropriate means of portraying this information is in spreadsheet form, linked to spatial information on a Geographic Information System (GIS) base, such as modelling and dose quantities, etc. Integration in such a way minimises the errors and maximises the efficiency of the exercise.

A review of the state of knowledge regarding environmental changes arising from the demonstration RIO (see Appendix C) demonstrates that this is sufficient to allow Coding and quantification of RIOs through pathways, including quantification of changes, both direct and indirect. "Gaps" will occur, where quantities of some RIOs are left untraced, and the approximate location and size of the RIO "loss" can thus be identified for further study. These unknowns are entered as explicit quantities and at specific destinations, with unique codes. For example, if an apparent quantity loss is established between A1/B3/B3RS/B3Lu and subsequent destinations on the pathway then, in addition to all subsequent known destinations being entered on the pathway inventory, a Code A1/B3/B3RS/B3Lu/? is added, with the quantified apparent loss entered. Where estimations or uncertainties occur in pathway inventory quantity data, caveats must be attached to the appropriate entries.

While data and knowledge shortages remain a challenge to the data-intensive pathway analysis method, it is clear that this can be met with current knowledge. Since the pathway analysis method points out very clearly where the gaps are, it is itself a useful potential vehicle for identifying areas of further research needed. Such gaps are expected to remain within the tracing process, with quantities of outputs, and therefore of their effects, remaining unknown for the foreseeable future. However, the pathway analysis method is not only designed to detect and illuminate these gaps, it is also designed to produce useful data outputs for use in valuation even where such gaps remain.

With improving knowledge and Information Technology, data are theoretically more widely available and model capacities are much greater. Therefore, data requirements will increasingly be met within practicality and resource constraints. The literature review for the demonstration RIO in Appendix C demonstrates that the reliability of modelling and monitoring data regarding the atmospheric dispersion and next pathway destination of anthropogenic sulphur from stack emissions is already sufficient to satisfy criteria requirements and is rapidly improving. Local and regional models capable of necessary accuracy to predict relatively low concentrations of pollutants around buildings are now available (for example, Owen, 1999). Sulphur transfer through the natural terrestrial environment and the water environment can also be accomplished using current knowledge. It follows that a Pathway inventory can be produced for the demonstration RIO. Furthermore, while the environmental changes from the demonstration RIO are undoubtedly complex and highly dynamic, with a multiplicity of overlapping change cycles, modelling is increasingly up to the task of predicting these with sufficient accuracy too, as demonstrated in Appendix C. Thus, even with complexities such as long lag times between doses and responses and variations in apparent recovery times following changes, a sufficiently complete and accurate Environmental Change Inventory can be produced for the demonstration RIO, given the current state of knowledge regarding environmental change measurement and prediction.

In summary, the capability and practicality of the pathway analysis method has been successfully demonstrated, via a review of the state of knowledge and data available, and a description and demonstration of data presentation and procedure. While data and knowledge shortages are a challenge to the pathway analysis method, it is clear that this can be met with current knowledge. Moreover, the only alternative would be to retain the current subjective approach to pathway and impact identification, which is unacceptable and illogical, since it presupposes the outcome of valuation. Therefore, not only *can* the pathway analysis method be applied, it *must* be, if impact measurement is to be undertaken within the bounds of acceptable accuracy, objectivity, rigour and transparency.

8.6 Pathway Analysis Method Rules and Criteria Compliance

Rules set out for the Pathway Inventory production process in Section 8.3 are met, ensuring that subjectivity is minimised in pathway identification. It allows differentiation between unlike materials and energy sources, as well as different receiving environments, and it is designed to operate from the generic to the specific. It also incorporates deliberately careful and inclusive consideration of uniqueness of each receiving environment before predicting next pathway points and environmental changes. Furthermore, the Pathway Inventory is an inherently transparent database, with pathway Coding showing origins and "ancestry" of each entry, through a Coding String. Likewise, rules set out for the Environmental Change Inventory method in Section 8.4 are met. Each environmental change is precisely defined and assessed for uniqueness prior to any aggregation and each environmental change is transparently attributed to the appropriate Pathway Inventory entry or entries. Both principal types of environmental changes (direct and indirect, or causal) are accounted for, with specific methods of identification, while unknowns, gaps and approximations of data are explicitly included in the Environmental Change Inventory. Finally, systematic comparison of items on the inventory is also incorporated, to detect potential system-level, synergistic, cumulative, or threshold effects.

While it is recognised that objectivity cannot ever be entirely eliminated, this does not preclude the aim of minimising it where it is unwarranted, and the basic structure of the pathway analysis method acts against subjectivity from creeping into the process. Examples of this problem include discussion of environmental impacts during pathway analysis, or scoping out of potentially significant (in human terms) RIOs by pathway analysts, both of which may occur in studies with less well-defined methodologies. Transparency, accessibility and consistency are maximised by including uncertainty, gaps and unknowns in data, adopting appropriate reporting and data presentation formats, and by using detailed and comprehensive Coding, accounting and inventory tools. Whereas current, subjective approaches to pathway analysis may or may not identify gaps and unknowns, they do not and cannot identify all of them. Since knowledge is always finite, identifying the size of potential unknowns is necessary if environmental impacts are to be identified with any certainty or completeness. The identification of gaps is central to the pathway analysis method, because the knowledge/data shortfall must be known, and because this information itself indicates where effort is needed to reduce them in the future. Similarly, efficiency and practicability are maximised through appropriate data processing, including aggregation where applicable.

Reliability can only satisfactorily be tested through well established methods of blind parallel trials and, ultimately, prolonged practical operation. However, validity is indicated by the extent to which what is actually measured corresponds to what it should measure in theory. Therefore, validity largely applies to the results of valuation (see Chapter 10). However, both theoretical and practical validity are established for the pathway analysis method, through compliance with other criteria and in the practical demonstration exercise in Section 8.5. Pathway Inventory and Environmental Change Inventory completeness and data sensitivity are established through the application of mass and energy balancing and sensitivity analysis, based on well-established principles of systems analysis. The only remaining criterion for compliance is whether the data requirements of the next task of the systematic framework, the valuation method, are met. The data outputs of the pathway analysis method consist of the Environmental Change Inventory (Step 6), and the first task of the valuation method which follows is the compilation of human consequence data from this (Step 7). This involves the production of four items of data for each consequence; risk, duration, number of people affected and quality of life outcome. The risk of the impact occurring, the duration of it, and the number of people potentially affected are relatively straightforward and further details are presented in Chapter 10. There is more difficulty with the measure of quality of life outcome for each human consequence, not least because a single scale of quality of life outcomes is required before measurement of this can be undertaken. This is discussed in detail in Chapter 9.

9. THE QUALITY OF LIFE OUTCOMES INDEX

Environmental changes arising from a product or service life cycle can cause a wide range of consequences for a person's quality of life. These include emotional, mental and physical outcomes. All must be incorporated onto a single scale, if measurement is to be possible using common units, which is the prerequisite to impact value comparisons. The aim of this Chapter is to describe in detail an index of Quality of Life Outcome States (QLOS Index), suitable for use in measuring the outcome of an environmental change in terms of resultant quality of life state. Three objectives must be met in order to achieve this aim. Clear criteria which the QLOS Index must comply with in order to meet the needs of the systematic framework are required. These are presented in Section 9.1. Definitions of the terms and concepts involved are needed to ensure clarity and preciseness in what is to be measured using the QLOS Index, and these are found in Section 9.2. A summary of the issues involved in measuring quality of life is also needed, to establish the current state of knowledge and the feasibility of a single scale of quality of life outcome states. This is dealt with in Section 9.3. Only then can the QLOS Index be presented in detail, before being reviewed for compliance against the criteria set. These tasks are contained in Sections 9.4 and 9.5, respectively.

9.1 QLOS Index Criteria

There are a number of requirements which the QLOS Index must meet, starting with the need to produce outputs which are suitable for use in the valuation method. As outlined in Chapter 6, Step 7 of the systematic framework involves producing the human consequences data required to undertake Step 8, impact value calculation, which is undertaken as follows;

Impact value = quality of life outcome x risk x duration x number of people affected

Of the four types of data required, the first, quality of life outcome, is of concern here. All possible quality of life outcomes for each human consequence must be included on the QLOS Index (with any aggregated data including range of response and mean response). The remaining criteria are five-fold, and are all linked to ensuring the suitability of the QLOS Index for producing appropriate data for use in the valuation method.

Firstly, the QLOS Index, and any scales which contribute to it, must be ratio scales. In other words, not only must the full range of quality of life outcome states be arranged in rank order of significance, but the spacing between each must be known, and relative to zero. Without this, comparability between points on the scale cannot be achieved with the necessary level of accuracy. Secondly, the QLOS Index and any contributing scales must be based on transparent, logical and rational (scaling) theory recognised in the literature. In this area – the objective measurement of subjective phenomena - more than any other, the validity of the scale relies upon its theoretical basis. Thirdly, all points on the QLOS Index must refer to generic human responses, rather than to specific ones, which will be of only limited use in measuring specific human consequences of specific environmental changes. A generic scale point is defined as one which is not object-specific, so it does not apply to the environmental impact(s) which give rise to the QLOS Index category, but only the outcome for the recipients quality of life. Fourthly, the QLOS Index and any contributing scales must be structured for optimum efficiency and with the capability to produce outputs that are sufficiently sensitive. Finally, clearly, the QLOS Index must contain the full range of possible outcomes, otherwise impact values for some consequences will not be achievable.

As established in Chapter 6, environmental impacts cannot be measured directly by examining the environment with some sort of quality/value toolkit. They can only be measured by calculating the value of loss of quality caused by the changes in the environment which stem from the RIOs released into it. This relationship is indicated in diagrammatic form in Figure 9.1. This indicates that physical changes caused by RIOs have magnitude, which can be (but are not always) measured essentially in an objective manner. However, magnitude of environmental change is only one determinant of the significance of its consequence —usually called its environmental impact. Significance is related to the quality assigned to the environment affected by the change.

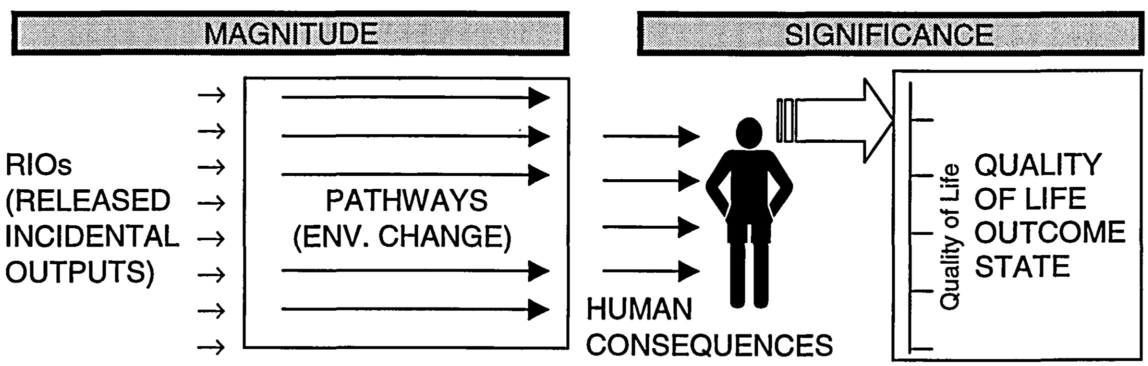


Figure 9.1 Summary Diagram Showing the Relationship between RIOs, Environmental Changes and Quality of Life Outcome State

Quality is assigned by people. Environmental quality is determined not just by the environment, but by people's perception of themselves within the environment, by the motivations that promote their appreciation of the environment, and by their economic and social values, culture, life-style and health state. In other words, environmental quality is experienced as a determinant of life quality – either directly, or through subjective perception or emotional feelings. Directly, the release of a pollutant into an environment may affect health, wealth or well-being. Emotionally, changes to a distant species or ecosystem may cause moral outrage or a feeling that something with intrinsic rights or value has been damaged. Both affect quality of life, damage to which is undesirable. Environmental impacts cause damage to quality of life and the extent of the damage indicates the significance of the impact. Therefore, an understanding and clear definition of quality of life is required, without which the term "environmental impact" lacks theoretical basis and practical meaning. Moreover, this definition is the basis of the appropriate scale for measuring environmental impacts, the QLOS Index.

While the quality of life concept has probably been around for as long as humanity, its modern application as a closely defined term is relatively limited. This may appear surprising, given that the desire to improve (and prevent damage to) quality of life is a fundamental personal motivation and a long-held social goal. However, this is where part of the problem lies. Quality of life is so fundamental that it is all-encompassing, and therefore presents ambiguity and tangibility problems to those who wish to define and measure it. Also, quality of life is fundamentally subjective, since quality is "in the eye of the beholder", and the task of objectively measuring a subjective construct is not straightforward. A common response to these problems is to seek to constrain the problem by adopting a measure and definition which is concerned with one area of quality of life – a part of the picture.

The largest and most well-developed literature, which encompasses a range of techniques and numerous methods and scales for measurement, is in the field of health-related quality of life (HRQL). There are many definitions of quality of life in the HRQL literature, a large number of which are not dissimilar to the following: "The term "quality of life" represents a broad range of dimensions of human experience, ranging from those associated with the necessities of life, such as food and shelter, to those associated with achieving sense of fulfilment and personal happiness. Cultural, psychological, interpersonal, spiritual, financial, political, temporal, and philosophical dimensions may be incorporated into various definitions. These widely valued aspects of human existence may not be thought of as part of personal health status and well-being. A safe environment, adequate housing, a guaranteed income, respect, love, and freedom all contribute to an individual's quality of life" (Walker and Rosser, 1993). There is a difference between quality of life, and HRQL, which is further defined: "Health-related quality of life is the value assigned to duration of life as modified by the impairments, functional states, perceptions, and social opportunities that are influenced by disease, injury, treatment, or policy" (Walker and Rosser, 1993). This definition is expanded on in tabular form to indicate the core concepts and domains it encompasses (see Table 9.1).

The critical difference between the quality of life and HRQL definitions is that the first is concerned with factors which contribute to an individual's quality of life. It is assumed that such factors, their content, size and relative weight should be reflected in the measurement of an individual's quality of life, *by the individual*, if the measure is operating properly. However, the HRQL definition is concerned not with appraisal of external phenomena, but with how changes (external or internal) affect the individual's quality of life. In each, the appraisal can only be achieved by the individual, since it is an inherently subjective measure. The difference lies in what is being appraised. The measure needed for the QLOS Index is that arising as a result of an environmental impact. It is not about measuring the environment or the individual's perception of the environment, it is about measuring the individual's perception of their changed life-state as a result of changes in the environment. Therefore, the HRQL definition is most suitable, and this has been adopted as the appropriate underpinning definition for the QLOS Index, with two qualifications. Firstly, all quality of life must be included (rather than merely "health-related"). In particular, emotional stimuli are included within the terms "perceptions" and "injury". Secondly, since the QLOS Index is only concerned with the measurement of environmental impacts, the HRQL definition must be embellished with the additional words, "as it relates to environmental impacts".

Concepts and domains	Definitions/indicators
OPPORTUNITY	
Social or cultural disadvantage	Disadvantage because of health; stigma; societal reaction
Resilience	Capacity for health; ability to withstand stress; reserve
HEALTH PERCEPTIONS	
General health perceptions	Self-rating of health: health concern/worry
Satisfaction with health	Physical, psychological, social function
FUNCTIONAL STATUS	
SOCIAL FUNCTION	
Limitations in usual roles	Acute or chronic limitations in usual social roles (major activities) of child, student, worker, independent householder
Integration	Participation in the community
Contact	Interaction with others
Intimacy and sexual function	Perceived feelings of closeness; sexual problems in performance
PSYCHOLOGICAL FUNCTION	
Affective	Psychological attitudes and behaviours including distress and well-being
Cognitive	Alertness; disorientation; problems in reasoning
PHYSICAL FUNCTION	
Activity restrictions	Acute or chronic reduction in physical activity, mobility, self-care, sleep, communication
Fitness	Performance of activity with vigour and without excessive fatigue
IMPAIRMENT	
Symptoms/subjective complaints	Reports of physical and psychological symptoms, sensations, pain, health problems or feelings not directly observable
Signs	Physical examination: observable evidence of defect of abnormality
Self-reported disease	Patient listing of medical conditions or impairments
Physiological measures	Laboratory data, records, and their clinical interpretation
Tissue alterations	Pathological evidence
Diagnoses	Clinical judgements after 'all the evidence'
DEATH AND DURATION OF LIFE	Mortality; survival; years of life lost

Table 9.1 Definition of Health-Related Quality of Life (after Walker and Rosser, 1993)

The concept of how the HRQL definition can be incorporated into a measurement system is also accepted (Walker and Rosser, 1993). The five broad concepts of the definition must be combined into a single continuum or index, which is anchored by an optimal value at the top ("maximum quality of life") and a minimal value at the bottom. Specific domains of survival, impairment, functional state, perceptions and opportunities therefore fall along this value continuum and are combined following individual examination and measurement of each domain. However, again, there is a proviso to adopting this approach in the QLOS Index. In HRQL, outcomes which are sought by measurement scales are often to assess relative achievement/failure of a particular procedure. Thus, HRQL practitioners may describe "need" and "outcome" as "two sides of the same coin" (Wilkin et al, 1992). In QLOS valuation, the change being measured is not due to medical intervention, but to environmental change. Apart from this, the scenario and definitions and theory of quality of life outcomes are identical. Hence, with appropriate caveats, a quality of life definition which is suitable for the QLOS Index is achieved, based on HRQL. This has the benefit of adopting much of the underpinning theory and practice of HRQL measurement which already exists in a vast and rapidly maturing literature.

9.3 Measuring Quality of Life

There are three possible approaches to measuring quality of life; personal judgement, collecting existing values from the literature (established by other authors or inferred from data), or measuring values directly from a sample of people. The first is the quickest and easiest, and could be combined with sensitivity analysis to test robustness of values. However, it is subjective and therefore runs contrary to the requirements of the systematic framework. The second is also potentially relatively simple. There is a large and rapidly expanding literature and, increasingly, relative values for different quality of life states are being established with reasonable confidence. However, where values are provided, methods vary, or the values are not in the form needed, that being transferable ratio data on the human responses to environmental impacts, measured in terms of outcomes for quality of life. Such requirements could be met by the third option, which holds out the possibility of providing necessary data quantity and accuracy. The QLOS Index is therefore based on the existing literature, while it is designed to facilitate the measurement of values directly from people.

The central problem with producing a single scale to measure quality of life outcome states is that quality of life is generally considered to be a multi-dimensional concept. The units of the various dimensions must be brought together into a common unit in order for a single scale to have meaning. There are only two ways of doing this. The first is to set a single "global" question, such as "How is your quality of life?" This is rejected, as it is well-established that this approach leads to unacceptably high loss of data and precision. The second is the development (by empirical means) of a very long list of descriptors encompassing, as a whole, all of the possible elements of the concept. This is often then compiled into a number of "domains" (usually between three and sixteen), within each of which, questions/descriptors can be identified as being of the same "group". The use of groups and domains is designed to assist in bringing concepts which may be measured in similar units together, stage by stage. The problem for then achieving a single index from these domains is that there are potentially millions of possible quality of life states, since theoretically, every possible state within each domain may exist at the same time as every other state within every other domain. In practice, there are improbable combinations, and multi-dimensional HRQL measures have thus been able to simplify and reduce the number of likely combination states. Nevertheless, the transition from a number of disparate domains to a single scale which combines all is difficult to achieve with a statistically acceptable level of validity. It is only in recent years that this desirable goal has become close to being achieved, through the use of scaling and weighting and by empirical work, improving on earlier attempts. Thus, the simplicity of a global score and the accuracy of a multi-dimensional quality of life status measure can only be combined in an index. While the term "index" is often used to mean different things, here, it is taken to mean a single ratio-scaled list of states measured in a single unit.

As a starting point, then, the measurement of quality of life can theoretically be achieved by asking the receptor (person whose quality of life outcome state is being sought) how their quality of life is or might be, given a scenario. This exercise is subject to the following conditions:

- Full information must be provided and understood on items for valuation and related information to ensure "rational" judgement;

- Impact values should only be elicited for impacts within the general sphere of experience of the receptor (if they have not experienced the state, they must at least be able to relate to it);

- Questionnaire and processing should be designed to "check" the rationality/consistency of the receptor;

- Tests should be designed to assess for the presence of any bias in the measurement method, and corrections for any which is detected should be made.

In general, existing HRQL measures satisfy these conditions. However, the cases of measuring quality of life as "is" are different to "might be" situations. Also, there are numerous other issues in measuring quality of life. These issues are presented in detail in Appendix D, although the key points are summarised here, grouped into four categories; appropriateness of the approach, issues in measurement bias, differentiation and distribution, and single scale construction.

Appropriateness of a "global" HRQL measure has been challenged on the grounds that the term means different things to different people, and that there are short term fluctuations in people's perceived quality of life. In reality, empirical data show that there is much common ground among people's relative evaluation of health states (Kaplan et al, 1993). While mood swings and so on may make quality of life perception transient, a statistically significant survey sample provides an accurate mean for impact assessment. The suggestions that the quality of life concept is too complex or poorly understood, that people are irrational, or that information requirements are too great, are all refuted by the fact that numerous HRQL measures have been successfully measuring quality of life outcomes for well over a decade. Furthermore, the only alternative is the continuation of the current situation, where quality of life outcomes are either ignored or subjectively viewed, where inaccuracy and flawed value judgement is accepted and its extent unknown. Even a partially successful Index-based valuation would be preferable to this.

Accepting that a global measure is both appropriate and necessary, there are also numerous issues involved with bias in scale construction and operation. Parallels can be drawn here with the problems in environmental economics valuation methods discussed in Chapters 4 and 5 and Appendix A. While content and scale bias can theoretically affect HRQL scales, in their application to a QLOS Index, they are used in entirely negative terms (only negative outcomes are sought, since environmental impacts are defined as negative). Thus, no relative bias can accrue from framing questions inappropriately, or from disparities between "willingness to pay" or "willingness to accept", as occurs with the direct environmental economics (contingent valuation) method. Design, method, hypothetical, regression and response biases occur when the administration of the measurement method affects results, and these are well-known, as are measures for identifying and correcting for them. Where an identified group has a significantly lower response rate to a survey than the mean, for example, because they have a particularly chaotic lifestyle and so do not respond to a survey, adjustments can be made to reflect this bias in the data.

The third category of measurement issues is concerned primarily with differentiation of quality of life within or between populations. Assessing quality of life across cultural and national boundaries is difficult, as is obtaining measurements for people who do not understand the process, such as the very young or mentally ill. However, QLOS data are generic and cross-cultural, in as much as they are only concerned with outcomes of environmental change on emotional state and health. If the environmental change and its implications can be understood, then its implications can be expressed in generic terms. Proxies are an established way of measuring outcomes for those who cannot express understanding but experience impacts. There is an open question as to how transferable QLOS data may be between cultures, classes, or populations and their sub-groups. However, this can only be solved by survey and cross-comparison, to establish the limits of transferability.

The production of a single index brings specific scaling and measuring problems, particularly related to the issue of aggregating or combining dimensions of quality of life. While it may be argued that each dimension warrants a different, distinct unit of measurement, the fact remains that these units can be combined into a single scale with an aggregate unit. Indeed, people do make judgements based on their quality of life as a single entity, so it can be (and is) done every day. Hence, the problem is a practical one, rather than a conceptual or theoretical one. In HRQL research, there is broad consensus over the definition and description of quality of life and increasingly, over the issue of producing single scores. Indeed, the problem of expressing several dimensions (or units) in a single-unit, single-dimensional currency is not new, nor is it confined to subjective, quality-based phenomena. By way of comparison, a single currency – money – is used to measure multi-dimensional physical things to which people are accustomed to assigning property rights. Land and Property Valuers use a "red book" which gives guidance on how to combine and adjust figures and where to get figures from. A property value relies on bringing various disparate elements, such as size, level of equipment, location, services, etc., together and combining them into a single value - in this case expressed in money terms. These elements would not be intuitively comparable. Location, stigma, etc., do not lend themselves to expression in recognised quantities or objective measurement. Observation, comparison and empirical measurement are the means by which approximate value is derived.

In summary, quality of life measurement issues are numerous and complex, which is not surprising, given the nature of the concept. However, a great deal of research effort has been expended and much progress has resulted in developing measures which overcome these issues, particularly in the field of HRQL scaling. Consensus over the need for and accuracy of recent HRQL measures is now established. Furthermore, QLOS is concerned only with the application of HRQL scales to impact assessment, which avoids the outstanding HRQL issues of how reliably the measures indicate ill-health or can be used for diagnostic purposes. Provided the scale can be applied to produce data which, aggregated for the receptor population, provide accurate measures of quality of life outcomes of environmental changes, then measurement is possible, as well as being a desirable alternative to current approaches. Given the criteria presented in Section 9.1, the definition of quality of life given in Section 9.2, and the summary of measurement issues presented above, it is now possible to present the detailed form of the QLOS Index.

9.4 The Quality of Life Outcomes (QLOS) Index

Given that an extensive HRQL literature exists which is closely associated with the needs of the QLOS Index, the most efficient way of generating the Index itself is to adopt a suitable index (or number of scales or indices) from this field. In accordance with this approach, a critical review of twenty existing HRQL scales against the criteria given in Section 9.1 above is presented in Appendix E. The result is that some theoretical weakness is apparent in several scales. Content bias is a factor in some quality of life studies, where they have been developed in a short space of time by a select group of researchers working with the same few theoretical texts. Instruments that are weighted and produce an overall score are potentially useful, while others which are statistically more valid and reliable assess multiple HRQL domains and fail to present single or global scores. The most suitable scales for use in the QLOS Index are the Quality of Well-being Scale and Rosser Scale of Illness States. Some questions remain with both, particularly, their sensitivity to less significant effects on quality of life state. The former scale has some weaknesses in terms of validity and reliability (though it is more sensitive), and the latter scale is more adaptable and flexible. Therefore, on the basis of closest fit with criteria needs, the Rosser Index is used as the basis of the QLOS Index, with additional categories added as required.

The resultant QLOS Index is presented in Table 9.2. It contains values extending from "worst possible state" (for example, states worse than death) to "best possible state" (e.g. no negative effect on quality of life) in ratio scale form. The origin and development of the scale is discussed in detail elsewhere (Horne, 1999a, 1999b). The QLOS unit is clear, definable and, although necessarily multi-dimensional (as discussed in Section 9.3), it is an identifiable single unit of measure, and is established as the unit which will be used in valuing environmental impacts. It is concluded that remaining barriers to the approach are few, and they are far outweighed by the potential benefits. In particular, once incorporated into a proprietary valuation method, it allows highly sensitive, reliable and valid values of environmental impacts to be calculated.

The main part of the QLOS Index is in the 0-100 range, with two reference states; 100 as death equivalent, and zero as unimpaired QOL. The reason for this is that a standard utility HRQL scale exists from 0-1, and the two orders of magnitude are simply added to provide more sensitivity for integer-based scale scores. The following transformations have been performed on the Rosser Index to achieve the QLOS Index as presented in Figure 9.2:

- Inversion (so that "unimpaired state" is at the bottom of the Index, with a value of zero);

- Multiply all points by 100 (to provide for less decimal points and more sensitivity potential at integer level).

The Index units – QLOS Index numbers (I) – reflect the function of the Index as a means of measuring quality of life outcome state. The ratio nature of the Index is retained, and it contains two anchor points - "unimpaired state" at 0 and "death" at 100. However, it also extends beyond 100 to accommodate states worse than death. The worst outcome state corresponds to an Index number of 248 (although states worse than this cannot be ruled out). Further categories can be added to the QLOS Index, provided the theoretical axioms of the method used are the same as those for the original Rosser Index, the validity and reliability of the method used is proven and acceptable, and the scale anchor points are the same. To provide an example of this, five extra categories not included in the Rosser Index, but provided by using economic Time Trade Off valuation methods and direct ratio-scaling methods, are included on the QLOS Index in Table 9.2.

QLOS Index Number (I)	QLOS category
248	Severe distress, confined to bed
~	
202	No Distress, unconscious
200	
~	
100[death	Moderate Distress, bed-ridden ~ Severe Dist, in wheelchair ~ Confined to bed, severe pain (ratio-T87)
98-99	
96-97	
94-95	
92-93	
90-91	
88-89	
86-87	
84-85	
82-83	
80	
78	
76	
74	
72	
70	Mechanical aids to walk and learning disabled (TTO-T87)
68	Hospital confinement (TTO-T87)
66-67	
64	
62	
60	
58	
56	Anxious/depressed and lonely much of the time (TTO-T87)
54	
52	
50	
48	
46	
44	Mild Distress, confined to bed
42	
40	
38	
36	
34-35	Some physical and role limitation with occasional pain (TTO-T87)
32-33	Moderate Distress, confined to wheelchair ~ No distress, confined to bed
30-31	Severe Distress, Unable to work/housebound
28	
26	
24	
22	
20	
18	
15-16	Mild Distress, confined to wheelchair
13-14	No Distress, confined to wheelchair ~ Severe distress, limited phys ability
11-12	
9-10	Mod Distress, unable to work/housebound ~ Severe distress, severe social disab.
7-8	Mild Distress, unable to work/housebound ~ Severe distress, slight social disab.
5-6	Mild dist, lpa ~ Mod dist, severe social disab. ~ No Distress, unable to work/housebound*
3-4	No Distress, limited phys ability (lpa) ~ Mild Dist, severe social disab ~ Mod dist, slight social disab**
1-2	No Distress, slight or severe social Disability ~ Mild Distress, No/slight Disability ~ Mod Distress, No Disab.
0	No Distress, No Disability

Note 1: Calibrated to 100 = death equivalent, 0 = healthy/unimpaired QOL
Note 2: All scale points from Rosser Index (Rosser and Kind, 1978, Kind et al, 1982), except: T87 (Torrance, 1987)
Note 3: TTO = Derived from application of Time Trade-Off method (Horne, 1999a)
Note 4: See elsewhere for detailed derivation of the QLOS Index (Horne 1999a, 1999b)
* Also: ~ Moderate distress, limited physical ability
** Also: ~ Severe distress, no disability
Abbreviations: dist – distress, mod – moderate, phys – physical, disab – disability, lpa – limited physical ability.

Table 9.2 QLOS Index

Use of the Index to score quality of life outcomes involves a simple exercise of correlating the data for each environmental change and its subsequent human consequence to the appropriate outcome state category on the QLOS Index. For example, for the human consequence "aesthetic impact of noise from power station", the response might be recorded as "no disability, mild distress" on the QLOS Index, to give an Index number of 0.005. Ideally, each receptor should score their own response, since only they can decide which outcome state category is appropriate for them. Noise may cause "moderate distress" rather than "mild distress", either because of the definition of mild, moderate, etc., or because of the potential for differential outcomes on different people. However, detailed definitions, population norms and sub-population "sensitivities" must be established by conducting surveys and deriving distributions of Index numbers for different impacts, in order to improve definition precision.

It is possible that unforeseen outcome states exist, for which there are no appropriate QLOS Index categories. However, new categories can be entered on the Index between the existing ones, by establishing population norms for the specific outcome in the same way that the existing categories on the QLOS Index were established (Horne, 1999a, 1999b). This is only likely to occur at the bottom end of the scale (for values very close to zero). Where a less sensitive result will suffice, a rough approximation of the appropriate QLOS Index number could be established by inferring the existence of a category between the two closest existing known categories on the QLOS Index.

9.5 QLOS Index Criteria Compliance

All of the criteria presented in Section 9.1 are complied with by the QLOS Index. The output data of the QLOS Index are in the units of "quality of life outcome states", as required by the valuation method. The QLOS index is a true ratio scale and is based on transparent, logical and rational scaling theory. Furthermore, it is derived directly from the Rosser Index, which itself has established theoretical and practical validity, and was developed from a wider HRQL measurement literature. The Index categories are all generic, in that they refer to emotional, mental or physical human responses that may stem from a wide variety of environmental changes and resultant human consequences. Without extensive application, it is difficult to prove that the criteria requirement that the QLOS Index must contain the full range of possible outcomes is met. However, as explained above, the QLOS Index goes further than this; it allows for the possibility of adding new scale categories should this ever be required.

The remaining criteria are concerned with efficiency and sensitivity. These are related, since efficiency relates to achieving required sensitivity against the effort and cost of data collection and processing. Sensitivity, the ability of the scale to differentiate sufficiently between points of value, is inherent in the ratio scale. While it would remain an issue if the Index data were to be used to differentiate between individual scores within a population, since the QLOS Index is designed to provide values for comparison with each other, sensitivity only becomes an issue if many impacts obtain apparently the same score. This is highly unlikely, given the number of categories and the range of possible outcomes identified. However, this cannot be proven until the QLOS Index is demonstrated in practice, as part of the valuation method. Presentation and demonstration of the valuation method is therefore the next logical task, and this is the subject of the next Chapter.

10. THE QLOS VALUATION METHOD

The outcome of environmental changes on quality of life can be measured using the QLOS Index, as presented in Chapter 9. The next task is to use such measurements in the process of valuing environmental impacts in a common currency. As has already been established, there is no question that valuation should be done – indeed, it is being done at present, albeit implicitly rather than explicitly, and partially and arbitrarily rather than impartially and with full information. Efficient and accurate valuation in suitable units is a pre-requisite to effective and appropriate decision making and regulation.

The aim of this Chapter is to present and demonstrate the QLOS valuation method, which can be used to provide accurate impact values in QLOS units. Two objectives must be met prior to achieving this aim. Clear criteria which the QLOS valuation method must comply with in order to meet the needs of the systematic framework are needed, and these are presented in Section 10.1. Step 7 of the systematic framework, human consequences, must be described, as this involves producing the data needed for valuation. This is presented in Section 10.2. Only then can the QLOS valuation method be presented in detail, before being demonstrated, and then reviewed for compliance against the criteria set. These tasks are contained in Sections 10.3, 10.4 and 10.5, respectively. Further detail on the demonstration is also contained within Appendix F.

10.1 Valuation Method Criteria

The QLOS valuation method must meet the systematic framework criteria requirements. In so doing, it must be designed to maximise objectivity and transparency in calculating reliable, accurate values for environmental impacts. It must also be practical, and sufficiently valid and sensitive.

Validity is the extent to which the results comply with those expected (e.g. by comparison against a "gold standard"). However, testing for validity is both problematic and complex, with subjective quantities such as human values, where no "gold standard" exists. Two means of ensuring validity are possible. Firstly, if the method is based on sound theory, then results can be expected to be valid. Theory has already been established at systematic framework level, in Chapter 6, and for the QLOS index, in Chapter 9. Secondly, if the valuation results are compared with those of another method, and they comply with predicted deviations from them, then again, validity is established. This is addressed in the demonstration (Section 10.4).

Sensitivity is the ability to differentiate sufficiently between points of value, in order for the results to be usable in the application. The point at which sufficient accuracy is obtained by using the least resources possible is the point of optimum efficiency. Therefore, there is an inevitable trade-off between sensitivity and practicality. Assessment of sufficiency of sensitivity can be examined by applying sensitivity analysis or simple sensitivity tests to assess potential ranges of values and margins of error. This is addressed in the demonstration (Section 10.4).

10.2 Step 7: Human Consequences

Human consequences are identified from the environmental change inventory, produced in Step 6 (see Chapter 8). For each environmental change, one or more human consequences may potentially occur. These consequences vary in terms of significance of outcome. Potential outcomes include a very wide range, as indicated on the QLOS Index, including from death, intense pain or distress, through depression or other clinical mental or physical illness, to social disability, aesthetic disturbance or annoyance (the QLOS Index is presented in Table 9.2 and explained in Chapter 9). Human consequences may also vary according to their likelihood of occurrence, duration of impact and number of people affected. Thus, in total, for each human consequence, four pieces of data are required to reflect the four variables involved. These data are the prerequisite to valuation in Step 8:

- QLOS Index data - surveyed population response including range of response and mean response;

- Risk data - probability of the human consequence;

- Dynamics data - time characteristics of the human consequence (duration, variation over time);

- Receptor data - number of people affected, including population distribution, density, and demographics.

If there is a possibility that the human consequence may not occur, then the probabilistic risk of occurrence is required (otherwise, if it is expected to occur, the risk is 1). Duration (and any variation over time) is also required, as well as receptor data to identify the number of receptors. Associated demographic data are important to check that the receptor group can be treated as homogenous, if not, the human consequence can be split into two, each assigned a different receptor group. In addition, a descriptor of each human consequence is required, to provide clarity in data presentation. The tasks of producing each of the four types of data needed are presented below, based on detailed development work documented elsewhere (Horne, 1999a, 1999b).

10.2.1 QLOS Index Data

The QLOS Index contains the full range of potential Quality of Life states which may arise from any given change (hence "Quality of Life Outcome States"), in ratio scale form. In concept, the process of producing QLOS Index data involves a simple exercise of classification, by matching human responses to environmental changes to points on the QLOS Index and recording the corresponding Index number. For example, taking the human consequence "impact of noise from power station", the response may be recorded as "no disability, mild distress" on the QLOS Index, to represent aesthetic disturbance, giving an Index number of 1. However, various outcomes are possible for some environmental changes, and where they vary widely it is important to differentiate them as separate consequences. For example, severe deafness may result from noise, and this outcome should be measured separately. To assist in differentiating between ranges of outcomes for the same impact and different impacts, it is useful to refer to a generic list of five general types of human consequences, as indicated in Table 10.1.

1	Mortality
2	Health; morbidity and illness
3	Damage/Loss of time/livelihood
4	Existence impact (e.g. remote biodiversity/heritage loss
5	Aesthetic; visual, noise, other emotional/annoyance

Table 10.1 Human Consequence Types

There are two issues which complicate the task of producing Index number data. Firstly, who decides which outcome state category is to be used? Noise may cause "moderate distress" rather than "mild distress", either because of the definition of mild, moderate, etc., or because of the potential for differential outcomes on different people. Therefore, detailed definitions, population norms and sub-population sensitivities must be established by conducting surveys and deriving distributions of categories for different impacts.

The second issue arises where an appropriate QLOS Index category does not exist for a given human consequence. This is possible, since the Index has been adopted for use from elsewhere. Where necessary, therefore, a new outcome state category may have to be measured in order to establish a new category between the existing categories. This can be achieved by establishing population norms for the specific outcome in the same way that those already on the QLOS Index were established (see Horne, 1999a). It is envisaged that this situation is only likely to occur at the bottom end of the scale (for values very close to zero). This is because the Rosser Index (see Kind et al, 1982), from which the QLOS Index is derived (see Horne, 1999a), was originally designed to measure Illness States, so it may be insensitive to emotional or less significant outcomes. However, as discussed in Chapter 9, where a less sensitive result will suffice, a rough approximation of the appropriate QLOS Index number could be established by inferring the existence of a category between the two closest existing known categories on the QLOS Index.

10.2.2 Risk Data

Impacts must be risk-weighted. In other words, they must incorporate the extent to which outcomes are likely to occur – the probabilistic risk. This is the probability of an event occurring and, where it is unknown, a best estimate is used. Probabilistic risk can be established by various means, although the preferred one is through assessment of primary data of a statistically significant number of similar cases which have occurred elsewhere. Where this is impossible, the risk assessment methods available are based upon a mixture of experience and prediction, often using iterative models.

10.2.3 Dynamics Data

The calculation of duration of outcome is conceptually simple, but involves some theoretical issues. Simply summing times in different impact states assumes Constant Proportional Time Preference, i.e. that days of illness/impact are of equal importance whenever they occur, proportionally however long they last for, and irrespective of following or preceding health state. This assumption has been criticised in the health-related quality of life literature as over-simplistic. Impact value may be affected not only by duration of time spent in the state, but also when the time starts. There is therefore a question of whether assuming Constant Proportional Time Preference is valid, or whether it will lead to significant loss of accuracy in results. However, the only alternative to assuming Constant Proportional Time Preference is to replace it with a more sophisticated model of time preference, which more accurately reflects reality. No such model currently exists and, while discounting is often applied to reduce the value of money in the future relative to now, there is no consensus over an appropriate means of adjusting Constant Proportional Time Preference as it relates to timing of environmental impacts or outcome states. If and when such consensus develops, alterations and adjustments can be incorporated into the dynamics data.

A more challenging issue is the selection of appropriate units for the time component of value. To use an absolute measure of the actual time period for which an outcome state will be experienced assumes that individuals view outcome duration in absolute terms. Reality is rather different, and there are also ethical reasons why duration of an impact should be measured in other than absolute terms. Firstly, while life expectancy unavoidably varies (and generally shortens) throughout one's life, there is a strong ethical consensus that the value of all lives should be viewed as the same, and constant. Each person should therefore be viewed as having a similar stock of quality of life available to be impacted. Everyone has all of their future life left in front of them, therefore, the time factor for outcomes must be normalised to the amount of life remaining. Thus, duration must be expressed not in absolute time (e.g. seconds) but as a mean proportion of conscious life remaining. For generic impact calculation purposes, this is the mean conscious duration of the human consequence divided by the mean expected conscious life-time remaining;

Proportion of conscious life remaining (T) = (conscious duration of the human consequence)/(mean expected conscious life-time remaining)

For example, an 80-year-old, faced with an impact of 25 years duration, may not be expected to live for this period of time, and thus may only be expected to experience it for, say, 5 years. However, to value the impact they experience on a 5-year basis would be to undervalue the amount of time one has left at age 80. The issue is one of how much stock of life one has left at any one time. If all lives are worth the same, then this stock should be constant. Therefore, while time is relatively constant, rather than consider life-span as a constant, it is life-remaining which is considered constant for the purposes of the calculations. For the 80-year-old, 5 years is the total amount of the person's expected remaining life. Therefore, to take the day of valuation as the start of the rest of one's life, the 80-year-old will be affected for all of their expected remaining life, whereas a teenager could be expected to be exposed to the impact for only, say, a quarter of their remaining life.

Given an established initial QLOS Index number for a human consequence, there are four possibilities for the subsequent development of this situation over time. Firstly, the QLOS Index number may remain constant; secondly, it may increase; thirdly, it may decrease, and; fourthly, it may follow a varied pattern over time. Clearly, the first possibility is the simplest. However, the other three possibilities can be incorporated into the QLOS valuation method through duration data, where assumption of the simple situation would lead to unacceptable loss of sensitivity. In such cases, the time profile of the environmental change is split into a number of sections, each representing a separate human consequence which is internally constant, as illustrated in Figure 10.1. Referring to this, for Impact A, a single QLOS Category Value is sufficient to approximate the actual outcome trace but, for Impact B, the human consequence outcome trace is split into four parts to approximate the actual trace, requiring calculation of time and QLOS Index number for three different QLOS Category Values.

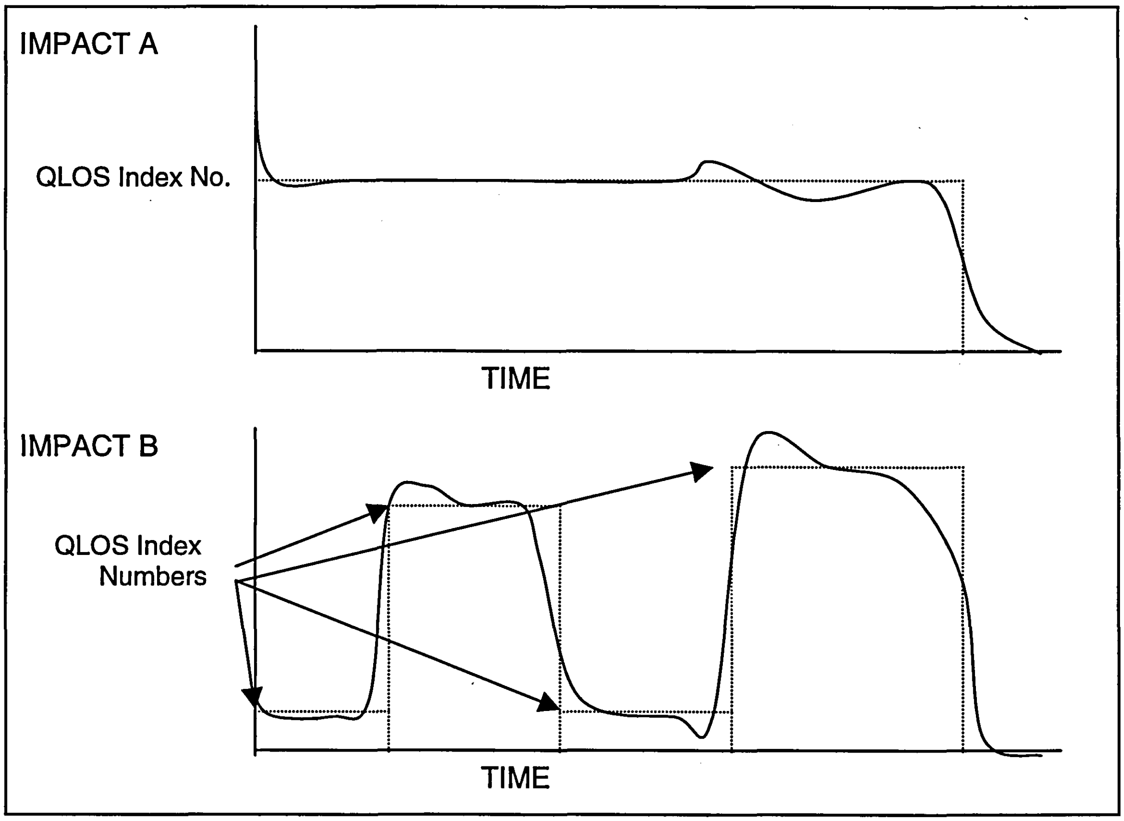


Figure 10.1 Incorporating Time Profiles for QLOS

Key: ——Actual trace of outcome state: ············QLOS Valuation trace

10.2.4 Receptor Data

A receptor is someone who is subject to a human consequence as a result of an environmental change. The number of receptors, 'n', is the total number of people affected by the impact. Although the views of children, babies, and incapacitated people have not been incorporated into the QLOS Index, it is assumed that their values, if they could be measured, would show a similar distribution to the population average, so numbers of people who receive impacts in these groups should be added along with the general population.

Adjustment of receptor data may be required to avoid double-counting. Where a particular human consequence *precludes* the receptor from experiencing other human consequences, they must be deducted from the receptor data for these other consequences. For example, if an individual was hospitalised as a result of a human consequence, they would not then be subject to local noise impacts while in hospital.

10.3 Step 8: Valuation

Once the data is available, valuation involves simply calculating the product of the four variables for each human consequence, as follows;

$$\text{Impact value (Qu)} = I \times R \times T \times n$$

where; Qu = QLOS units, I = QLOS Index number, R = Probabilistic risk, T = proportion of conscious life remaining, n = number of receptors.

It should be notes that R and T are dimensionless. Therefore, the QLOS unit (Qu) of impact value is "people-QLOSs". Thus, the impact value, in QLOS units, can be produced for each human consequence. The results can then be summed as necessary for comparative purposes or to calculate "Total QLOS" for a given production system, fuel cycle, or project, or part thereof.

10.4 Valuation Method Demonstration

Although valuation is conceptually simple, in practice there are various pitfalls and problems in data suitability, and ways in which these can be overcome. Furthermore, there remains the need to show that the method produces sufficiently valid and sensitive results. These two issues must be dealt with by a demonstration. The process of selection of an appropriate demonstration is dealt with in detail in Appendix F, where further detail on the calculations, data presentation and conclusions are also found.

Ten impacts are incorporated into the demonstration, covering a wide range of impact types. These are applied to three levels of comparison; between similar impact types in the same fuel cycle, between different impact types in the same fuel cycle, and between different fuel cycles. The data required are drawn from the same set of studies, in order to ensure that the results are not biased by variations in data on environmental changes, rather than values. It is important that QLOS valuations and comparisons are derived using similar environmental change data, to ensure that it is the difference in method rather than data which is being measured.

10.4.1 Calculations and Results

The derivations for each specific datum are presented in Appendix F, and summarised along with results in Table 10.2. Impact Numbers 1-7 are from the coal fuel cycle (Impact Numbers 1-5 being from traffic and emissions to air from the coal power station, and 6-7 from coal mining and coal handling, ExternE, 1995c), while Impact Numbers 8a-8b are from the nuclear fuel cycle (ExternE, 1995e), and Impact Number 9 is from the wind power generation cycle (ExternE, 1995f). To illustrate how data and calculations are undertaken, for Impact Number 1 ("Power Station related traffic; public accident deaths"), the QLOS Index number is 100, since death is rated as 100 on the

QLOS Index. For probabilistic risk, average expected deaths per year arising from the impact are 0.122, or 4.874 in the lifetime of the plant. Since the number of receptors at risk is not known from the data source (ExternE data does not include it), this risk-weighted figure is effectively a composite of 'R' and 'n'. Probabilistic Risk is therefore entered as 1, with 4.874 as the value for the number of receptors killed, for purposes of the demonstration. Finally, the duration in conscious life remaining (T) is 1, since death affects all of remaining life. Thus, the impact value for Impact Number 1 is calculated as follows;

Impact value = I x R x T x n

100 x 1 x 1 x 4.874 = 487.4 Qu

To normalise this value it must be presented as a ratio of power output;

Total output = 1710 x 8760 (hours per year) x 40 (lifetime) x 0.76 (load factor) = 455.4 TWh (or 455.4×10^6 MWh)

Therefore, the normalised value is calculated as follows;

487/455.4 = 1.07 μQu/MWh

	Stage 1			Stage 2		Stage 3		Stage 4		
Impact Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8a**	**8b**	**9**
Human Consequence Criteria	Power Station related traffic; public accident deaths	Power Station related traffic; public accident major injuries	Power Station related traffic; public accident minor injuries	Power Station SO_2 emissions; agricultural crop losses; loss of livelihood	Power station: Global climate change; deaths associated with crop loss and starvation	Coalworkers' occupational health; mortality from lung cancer/radon exposure	Coalworkers' occupational health; death from Progressive Massive Fibrosis (PMF)	Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; fatal cancers	Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; non-fatal cancers	Amenity loss and annoyance caused by noise from wind turbine operation
QLOS DATA										
Index number (I)	100	10	2.8	8.8	100	100	100	100	10	0.1
Risk (R)	1*	1*	1*	1	1*	0.002	0.21	$2.85x10^{-4}$	$2.85x10^{-4}$	0.99
Time value (T)	1	0.55	0.025	0.11	1	0.375	0.5	1	1	0.089
Receptors value (n)	4.874	55.113	250.821	3.75	884,615	4000	48	14,357	34,889	40
TOTAL (Qu)	487.4	303.12	17.56	3.63	$88.46x10^{6}$	300	504	414.3	99.43	0.3524
QLOS VALUE (μQu/MWh)	1.07	0.666	0.039	0.008	194,250	0.659	1.107	1.817	0.436	0.868
COMPARABLE VALUE (mECU/KWh)	0.029	$3.4x10^{-3}$- $26.4x10^{-3}$	$0.23 x10^{-3}$- $1.87 x10^{-3}$	0.022	5030	0.047	0.057	0.068 combined value		1.1

Table 10.2 QLOS Valuation Demonstration Results and Comparables

*Not the true value for risk, but since only a risk-weighted value for n is available, the value for R is embedded in n.

Note: QLOS value is normalised to output and expressed in μQu/MWh (Qu is QLOS unit).

Note: The Comparable Value for Impact Number 5 is the upper value from the literature, and is approximately 1.5 orders of magnitude above the median value (ExternE, 1995c).

Note: The Comparable Value for Impact Number 8 is the combined value of Impact Numbers 8a and 8b. It also includes minor additional elements for hereditary effects, early diseases and deaths (ExternE, 1995e) although the combined proportion of these elements is low.

Impact Numbers 1-3 show a similar rank for both the QLOS and ExternE values. Normalising both sets of values shows that there is close proportional correlation between major injury values and deaths, and between minor injury values and deaths, for the two valuation methods. However, extending the valuation to two further impacts (Impact Numbers 4-5), reveals a deviation from the close similarity seen in Impact

Numbers 1-3. Ranking order is still closely similar, with a maximum change of two places, which is that of Impact Number 4, ranked 5th for QLOS and 3rd for ExternE (for QLOS values, the rank order is 5-1-2-3-4, and for ExternE values, it is 5-1-4-2-3). With the inclusion of two further impacts (Impact Numbers 6-7), further deviation from similarity is revealed (rank order for QLOS values is 5-7-1-2-6-3-4, and for ExternE values, it is 5-7-6-1-4-2-3), although the maximum change is still two rank places. The addition of the last two impacts, one from the nuclear fuel cycle, and one from the wind fuel cycle, reveals yet further deviations from a simple parallel set of QLOS and ExternE values. For example, for Impact Number 9, under ExternE, this is ranked 2nd, whereas under QLOS, it is ranked 5th, showing the largest rank difference in the demonstration study. Discussion and rank and graphical representations of the full set of demonstration results are presented in Appendix F.

10.4.2 Interpretation

The first thing to note from the demonstration is that valuation is possible using currently available data. However, further conclusions can be drawn from the comparison between the QLOS values and corresponding ExternE values. In the latter, valuation is invariably undertaken by converting impact data into money terms, using, for example, generic value of statistical life (VOSL) quantities established in the literature, or quantities established by neo-classical environmental economics methods in other studies (which are then used as surrogate values and transferred to the ExternE reference situation). The methods used vary from indirect approaches (such as hedonic pricing) to direct approaches (such as contingent valuation). A description of these methods and of potential problems with them is presented in Chapter 4 and Appendix A, and elsewhere (Horne, 1995). Since every attempt has been made to use similar data for environmental change and human consequence scenarios in the QLOS calculations as those in the ExternE valuations, the principal comparison should be between the method of adopting surrogate environmental economics valuations and the QLOS valuation method.

Perhaps the single most striking observation from the comparisons is that non-injurious impacts (Impact Numbers 4 and 9) provide lower relative values under QLOS. Impact Numbers 4 and 9 are essentially "economic" and "annoyance" in type respectively, rather than physically injurious (or potentially so), as in the other impacts in the demonstration. This indicates that these human consequences have a relatively lesser negative outcome on Quality of Life than they do on utility, as suggested by the ExternE results, using a utility-based environmental economics approach.

Under both methods, Impact Number 5, deaths caused by global climate change-related crop losses and starvation, ranks at the top. Both ExternE and QLOS values are derived using 0% Discount Rates (applying even 10% Discount Rates would still leave this impact orders of magnitude higher than all other impacts in the demonstration). However, it must be noted that no single ExternE value is suggested, and that, on the basis of the results presented, the ExternE value quoted in Table 2 is approximately 1.5 orders of magnitude above the norm or median environmental economics-derived value quoted in this source. The value given in this demonstration is adopted because it is derived from a single human value for all humans (rather than differentiating for "Willingness to Pay" between developed world and developing world deaths), and this is therefore the most similar methodologically to the QLOS approach (where there is no "Willingness to Pay" effect and all human lives are valued equally). If a more norm-based value is used, the ExternE value would be around 10 mECU/kWh, and would still rank top, with approximately an order of magnitude difference between this and the next-ranking ExternE value in the demonstration. It would, however, affect the relative interval difference between the ExternE and QLOS values.

If the assumption is made that death is a suitable point of calibration between the QLOS Index and the monetary-derived estimates presented in ExternE, then a comparison can be drawn between values for levels of injury under QLOS and ExternE. Under this assumption, the results suggest that the QLOS valuation method gives rise to higher values for major injury than ExternE (20.3% higher), but only slightly higher values for minor injury (5.5% higher).

So, there is a significant difference between the economic values quoted for major injuries as used in the ExternE study, and the values suggested by the QLOS valuation. The question which arises here is to what extent this could be the result of the proxy system used in this demonstration for assigning QLOS Index categories to different impact states. In order for the proportional difference in Impact Number 2 to be accounted for by the proxy QLOS category alone, this would need to be varied considerably. Sensitivity tests (see Appendix F) show that a QLOS Index number of 8.3 would be required to produce a value proportionally equivalent to the ExternE value for Impact Number 2, which varies from the actual value used (10) by over 20%.

A further difference arises not from results comparison, but from application of the human consequence method. Impact Numbers 6, 7 and 8 only consider the "death phase" of the impact. However, death from lung cancer, for example, is often preceded by a lengthy period of morbidity. To make the comparison valid, the QLOS calculations are restricted to the "death phase", but illness phases prior to death would be important in any full QLOS valuation, particularly when death is delayed, because a value for illness during the period after suffering starts and before death occurs is integral to the value of the disease. Indicative calculations suggest that illness preceding death may provide a significant additional element to the overall impact value in some cases. For example, it may easily elevate Impact Number 6 above Impact Number 2, major traffic-induced injury (where no morbidity phase would apply), demonstrating that this factor makes a material difference to relative as well as absolute impact value.

10.5 Valuation Method Criteria Compliance

The QLOS valuation method is a theoretically robust and yet practically simple means of generating and combining appropriately the necessary elements which contribute to a given range of outcome states arising from a set of environmental changes. It is also transparent, logical, based on clear theory, accessible and adaptable. There are points at which potential exists to build more sophistication into the procedure, where a need is demonstrated for this, to avoid loss of accuracy. For example, it is possible that some instances may occur where maintaining the assumption of Constant Proportional Time Preference, or the assumption that all impact values act independently from each other without synergistic effects, may lead to loss of accuracy. In such cases, adjustments can be made to correct for such effects once they are proven and understood.

Given the sound logical and theoretical basis, the key tests which determine the efficacy of any value measure are validity, reliability, sensitivity, and efficiency. Validity is established in theory, and this is further supported by the comparison with other results in the demonstration. Reliability is the extent to which results can be reproduced under varying circumstances, and it exists when the variables being recorded are those which are intended to be recorded, and not forms of bias. Bias is minimised by maximising simplicity, clarity, and transparency – all of which are prominent within the QLOS valuation method. Sensitivity is established through the variation in impact values obtained, while efficiency and practicality are apparent from the success of the valuation demonstration using existing data from a single research project.

In summary, the QLOS valuation method has been developed and described, and successfully demonstrated. It is theoretically and practically sound, and meets the key criteria requirements of validity, reliability, sensitivity and efficiency, suggesting that it will produce useful, relevant, accurate measurements of environmental impacts in QLOS units. Valuation can be carried out simply and effectively, assuming data is available. The next and final method in the systematic framework following valuation is the application method, Step 9, which involves applying the valuation results to decision-making about planning, design, construction, operation and decommissioning of projects, policies or programmes. Therefore, it is this issue which must now be focused on in Chapter 11.

11. APPLICATION OF QLOS VALUES

In Chapter 3, it was established that considerable gaps and inconsistencies exist in the current environmental regulatory framework. One of the main problems is lack of objective information about impacts. Steps 1 to 8 of the systematic framework, as discussed in Chapters 6 to 10, provide a means of addressing this shortfall by providing objectively-based impact valuation information. The application method, Step 9, is the means of determining an appropriate way or ways of applying this objective-based valuation data in practice. Invariably, it involves changing existing regulation(s) or creating new ones. The term "regulation" is used here in its widest sense, meaning any mechanism by which QLOS values can be reflected appropriately in project decision making and operation.

The application method is simple in form, and involves three main tasks, as follows:

- Set out the criteria which the new/improved regulation or application device must meet;

- Set out the range of possible options;

- Select the most appropriate option(s), based on fitness with the criteria.

This method has been applied for Total QLOS value data for a generic ESI project, and the results are presented below, drawn from elsewhere (Horne, 2000c). The criteria which any new regulation must meet are laid out in Section 11.1. The range of possible regulatory options are identified, and this is presented in Section 11.2. The proposed regulation is outlined in Section 11.3, and assessed for fitness against the criteria requirements in Section 11.4.

11.1 Regulation Criteria and Approach

The following points have been proposed as determinants of constructing good regulation (DTI, 1994):

- Identify the issue and match the regulation to it closely;
- Keep it simple ("goal-based");
- Provide flexibility for the future, by setting objectives rather than details wherever possible;
- Keep it short;
- Minimise costs of compliance;
- Integrate with previous regulations;
- Make sure the regulation can be effectively managed and enforced;
- Make sure the regulation will work and that you will know if it does not;
- Allow enough time for consultation, drafting and phasing in.

Based on this, the following criteria are applicable to the selection of regulatory measures which are suitable for applying QLOS values:

- Addresses weaknesses and inadequacies identified as current and ongoing;
- Tried and tested, and good 'fit' with current policies and regulations;
- Able to deal with and fully utilise QLOS value data and, therefore, the impacts they represent;
- Minimise free-riding (i.e. the practice of causing/contributing to impacts but not bearing the consequences);

- Simple, in terms of ease of introduction and enforcement;
- Reliable, in terms of likelihood of compliance and, therefore, of success.

The optimum regulatory option for applying QLOS values is that which best matches the criteria above. Any application must be able to utilise the total impact value figure which is the sum of all impact values for a given project, and is expressed in μQu/MWh. Also, it must be capable of taking into account other aspects of Total QLOS data, including that relating to potential issues for which no or poor information and/or knowledge exists. Furthermore, it must be able to utilise the transparency of the QLOS data to ensure quality control of the QLOS exercise and the results, and to enable QLOS values for individual impacts to be examined where necessary in the regulatory process. Meeting these requirements will ensure that the application makes full use of the QLOS data and, therefore, of the resources which are required in the QLOS exercise.

The correct approach to selecting and developing appropriate regulation which meets the criteria is determined by the needs of practicality. As with the output analysis, pathway analysis and valuation methods, they are not being developed in a vacuum, but in a situation where current practice already exists. Any new regulation should be based wherever possible on existing regulation, since this provides familiarity and experience, both of which are important for successful application. The review of the existing environmental regulatory framework in Chapter 3 is directly relevant here, since it provides an overview of the regulatory options available within current practice. This review is also important because, irrespective of the form or type of any new regulatory mechanism(s), it/they must be designed to operate within this existing wider framework. Furthermore, the review provides a summary of current weaknesses, knowledge of which should assist in avoiding the creation of further such weaknesses in any new regulation.

One of the biggest current problems is the number of gaps and inadequacies, where impacts (or residual impacts) continue either unregulated or partially regulated. The main solution to many of these gaps, which has been discussed at length in the literature, is the application of market mechanisms, such as taxes, subsidies or other money-based adjustments, to correct for market distortions created by the lack of incorporation of environmental values in the market. It has been established in Chapter 3, and in the literature, that the basic problem with this proposal is that while theoretically, the approach seems simple and effective, in practice, it is hampered by one fundamental flaw – some impacts cannot be measured in money terms, at least with acceptable consensus. Therefore, the externality monetisation and market mechanism approach is largely eliminated from the list of possibilities. However, there are aspects to this approach which are attractive, such as the implicit drive towards optimum efficiency, through identification in the market of "optimal pollution" – the level of impact at which benefits from the goods/services produced are balanced with the total costs (including impacts). As technology develops, this point changes, and an efficient regulation which incorporates such a mechanism is desirable. However, it must use non-monetary, quantified, objective data, as produced by the QLOS valuation method.

11.2 Assessment of Options

The following options have been proposed as general possibilities for approaches to regulation (after DTI, 1994):

- Do nothing;
- Review existing law and improve compliance;
- Licensing;

- Legislation;
- Economic instruments;
- Self-regulation;
- Use Codes of Practice which appear to have force of law;
- Use voluntary schemes such as Codes of Practice;
- Improve information and retain existing framework.

In reality, the first two of these options can be discounted for the application under consideration here immediately. "Do nothing" is not an option, as it leaves the inefficient status quo in place. "Improve compliance" is not sufficient, since the regulatory framework is generally lacking, as outlined in Chapter 3. Impacts are not being taken into account appropriately and there is a potential for significant free-riding, thereby rendering these options unfeasible. The four remaining approaches are therefore identified as general possibilities. Taking the review of the existing environmental regulatory framework in Chapter 3 as the starting point, the merits of each are presented below. Command and control options are considered in Section 11.2.1; market mechanism options are considered in Section 11.2.2; self-regulation options are considered in Section 11.2.3; and planning framework options are considered in Section 11.2.4.

11.2.1 Command and Control Options

Command and control measures may be set in each industry or plant type or project (for example, through planning or licensing controls), or centrally determined (international or national) legislative limits on pollution. Licensing is an option to control or restrict unsuitable operators or operations. There are two stages of licensing; at new development stage, prior to construction, and at operational phase, during updating and/or monitoring. Considerable scope exists for improvements to current licensing, in the light of particular impacts as identified and measured by QLOS valuation. Particular RIOs excluded from the current system could be controlled by licensing. However, there is limited potential here since, firstly, the impact can only be reduced and never eliminated by this method and, secondly, the range of impacts which can be controlled in this way is limited by the availability of technical options.

Legislation has the apparent advantage that standards can be laid down directly. Operation and compliance is therefore relatively simply achievable, although setting up legislation is generally expensive and time-consuming. For QLOS values which show clear regulatory gaps, for example, impacts associated with human-induced global climate change, direct legislation may be appropriate, particularly if the impact is reasonably well-understood (clear, low-risk), technologies to provide alternatives are available, and retrospective regulation is desirable or feasible. Similar constraints apply as those to licensing, in that residual levels of impact are likely to be left which are allowed, not legislated for, or not complied with.

One of the major problems with command and control limits is that QLOS data is much more sensitive, detailed and instructive than the "yes/no"-type structure of such limits. They do not allow any differentiation to be made between projects other than whether they meet the criteria or not, and so they do not encourage developers to move towards an optimal level of (lowest achievable) QLOS. This would be left to the regulator, in setting the required level of control, which would be a potential point of weakness in a command and control approach in the case of QLOS values. Caps on pollution or risk must be set which are acceptable in environmental and societal terms, by keeping or reducing impacts and risks within known limits, but which do not reduce the welfare produced by economic activity, by setting in place prohibitively punitive measures on polluting industry. Such an approach to regulation meets the demands of simplicity, but it may not provide sufficient sensitivity in utilising QLOS values, and would invariably leave residual impacts.

11.2.2 Market Mechanism Options

The only applications of economic instruments in the recent regulatory regime are taxes and a form of subsidy (the Non-Fossil Fuel Obligation). Both are immediately problematic for QLOS application because they theoretically rely upon the notion that the environmental impacts which need regulating for can be expressed in money terms, and that the market will then deliver optimal pollution in reality as well as in theory. Placing the costing issue momentarily on one side, the goal of market efficiency is not necessarily achievable by application of market mechanisms, due to numerous imperfections in the market. Inertia occurs in both industrial planning and technology lead-in as well as in consumer habits and electricity use. It has been suggested that 50% surcharges are needed to reduce consumption by 20% and almost 3 times this to get a 50% reduction. Thus, simple internalisation may be insufficient to achieve environmental goals (Jones, 1990).

More fundamentally, the use of market mechanisms in the classical sense is not possible for applying unmonetised impacts such as those represented by QLOS values. Since QLOS impacts are not monetised, the notion of theoretical optimisation by reflecting environmental costs in prices is lost. However, it is relevant to discuss some of the issues raised, because some of these mechanisms may be more dynamic, effective and optimal than, for example, traditional command and control limits. Therefore, elements or structures of market mechanisms may be relevant in developing QLOS regulation, even though monetary quantities are not applicable.

The various types of market mechanisms (or economic instruments) were reviewed in Chapter 4 and Appendix A, as well as elsewhere (for example, Horne, 1995, OECD, 1994, EC, 1994, Tietenberg, 1994). Market mechanisms can be more efficient but less predictable in effect than conventional regulations, although they hold out the apparent possibility of being sophisticated enough to taper at the edges. In other words, provided a sliding scale mechanism (or surrogate market) is applied to reflect the sliding scale of impact (temporally, spatially, etc.), then taxation/compensation (whatever the mechanism) can theoretically be developed to internalise impacts according to their severity. This improves upon licensing or legislation, where traditional cut-offs invariably apply, below which impacts occur and the polluter can free-ride.

The view in favour of shifting the balance of taxation away from classical income and value added type taxes, designed purely for state revenue-raising purposes, and towards those undesirable elements of life such as environmental impacts and use of non-renewable resources, is fundamentally simple. The logic is that taxes should be applied to things which are not wanted, rather than things which are. The theory is that, since taxes act against the activities they apply to, this shift will result in more of what is wanted (wealth creation) and less of what is not (environmental degradation, in this case). Incidentally, there are wider policy issues, including that resource taxes are more punitive in a society where wealth is unequally distributed.

Strictly in terms of market efficiency, the logic of environmental taxes is appealing. The current "command and control" based legislation and planning system tends towards the dominant business view, which is that if the damage is outweighed by the benefits of the activities, then the project (and the damage) should go ahead. Green taxes could potentially be much more sophisticated, in making projects expensive in proportion to their impact. Thus, activities are not banned, but become prohibitively expensive, the greater the potential for environmental damage. In theory, marginal rates of pollution are automatically generated (Gee, 1993, Geller et al, 1993). The benefit is two-fold; not only are activities not banned, but they are also taxed below the margin, albeit at a lower rate. So relatively minor levels of impact are reflected in the costs of the development, via the sliding green tax scale.

However, there are problems with applying market mechanisms to impacts. With green taxes, the business polluting pays, although it is the society as a whole that suffers the impact. Therefore, it does not strictly correct the inequality, although it does correct the market distortion. Consumers pay more for receptors to receive less impact. It would be a mistake to assume that these two groups are the same, so that residual impact is evened out by lower prices. The fact is that those who can afford to avoid avoidable impacts will do so, while those who cannot, will be forced to accept the residual. Nevertheless, partial correction of the current market distortion is better than none. Even if all environmental impacts were unlikely to be captured, a market with fewer externalities is better than one with more.

Further problems for market mechanisms arise from the inherent inertia and irrationality of the market. Decisions to pollute or not are not made purely on the basis of new-build costs, to which environmental taxes could be simply added. Indeed, it is widely accepted that to correct market distortions, taxes need to be applied at up to double the strict correction level, in order to overcome inertia, as well as business and technology barriers. This suggests that market mechanisms may not be as effective in practice as they are in theory. A further argument is that taxing business for polluting is problematic, unless consumers and others are also subject to similar taxes when they pollute, so that responsibility for pollution is shared amongst all those involved in the process. Also, if environmental taxes are applied, then they will inevitably be aimed at particular, currently free-riding, polluting industries, such as the ESI, and this will lead to major restructuring and change across the economy. If similar taxes are not applied in other countries, then electricity will be relatively more expensive, so industries which use large amounts of it may be forced to relocate to another jurisdiction with a different tax regime.

Apart from taxes and their counterparts, subsidies, there are two other possible forms of economic instruments; market creation and enforcement incentives. In market creation, the aim is to internalise by generating a working market in externalities. Values or rights are applied to the externalities, so that market-based choices can be made about whether, where, when, or how much environmental damage is created. An example is emissions trading, where the total pollution is set and split into quotas. Permit holders can then trade their pollution rights as they wish. In enforcement incentives, returnable bonds are created which are recoverable on demonstrating compliance, or fees are imposed on non-compliers.

The current proposals for the Climate Change Levy involve an element of enforcement incentive, since there are various linkages which allow for waiving of the tax in exchange for investment in non-carbon initiatives. Incidentally, the form of the emerging Climate Change Levy provides an example of the worst kind of market mechanism. There is no explicit application of monetised or unmonetised impact value in the Levy. Indeed, it is not even related to the amount of carbon emitted (hence, it is not a "carbon tax"). In short, in the absence of reliable values for impacts, a mechanism has been created which is arbitrary in that it does not attempt to reflect impact values in the Levy which is designed to address the market distortion. Hence, it can never lead to market optimality.

In summary, one problem for all market mechanisms affects the potential for QLOS application; QLOS values would need to be converted to money terms. This is contrary to the well-established fact that many impacts do not lend themselves to monetary valuation. Also, a general principle of the QLOS valuation method is that the quantities produced are used to provide decision makers with clear, transparent and reasonably objective data with which to make decisions, rather than having decisions taken away from them by an arbitrary monetisation process and incorporation into cost benefit analysis. Therefore, market mechanisms in the monetary sense are, by definition, not appropriate regulatory tools for QLOS values.

11.2.3 Self-regulation Options

Self-regulation, "use Codes of Practice which appear to have force of law" and "use voluntary schemes such as Codes of Practice" are fairly closely related, in the sense that they involve a degree of voluntary initiative from business (beyond compliance). Voluntary regulations have received a great deal of attention in recent years, particularly in relation to environmental control (OECD, 2000). Codes of Practice must by necessity be clear and simple, but QLOS impacts typically are not. This is not to say that some may not be suitable for application via a Code of Practice, such as further reduction of currently regulated noise impacts. In this case, a Code of Practice which ensures further reduced noise annoyance would reduce the impact beyond that currently occurring with regulatory compliance. However, for Total QLOS values, which encompass a wide range of QLOS values with different variables, a more complex self-regulation tool is required, and this is currently available in the form of an environmental management system, as introduced in Chapter 3. It should be noted that in order for it to be within business interests to set continuous improvement targets for reducing QLOS values, the impact must have considerable potential for reduction, which is usually through technical means. Therefore, only a restricted range of QLOS impact types could be expected to be addressed and, due to the nature of the system, it is likely that these may involve continued impact, albeit reduced.

11.2.4 Planning Framework Options

Improving existing regulatory tools is an attractive option because it provides the benefits of being built on existing experience and is therefore both familiar and simple. The planning framework has been developed periodically throughout its history. The single most significant strengthening in the area of environmental appraisal occurred in the form of the EIA regulation, which came into force in 1988. The EIA process is an important environmental development control element, and it is not currently quality-controlled by an overseeing body. Also, it generally suffers from a lack of transparency and objectivity, partly due to the nature of the process and partly due to historical factors in the way in which the process has evolved. Thus, Total QLOS values would be directly beneficial to the planning process, since its standard, explicit, transparent and accessible methodology would ensure a higher overall level of quality of information for decision makers, which could then also translate into more transparent accountability and improved decision making.

The planning system has developed within the area of judgement, comparison and trade-off of differing values. It is therefore geared towards dealing with complex data, and Total QLOS values are not simple mathematical quantities. Firstly, they convey information about human health and well-being, which is not a simple construct, and secondly, they are expressed as "greater than or equal to" rather than simply "equal to", simply because QLOS will inevitably always potentially require knowledge and information which does not exist. Even if 100% mass balance is achieved, and 100% certainty is achieved in pathway analysis and environmental change measurement, human consequences may vary over time as human values do.

Total QLOS values could be utilised in the existing development control system, and could be expected to improve it. One reason for this is that the former consists of values data about what people perceive, whereas the latter concentrates on what politicians and decision makers consider society perceives through the planning, policy and legislative process. Therefore, QLOS is a more direct approach to providing data on impacts than the traditional planning approach. By way of example, the Public Inquiry system is highly adversarial, partly due to the lack of clarity and availability of information about the impacts of the project under review. QLOS values provide direct and transparent data on impacts, and so could be expected to be of direct benefit to the Public Inquiry system.

The most suitable areas for application through alterations to the planning system are impacts which are site-specific, have a varying dynamic/time profile, and potentially involve high risk and/or a low state of knowledge. This is because these are more complex and can best be dealt with by local level decision making on a case by case basis, within a wider QLOS policy framework which sets out concepts and models for planners to use. The idea of incorporating a specific quality of life based evaluation system into the planning system is not new (for example, Drewnowski, 1974), but it has not been attempted in practice. More recently, guidance on sustainability indicators have been produced, which incorporate the term "quality of life" and are designed to assist in sustaining quality of life at national and local level (DETR, 1999, 2000). The critical difference between both these examples and the QLOS approach is that, in the latter case, the appropriate measurement method was designed first, followed by assessment of the most appropriate means of application, rather than the other way around. No theoretical basis is provided for the choice or importance of different sustainability indicators, and no method is provided for comparing significance. It is also an example of the typical problem centred approach, where problems are acute before action is taken to control/prevent. The QLOS approach marks a shift away from this to a more pro-active approach, where impacts are sought and measured at the initial, pre-impact stage.

11.3 The QLOS Tranche System

It is established that a regulatory mechanism for applying Total QLOS values must start from that which is currently in operation, and be practically achievable. It must minimise economic pain but allow full incorporation of effects. Although economic activity and environmental objectives are not necessarily diametrically opposed, the best decisions cannot be made where two or more differing units are involved in the quantities to be considered. The planning system does largely address the problem of balancing impacts and financial criteria, by leaving the latter to the developer and incorporating assessment of impacts directly into planning decision making. However, there is a strong benefit in a mechanism which balances costs, impacts and benefits to achieve optimal pollution, and the planning system currently does not incorporate this. Furthermore, QLOS data contain information on gaps, etc., in addition to known impact values. The current planning system would be less able to use this, other than in setting planning conditions.

A solution is a system which separates financial considerations and leaves them to the developer, but also incorporates a mechanism which should ensure optimal pollution is reached. A driver is required where developers can be competitively engaged in reducing the environmental impacts of their operations. This can be achieved by allowing comparison between proposed schemes on the basis of µQu/MWh and other QLOS data, and incorporating such a comparison into the criteria for decision making about whether schemes should proceed or not. Thus, the basic outline of the proposed system for incorporating QLOS values is as follows:

- Developers must apply for permission at national level, in a competitive process with a regular cycle of tranches (for example, every 2 years);

- Developers must supply a full QLOS valuation with their application and must make life cycle, output analysis and pathway analysis data available on public registers;

- The industry regulator, OFGEM, applies "will secure" checks to all applications, to ensure that each proposed scheme is viable;

- Decision makers have the opportunity to clarify, scrutinise and request more information on individual QLOS valuations (planning professions and the Environment Agency will be involved in this);

- Decision makers rank all the competing schemes in the tranche, with the cut-off being established according to national electricity requirements at the time;

- All schemes above the cut-off (i.e. with the lowest μQu/MWh ratios, and accounting for other data and non-environmental criteria) will automatically gain planning permission.

The result is a system which does not internalise in the strict economic sense, but does optimise in terms of ensuring the electricity service is provided with the minimum of total QLOS impact.

11.3.1 Practical Implications

A number of practical implications flow from the basic form of the proposed QLOS Tranche System. Firstly, the issue of lack of knowledge and its role in the regulatory process is important. In general, it is desirable for unknown impacts with potentially large risks to be eliminated at source and regulations should be designed to achieve this wherever possible. Where this is unachievable, producers of impacts must be held fully accountable for them. The essential approach to regulations for this type of effect is one of protecting the environment against unknown and unacceptable risk, accepting that this may not deliver optimal welfare, but erring on the side of environmental protection rather than on providing economic welfare combined with unknown risk. This approach could be described as precautionary and is compatible with sustainability criteria. Practically, this means that high QLOS estimates, or high weighting generally, must be applied to issues where knowledge is lacking by decision makers, in altering ranking or otherwise accounting for these in selecting projects for development permission.

Transparency of the QLOS valuations of past proposed projects, and other current proposed projects, is another important practical issue. Since one of the most common reasons given for having weak environmental regulation is that knowledge about environmental impacts is poor, it is clear that all QLOS information should be in the public domain (on public registers). This will help speed up the process of understanding and knowledge about environmental issues and the means to reduce impacts, and will clearly improve the flow of information required in QLOS valuations.

It should be noted that other, non-environmental criteria may be used in addition to Total QLOS values, in the process of decision making. Such criteria are clearly outside the scope of the QLOS approach or environmental regulation, but it should be noted that the relative weighting of such issues will be left to decision makers. There may, for example, be a need for social criteria, or for some equity of distribution of projects by region/locality. Regarding the latter, with the use of standard Geographic Information System (GIS) based QLOS software, a national picture of QLOS distribution for an entire proposed tranche can be established by merging data from different proposed schemes, thus giving a clear picture of the QLOS concentration by locality. This will form the basis of expanding the QLOS-based planning process from the project level to the programme/national level.

Another practical implication of the new proposed regulation concerns its development over time. The basic proposal is dynamic, in the sense that each successive tranche will incorporate improved QLOS accounting methods, based on current knowledge and experience. Comparability will be established by applying the latest generation of QLOS valuation techniques equally across each tranche. There will also be feedback from the operation of the QLOS Tranche System. Once in operation, post-auditing of actual QLOS (residual impact) can be used to feed information back into the policy process, particularly regarding the development of future projects. Thus, cumulative and synergistic tendencies at and above project level could be monitored, and the results built into additional criteria in future tranche assessments.

The issue of geographic scale of QLOS regulations is also important, mainly due to the problem of free riding. Ideally, the QLOS Tranche System should apply across the entire industrial sector, in order to maintain a level playing field in the market. Although there is no direct burden on developers arising out of the regulation (apart from the resource costs of the QLOS exercise), it is inevitable that developers outside the jurisdiction of the QLOS Tranche System could theoretically produce higher-impact, cheaper electricity. In reality, the ESI is a good example of a closed system, since there is only marginal import-export of electricity into and out of the UK. Thus, the ESI itself will not be subject to unfair competition. However, as pointed out above in the discussion on market mechanisms, if electricity becomes relatively cheaper outside the jurisdiction of the QLOS regulations, then electricity-intensive industries may be forced to relocate. However, the system can be effectively governed by decision makers at the point of granting planning permission. The judgement to be made is between the level of residual externality (which is minimised in the lowest-QLOS schemes) and the potential for higher prices, due to the bidders into pool reflecting the internalised costs of their low-QLOS schemes. In short, the risk of industry relocating to countries which tolerate higher electricity-related environmental impacts is no reason to abandon the QLOS regulation, since this pressure will apply to any attempt to address external environmental impacts in the industry.

One final issue relates to how and where the cut-off is set in each tranche. Currently, a range of demand-based criteria are used in the planning process, and these can readily be applied to the issue of deciding how much capacity is given permission in any given tranche. The current pool pricing system of the ESI will also be unaffected in its operations. Therefore, the general principles of the demand planning process will remain intact, with allowances made for spare capacity and security of supply, etc., to ensure the continued smooth-running of the industry.

11.3.2 Possible Additions

The basic form of the QLOS Tranche System as presented above, is a workable and practical proposal. However, it is also possible to add some more tentative proposals which could provide further additions to the system. Subject to further analysis, these could conceivably improve further the optimality of the proposed mechanism.

Firstly, developers could link the closure of their existing power stations with the development of new ones in a single project, thus benefiting from an overall QLOS reduction given new technology. A generally established principle is to avoid retroactive regulation wherever possible since it provides an unpredictable burden on developers, and this would also contribute to addressing this issue.

Secondly, a similar system to that proposed here for the ESI could be envisaged for other industry sectors and other permissions. However, it is expected that other issues would arise. The ESI is a particularly useful candidate for trialing QLOS application, since the planning regime is already effectively centralised to national level (decisions are invariably made by the Secretary of State).

Thirdly, the system could be monitored for ongoing compliance of predicted QLOS values with actual values, by an environmental regulator such as the Environment Agency. If a project exhibited a significantly larger QLOS than predicted, the responsibility for this would be shared between the regulator, who has responsibility for checking the efficacy of the QLOS valuation, and the developer/operator, who has responsibility for enforcement of the standards and specifications in the design which are predicted to produce the necessary QLOS values. This is the case generally with compliance with pollution authorisations at present.

Fourthly, while issues of social and economic equality are excluded from environmental impact valuation, it is possible to envisage how adjustments could be made to the QLOS Tranche System in order to address the problem of equity of impact and use of electricity. One way of minimising the inequalities associated with this arrangement would be to allow for ranking adjustments in situations where project impacts overlap to create a large QLOS burden in a local or regional population. In other words, projects could be chosen so that the residual QLOS is spread across the population rather than being concentrated in one area with a large number of ESI facilities. This would effectively equalise externality effectively, since the vast majority of people are electricity users and so are also receiving the benefits. Another means of maximising equity would be to look into the individual QLOS values for projects which have similar QLOS totals, and select the scheme with the fewest deaths and other high QLOS intensity outcomes. Since very many lower QLOS annoyances are required to equal one death, by definition, choosing the project with the wider, flatter QLOS profile (of many annoyed people) rather than the narrower, sharper one (of a few deaths), would ensure greater equity in the spread of the residual QLOS impact.

11.4 Regulatory Criteria Compliance

The requirement to "Identify the issue and match the regulation to it closely" (see Section 11.1) has been addressed by examining the four main types of regulatory tools, as reviewed in Chapter 3, and assessing their weaknesses and potential for applying QLOS values. It is concluded from this that the QLOS regulatory mechanism should be applied to strengthening the project planning system. In addition, three other regulatory options appear to have limited potential in internalising QLOS values; self-regulation, strengthened licensing for pollutant discharges, etc., and legislation. Self-regulation could be used, through an environmental management system which incorporates public statements, such as EMAS, but it is relatively new and so fails to meet the "tried and tested" criterion. Furthermore, it may not avoid free-riding, as it is voluntary, so some businesses may choose not to participate. Licensing and legislation have relatively limited potential to assist the application process, and the main opportunities are limited to regulatory gaps in the current system, where impacts are well understood. Market mechanisms have the potential to be more efficient and more sophisticated, but they suffer the inevitable problem that QLOS values are not monetised. Hence, a competitive element is incorporated into the QLOS Tranche System, in order to incorporate the benefits of the market based approach, without the need for monetisation.

The proposed QLOS Tranche System fits the "tried and tested" requirement, since the development control system within which it is incorporated has an established history. Also, a not dissimilar structure of timed tranches has been tried and tested in the Non-

Fossil Fuel Obligation, albeit in this case on a cost basis; but it establishes that a system of competitive pooled tranches can work. It also exhibits good fit with current policies and regulations, since the permission system for projects within the ESI is already effectively operated on a national level. Furthermore, there is no conflict between the QLOS Tranche System and the logical means of internalising impacts which can be monetised, that is, by incorporating them directly into ESI costs, for example, by externality taxation. There is no risk of double-counting, where impacts are both internalised and included in the QLOS Tranche System, since the QLOS valuation method allows for recognition of such existing regulations.

The third criterion, to minimise free-riding, is met, since only those developments with the least uninternalised impacts will proceed. However, it should be noted that some potential still exists for free-riding, given that a permission is granted on the balance of factors, and that this balance includes impacts, some of which will go ahead uninternalised once permission on the basis of overall ranking is obtained. Possible additions to the QLOS Tranche System, as discussed in Section 11.3.2, could deal with this residual free-riding potential.

The need for simplicity and reliability drive the two remaining criteria. These are critical to the success of any regulation and, as the discussion in Section 11.3.2 demonstrates, the proposed QLOS Tranche System fully meets these requirements. The structure of the tranche system is simple, in terms of ease of introduction and enforcement, and is reliable, in terms of likelihood of compliance and, therefore, of success. Indeed, the decision to base the application of QLOS values around the existing project planning system is largely a result of its fitness with reliability criteria, ability to deal with the wide range of variable attributes of QLOS impacts, and simplicity of function. Also, coincidentally, the existing system is in need of a non-monetary, quantitative, objective, quality of life based valuation system in order to improve the data upon which decisions are currently made. The project environment appraisal elements of the planning system are currently lacking in objective data with which to make decisions, and therefore QLOS data will fulfil a need. It will provide an objective counter-balance to the wider political influences within the development control system, and provides a means of obtaining objectively measurable sustainable decision making. Consequently, it will have benefits for decision making within the development control system. It also clearly utilises the complexity and uniqueness of Total QLOS data fully and, therefore, the impacts they represent.

In summary, the QLOS Tranche System meets the criteria set, and is a viable regulatory option for applying Total QLOS values in the ESI. It is based on the existing planning system, which is structured appropriately and has current needs for data sources, such as those that QLOS values can provide. The tranche format brings market-based optimal impact efficiency to the proposal, ensuring that projects adopted are those where developers have successfully competed to produce projects with the lowest possible μQu/MWh ratio. Upon application, improved decisions can be expected, since they will be made with more and better information. Given that the QLOS approach is both transparent and accessible, the legitimacy of the decision makers' work will also be enhanced in the eyes of those who will be affected by the decisions made.

The systematic framework is now complete. Over Chapters 6 to 11, the framework and each of the nine Steps have been presented in detail. It remains now to turn to a comparison between what is proposed in the systematic framework, and what exists currently. Although some references have already been made to current practice, this will be the specific focus in the next Chapter, providing an opportunity for similarities, differences and improvements to be highlighted.

12. COMPARISON OF METHODS

The systematic framework developed here represents a new set of methods which, when linked together, provide a means of measuring and regulating for environmental impacts associated with the ESI. It draws on varied and developing academic disciplines and provides them with a new application. However, it also has links with existing methods. It is important that the similarities and differences between these existing methods and the systematic framework are highlighted, in order to indicate where improvements are possible by adopting the latter. Such comparison is the subject of this Chapter.

In the 1980s, the main impetus for measuring environmental impacts arising from the ESI came from a perceived need to reduce failures and imperfections in the market, by internalising external costs. This led to a strong reliance upon the neo-classical environmental economics approach to impact assessment. While problems with this were discussed in Chapter 4, a comparison with the proposed systematic framework approach is needed, and this is presented in Section 12.1. The monetisation-led approach culminated in a number of external costing studies, as discussed in Chapter 5. A summary comparison of these with the systematic framework is provided in Section 12.2. A method of impact assessment with more initial similarities to the systematic framework is Life Cycle Analysis (LCA), and comparisons with this are given in Section 12.3. The ExternE project is the most significant study to date as far as the valuation of ESI environmental impacts and the systematic framework is concerned, and comparisons with this are considered in Section 12.4. Finally, regulatory comparisons are examined in Section 12.5.

12.1 Environmental Economics Valuation Methods

In Chapters 4 and 5, problems with environmental economics methods were discussed in both theoretical and practical terms. Two principal issues arise from environmental economics methods as typically applied at present. Firstly, the means of identifying environmental impacts and collecting data about them is not rigorous enough to ensure that the values obtained are sufficiently accurate and that all impacts are included. Secondly, there is evidence that monetisation is not necessarily the most appropriate means of valuing impacts. In particular, a different method is needed to value those which are not normally experienced or thought of in money terms.

Numerous problems also arise within monetary valuation methods themselves. In theoretical terms, there are discrepancies between different methods, as well as various biases. Most fail to capture Total Economic Value, since they cannot accommodate existence values or other aspects, so valuations are generally underestimates. The Contingent Valuation Method, where values are elicited directly from people, is arguably the most theoretically sound in this regard. However, even this method has considerable potential for biases, including starting point and design biases, where conduct of the survey or bidding process leads to conflicting values, and strategic bias, where people deliberately give incorrect values because they feel that they may benefit in some way (for example, by free-riding). Furthermore, it was found that a large discrepancy often exists between values of Willingness To Pay and Willingness To Accept, as discussed in Chapter 4.

Since the systematic framework has been developed partly in response to problems identified with environmental economics, it is not surprising that there is a major difference in approach, particularly in ensuring data accuracy and completeness. In practical terms, environmental economics methods are typified by incorrect, insufficient or restrictive scoping and selection of the environmental impacts to be valued, plus partial impact selection processes, with resultant lack of transparency and completeness. Many studies use poor, incomplete, or over-aggregated data in an attempt to obtain results, while also displaying a lack of attention to difficult or unquantifiable impacts, effectively valuing them at zero. In short, there is a general lack of rigour. Added to this, the use of money as the sole unit of value also incorporates questionable assumptions about public perception and experience of money, including the implicit belief that people make rational choices when faced with complex situations and partial information.

In contrast, the systematic framework is rigorous as well as being both theoretically and practically valid. It avoids the weaknesses associated with environmental economics methods, by providing a sequential set of steps for producing objective data and undertaking valuation with these data. This standard and objective approach, incorporating four linked methods, ensures that arbitrary selection of impacts and use of inappropriate, poor, incomplete, or over-aggregated data is avoided. Furthermore the lack of transparency in many environmental economics-based studies is replaced by methods in which both transparency and demonstrable completeness of data are maximised. Thus, the problem where difficult or unquantifiable impacts are omitted, effectively valuing them at zero, is overcome, since, according to the systematic framework procedure, these must be entered as gaps in the analysis. The selection of impacts to be valued is no longer left to chance or to arbitrary and subjective judgement of researchers; it is determined by the application of the output and pathway analysis methods, and subsequent production of the RIO, Pathway and Environmental Change Inventories.

The problems of attempting to achieve monetary values for phenomena which are not normally valued in monetary terms is overcome by use of QLOS units in the systematic framework approach. The logic of this is based on the idea that effects of environmental impacts on quality of life state provide a more valid measure of value than money, particularly in the case of impacts which are not generally thought of in money terms. Furthermore, the methodological problems of individual environmental economics valuation methods are overcome where the QLOS valuation method is used. Thus, problems such as starting point and design biases, where conduct of a Contingent Valuation Method survey or bidding process leads to wrong values, are avoided, as are strategic bidding bias and the theoretical problems over the discrepancy between values of Willingness To Pay and Willingness To Accept. Various other weaknesses of specific methods exist, as discussed in Chapter 4, and are overcome by use of the QLOS valuation method within the systematic framework.

It should be noted that the systematic framework is designed to provide a means to value non-economic impacts and apply them through an appropriate regulatory mechanism. However, it does not discount the possibility of impacts existing which can be most appropriately valued in monetary terms, for the purposes of application through current economic-based regulatory mechanisms. Therefore, the two approaches to valuation can theoretically co-exist, each providing measures for different elements of value. The major caveat here is that any monetary-based approach must address the issues of rigour in ensuring data accuracy and completeness prior to valuation. This could be achieved by preceding any monetary valuation with application of the first two methods of the systematic framework.

Following valuation, environmental economics based methods implicitly or explicitly assume that regulation will be automatically taken care of through market mechanisms. This amounts to insufficient regard as to the needs of regulators and decision makers, leading to lack of confidence in valuations by those intended to use monetary values in the regulatory framework. It also indicates an overconfidence in the theoretical validity of market mechanisms, and a lack of recognition for what happens in practical reality, such as inertia. By incorporating application into the systematic framework, such problems are avoided, since practicality is a central test of any proposed regulation, and all environmental impacts, once valued, are included. Above all, the rigour of the systematic framework ensures that valuations will be accurate, thus allowing regulators and decision makers to have the necessary confidence in them.

12.2 Pre-1995 ESI External Costing Studies

The review of eight costing studies in Chapter 5 (and Appendix B) illustrated the approaches taken to environmental impact valuation and regulation for the ESI up to the mid-1990s. One of the first large studies of ESI external costs was undertaken in 1988 (Hohmeyer, 1988), and was based on fuel cycle comparisons using aggregated data. Unfortunately, aggregated data cannot provide sufficient detail to take account, for example, of impact value variations based on site specific factors. In short, such aggregated data-based studies were sufficient to highlight the problem of external costs, but are not sufficiently accurate to allow valuation of these externalities with adequate confidence.

Other studies based on aggregated data followed, using various methods. One notable study attempted a more detailed level of assessment (Ottinger et al, 1991). However, all these studies share a similar weakness in their initial approach to identifying the main external costs for valuation. There is no objective method for identifying suitable impacts. Therefore, there is no evidence that all significant impacts have been valued. Indeed, this is unlikely, given the arbitrary approach. In contrast, the systematic framework provides a clear means of identifying all possible sources of impact, through the output analysis method and production of the RIO Inventory (Steps 1-3, in Chapter 7).

A further weakness was found to exist with this group of studies. They shared the same aim; to achieve values in monetary terms using neo-classical environmental economics methods, prior to application into regulation using market mechanisms. This differs with the aim of the systematic framework, which is to value environmental impacts in non-monetary, QLOS units, prior to application into the wider regulatory framework. As previously established, there are problems with environmental economics approaches. Hence, in assuming the environmental economics approach, these studies inherited the same problems.

12.3 Life Cycle Analysis

A brief history of the development of Life Cycle Analysis (LCA) is presented in Appendix G. At its heart, LCA consists of tracing energy and material through complex systems. It involves compiling an inventory of life cycle components (the Life Cycle Inventory, LCI) and then assessing the environmental impacts which result. Thus, it has an initial similarity with the first three methods of the systematic framework (though it is notable that the fourth is excluded from consideration within LCA). LCA is generally a more systematic approach to measuring the environment than neo-classical environmental economics, although it has limitations in this regard when compared to the systematic framework.

LCA is the most recently developed tool related to the output analysis method. However, commonly, the aims of LCA studies are different from those of the output analysis method and the wider systematic framework. There are also differences in approach across LCA studies, whereas the detail of the output analysis method provides for uniformity of approach.

Goal definition and scoping are the means by which the aims and objectives of the LCA study are set, along with the types of impacts which are to be considered. This differs from the output analysis method. The goal of the output analysis method is the RIO inventory, and this is pre-defined within the systematic framework, and not as a part of the output analysis method itself. Therefore, due to the nature of the step-wise approach taken in the systematic framework, there is no need for the output analysis method to look beyond the RIO Inventory. Indeed, it is important that it does not. Also, it follows that, unlike the early stages of LCA, the output analysis method is not concerned with scoping, which itself creates problems of subjectivity (see Appendix G).

For similar reasons to those raised against scoping, unlike in LCA, the output analysis method is not concerned with setting out the types of impacts which are to be considered. At this point in the process, all RIOs must be viewed as potentially the source of significant and relevant impacts. The only decision which is made about each RIO is how accurate the measurements (and balance) of primary process inputs and RIOs should be. This is unrelated to the impacts which may be caused, although inevitably, since some materials have higher potential toxicity than others, there will be a need for different RIOs to be measured to different levels of accuracy. For example, only small quantities of radioactive compounds may be used in the production process, in comparison to the quantities of concrete or steel. However, because of the potential for impact, the quantities of the radioactive compounds may need to be determined with greater accuracy. The starting point for deciding to what level of accuracy RIO quantities are required is the current level of data or data measurement capability.

The RIO Inventory is a list drawn at the system boundary, whereas the LCI is a list of pollutants, etc., drawn at the receiving environment, and so apparently incorporates the initial pathway in the environment. It may also include quantities which are not outputs quantities, but environmental burdens. Hence, an important difference is that the LCI may contain pathway and burden information, whereas the RIO Inventory is limited to a list of materials and energy as they cross the boundary out of the activity under consideration. Thus, the RIO Inventory is more precisely determined than the LCI, and this preciseness is important in ensuring objectivity and clarity.

The comparative lack of clarity in the LCA methodology increases as it progresses towards impact assessment. Also, the boundaries which are drawn in the systematic framework between the output analysis method, the pathway analysis method and the valuation method are not mirrored with the same preciseness, rigour and objectivity in LCA. In particular, the equivalent to the three steps of the pathway analysis method are, broadly, classification and part of characterisation (see Appendix G). These involve grouping impacts by type and size of changes. There is a potential overlap within characterisation between production of objective, magnitude-based environmental change information and subjective, significance-based human consequence (impact) assessment. The systematic framework approach avoids this in the interests of locking subjectivity into its rightful place – the valuation of consequences, by separating these parts of the process into two methods (pathway and valuation).

While the systematic "cradle to grave" materials and energy balance-based approaches are common to both approaches, LCA is often simplified prematurely by excluding some parts or components of the system, through scoping or arbitrary selection or rejection of potential impacts. Also, the LCA equivalent to the human consequences data (Step 7 outputs) is a set of impact burdens expressed with regard to various categories, such as social welfare and ecosystem quality (the determinants of which being clearly a matter of subjective debate). In comparison, the pathway analysis method presents greater transparency and clarity in impact assessment by separating it into discrete Steps, each with clear data requirements. The result is that the systematic framework avoids the weaknesses of many LCA studies, which include questions over efficacy and completeness of data and findings and - perhaps the largest single weakness of LCA - the lack of a clear and theoretically sound single unit of measurement for impacts, which causes problems with forming comparisons across studies.

In the absence of common impact units, LCA practitioners generally borrow values from neo-classical environmental economics to derive monetary estimates for LCI quantities, which is generally referred to as life-cycle costing (the general method adopted is to use generic external costs from the literature and apply them to quantities of materials/energy in the specific LCA being undertaken). In so doing, such LCA approaches borrow from, and therefore incorporate the weaknesses of, environmental economics. Clearly, this differs from the systematic framework, where the valuation method incorporates QLOS valuation in single, specified and clear QLOS units, by calculating the product of four items of data for each human consequence, without recourse to pre-existing monetised environmental costs data. The most thorough project on ESI environmental impacts to date, which uses an LCA approach, also seeks to achieve values in single unit terms, albeit on a life cycle costing (monetary) basis. This is the ExternE project, and this warrants specific comparison with the systematic framework.

12.4 The ExternE Project Approach

As reviewed in Section 5.2, the ExternE project first reported in 1995 (ExternE, 1995a, to 1995f) and has subsequently produced updated summary, methodology and results (ExternE, 1998 and 1999). The original aim of this project was to develop an accounting framework and produce damage cost estimates for ESI case study projects. The method uses a fuel cycle approach which is based on LCA principles. The

ExternE project is one of the most detailed and involved exercises which has attempted to apply the LCA methodology in practice. The equivalent stage of the ExternE methodology to the output analysis method is dealt with in the task described as "Defining the Boundaries of the Analysis" (ExternE, 1995b). There is considerable commonality between the approaches here, which is not surprising, given the correspondence between the underlying principles of the two methods. Both emphasise the need for transparency and consistency, and both use a systems approach to fuel cycle description.

Examples of stages considered are laid out in the ExternE methodology (ExternE, 1995b), and these broadly correspond with many of the stages which would be produced in Step 2 of the systematic framework. The ExternE method states that "analysis of each stage is often not necessary in order to meet the objectives of the analysis ... However, the onus is on the analyst to demonstrate that this is the case". The output analysis method precludes such judgements by the "analyst", since this introduces subjectivity and, anyway, the "analyst" should not consider impacts at this point in the systematic framework. The ExternE method allows for order of magnitude calculations at this stage, to demonstrate that, for example, emissions associated with production of steel to make wagons which transport coal to power stations are negligible and can be discounted. The output analysis method holds that this would breach the principles of the systematic framework, and that if "gaps" are left in the RIO Inventory, these must be recorded as inputs "not accounted for".

While the ExternE project provides a major step forward in detailing the fuel cycle aspects of the process being assessed, and in requiring transparency in reporting of this process, the output analysis method of the systematic framework goes further. No major valuation studies have documented how impacts were selected for valuation in detail. Where this is mentioned, the most common method is some form of peer and/or literature review. In either case, it cannot be claimed with any confidence that all potentially significant impacts have been identified for valuation. At the heart of the problem is the implicit assumption that any impact of significance is already being felt by humans and so valuers know about it, and that the time at which it is felt is the time to value it. However, by the time many environmental impacts are noticed, irreparable change can have already occurred, so prevention (which may be better than cure) is no longer a regulatory option. Only by using an approach like the output analysis method can the source of such latent impacts be identified.

Apart from these differences, the initial stage of the ExternE methodology is broadly similar to the output analysis method. Both have affinities with the general LCI method. There is also broad similarity between the pathway analysis method of the systematic framework and the equivalent parts of the ExternE methodology. Both reject top-down or macro analysis using aggregated data, as in former fuel cycle costing studies elsewhere (for example, Hohmeyer et al, 1988, Ottinger et al, 1991). They also reject the control cost method, which adopts the current regulatory framework as a surrogate for what should be done (for example, Tellus, 1991). These similarities are encouraging, since they indicate that the practical-based work of the ExternE project may broadly be expected to meet the theoretical-based requirements of the academic-based approach of the pathway analysis method, and vice versa. Such convergence adds to the validity of both methodologies. It also allows subsequent comparison here to concentrate on differences of detail, rather than broad, common, framework-level principles, such as transparency and the systems approach, which are now well established in preceding Chapters.

Regarding comparisons with the pathway analysis method, ExternE establishes an "Impact Pathway Approach", broken down into seven stages (ExternE, 1995b):

Stage 1. Fuel Cycle Activity (e.g. emissions);

Stage 2. Pollutant transport and chemistry;

Stage 3. Deposition/uptake;

Stage 4. Intermediate processes within the ecosystem (soil acidification, etc.);

Stage 5. Biological/physical/chemical startpoints;

Stage 6. Biological/physical/chemical endpoints;

Stage 7. Valuation.

Stages 2 to 6 can be recognised as the ExternE equivalent to the pathway analysis method. A practical case study approach is adopted, using closely specified technology options, although it is noted that "extrapolation of the results to other options can often be achieved with little additional effort" (ExternE, 1995b). All analysis is conducted on a marginal basis. The stages in the ExternE methodology which are equivalent to the pathway analysis method can be summarised as follows:

- Impact categorisation.

  Impact categories are identified in terms of burdens; factors which are capable of causing an impact of whatever type. ExternE states that the following broad categories of burden have been established (although there is no explanation of how); solid wastes, liquid wastes, gaseous and particulate air pollutants, accidents, occupational exposure to hazardous substances, noise, heat, presence of human activity, others (e.g. exposure to electro-magnetic fields). Each burden is examined and a list of impacts produced (no consistent and inclusive method for doing this is stated).

- Prioritisation of impacts.

  To "be sure that the analysis covers those effects that will provide the greatest externalities. ...Wherever possible, scoping calculations were made to gain some idea of the likely magnitude of impacts, during the identification of the priority impacts" (ExternE, 1995b). General principles adopted were that local impacts tend to be less important than regional/global ones, and that selection of priority impacts is achieved by iterative means and order of magnitude calculations.

- Description of priority impact pathways.
  Where complex burden-monetary cost links are identified, impact pathway listing is used to check the extent to which the impact has been considered. As stated, "much of the analysis presented by the ExternE project is incomplete ... (it) ...can easily be extended once further data become available" (ExternE, 1995b).

- Quantification of burdens.
  This is generally achieved by modelling or otherwise adopting a means of predicting the dispersion of pollutants, etc. It is noted that site specific data are often unobtainable or do not exist.

- Description of the receiving environment.
  This involves scenario definition and description, including meteorological conditions affecting dispersion and deposition, location, age and health of the population, status of ecological resources, and value systems of individuals.

- Quantification of impacts.
  "In some cases externalities can be calculated by multiplying together as few as 3 or 4 parameters. In others, it is necessary to use a series of sophisticated models" (ExternE, 1995b). Again, common approaches involved dispersion modelling and the use of a dose-response function.

As discussed above, the values of the "analyst" can impinge upon the choice of what to measure, since scoping and prioritisation of impacts is inherent in the ExternE methodology. However, it is also stressed that no impact should be ignored for convenience, and that emphasis is placed on establishing gaps, where potential impacts may occur but are currently unknown. A further element is the description of uncertainties, which are listed as variability in data, extrapolation from laboratory to field, extrapolation of data from one location/situation to another, assumptions regarding thresholds, lack of information on human behaviour and taste, political and ethical issues, such as choice of discount rate, assumptions about future scenarios, and "the fact that some types of damage cannot be quantified at all" (ExternE, 1995b).

Scoping occurs too early in the ExternE methodology, and the definition of significance and the grounds for rejecting further study are not sufficiently defined. As established in Section 12.3, scoping of impacts, in the sense of rejecting potential impacts from further study, should only take place when human consequence data are established. The pathway analysis method requires as a prerequisite the identification and listing of all life cycle outputs and the Pathway Inventory contains all possible quantities, with gaps identified where there are unknowns, or where quantities have not been calculated. The logic is that these stages of the analysis are objective, and that until subjective valuation is introduced, no quantities can be left out, since only magnitude is being measured at these stages and this is not the only determinant of impact significance. Only quantities in similar units can be compared. Therefore, to leave out impacts at this stage is to introduce potential error, the significance of which cannot be known until valuation. Scoping of magnitudes, in the sense of rejecting some elements of an impact or its precursor because it is negligible in quantity, is less likely to cause major inaccuracies or uncertainties in results of subsequent valuation. This is because the impact is to be valued, but an element, defined as negligible, is missed out. Here, the comparison is between quantities of the same units. However, the problem remains that, without tracing a RIO through all its pathways, there is no guarantee that other impacts it may cause may not also have been discounted from further consideration, as well as the negligible, comparable one. These others may be less negligible.

Regarding transparency, there is some evidence for a difference between principle and practice. Establishing the difference between small quantities and unknowns for all RIOs is essential. It is understandable that the ExternE project, with the inevitable constraints of resources, time, and the need to reach contract goals, found it necessary to scope out some RIOs and some pathways from consideration. Often, but not always, explanations are given. However, the only way to demonstrably include all RIOs in pathway analysis is to start from a complete RIO Inventory and end with a complete Pathway Inventory (complete with gaps).

Critically, the pathway analysis method provides a means to maximise objectivity. The tracing of pathways using pathway identification, and production of the Pathway Inventory, are inherently objective processes, and subjectivity is eliminated, provided there is compliance with the rules laid out in Sections 8.3 and 8.4. Without the rigorous approach of the pathway analysis method, impacts which are now considered paramount, like carbon dioxide-induced global climate change, would not necessarily be detected if a less rigorous approach had been adopted previously. If tomorrow's environmental impacts are to be detected at the earliest opportunity, a method with the rigour of the pathway analysis method is required. Practicality must be judged within the bounds of the other criteria. Without satisfying these, the exercise cannot be completely or satisfactorily achieved.

While the LCA methodology is well-established in its initial stages, it becomes increasingly contentious in impact assessment, the culmination of which is in valuation.

Therefore, it is perhaps unsurprising that the greatest differences between the ExternE approach and comparable Steps of the systematic framework occur within valuation. ExternE typically attempts valuation by drawing generic values from the neo-classical environmental economics literature and applying them to the quantified impacts. As already established, such values are both problematic and insufficient to reflect non-economic elements of value. The alternative is presented in detail in Chapter 10, where the Environmental Change Inventory produced in Step 6 is used to derive the data on human consequences in Step 7. Valuation, in Step 8, also requires data on the quality of life outcome state associated with each human consequence, using the QLOS Index, which contains all outcome stages on a single, ratio scale. This is a complete departure from the ExternE life cycle costing approach and, incidentally, is in accordance with the original intention of the LCA methodology.

In summary, the ExternE project provides the best example of a practical attempt to identify and quantify pathways for impact assessment in the ESI to date, and it provides a major advance in methodology and approach. However, the pathway analysis method of the systematic framework is a further development. The pathway analysis method has broad similarities with LCA and the ExternE methodology, but provides improvements in a number of key areas, including consistency, transparency, completeness and the potential for feeding data directly into a valuation method and regulation tool, both of which are integral to the complete systematic framework. Significantly, valuation involves use of a single scale designed to capture non-economic elements of value. There are also differences in the application of values in regulation as will be seen below.

Most impact valuation studies do not consider the incorporation of values into the market nor the regulatory implications of values in any detail. Typically, for studies where valuation in monetary terms is the goal, it is assumed that one or other market mechanism(s) should be used to achieve this. However, since the systematic framework involves the production of values in non-monetary terms, a simple market mechanism approach is invalid. Thus, for completeness of the systematic framework, it incorporates the application method (Step 9), and this is applied to produce the QLOS Tranche System for Total QLOS values, in Chapter 11.

As stated in Chapter 3, the principle reason for the current unacceptable situation where impacts are not sufficiently regulated, is that regulators have insufficient information about impacts. There has been a lack of confidence in valuations by those intending to use impact values in the regulatory framework. Therefore, it follows that many studies have shown insufficient regard as to the needs of regulators and decision makers. There is some indication that this may change in the case of the ExternE project. Due to the level of detail and transparency about uncertainty, decision makers may take the values obtained more seriously than earlier studies. Indeed, as noted in Chapter 5, even with shortcomings and problems in valuation, ExternE results provide important provisional implications for energy policy and regulation (Eyre, 1997).

Such developments in confidence are critical to the successful incorporation of values into the regulatory framework, and this is badly needed. Otherwise, wrong decisions about energy technology investments are likely to continue, given the recent finding that "cleaner technologies, such as renewables, gas or nuclear, or pollution abatement technologies, are always profitable from the social point of view, even though not all their environmental benefits have been assessed yet" (ExternE, 1999). In the same study, it is stated that the damage costs from the UK power sector are around 2% of Gross Domestic Product, and that these are dominated by the effects of air pollution on health, and of greenhouse gas emissions from the use of fossil fuels on global climate change. Nevertheless, many problems remain with application by simple market mechanisms, and current evidence of recent regulation suggests that the much-desired perfect market, where optimal levels of environmental damage are achieved, is far from becoming reality, and closer to being confirmed as a myth.

There are ongoing problems with monetisation, particularly where impacts are not generally viewed in monetary terms, and regulators appear cautious about incorporating the results of monetisation studies into market mechanisms. Furthermore, there are theoretical and practical reasons why this approach may not provide the optimal regulatory solution anyway. Hence, the case for producing impact values in monetary terms is certainly questionable. Indeed, it is weak when compared to the approach advocated in the systematic framework, where the unit of impact value has been developed through theoretical application. Measurement in QLOS units is preferable for impacts which are more appropriately considered in terms of quality of life than in money. Moreover, this approach is practical, in that mechanisms can be designed which will allow the application of the values produced in decision making, as demonstrated in Chapter 11. Indeed, it is also timely since, coincidentally, recent government policy and guidance acknowledges the quality of life concept as a viable basis for indicating progress towards sustainability (for example, DETR, 1999, 2000).

In summary, the systematic framework is set apart from all other impact assessment methods to date in its approach to regulation. While other studies either ignore or assume that regulation will take place, the systematic framework specifically incorporates an application method for this purpose. This approach, combined with the rigour of the earlier methods and consequent credibility of their results, provides regulators and decision makers with a package designed to overcome their reluctance to act on the results of earlier studies. The systematic framework provides convincing results and a specific means of applying them. The challenge is clear. The problems with impact assessment can now be solved, and the best means of dealing with them identified. It is now down to the political will of the regulators and decision makers to apply. In so doing, they can be in no doubt that the result will be a more efficient and environmentally sound ESI, where the winners will be both the environment and those currently receiving the impacts which are currently adversely affecting their quality of life.

13. CONCLUSIONS AND RECOMMENDATIONS

13.1 Summary of Conclusions

Despite government targets on renewable energy utilisation and policies to achieve sustainability, major environmental issues associated with the electricity supply industry (ESI) still exist. Although initiatives have been taken which may be expected to have environmental benefits, the most recent being the Climate Change Levy, these are not linked directly to the environmental impacts concerned. While "green" energy is currently being promoted, there remains the problem of establishing how "green" this is and how objective the means of measuring environmental impacts are. Meanwhile, the problem of optimising energy efficiency in an industry driven to maximise consumption remains. In short, there is a pressing need for a means to measure and regulate for environmental impacts.

The ESI is a critical industry with a varied history. It also encompasses a wide range of technologies and these factors, plus its sheer size and importance, mean that two complications are faced in attempting to develop the environmental regulatory framework. Firstly, technological and process variations mean that any approach to measuring and regulating for environmental impacts must be versatile, if it is to be applied across the ESI. Secondly, there are numerous historical and political reasons why the industry has hitherto failed to achieve optimal economic, social and environmental efficiency and if any proposed remedy is to be practical, it must take account of these.

Better regulation starts with better understanding of the issue(s) to be regulated. In this case, it requires appropriate data about values of environmental impacts. The main technique which has been promoted to provide this over the last decade has been the valuation of environmental impacts in monetary terms, using neo-classical environmental economics methods. The monetary values produced can then, in theory, be used to internalise environmental externalities, by applying market mechanisms to correct for the market inefficiency. In the early 1980s, regulators in the western world initially embraced environmental economics as a potential solution to the problem of environmental regulation. However, subsequently, numerous objections have been raised and weaknesses identified, including the lack of a systematic approach. It is concluded that some of these weaknesses could be addressed by adoption of Steps 1 to 6 of the proposed systematic framework. However, other weaknesses cannot be addressed using environmental economics, since they point to the existence of non-economic environmental impacts - those elements of impact which cannot be sufficiently captured using monetisation approaches.

While environmental economics is not rejected outright, further improvements are required and, in any event, it must be supplemented by the systematic framework approach, which encompasses a means of valuing non-economic elements of value. Furthermore, any and all methods must conform to the requirements for objectivity, transparency, versatility, and practicality, all of which are met by the systematic framework. Also, given the problems arising from the piecemeal approaches to date, it must be systematic, rigorous, clear and complete, in the sense that all impacts must be verifiably captured in valuation. These requirements are also met by the proposed systematic framework, which fulfils two basic functions. Firstly, it breaks this process into its component parts. Secondly, it provides a single integrated process for regulating environmental impacts, from the point of origin, to the point of applying regulation.

One fundamental difference between the systematic framework and most neo-classical economics-based valuation studies to date is that the former follows the impact valuation and regulation process in the same direction as the impacts originate. Thus, it starts with human activity to exploit natural resources, progresses by tracing materials and energy through the production process and out into the environment, and ends with quantified impacts stated in comparable terms, with an appropriate means of reflecting values in regulations and decision-making. By starting with the life cycle, examining the entire life cycle-pathway-environmental impact relationship and breaking it down into its constituent parts, understanding of the process and information about which parts of it are poorly understood can be improved. Constituent parts can be examined and further broken down to reduce confusion and uncertainty, highlight where and which data need improving, and where gaps in knowledge exist which require filling.

This approach has its closest recognised method in Life Cycle Analysis (LCA). However, LCA has been fraught with difficulties, not least over data standardisation, and availability and comparability of impact assessment practices and results. The systematic framework solves such problems by prescribing transparency, rigour and clarity in a detailed stepwise procedure. In particular, it identifies two critical points in the impact assessment process, where the transfer of data must be precisely defined to avoid subjectivity and inaccuracy. This is achieved by splitting the systematic framework at these points, resulting in separate methods being defined. LCA currently fails to bring such clarity to these critical points in the process.

The systematic framework has been developed with non-economic impacts in mind, and with the ESI as the industry within which improved environmental regulation is sought. However, given the findings established in the course of this thesis, it is now possible to widen the systematic framework to incorporate environmental economics and regulatory mechanisms based on internalising externalities. Moreover, although it has been developed for application specifically to the ESI, it could also be applied to other industry sectors. It is proposed that the resulting wider, generic framework is called the Environmental Analysis, Valuation and Application (EAVA) Framework.

13.2 The EAVA Framework

The EAVA Framework is illustrated in Figure 13.1. It consists of four methods, each with clearly defined input and output requirements. The first method, the output analysis method, comprises Steps 1 to 3, and is entirely concerned with the production system life cycle. Specifically, it involves defining all quantities of matter and energy, leaving the system involved with provision of the product or service in question. Environmental impacts arise exclusively as these unintended or incidental outputs of the life cycle, referred to as Released Incidental Outputs (RIOs). A potential problem with gathering RIO data is commercial confidentiality, and it is concluded that RIO data should not be regarded as confidential (commercial) information and this should be enforced by regulation. The output analysis method culminates in Step 3; the RIO Inventory, which is a quantified list of all RIOs, checked for completeness.

Steps 4 to 6 comprise the second method, the pathway analysis method, which involves tracing quantities of each RIO through the environment (including humans), and recording all environmental changes which result. By measuring quantities of each RIO at each point on each pathway, gaps where quantities of RIO are unaccounted for can be identified. This is not attempted by any existing methods and it is critical in providing the means to identify potential future impacts. It is generally accepted that global climate change has been identified as a problem too late. With application of the EAVA Framework, it would have been identified earlier.

ANALYSIS

OUTPUT ANALYSIS METHOD

Step 1. Life cycle definition

Step 2. Stage definitions

Step 3. RIO Inventory

↓

Outcome: A list of all material and energy which are released incidentally from the process being analysed

↓

PATHWAY ANALYSIS METHOD

Step 4. Pathway identification

Step 5. Pathway Inventory

Step 6. Environmental Change Inventory

↓

Outcome: A list of all changes in the environment caused by the RIOs listed in Step 3. Also a list of 'gaps' - quantities of each RIO not traced to sinks and point at which tracing stops

↓

VALUATION

For impacts which lend themselves to valuation in money terms

↓

Value and monetise using proprietary neo-classical environmental economics valuation method:
e.g. Contingent Valuation Method

For impacts which do not lend themselves to valuation in money terms

↓

QLOS VALUATION METHOD

Step 7. Human Consequence Inventory

Step 8. QLOS valuation

↓

Outcome: Impact values in QLOS units

↓ ↓

APPLICATION

Market Mechanisms for Monetised Impacts:
e.g. Cost-related Pigouvian Taxes

↓

QLOS APPLICATION METHOD

Step 9. Regulatory measures for unquantifiable, unknown and QLOS-valued Impacts

↓

Outcome: Optimised production system by application to the regulatory system to ensure highest net quality of life outcomes

Figure 13.1 Environmental Analysis, Valuation and Application (EAVA) Framework

While it is recognised that objectivity cannot ever be entirely eliminated, this does not preclude the aim of minimising it where it is unwarranted. The basic structure of the pathway analysis method acts against subjectivity creeping into the process. Therefore, it overcomes a major problem which affects all other methods used at present (see Chapter 8). The outcome of the pathway analysis method is a complete list of objectively quantified pathways and environmental changes (including quantified gaps), attributed individually to source RIOs. The result is the Environmental Change Inventory. Systematic comparison of items on this Inventory is also incorporated, to detect potential system-level, synergistic, cumulative, or threshold effects.

Whether impacts are to be valued in QLOS or money, they should first be identified, quantified and analysed using the output analysis and pathway analysis methods. Only then can one of the two possible subsequent routes be identified for a given impact; monetising and internalisation through market mechanisms, or QLOS valuation and application. Whichever of the two is chosen, the third method in the EAVA Framework involves valuation.

In the case of impacts which can be valued successfully in money terms, the Environmental Change Inventory is applied to generate impact data for valuation using appropriate methods from environmental economics. For non-economic or less well-known impacts, a two-Step approach is needed. Step 7 involves the production of four parameters for each consequence; the risk of the impact occurring, its duration, the number of people potentially affected and the quality of life outcome state (QLOS) resulting from it. The latter requires use of a single scale of quality of life outcomes adapted from the well-established health-related quality of life literature, as demonstrated in Chapter 9. The other three parameters can be provided in a more straightforward way, often using existing data, as discussed in Chapters 8 and 10. Once these parameters are provided, Step 8, consisting of valuation itself, is primarily a simple mathematical exercise, resulting in a quantified value for each environmental impact, expressed in standard, comparable QLOS units.

The fourth method and final Step in the EAVA Framework is Step 9, where the valuation results are applied to appropriate regulation. The application method involves identifying criteria to be met by regulation, considering the options available, and selecting from these the most appropriate regulation (or combination thereof), based on fitness with the criteria. Through applying this method, it was concluded in Chapter 11 that, for Total QLOS values for proposed ESI projects, an appropriate regulatory mechanism is a proposed QLOS Tranche System, where the valued impacts are used to compare project proposals in a strengthened project planning system. All other decision-making criteria being equal, the projects with the lowest QLOS should be given permission over others. Thus, a competitive element is incorporated into the QLOS Tranche System, in order to incorporate the benefits of the market based approach, without the need for monetisation. The structure of the proposed mechanism is simple, in terms of ease of introduction and enforcement. It is also reliable, in terms of likelihood of compliance and, therefore, of success. Furthermore, the project environmental appraisal elements of the planning system are currently lacking in objective data with which to make decisions and so the QLOS Tranche System will fulfil a wider existing regulatory need. Although the QLOS Tranche System has been specifically designed to apply Total QLOS values for proposed ESI projects, similar mechanisms may be envisaged for other industry sectors as the EAVA framework is applied to them.

Within existing studies, the ExternE project provides the work which is closest in approach, topic area and method to the EAVA Framework. Although the EAVA Framework has not been developed out of the ExternE work, it is timely, since there are broad similarities between the studies. ExternE takes a practical approach to deriving results, while the EAVA Framework takes a more theoretical approach to identifying what and how the task of valuation and regulation should be undertaken. Thus, the latter is freed from the constraints of having to produce results within tight deadlines, as required in commercial research. Given the differences of starting points, it is notable that there are some common basic elements. However, there are also clear differences of substance, as discussed in Chapter 12. In essence, while both share a common methodological heritage in energy and mass-balance analysis and LCA, the EAVA Framework provides a further methodological development beyond that practised within the ExternE project.

In summary, the EAVA Framework provides an explicit set of rules based on a clear set of principles, each element of which contributes to a rigorous means of establishing relationships between and within human and natural systems. The sequential approach provides precisely isolated and manageable tasks with specified inputs and outputs for each Step, allowing a large research effort involving disparate disciplines and working in different places and times to produce data which are useful, recognisable and appropriate within the wider purpose of the framework. Most importantly, it provides a means by which the issue of outstanding unknown, unquantified and unaccounted for impacts can be addressed for the first time in practice, ensuring that the most environmentally acceptable schemes are given permission for development.

13.3 Recommendations for Further Work.

The EAVA Framework is complete, and it has been demonstrated that each Step within it is both rigorously defined and practical at demonstration level. The next stage is full application. This process can be broken down into phases, giving rise to three specific recommendations. Firstly, a complete exercise must be carried out for a single, real project. This will serve the purpose of providing implementation experience along with real results for application to regulation. It will also provide an opportunity to develop and apply GIS-based QLOS valuation software. Secondly, following the detailed application, it will be necessary to extend this into a full pilot study, involving implementation for a further range of different projects, with varied technologies, scales and locations. A similar exercise was undertaken as part of the ExternE project, and it may be possible to use some of the databases developed for this project in the pilot application to QLOS assessment. Thirdly and finally, a detailed form of regulation needs to be developed following the basic method and outline proposed in Chapter 12, with subsequent trial and application within the project planning system in England and Wales.

Part of the process of EAVA Framework implementation will involve clarification of the role of environmental economics within it. Two areas of further work will be needed here. First, it has been demonstrated that valuing impacts using neo-classical environmental economics methods involves numerous problems. While these problems do not preclude further use of this technique, improvements are required. There is potential for the EAVA Framework to assist in this regard. In Steps 1 to 6, a means exists for identifying life cycle outputs, RIOs, pathways and environmental changes in an objective, transparent and rigorous manner, thus addressing some of the problems. A study is therefore required to establish how the EAVA Framework can assist in improving the efficacy of environmental economics methods in practice, and how any remaining problems can be overcome. Again, detailed reference could be made to the ExternE project in this study, since this attempts to achieve monetisation-based outcomes. Secondly, a study should examine the inter-relationship between non-economic and economic impacts, including the establishment of theoretical and practical boundaries. In particular, this should focus on developing a test to determine whether money or QLOS is the appropriate currency for any given impact, and on a means of demonstrably avoiding double-counting and/or underlap, wherever the two currencies are used to measure different aspects of related impacts.

Two other areas of further work relate to the application of the EAVA Framework outside the ESI and outside England and Wales, respectively. The results obtained in the demonstration work suggest that it could be applied to other sectors. However, there will be differences which will affect the emphasis and outcome of different parts of the exercise. For example, there is only considered to be one end-product in the ESI process and this product, being electricity, has no material to be disposed of after use. Also, the project planning system as it relates to the ESI is unique, and so the application method may be expected to produce different regulation proposals in other sectors. The logical means of extending the EAVA Framework will be to select another sector and apply it at pilot level, prior to further extensions into other sectors. An appropriate initial sector will be one which produces material products, since this will require RIOs arising from product use and disposal to be identified, as well as production. Regarding the implementation of the EAVA Framework outside England and Wales, it is expected that this will be readily achievable as a research study, within other countries in western Europe. However, there may be more potential for adjustments to be required when implementation is attempted in the less developed world. This warrants specific study.

In applying the EAVA Framework, as indeed, in any method involving measuring environmental impacts, the priority will always be more knowledge in improving understanding of environmental impacts and what causes them. It is implicit that further knowledge will always be required in order to improve data. Application of the EAVA Framework will itself result in the identification of such areas. Since it specifically identifies gaps in current knowledge, it can be used as a tool for identifying further work requirements. Therefore, the final recommendation is that these gaps in data from any and all applications of the EAVA Framework should be used in setting the agenda for further research in the subject areas from which each gap originates. In this way, gaps in current knowledge will be both identified and acted upon, thus reducing the gaps and improving the completeness of human knowledge about the environmental impacts which are being created.

APPENDIX A. REVIEW OF NEO-CLASSICAL ENVIRONMENTAL ECONOMICS THEORY AND METHODS

Neo-classical environmental economics methods can be split into two types; direct and indirect. Six methods are discussed in Section A1 below, three falling into each type. Five potential applications of these methods are discussed in Section A2. The related concepts of value and risk are then discussed in Section A3, before summary and conclusions of the review are drawn in Section A4.

A1. Direct and Indirect Valuation Methods

Although there is contradictory usage of the terms direct and indirect in relation to environmental valuation in the literature, they can be used to describe two general sets of approaches. Direct valuation techniques aim to establish revealed preferences either through eliciting responses from individuals (thereby creating imaginary markets) or by studying suitable surrogate markets. However, indirect costs are obtained by summing individual effects, sometimes expressed in terms of replacement values. At least one author (Smith et al, 1986) uses the terms differently and differentiates between indirect methods which rely on household behaviour to reveal valuations of non-market goods, and direct methods where individuals' valuations are elicited directly through surveys.

A1.1 Direct Valuation

There are three direct valuation approaches included in this review; Contingent Valuation Method, Household Production Function and the Hedonic Method.

A1.1.1 Contingent Valuation Method

First used in the 1960s (Davis, 1963) for valuing environmental resources, the use of the Contingent Valuation Method (CVM) since has been almost exclusively for environmental valuation. Its wide applicability combined with contentious elements in methodology have ensured the growth of a broad and extensive literature, although since it is still a relatively new technique in Europe, most of this reflects experience in the USA. Important reviews include; Cummings et al, 1986, and several publications by Pearce with fellow researchers.

Despite the range of applicability, the approach is standard. The application of the method is discussed in Hanley (1990). In short, people are asked for their Willingness To Pay (WTP), and/or Willingness To Accept (WTA), a particular environmental change or range of changes through a questionnaire. A bidding system is used to elicit answers from respondents by offering various options and asking whether bids should be raised or lowered until a satisfactory monetary value is reached. Analysis of the questionnaire results then leads to establishment of a mean value.

An important distinction is drawn between the concepts of Compensating Variation (CV) and Equivalent Variation (EV). In the latter, it is assumed that the project is carried out, and so WTP for avoiding and/or WTA deterioration in the environment is elicited. In the former, the project is at the planning stage (that is, in the present rather than the future), so WTP for environmental improvements and/or WTA for environmental losses under the project are sought. Only in an infinitely large market with zero transaction costs and perfectly divisible goods would the results be the same (Brookshire et al, 1980). One author states: "For a decrease in welfare, WTP is equal to EV and Willingness To Sell (equal to WTA) to CV. For an increase in welfare, the situation is reversed" (Hanley, 1988). He develops his ideas further through a broad discussion of the problems and a review of CVM studies (Hanley, 1989).

Two of the most significant advantages of the CVM over other valuation methods are that it has a very wide range of applicability since it can be used to measure almost any aspect of the environment, and that it is the only method which can comprehensively measure non-use values. The latter point is particularly important since it has been shown that such values can form a significant proportion of TEV (Madariaga and McConnell, 1987). Indeed, in at least one study, non-use values of freshwater fish in Norway were found to be an order of magnitude higher than use values, posing the question that if such an asset (with high use value) reveals these results, environmental assets with low use values, for example, Sites of Special Scientific Interest (SSSIs), may reveal proportionately much higher non-use values (Navrud, 1989).

Various shortcomings of the technique have been identified and discussed extensively in the literature (for example, Pearce et al, 1992). Potential errors have been summarised elsewhere (Turner and Bateman, 1990). Although they are clearly potentially significant, biases can often be mitigated by careful questionnaire wording. Strategic bias occurs when respondents state untrue values, often if they feel they are in a “free rider” situation, where they can gain the benefits without paying the costs. Indeed Green et al (1990) have noted that Samuelson's (1954) free rider hypothesis that people would systematically lie in response to a CVM survey led to the rejection of CVM as a technique by some economists (Feenberg and Mills, 1980), although subsequent work has shown that it is not a significant problem (Marwell and Ames, 1981). Hypothetical bias is more structural in nature, arising because the transactions taking place in the questionnaire are not real; Design Bias includes the layout/type of information or type of bidding offered (bid vehicle) and Starting Point Bias occurs where the starting bid offered to respondents affects their final bid due to their impatience or the suggestion of an appropriate bid size. A detailed discussion of biases is presented in Schulze et al, 1981. CV has evolved several variants in efforts to overcome the various biases and resultant improvements in survey design and sampling have increased the reliability of CVM estimates. In general, the quality of results is heavily dependent upon how rigorously the method is applied. This is particularly the case with CVM as careful survey design, implementation and interpretation of results is clearly critical in avoiding bias or other inconsistencies.

A1.1.2 Household Production Function

This method is most often used where the output of the environmental good is marketable, so the environment is part of the production function. The basic approach here is to value the costs which households are prepared to undergo to avert/substitute for environmental damage or to experience an environmental benefit. The former results in the calculation of Avertive Expenditures and the latter results in calculation by the Travel Cost Method.

Calculation by the Avertive Expenditures method has not been used extensively and thus the accuracy is not fully established. It is clearly only relevant to cases where households spend money on measures to offset environmental impacts, such as noise insulation. Non-user values (values assigned by those who are not directly consumers) are unlikely to be covered due to the lack of market, although it could be argued that payments to conservation groups are Avertive Expenditures to protect the natural environment.

The Travel Cost Method (TCM) is based on the premise that the value people place on the environment is inferred from the time and cost they incur in travelling to it. Usually, the valuation is conducted at the destination environment under consideration, so the site is chosen and people visiting it questioned, rather than enquiring about destinations chosen by individuals. While it has been used in several major studies, the TCM does have clear limitations in application. It can only be used to estimate environmental goods which involve travel costs, and is usually applied to sites with minimal or no entry fees, the travel cost being analogous to the entry fee. The main application in the UK has been for valuing recreational sites (Willis, 1990). Non-use values are clearly excluded from TEV calculations since only users travel to the site.

Various technical problems are associated with TCM (Kealy and Bishop, 1986). At least one study has shown that care must be taken in choosing appropriate applications for the method, and, furthermore, that "the method should not be used unless there is evidence for the site in question that the key relationship (enjoyment increases with distance travelled) is approximately correct" (Green et al, 1990). The main shortcomings in practice are that visitors (travellers) are assumed to not enjoy travelling aspects of the trip (see also Winpenny, 1991), some trips are multi-purpose (the traveller may visit several sites in one trip or be on holiday and so have completed part of the trip already), and there may or may not be other similar sites nearby, thereby affecting trip length. A further difficulty is in choosing an appropriate rate for an individual's travelling time; if work time is given up to travel then the work rate of pay is appropriate, but more often it is leisure time and this is more problematic. The Department of Transport have produced their own figures for working and non-working time, based on attempts to establish shadow prices (see Hanley, 1990, who also discusses statistical problems associated with the technique). A further point is that the travel cost must be regarded as the minimum cost a traveller is prepared to pay; he or she may be prepared to pay much more and, thus, the site value may be higher.

A1.1.3 Hedonic Method

Like the TCM, the hedonic method is a revealed preference approach, where valuations are obtained through the study of surrogate markets, although here it is Hicksian CV or EV that is being measured. Based on pioneering work in the 1960s (Lancaster, 1966), in hedonic methods costs are implied through the analysis of demand in existing markets where environmental commodities are traded. It therefore assumes that people choose the amount of an extra-market good they use by altering their consumption of a marketed good. The most common use is in house price methods (sometimes called the Property Value (PV) approach) where environmental aspects of location such as noise or pollution are valued by using complex analytical techniques to remove all other influences on property prices to leave a value for the aspect in question. Since the method can only address those involved in determining market prices it cannot measure non-use values. However, in terms of values measured, a hedonic valuation would be expected to exceed a CVM WTP valuation since the former reflects the rent difference as demonstrated by the *most sensitive* individual, in terms of the extra-market (environmental) good being valued, whereas the latter represents *average* preference for the study population (Kneese, 1984). Also, it is only relevant where there is awareness on the part of property owners of the environmental variable(s) being measured. The comparison of data derived using this method with more comprehensive TEV calculations derived from CVM has shown broadly similar results indicating the relevance of this technique.

However, there are limitations peculiar to the PV approach, including that the assumption of a well-functioning property market is often not valid since most housing markets are segmented, rent controls affect prices, there is not usually a continuous and smooth supply of houses of all desired characteristics, transaction costs may be high enough to increase inertia to movement, and property values reflect expectations of future environmental quality as well as present (whereas only the present is required). Further difficulties stem from the fact that different social groups may have different housing preferences, calling into question the homogeneity of the sample. Apart from these shortcomings, the method is limited in its range of application.

A further hedonic method is the Wage Risk (or Wage Differential) Method which measures the employees' willingness to accept a risk, such as dangerous, unhealthy or disagreeable working conditions (see Section A1.6.1). This involves analysis of wages in different occupations and elicits values for risks of morbidity and mortality through comparison and calculation of the "wage premium" associated with risk. The relevance to the electricity supply industry (ESI) is apparent although it should be noted that since these wage elements are reflected in the market and paid by employers, such occupational risk does not constitute an external cost.

A1.2 Indirect Valuation

Indirect techniques are appropriate for cases where individuals' knowledge of environmental impacts and their causal linkages is limited. Thus, revealed preference cannot be used to elicit values from individuals.

A1.2.1 Conventional Market Approaches

There are various methods for establishing market prices for environmental effects, or, if market prices are inappropriate, shadow (substitute) prices may be used. Although other approaches use market values, the main methods discussed here are Dose-response and Alternative or Replacement Cost. A clear limitation of all conventional market approaches is that they cannot be used to estimate non-use values. However, the calculation of approximate replacement, compensation and alternative costs can provide useful approximations for valuation purposes in the absence of other more detailed surveys.

With the Dose-response technique, a linkage is established between the source and magnitude of the causal human activity and the resultant environmental impact. This impact is then measured and valued at market or shadow prices. The approach is, therefore, only relevant to situations where the causal relationship is well known and both the pollution (dose) and its impact (response) can be measured satisfactorily. Thus, most of the effort involved is usually concentrated on the establishment of the dose-response relationships; the functional form of the relationship often being difficult to identify. In particular, the dose-response relationship is often much more complex than a straight line positive correlation, for example, with thresholds at which major damage occurs. Consequently, a typical approach to establishing a Dose-response relationship might involve field research, controlled experiments to mimic the effect and to allow observation of receptors in isolation, and the use of statistical regression techniques to allow the separation of one cause-effect relationship from others, which is a particular problem for most pollutants. The problem of relating, for example, visible leaf injury to reductions in growth, is worsened where no visible effects occur, and compounded when provisional estimates of future damage are to be drawn from such data. Thus, data requirements are always a primary concern and, in most cases, assumptions have to be made, while often data is used from relationships observed elsewhere. Once a physical damage function has been established, the second stage of the technique is to apply an appropriate monetary damage function to multiply the physical damage by a unit price. Typical applications of the Dose-response approach include the assessment of the effect of pollution on health and depreciation of material assets such as buildings, aquatic ecosystems, or vegetation (Pearce and Markyanda, 1989).

Alternative Cost is a related method which investigates defensive expenditure necessary to remove the environmental damage, such as costs of double glazing to reduce noise in buildings (Pearce and Turner, 1992). This is also known as the Replacement Costs technique; the cost of replacement or repair of the environment is calculated. As with Dose-response, non-user values are not reflected. The problem of restoring the environment to its original state, however, creates the most significant problem, not just physically but also in defining the original state, unless there are agreed standards as to the level of restoration required.

A1.2.2 Delphi

The Delphi technique is not a market economic approach. It involves eliciting views from a panel of experts as to the value of environmental benefits and costs. As such, while it may have some merit as a method of producing quick, cheap, initial rough estimates, it clearly cannot reflect preferences or market forces.

A1.2.3 Human Capital

This approach relies upon the quantification of effects on human health of impacts on the environment. People are treated as units of economic capital, and effects on health are quantified through loss of earnings and resource costs of health care. However, this approach is only likely to be reliable under specific conditions, where a direct cause-effect relationship is established, the illness is not life-threatening or long term, the economic value of lost production is calculable, and the costs of health care are known (Winpenny, 1991).

A2. Applications of Valuations

In order for monetary valuations to be useful they need to be applied in the decision making process. This invariably means incorporating them into existing frameworks for economic analysis, the purpose being to assess the economic rationality of investing in an environmental improvement. Types of framework widely advocated have been reviewed in the literature (Pearce and Markyanda, 1989). A summary of the main types is given here.

A2.1 Cost Benefit Analysis

Cost Benefit Analysis (CBA) is concerned with balancing costs and benefits expressed in monetary terms. A cost is a foregone benefit and a benefit is a foregone cost. Both costs and benefits are discounted at an appropriate rate and the result is usually expressed as a net present value number, or a discounted benefit-cost ratio (Pearce and Turner, 1992). Where the former is positive or the latter is greater than unity, the project, policy or programme is economically efficient. To the extent that benefits must outweigh costs before a decision to develop is taken, CBA is a rational decision-making aid which avoids judgmental assessments associated with some other decision-making aids. It was originally developed in the US (where it is normally referred to as benefit-cost analysis) in response to the Flood Control Act 1936. Subsequently, major advances were made by researchers at Resources for the Future in the 1960s.

For CBA to include environmental considerations, it can be extended as follows;

$$\text{option value} + \text{existence value} + \text{use value} + \text{pollution damage cost} + \text{project capital cost} \Leftrightarrow \text{project benefits;}$$

where the first four elements constitute the extension of conventional CBA, to ECBA, Environmental Cost Benefit Analysis.

CBA is rigorously objective in its theoretical approach. However, in common with other decision making techniques, CBA is not always applied objectively by any means. For example, it may be misused to demonstrate that a pre-conceived outcome is preferable to the actual outcome had the technique been applied objectively. Even where deliberate or subconscious bias is not applied, CBA must fully incorporate disparate and difficult questions of sustainability, environmental risk and uncertainty if it is to be valid. Although this is not a direct potential failure of valuation methods, it is a further shortcoming of the whole approach of monetisation as a sole decision making tool. A further problem with CBA and all techniques involving valuations is that they clearly inherit all the shortcomings of valuation. Various approaches have been developed to overcome objections to CBA. For example, the Krutilla-Fisher approach (Krutilla and Fisher, 1985) aims to overcome the problem of irreversibility of environmental damage.

A2.2 Cost Effectiveness Analysis

Cost Effectiveness Analysis (CEA) aims to select the most cost-efficient option, and is therefore most applicable to selecting the best use of a fixed budget, for example, in maximising environmental quality. It can also be used to compare options in determining best/most efficient choice. Each project is compared against a criterion or number of criteria, the most suitable being that which satisfies these at least cost. The criteria (for example, a set of environmental standards) constitutes an extension of conventional CBA to CEA, where benefits are not measured in money term but costs are.

A2.3 Multi-Criteria Analysis

Multi-Criteria Analysis (MCA), sometimes known as Multi-Criteria Decision Analysis (MCDA), encompasses several dozen evaluation techniques, all aimed at overcoming the inherent conflicts in environmentally sustainable economic development by including a number of choices for a number of criteria with conflicting aims and priorities. The main techniques involved are discussed elsewhere (Pearce and Turner, 1992).

A2.4 Risk-Benefit Analysis

Risk-Benefit Analysis (RBA) fits closely into the CBA framework. It involves the same approach, but takes into account risk events. Thus, some of the costs are expressed in probabilistic form.

A2.5 Environmental Impact Assessment

Environmental Impact Assessment (EIA) is a process leading to an evaluation of the environmental effects of a planned project. The methodologies used for identifying effects and assessing their significance are many and varied. Early EIA practice in Canada and the US (for example, Leopold et al, 1971) established non-quantitative or semi-quantitative methods for evaluation and such methods have been used to the present. However there is potential for the increased use of quantification. A further area of development for EIA is to Strategic Environmental Assessment (SEA), where the assessment is undertaken at policy/programme level (for example, Wilson, 1993).

A3. Concepts of Value and Risk

A3.1 Concepts of Value

Although various methods exist for determining the value of environmental effects and these have been extensively discussed in the literature (for example, Pearce et al, 1989) three different approaches are immediately apparent. The first is to take the sum total value of all damage expected to be caused by the effect, the second is to determine the cost of mitigating the effect and the third involves calculating the cost of control measures which would prevent the environmental effect occurring (sometimes called the abatement cost). The latter is problematic because even if technology exists to control one environmental effect, it is likely that other effects will remain and/or be created so that one of the other approaches will have to be employed to value other elements of the total value, i.e. other environmental effects. Mitigation and damage costs are more likely to represent more fully the total value of environmental effects (Ottinger et al, 1991). However, the former approach also has associated difficulties, not least that the actual effect is dependent upon the assimilative capacity of the total environment affected, and that marginal damage costs often rise steeply once a critical threshold has been reached, so that the knowledge of effects and the complexity and diversity of damage to be predicted is often likely to lead to inadequate values.

An attraction of methods which set out to value the environment is that they can allow benefits to be highlighted which otherwise remain intangible. For example, the concepts of existence, option and bequest values can be "hardened" into quantifiables using various valuation techniques. By way of definition, existence (or "intrinsic") value is the value of knowing that something exists although direct benefit may not obtained from it, option value is the value of retaining or conserving the environment for future use, and bequest value is that element of option value representing a value assigned to the need to pass the environment on for future generations, i.e. not for future use for the present generation. All such values are additional to use value which is the more commonly held view of value as being its direct use in economic terms. It should be noted that different authors use different definitions for these elements of value and that it is, therefore, possible to underestimate total value (by missing some element out) or overestimate it (by double counting through overlapping elements). Some authors distinguish differently between elements of TEV, for example, between consumptive use values where the economic system interacts directly with the environment, such as farming and forestry, and non-consumptive use values where there is no direct interaction, for example, aesthetic appreciation of the environment (Johansson, 1991). They also recognise the value of indirect appreciation of the environment, for example, through books and pictures, and the satisfaction associated with the knowledge that an aspect of the environment exists - equivalent to the existence value described above. Another possibility is to defind a special type of option value. For example, quasi-option value is concerned with the value of preserving the environment for the future, as with option value, but specifically emphasises the irreversibility of environmental change and complexity of ecosystem linkages. By conserving an environment as yet unknown but potentially useful genetic information will be available in the future, and uncertainty about the effect of removal of part of the ecosystem on its interrelations will be avoided. Thus, the value of the expected benefit of delaying development until uncertainty about effects is resolved, or the expected value of perfect information, is the quasi-option value.

A3.2 Concepts of Risk

The term risk encompasses uncertainty, known consequences and probability. Quantified risk must include the set of all possible responses to an environmental intrusion and the associated probabilities of those responses. The value of risk is the environmental cost, as determined, for example, by society's "Willingness To Pay" (WTP) to avoid the risk. Therefore, as the environmental effects and their costs and the receptors involved are always unknown at the project planning stage, the associated risks of possible environmental changes must be valued. Option value reflects this risk.

A3.2.1 Perceived Risks, Comparable Risks And Benefits, Avoidable Risks.

Risk arises from the existence of a perceived or predictable outcome, such as a danger or loss. Perception is one of the most difficult aspects of risk to quantify. People are more willing to tolerate a hazard over which they feel they have some control, than that over which they have none, almost regardless of the statistical risk involved. The comparability of risks is, therefore, difficult. A clearly untenable comparison is where the benefits accruing from the risk are unequal, and/or where there is a difference in the avoidability of the risk. For example, the probability of a meltdown at a nuclear power station may be smaller than an earthquake, but the benefits from accepting the risk of earthquakes is equal to the total economic benefits arising from a given area of land whereas that accruing from a nuclear power station is simply electricity. Clearly also, the latter constitutes an avoidable risk whereas to avoid the former would involve loss of economic value of the land, that is, to leave it undeveloped.

The statistical deviation from the expected outcome associated with a given risk varies between ordinary and rare events to the extent that they are often considered to be incomparable for aggregation purposes. Indeed, rare events are difficult to provide for using traditional risk assessment techniques since the mathematical methods used to derive expected value (probability multiplied by magnitude of consequence) are inappropriate given the high statistical deviation from the expected outcome. There is also a low statistical base from which to evaluate probabilities reliably. This leaves rare events as a key problem area. The need to place approximate values remains, however, since not including then in the valuation process assigns them a zero value which is clearly unacceptable.

A3.2.2 Net And Gross Risk

There is also the question of whether to consider gross risks or net risks. The former is the total risk associated with a particular activity. However, this effectively compares the activity with zero risk, a situation which is unreal. A more meaningful assessment of risk would be arrived at through calculating the difference between the gross risk and the risk of undertaking some other alternative activity.

A3.2.3 Average And Marginal Risks.

Whether average or marginal risk is the most appropriate for consideration depends upon the purpose of the assessment. Average risk is appropriate for impact assessment of past and present energy use while marginal risk should be used for assessing future technology choices (Rowe and Oterson, 1983).

A3.2.4 Occupational and Public Risk.

Occupational risks are borne by the workforce, whereas public risks are usually involuntarily borne by the general public. This is an important distinction, since if it is assumed that occupational risks are voluntary and vice versa, then occupational risks do not constitute externalities since they are included in remuneration packages offered to employees. However it should be noted that voluntary employee acceptance only holds true where alternative employment exists and compensation reflects risks involved. Also, although it is clear that involuntary public risks are imposed without regard to benefits, if the public were to be involved in decision making then they may be expected to allow for known risks in this process.

A3.2.5 Risk and Valuation

Assessment of the risks of aspects of energy production and supply results in a large number of evaluated risks, each related to a particular part of the overall process. The aggregation of these risks provides further problems. An overview of problems in quantitative risk evaluation is given in the literature (Rowe and Oterson, 1983). Comparison of disaggregated results remains elusive unless they can be transferred to a common scale by valuation. In order to do this a common basis for risk measurement must be formulated, such as the number of deaths or injuries per unit energy output.

A4. Summary and Conclusion of Review

The technique of monetisation using the methods of neo-classical environmental economics attempts to overcome comparability problems by using established underlying theory and the single scale of money. However, it has weaknesses, particularly for impacts which are not normally considered in money terms. It particular, there is a lack of rigour and systematic approach; it more often takes a problem-centred approach which leads to the subjective selection of impacts to value. Therefore, what is needed is a technique which addresses the weaknesses of attempting to value in monetary terms. This technique must have as its foundation the pre-requisites of rigour and systematic approach - the main things missing from current approaches. In other words, the weaknesses in current approaches give us the criteria which any alternative approach must satisfy.

APPENDIX B. REVIEW OF ESI IMPACT VALUATION STUDIES

Eight studies which have attempted to value impacts associated with electricity supply industries across the western world are reviewed below. The studies are considered in reverse order of date of publication. Generally, it may be expected that more recent studies will be less problematic, having drawn on the experience and solved problems which have occurred in earlier studies, although this is not always the case. The literature is expansive, and these studies have been chosen not for their accuracy, but for their range of approaches and general thoroughness, as a cross-section of the literature.

B1. ExternE Project

The EC and US Department of Energy launched a joint research project to assess the external costs of fuel cycles in 1991 (ExternE, 1995a to 1995f). The work has subsequently been developed as part of the EC's JOULE programme and several member states, as the ExternE project. This major study has attempted to develop an accounting framework for all the major fuel cycles/options, subdivided into separate studies as follows; Coal, Oil, Natural Gas, Nuclear, Wind, Photovoltaics, Biomass, Small Scale Hydroelectric, and Energy Conservation.

The ExternE project was planned to be undertaken in 3 phases:

Phase 1. Development of the accounting framework, which was accomplished collaboratively between the US and EC teams. The methodology was applied to a range of fuel cycles (the EC team concentrated on coal and nuclear) as part of its development and the phase was completed in June 1993.

Phase 2. This encompasses three main elements and includes the application to all fuel cycles, where the EC team was expanded to include a range of centres with expertise in other fuel cycles, particularly for renewables and energy efficient technologies. It also includes national implementation of the accounting frameworks in a number of member states.

Phase 3. This was planned to involve the development of methods for the aggregation of accounting framework results in a way which makes them of value to policy and decision makers.

The ExternE project goes beyond earlier studies in several respects:

- Use of original data sources (both recent studies and modelling results generated within the project) to undertake more thorough characterisation of impacts on a site specific basis;

- Examination of all stages of the fuel cycle (previous studies have invariably concentrated on electric power generation);

- Identifying cases where externalities may be partly or wholly internalised.

Further review, detail, and discussion of the ExternE study is included in Chapters 5 and 12.

B2. Ferguson

From preliminary work on calculating costs of electricity generating technologies in the UK, Ross Ferguson has developed a rigorous approach to quantification of generic type impacts at an order of magnitude level (Ferguson, 1993, Ferguson, 1994a, Ferguson, 1994b). Results indicate that, apart from the effects of global warming (and possibly acid deposition) through burning fossil fuels, human health costs are only significant where weightings are applied to reflect society's aversion to occasional major accidents, although some significant occupational health costs are associated with coal mining and energy crop production. Effects on the natural environment are not quantified to the same degree since they are site specific and so cannot generally be quantified in generic terms. In general, electricity generation gives rise to the largest external costs, with fuel extraction, processing and transport being less significant. Possibly the most valuable aspect of this approach is the clarity of conclusions, derived from the rigorous nature of the approach and the diligent scoping calculations. Thus, discussions over the efficacy of individual numbers from various methodological variations in costing are foregone in favour of concentration on key issues, identified through order of magnitude calculations. Some of the clear conclusions are;

- Fossil fuels have huge health costs, particularly warming-induced famine;
- Nuclear has huge public aversion to accidents' costs;
- Renewables have potentially large environmental amenity costs if sited in the wrong places (site specificity is only considered in these general terms).

B3. CSERGE

The UK Department of Energy commissioned a team from the Centre for Economic and Social Research on the Global Environment (CSERGE) to survey the available literature on the monetary estimation of the social costs of energy production and the report was published in 1992 (Pearce et al, 1992). Since it is purely a review of previous studies the figures quoted incorporate the inaccuracies and shortcomings of those studies. Although a wide range of work is quoted, the report draws particularly heavily on that of Pace (see Section B4), with its incumbent assumptions. In general, standardisation of figures gathered is undertaken and broad comparisons are made, with conclusions being drawn on the basis of the authors' perceived reliability of studies and their comparability.

While both voluntary and involuntary health risks are considered, and, for example, routine radiation costs are quantified, the critical area of nuclear accident costs is avoided "owing to the absence of suitable literature". Likewise, although figures are quoted for damage, particularly in the areas of agriculture, forests, biological diversity, and buildings and materials, omissions are pointed out and, in general, where these are considered significant or where there is considerable disparity between figures, overall external costs are not reported in the final analysis. Global warming costs are quoted, on the basis of tonnages of fossil fuels burned, using a simple average of figures produced by Nordhaus (Nordhaus, 1991) and Cline (Cline, 1992a) for tonne-carbon-equivalent greenhouse damage. The approaches of these authors vary and questions can be raised about the efficacy of the figures produced. These are not answered by simply taking the mid-point between the two figures. However, the principal problem with the estimates is that they exclude "catastrophic" effects from global warming, which are considered to be likely. It is envisaged that an "insurance premium" be added to the estimate to reduce this risk but this is not done as "the current state of the art does not permit any reasonable guess at what this premium would be". The report does not suggest what should be done to internalise unknown externalities. The danger is that such an approach leads to confusing things that are countable with things that count.

B4. Pace

The Pace University study of 1991, although clearly based on the US ESI, is valid because it was the most detailed externality study for several years and has been used as a reference study by many authors (Ottinger et al, 1991). It was the first major study to use the damage cost approach at this level of detail. In common with many other studies, the costings are based entirely upon numerical estimates from earlier work. However, they do attempt to allow for site specificity and, through the use of dispersion modelling, determine reference areas of impact. The main omissions from the work are the lack of original raw data, the lack of inclusion of all stages in the fuel cycle prior to generation of electricity, and the lack of damage costs (and their replacement by control costs) in a few cases, namely global warming and certain water pollution impacts, where damage costs were found to be too difficult to determine.

B5. Tellus

The Tellus Institute uses the control cost approach and costs are based upon the existing legislative framework (Tellus Institute, 1991). This is considered to be a reflection of the societal value of the environment and therefore assumes that policy makers generate legislation on the basis of the wishes of society. This represents an alternative approach to that considered by the majority of recent studies, but its validity is questionable since the assumptions are dubious. Also, the approach is liable to generate a static policy mechanism where valuation is used to input into the decision making process, as current regulations are being used to generate values from which future regulations will be drawn. Since internalising externalities through the implementation of appropriate policies based on valuations is the eventual aim of valuation work, the approach is flawed and the results are of limited value.

B6. Hall

On the author's own admission, the results presented in this 1990 study are a "first rough cut" culmination of recent estimates (Hall, 1990). Air pollution effects are included (impacts on health and property) for coal, oil and gas. Other fossil fuel impacts calculated include;

- Acid deposition damage costs and cost of reducing acid rain;
- Costs of delaying global warming effects;
- Benefits of oil security and security costs of importing oil.

Omitted from fossil fuel costs are external costs of water pollution and solid waste. All effects of acid deposition and fog are omitted from oil and gas externalities, as are methane contributions to global warming from the latter. On the other hand the benefits of reducing air pollution from gas and oil are also underestimated. Only two pollutants are considered, and no botanical and property damage is calculated.

Nuclear externalities considered include insurance subsidies for nuclear accidents, storage of waste, loss of reactors from accidents and safety risks, benefits from additional safety regulations and security costs of minimising damage from terrorist attacks on power plants. The value of the subsidy derived from liability limits as established in the Price-Anderson Act is taken from Dubin and Rothwell (Dubin and Rothwell, 1990). This is conservative, since it relies on an assumption of maximum damage equal to $10bn, which compares to the cost of the Chernobyl nuclear accident at $41bn-$673bn, depending on the exchange rate used (Dubin and Rothwell, 1990).

B7. Bonneville Power Administration

The Bonneville Power Administration (BPA), a major US electric utility, sponsored a study by Hinman et al of Washington State University which surveyed public attitudes towards various risks using the Contingent Valuation Method (CVM, Hinman et al, 1990). In early 1990, 1600 random households were canvassed across Washington, Oregon, Idaho and the part of Montana where BPA operate. The response rate was 52% and, of ten environmental issues, coal-fired air pollution, water pollution, nuclear waste storage, radioactive material transport, and ozone depletion were identified as the main concerns. Of less concern were global climate change risks, new nuclear generation capacity, fish losses, radon in homes, and new dam construction for hydro schemes.

The BPA study did not examine the cost of different options, rather, it sought only valuations of willingness to avoid three entire technologies; hydro, fossil fuel and nuclear. Thus, the costs are in the form 'total you would pay to avoid...such and such technology..'. No information was given about these options, or other pertinent information such as potential locations. The results cannot therefore be used to inform choices regarding specific capacity-based or cost-based decisions on specific proposals.

A feasible way of using CVM in decision making may be to take a number of options, for example, to fill a 200MW capacity gap, and elicit values to allow a total environmental cost benefit analysis for each option to be undertaken. Furthermore, information bias is possibly one of the greatest problems with CVM studies. If people are not given information, it relies on what they have been fed by media, which is usually inadequate. Put crudely, if global climate change risks are given a low media profile (not necessarily deliberately suppressed, but maybe they are considered less news-worthy by media, but not the general populace) then they will incur lower costs. Information provision and type is an issue in all CVM studies and this links with the problem of how an impact or risk can possibly be valued if it is substantially unknown, whether this ignorance lies with the public at large or with the total level of scientific knowledge.

B8. Hohmeyer

Hohmeyer has undertaken several studies of externalities which have been published. The source considered here is the original study (Hohmeyer, 1988). As an initial low-cost estimation exercise, it attempts a comparison of fuel cycles on a general level. However, the approach, where existing aggregated air pollution data is apportioned to give the proportion attributable to electricity generation, has clear limitations and could not be used as a basis for the detailed assessment of external costs. As stated in the discussion above, aggregated data cannot effectively inform the decision making process where there is a possibility that site specificity and the type of technology are important factors. The data sources are approximate and generalised, while the range of impacts considered is limited and no consideration is given to stages in the fuel cycle other than electricity generation.

Specific criticisms of the work have been made (Jones, 1990). Premiums allowing for resource depletion (especially uranium) and research and development costs may be too high, and construction impacts are omitted, which could affect the relatively low estimates for renewables. However, it must be noted that the work was undertaken in order to highlight areas for further consideration, and to show the approximate potential size of some of the major environmental externalities, which the work does adequately, with some omissions.

APPENDIX C. DEMONSTRATION OF THE PATHWAY METHOD: SULPHUR DIOXIDE STACK EMISSIONS

The pathway analysis method has been developed as a means to identify all changes in the environment which result from the released incidental outputs (RIOs) of a given study project. The starting point is a complete list of RIOs, the RIO Inventory. Each RIO is then traced along the possible pathways it may take through the environment, until it reaches a sink or state of equilibrium. At each point it reaches (called a pathway destination) the changes it may cause are measured, predicted and recorded on an Environmental Change Inventory. Here, the pathway analysis method is applied to a single RIO from a generic case, in order to test the practicality of the method.

C1. Introduction to the Demonstration

One of the principal challenges facing environmental impact assessment, methodological development aside, is lack of data about the interaction of natural systems. While this is a legitimate concern for the pathway analysis method in application, for this demonstration, it simply provides a potential barrier to the successful demonstration of the method. Therefore, it is appropriate to select a RIO which is relatively well understood, and has been extensively studied. It is also important that the RIO selected is not a special case or a particularly simple case, since this would raise a question mark over whether and how the method could be applied successfully to more complex RIOs.

Some of the most complex RIOs are the most mobile ones, because they can travel and disperse over a wide area, thereby affecting numerous natural systems in many ways. Further complexity is provided by RIOs which can become chemically unstable in particular environments, or otherwise can affect secondary reactions. Even greater complexity is introduced where these substances' effects may trigger changes in natural systems which are cumulative, or where thresholds exist which, when exceeded, major shifts in equilibrium are generated. Gaseous emissions due to UK fuel cycles are reasonably well understood and are a major contributor to global, regional and local air pollution impacts, so it is appropriate to select a RIO from this group.

A RIO which satisfies the requirements of this demonstration is the sulphur dioxide (SO_2) stack emission from a coal fuel cycle electricity generation plant, or, more specifically, the sulphur (in whatever form). It directly enters the atmosphere on leaving the production system, it is mobile, and it is dispersed over a potentially wide area. Furthermore, it is known to take on various forms, influence various chemical and natural processes, and have the potential to influence delicate balances which exist within natural systems, along the lines indicated above. While many other examples of RIOs exist which satisfy this criteria, there is no reason why SO_2 stack emissions should not perform the demonstration function required here.

Clearly, SO_2 acts in concert with other pollutants. In particular, SO_2 is only one of the primary acidifying pollutants; others include oxides of nitrogen, ammonia and hydrocarbons, while secondary acidifying pollutants include ozone and acid rain. The effects of SO_2 may be offset, exacerbated or synergistically accelerated by other combinations of materials, energy and events. Such effects are outside the scope of this demonstration. However, it is important to note that the pathway analysis method does provide specific provision for such considerations (Horne, 2000b).

'Anthropogenic' SO_2 also meets the need of a substantial background of relevant research. Much of this is due to the intense acid rain debate at international level over the past 3-4 decades, and the resultant need to establish the facts around the issue of sulphur transport, deposition and effects.

In fact, anthropogenic sulphur emissions have been affecting local air quality to noticeable extents for hundreds of years. Since the 1850s, it has increasingly been recognised as a major potential problem. A reliable predicted emission inventory of sulphur emissions since this period has been established (Lefohn et al, 1999). In the 1850s, rains were analysed around Manchester, noting that in the city itself they were acidified by sulphuric acid resulting from coal combustion and a book was subsequently published which heralded the beginning of modern air pollution climatology (Smith, 1852, Smith, 1872, cited in Gorham, 1998). After the turn of the century, acid rain was established as having a causal relationship with impaired plant growth, seed germination, microbes in the nitrogen cycle, and soil quality. In 1955, a correlation was established between hydrogen ions and sulphate ions in the observation of precipitation in rural Scandinavia. Thus, it was established that the main problem is the contribution of anthropogenic SO_2 to acidity, that is, the total hydrogen ion loading on a given deposition area, since it is shown to be a major precursor. It should be stressed at this point that the pathway analysis method is not concerned with acid effects; the starting point is simply measurement and tracing of sulphur atoms (in whatever form) from the flue through the environment to a steady state or sink. There may be other effects of this process other than acid deposition. Nevertheless, it is this phenomenon which has driven the vast majority of sulphur cycle research.

The episodic events of localised acidic smogs and fogs from sulphur emissions which caused increased deaths in several industrial and metropolitan areas up to the 1960s were generally dealt with by bringing clean air legislation and raising point source stack heights to increase dilution rates. This is not to say that such acute problems have been eliminated, since strong localised acid aerosol events do still occur in relation to sulphur (for example, Thurston et al, 1994), and are more widespread in relation to other acidifying pollutants which are outside the scope of this study.

Long-distance air pollution emerged as an identified major problem during the 1960s. During this time, Scandinavian countries, in particular in a European context, began to investigate transboundary sulphur and its impact on the environment (for example, see anon, 1971). Two environmental problems triggered the increased focus on acidification as a major pollution issue. Acidification of fresh waters in Scandinavia was increasingly recognised as being associated with acid deposition, and a previously unidentified type of forest damage was recognised in a number of central and northern European countries. Initial research identified that the acidity of precipitation in Europe apparently increased by a factor of 10 in the century leading up to the 1970s. By the 1920s, sulphate ion concentrations in ice core samples from Greenland had increased by a factor of 2 over pre-industrial levels, by the 1950s by a factor of 3, and by 1980 by a factor of 3.5 (Alcamo et al, eds, 1990). As concerted research progressed, better understanding of the sulphur system and the impacts of 'anthropogenic' sulphur developed. It is now possible to describe in detail the range of states and reactions involving anthropogenic sulphur, as shown in Figure C1.

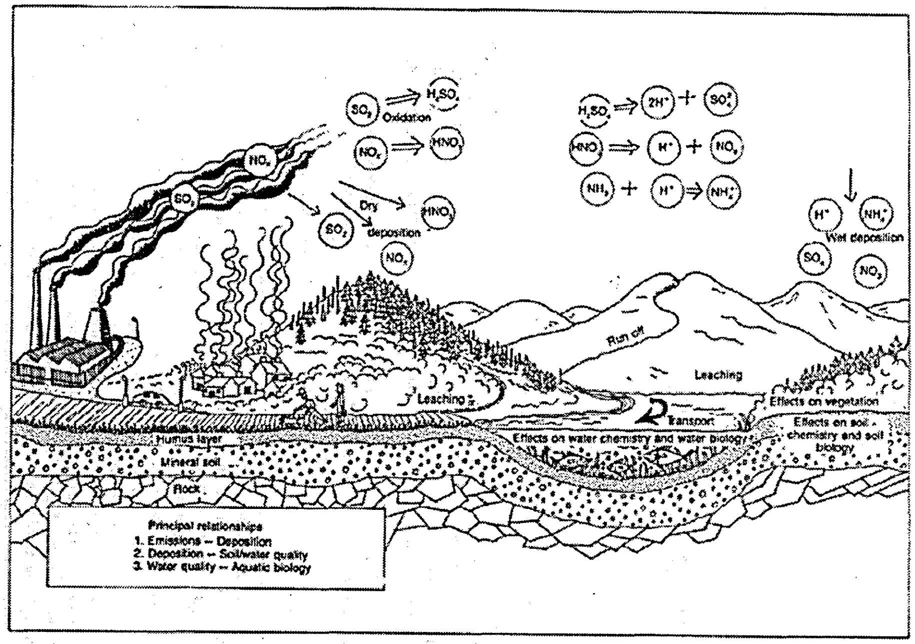


Figure C1. Acid-Generating Emissions, Atmospheric Reactions, and Targets (after Howells, 1990)

International deliberations on co-ordinated policies started in the 1970s, particularly towards the end of the decade. The UN Economic Commission for Europe adopted the Convention on Long-Range Trans boundary Air Pollution in November 1979, and came into force in 1983. In 1987, the Protocol on the Reduction of Sulphur Emissions or their Transboundary Fluxes by at least 30% came into force. The development of policies and regulations was matched by a corresponding increase in monitoring of pollutant dispersal and its effects. More recently, the second sulphur protocol has led to more developments in research to meet policy-making requirements. Thus, detailed tracing and modelling studies continue to be undertaken.

C1.1 The Sulphur Cycle

The RIO (and indeed all 'anthropogenic' sulphur) will affect the natural sulphur cycle, which, in perhaps it most simplest form, may be indicated as in Figure C2.

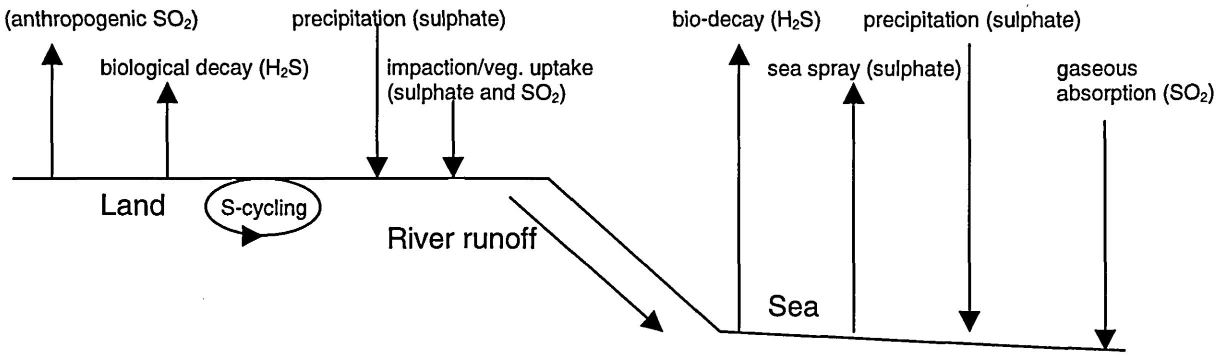


Figure C2. Generalised sulphur cycle (adapted from anon, 1971)

The accuracy of quantities of transfer is still questionable, and it is likely that all the transfers shown above are within 2 orders of magnitude of each other. One quantified sulphur model is shown in Figure C3. The other component type of the sulphur system is the sink, a special type of destination with a long residence time. While interim storage of sulphur does occur at the terrestrial surface, the vast majority of sink sulphur occurs in two forms; in solution in water, particularly marine waters, and in sea-bed deposits of precipitated gypsum (including fossilised versions, in the form of geological formations).

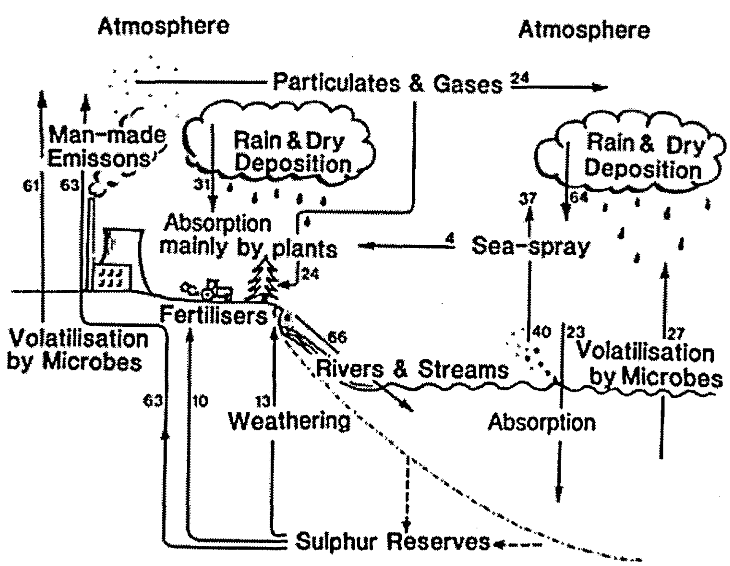


Figure C3. A Quantified Sulphur Model (Wellburn, 1988)

Note: quantities are Mt y^{-1}

C2. Pathway Identification: Dispersion and Deposition

All sulphur emitted from the power station stack will enter the atmosphere, usually initially as SO_2. Therefore, all of the first destination pathway for the RIO will be the mixed layer (roughly up to 1km height) of the atmosphere. The atmospheric conditions determine what happens to the sulphur next. Therefore, it is necessary to gain a detailed understanding of these in order to be able to predict where the second pathway destination is. For each S-atom, there are three possibilities; continued residence in the mixed layer; upward mixing into the troposphere, and downward deposition out of the mixed layer onto the earth's surface. This Section is therefore concerned with pathway destinations 1 and 2, which can be illustrated diagrammatically as in Figure C4.

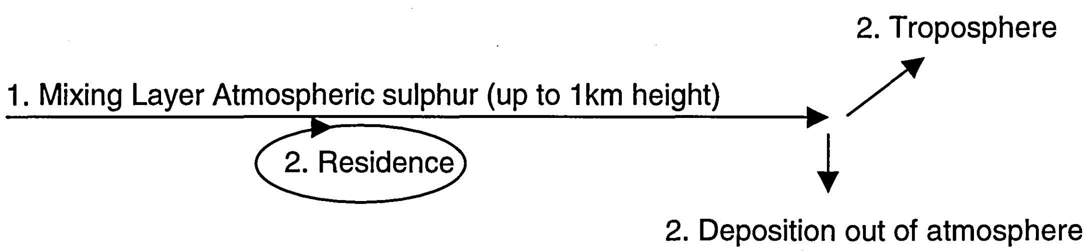


Figure C4. Pathway Destinations 1 and 2

Sulphur is mainly found in the atmosphere as SO_2, H_2S and bound to particles, usually as sulphate or sulphuric acid. A significant mode by which sulphur is transported into the atmosphere naturally, is in the form of H_2S from biological decay - another is in the form of sulphate aerosols emanating from sea spray. However, knowledge about the modes of transfer of these is still developing. Table C1 provides some figures regarding the lifetime and quantity of sulphur species in the atmosphere.

Sulphur Compounds	Concentration*	Pool Size Tg(S)**	Residence Time (Days)
SO_4^{2-}	0.1 - 0.56	0.26	7
SO_2	45 - 340	0.25	4
COS	100 - 560	2.20	160
CS_2	70 - 370	0.60	45
$(CH_3)_2S$	58	0.021	0.75
H_2S	0 - 320	0.041	1.5

* SO_4^{2-} concentration in $\mu g\ m^{-3}$; others in parts per trillion (ppt) by volume. Values reported for remote locations.
**1 Tg = 10^{12}g.

Table C1. Lifetime and Quantity of Sulphur Species in the Atmosphere (Legge and Krupa, 1990)

The same authors provide a schematic representation of the oxidation processes of volatile sulphur in the atmosphere (reproduced in Figure C5).

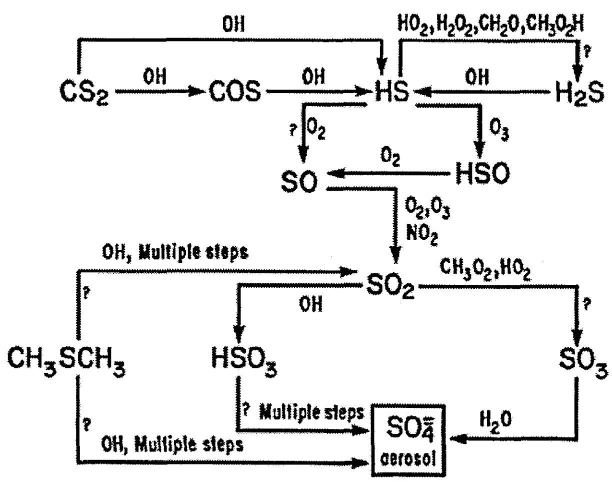


Figure C5. Oxidation Processes of Volatile Sulphur (Legge and Krupa, 1990)

SO_2 and H_2S are therefore gradually oxidised in the atmosphere, the speed of the reaction being dependent upon presence of catalysts, drops of water, ammonia, particulate matter, etc. As a result, the two forms become sulphate, and hydrogen ions are formed at the same time. The neutralisation of this acid within the atmosphere is dependent upon the local presence or absence of free bases.

The photochemical oxidation of SO_2 to form SO_3 in clean, dry air occurs within a period of 24 weeks, a period which can be shortened in the presence of moisture, ammonia and metals and hydrocarbons that can act as catalysts. In polluted air, the conversion can take place in a matter of hours, and the sulphate, in misty or foggy weather, can be rapidly returned to the ground. So, the residence time of sulphur in the atmosphere can range from hours to weeks. In Northern Europe, the mean residence time has been estimated as 24 days (anon, 1971).

C2.1 'Anthropogenic' Sulphur in the Atmosphere

Sulphur emitted into the air by human activities represents the vast majority of the sulphur found in the atmosphere over NW Europe. Emissions of sulphur are now generally well known, and at least one global inventory has been produced (Spiro et al, 1992). Emissions of sulphur from a point source such as a power station stack are carried away from the source by wind and mixed rapidly with clean air. The dilution rate depends upon prevailing weather - in winter, for example, vertical temperature stratification often depresses mixing and dilution rates. In the air, direct effects of 'anthropogenic' SO_2 are confined to the immediate vicinity of emission sources. However, on average, sulphur remains in the atmosphere for 24 days, during which time it may travel over 1000km. Upon deposition, either as rain or as dry particles, it may then oxidise to form acid. At this point, the newly formed acid may be more or less neutralised by bases, depending upon their quantity, state and chemical conditions prevailing.

There are two main phases of products of 'anthropogenic' sulphur in the atmosphere; dry, as particles containing sulphates and wet, as solutes containing sulphuric acid. Regarding the latter, this may be "occult" as in fog/mist which is relatively stable in suspension, or as larger water droplets in rain. A model has been successfully produced to predict the process of acid rain formation by SO_2 (Pal, undated). Rain is naturally acid, since it contains dissolved CO_2 in equilibrium, giving it a pH of 5.6 at normal atmospheric temperature and pressure. Even this figure has been questioned (for example, Legge and Krupa, 1990), since natural sources of sulphur and nitrogen could locally result in rainfall of below pH 5. Sulphur compounds and others, including oxides of nitrogen and ammonia, whether natural or anthropogenic, scavenge hydrogen from the air into rain and so further reduce pH. In so doing, they reduce atmospheric acidity, but on oxidation following rainfall, they can lead to acidification of soils and waters or other deposition surfaces. While dry air particles are not all acid forming, sulphates which form from the oxidation of SO_2 are.

Rates of oxidation of sulphur after emission are clearly important in determining rates of mixing and potential for deposition. It has been estimated that, for a 220m stack, oxidation of the sulphur plume occurs at the rate of 1-3% per hour, with a dilution factor of 10000 being achieved within 10-20 km of the stack (Howells, 1990). In dry weather, around half the SO_2 remains in the air for at least 24 hours, by which time it typically travels up to 600km (7m/s). Quantities are also important. Figures for 1976 suggest that of 174 Mt S y^{-1} total northern hemisphere atmospheric emissions, 98 Mt S y^{-1} came from 'anthropogenic' sources (Legge and Krupa, 1990).

C2.2 Atmospheric Sulphur Dispersion and Deposition Modelling

The general pathway approach and its use in informing the development of models to predict dispersion and effects of atmospheric pollutants has been advocated as far back as the 1970s (see, for example, anon, 1979) and probably beyond. A major international study of long range transport of sulphur pollutants on a scale of over 1000km was initiated by the Organisation of Economic Co-operation and Development in 1971. While the approach is not new, recent technological advances and improvements in understanding mean that more recent models are more accurate. Indeed, the major research effort across the developed world has led to rapid improvement in knowledge of the dispersal patterns of pollutants, including sulphur from point sources within and across Europe. The summary presented here is an illustration of some of this knowledge rather than an exhaustive review of it.

The principle method of measuring dispersion of pollutants over a wide area is through the development of computer-based models, using algorithms to predict the pollutant loading in any given place over time. Typically, the early models used Gaussian plume dispersion equations, and typically they showed poor performance in predicting lower concentrations with pinpoint timing and location accuracy, but perform well for average, large dose predictions over longer time periods. Many current models are based on the basic Gaussian model (Jennings and Kuhlman, 1997).

To predict deposition at any one place or into any given pathway destination sub-system - an accurate model which is designed to predict long-distance deposition rates of sulphur is required, in this case, one which is designed to predict dispersion and deposition from a point source. Early models included the CDM (Climatological Dispersion Model) developed in the USA, which used sector-averaged Gaussian plume dispersion equations, and was subsequently modified to include more variable input and improve accuracy. Subsequent tests showed that, in SO_2 vegetation response studies, these improvements were sufficiently accurate to be "a useful substitute for measured pollutant data" (Legge and Krupa, 1990). The major National Acid Precipitation Programme (NAPAP) was initiated in the US (NAPAP, 1987). Numerous models have also been developed and refined in Europe and elsewhere.

A "single layer" basic model was developed and subsequently improved, under the major Co-operative Programme for Monitoring and Evaluation of the Long Range Transmission of Air Pollutants in Europe (EMEP), which resulted from the 1979 Convention on Long Range Transboundary Air Pollution in Europe (Sandnes and Styve, 1992, Mylona, 1989). This provided calculated annual budgets for acidifying pollutants, and used real dynamic weather data on humidity, wind, height of mixed layer, precipitation, pressure, vertical velocity, temperature, turbulence, cloud cover and temperature of mixed layer. The model suggests that 15% of SO_x is redeposited in the local area, 5% is emitted as particulate sulphate, and 80% is emitted as SO_2. Verification tests showed that predictions were generally good, with over estimations for late 1980s winter values (due to warmer winters) and a few local unexplained anomalies. SO_2 modelling is easier that NO_x modelling, since it is efficiently dry-deposited, does not transport as far and tends to be more point-source based.

Although dispersion describes the mechanism which is sought in the predictive part of the atmospheric model, it is, of course, the deposition pattern which is the required end-point of the exercise and, more specifically in this case, the loadings of sulphur into each pathway sub-system. Hence, the term deposition is also used in many models. The Regional Acidification INformation and Simulation (RAINS) Model, was developed by the International Institute for Applied Systems Analysis (IIASA) in the mid- to late-1980s as a tool to assist environmental policymaking. RAINS uses country scale emissions of SO_2 and NO_x derived from mass balance using data on energy consumption, together with information about calorific value and sulphur content of fuels and ash. Transport and deposition is predicted using data on winds, precipitation and other meteorological and chemical variables. Acidification is estimated specifically for forest soils, and the possible impact on forest health is assessed by setting a simple threshold of pH or base saturation where this risk is assumed to occur. A 50cm average soil profile is assumed, as are standard buffering characteristics, and standard rates of filtering by forest foliage and neutralisation by base cation deposition, such as magnesium and calcium (Alcamo et al, eds, 1990). RAINS also provides a quantitative overview of lake acidification, through basic surface and groundwater modelling of runoff and throughflow in catchments using data from Sweden, Norway and Finland.

Further discussion of changes in forest, lake and other ecosystems is given in Section C3. However, one of the primary purposes of modelling is to allow accurate prediction of pollutant loading, that is, the amount of pollutant deposited in a given area. The RAINS Model was tested against observations and found to be accurate within a factor of 2 in predicting sulphur transport and deposition (Alcamo et al, eds, 1990). The latest incarnations of such integrated assessment models demonstrate even greater validity and accuracy, for example, RAINS 7 Europe (Schopp et al, 1999). For the first time, the Second Sulphur Protocol made use of an alternative, effect-oriented approach, in which the extent of emissions reductions is guided by the impacts the known source of emissions will have on known sensitive ecosystems. In other words, they are more efficient and better targeted than the earlier, less accurate, broad-brush models.

In one study, three dispersion models were used to estimate concentrations of SO_2 arising from different sources at 4 locations across the UK (BERG, 1989). These were developed by the Central Electricity Research Laboratories (CERL) the British Coal Corporation (BCC) and Warren Spring Laboratory (WSL). In was concluded that each provided “reasonably accurate” estimates. One recent comparison of 20 acid rain models (Hordijk and Kroeze, 1997) showed that all had basic similarities in structure, and all suffered from high levels of aggregation, difficulty in validation, and indicated by inference potentially missing sulphur. In other words, they fail to predict actual measured sulphur deposited perfectly. Another recent comparison of three different receptor-oriented models was made using EMEP-based emissions and good correlation was found in each case (Charron et al, 1998). A combination of modelling and some monitoring to verify predictions allows detailed and confident predictions to be made over wide areas (for example, Reynolds et al, 1999, ETSU, 1999).

In 1991, scientists and regulators in Europe launched the initiative for harmonisation within Atmospheric Dispersion Modelling in Europe. The aims of this were to improve dispersion models and their evaluation, improve data and tools, and thus raise the benefits of using models in decision making. The range of models which have benefited from this initiative extends from local (urban) models of dispersion around buildings to those operating at the regional level (Cosemans, 1996).

Central issues in all models are calibration and validation, which become progressively more difficult to achieve with increasing size, complexity and number of variables involved in the model. Hence, numerous models have been developed for specific purposes. For example, a small scale dispersion model has been developed for sulphur modelling in urban areas (Owen et al, 1999), another was combined with a soil model for determining deposition and impacts in high Alpine forests in Switzerland (Graber et al, 1996), and another for modelling bulk deposition from a point source in NE Spain (Alastuey et al, 1999). Indeed, given the expansion in computational power and explosion in the number of dispersion models, initiatives towards standardisation and comparison are likely to be a continuous requirement for the foreseeable future. Incidentally, verification by monitoring is not always possible, not only due to resource requirements and number of locations required, but also because the accuracy of some measurements has been questioned. One way to avoid direct collection of pollutants and the dilution/evaporation problems they may involve is to use biomonitoring tools to indicate depositions of pollutants, such as through analysis of pine tree bark (Schulz et al, 1999).

The current state of the art does still include very large dispersion models, since they are required by policy makers, for example, in connection with recent and proposed legislation. The UK national acid deposition network provides co-ordinated predictions of deposition of nitrogen and sulphur compounds across the UK with reasonable consistency, although it does not accurately account for unusual weather patterns (for example, Beverland et al, 1998). Particulate and sulphate aerosol modelling is also developing rapidly (for example, Seland and Iversen, 1999). The latest generation of promising atmospheric pollution models are based upon artificial neural networks (for example, see Reich et al, 1999).

C2.3 Second Pathway Destinations and Quantities

As stated above, there are two possibilities for sulphur leaving the approximately 1km thick near-surface “mixed layer” of the atmosphere. It could go up into the troposphere, or down and be precipitated out as dry, wet or occult deposition at the earth's (or hydrosphere's) surface. While tropospheric mixing of sulphur is not well understood, an inability to perform effective mass-balance suggests some small fraction escapes upwards, much of which must re-enter the mixed layer elsewhere and at some undetermined future date (Crane and Cocks, 1987). It has been suggested that tropospheric SO_2 levels are about 1ppb above remote areas, 1-30ppb above rural areas, 0.03-0.2ppm above moderately polluted areas, and 0.2-2ppm above heavily polluted areas (Legge and Krupa, 1990). However, the bulk of the anthropogenic sulphur is, at least at some point, deposited downwards.

Typically over Europe 0.1-10gS/m^2 are deposited annually by dry deposition, and 0.2-2gS/m^2 by wet deposition (Smith and Hunt, 1979). Ambient air concentrations of SO_2 were typically 50ppb in urban areas, 15ppb in suburban areas, and 7-12ppb in rural areas in the late 1970s (Roberts, 1984, cited in Howells, 1990). The deposition of sulphate with rain (wet deposition) accounts for 60-70% of the sulphate deposition in Europe (Dunderdale, 1990). The composition of precipitation and air has been measured in the EMEP network throughout Europe since 1978, over which period there has been a significant decrease in sulphur compounds in air which has passed over the British Isles. There are observed disparities between modelled and measured sulphur over even annual cycles. However, on a long term basis and averaged over a large area (for example, for the annual mean over Europe) there is a linear relationship between the change in SO_2 emissions and the depositions of S-compounds (Hov, 1990). The first and second pathway destinations for the demonstration RIO are therefore well-established. A simple schematic of pathway destinations established so far is shown in Figure C6.

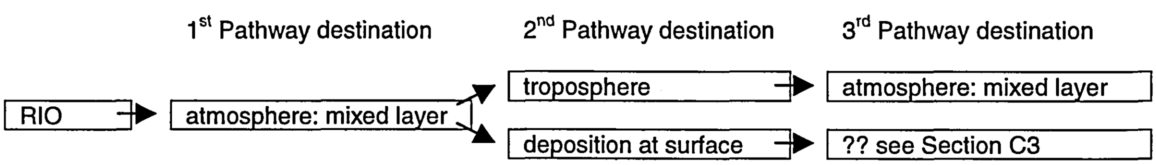


Figure C6. Simple schematic showing pathway destinations

Note that if a modelling exercise was carried out for a point source such as the demonstration RIO, the quantities could be established for all pathways (shown as arrows) and therefore the dose to each sub-system (shown as boxes). The third destinations can also be iterated for the tropospheric portion of the RIO, since there is good evidence (from mass-balance over time and over large areas) that it returns to the mixed layer and is eventually deposited at the surface (which will, in fact, be its fourth destination, provided it does not do another troposphere exchange cycle). This troposphere element to the circulation model will not therefore be given further attention in this demonstration. However, the third destination for the bulk of the RIO is what happens after deposition at the surface, and this is the subject of Section C3.

C3. Pathways and Environmental Changes: Terrestrial and Water Environments

C3.1 Initial Pathway identification

As the review of dispersion and deposition modelling in Section C2 demonstrates, it is reasonable to assume that all initial deposition rates can be predicted for the RIO. In other words, the plume dispersion can be modelled and the amount of RIO being deposited in any given geographical area can be predicted. Therefore, we know what proportion of RIO will enter which area.

The “area” is the next issue. What is the most appropriate sub-system to use? The generic inventory for coding must here be qualified to the extent that each receiving sub-system/area can be defined with easily identifiable boundaries. These will inevitably be best described and devised by specialists within each area being studied, although agreement must be reached over what the specific boundaries are. Fortunately, with the development of GIS base-maps, physical boundaries are easily delineated with precision and, provided a team of workers use the same subsystem divisions for the same RIO pathway analysis, the pathway Codes for each can be agreed on a case by case basis.

It should be stressed that in pathway identification, each sub-system will be a site-specific physically delineated volume of space (for example, “Duddon catchment”) rather than a generic environment or ecosystem “type”. For example, numerous subsystems will be required for individual study within the natural terrestrial environment. Pathway Codes can be assigned according to pathway identification steps, with each Code incorporating the Code from the parent/previous sub-system. For example, a given eco-homogenous forest area such as north Sherwood could be assigned the Code A/NTE/NS, to indicate it was receiving sulphur material from A (mixed layer of the atmosphere) via the natural terrestrial environment subsystem (NTE) of the specific location subsystem of North Sherwood (NS). Components, species, or areas within north Sherwood can then be assigned further Codes, also incorporating all previous Codes to indicate materials flow. Coding allows immediate and clear transparency and tracing of particular materials flows through particular pathways. Quantities need to be assigned to each pathway destination (which, by definition, has a unique Code).

As a means of demonstrating whether some of the more sensitive and complex subsystems are well understood enough to provide valid RIO-tracing information, it is necessary to examine general sub-system types, to indicate the state of knowledge. This is not intended to be an exhaustive review of effects, but an illustration of them as drawn from a small sample of the literature. Many exhaustive reviews have been conducted for various purposes (for example, TERG, 1988). This review is included to inform the process of pathway analysis.

The structure of the following discussion is based on generalised subsystem elements of the terrestrial and water environments, preceded by a short discussion about general sulphur model and the critical loads approach to pathway analysis and environmental change assessment.

C3.2 Critical Loads

Aspects of the general sulphur model have been introduced above, including the agents of bulk circulation of sulphur between the atmosphere, land and oceans. However, what happens on land is important and worthy of closer study, since it involves more complex potential range of states and transfer mechanisms for sulphur than in the other two major system components.

Sulphur is present in a wide range of rock types, from which it is weathered out, invariably as sulphate. The majority is then taken into solution and transported via fluvial and groundwater systems to the sea. Therefore, the introduction to the terrestrial environment of sulphur occurs from numerous sources; principally rocks, sea spray/salt precipitation (mainly near coastlines), and atmospheric deposition. The natural exit of sulphur from continents is primarily through biogenic breakdown (to the atmosphere as SO_2) and solute runoff via fluvial systems to oceans.

The task, therefore, is to quantify and trace the anthropogenic element of sulphur from atmospheric deposition through the continental system and identify where it causes critical build-ups or changes to the existing flow of materials. Definition of "critical" is largely a matter for environmental change assessment, which should and does follow rather then precede the assessment of pathways and flows. However, the QLOS approach is not being developed and demonstrated in a vacuum, and the so-called "critical loads" approach has been both advocated and used by major studies on the effects of anthropogenic atmospheric sulphur. Therefore, before proceeding to the reviews of the state of knowledge of flows through various parts of the continental system, a brief overview of the critical loads approach is useful in informing the sorts of approaches which have led to the current state and type of knowledge and data available on the subject.

The critical load of sulphur is the contribution of sulphur to the computed amount of acidifying material which causes the threshold effect identified as critical. The critical load concept was first used as a policy tool by the Canadian Government in the early 1980s. It is based on the establishment of dose-response relationships, and from these, of critical levels of pollutants for direct effects to occur on various receptors, such as forests crops, materials and humans. This then provides the possibility of establishing critical loads for sulphur compounds, above which, particular effects can be expected to occur, and below which, significant harmful effects do not occur according to present knowledge (Nilsson and Grennfelt, 1988). Critical loads were defined as those required to maintain a surface water pH of 5.3-5.5. It has been estimated that these loads ranged from 7.2-19 kg ha^{-1} per year, depending on soil tolerance factors. European wide studies have concentrated on establishing methods for mapping and producing maps showing critical loads (Hettelingh et al, 1991, anon, 1993), and incorporating these spatial data into dispersion and deposition models to show areas where critical load are or may be exceeded in the future under various emission scenarios.

The methods used to estimate critical loads vary, but a common approach is "steady state mass balance", which involves computation of loadings at which known threshold levels of acidity will be reached, taking into account the main acid-determinants, including base cation availability, secondary reactions, precipitation and evapotranspiration, elemental contents of soil and plants, and levels of plant growth. These data are for forest soil critical loads, but similar data could be used to construct mass balance equations for other natural systems, and has been done for water chemistry in the steady state water chemistry method which was also used in mapping critical loads for Europe (Hettelingh et al, 1991). Closely based on this approach, the following steps could be undertaken to compute similar critical loads for use in pathway analysis to assist in identifying areas of particular environmental changes, as follows;

Step 1. Define map parameters including target physical areas (the EMEP programme grid squares have been used by Hettelingh et al, 1991, to fit with dispersion model computations, although these could potentially be delineated by specific areas which have been given site specific Codes);

Step 2. Compute critical load using mass balance equations for locations across Europe (these could be done at sub-system level, incorporating ecosystem and base cation availability data, and data on other S-cycle factors such as natural inputs from marine salts);

Step 3. Assign dispersion modelling data to each area across the entire deposition area (this could also be done on sub-system level);

Step 4. Compute percentiles from combining the critical load data and cumulative distribution data, to show areas (subsystems) where excedence of loads (and the rates and times of excedence) will occur;

Step 5. Calculate sulphur fraction (for general acidity dispersion models, or combined S and N dispersion models, the sulphur fraction will need to be established to allow calculation of the percentage of critical load excedence attributable to sulphur);

Step 6. Assess the homogeneity of critical load effects across each area (sub system), to check that critical load levels calculations at subsystem level do not conceal local variations within the subsystem (if they do, the subsystem should be split into parts of homogenous critical loads and dispersion loading).

The critical loads approach has been applied across Europe using aggregated emission data and large scale modelling. It has also been applied to more localised point source cases, and numerous studies have been undertaken of localised sensitive ecosystems, both in northern and western Europe and elsewhere in Europe (for example, Camamero and Catalan, 1998).

Although the critical loads approach does not explicitly assist the calculation of environmental changes, it does form a method by which areas where given environmental changes could be expected can be easily identified under given loads. Thus, if different effects are found to be detectable at lower S-loads in particular sub-systems, then adjustments of critical loads equations can be easily achieved and the model re-calculated to show areas in which these effects could be expected. Thus, rather than using the critical loads approach in a simplistic way as under the first Sulphur Protocol, where emissions were targeted to try to reduce loads to below the critical threshold in the vast majority of areas (which meant also that in many areas they were reduced to well below critical levels), the same approach could be used to show a range of expected effects across the entire area, provided dose-response relationships for each effect could be built into the model.

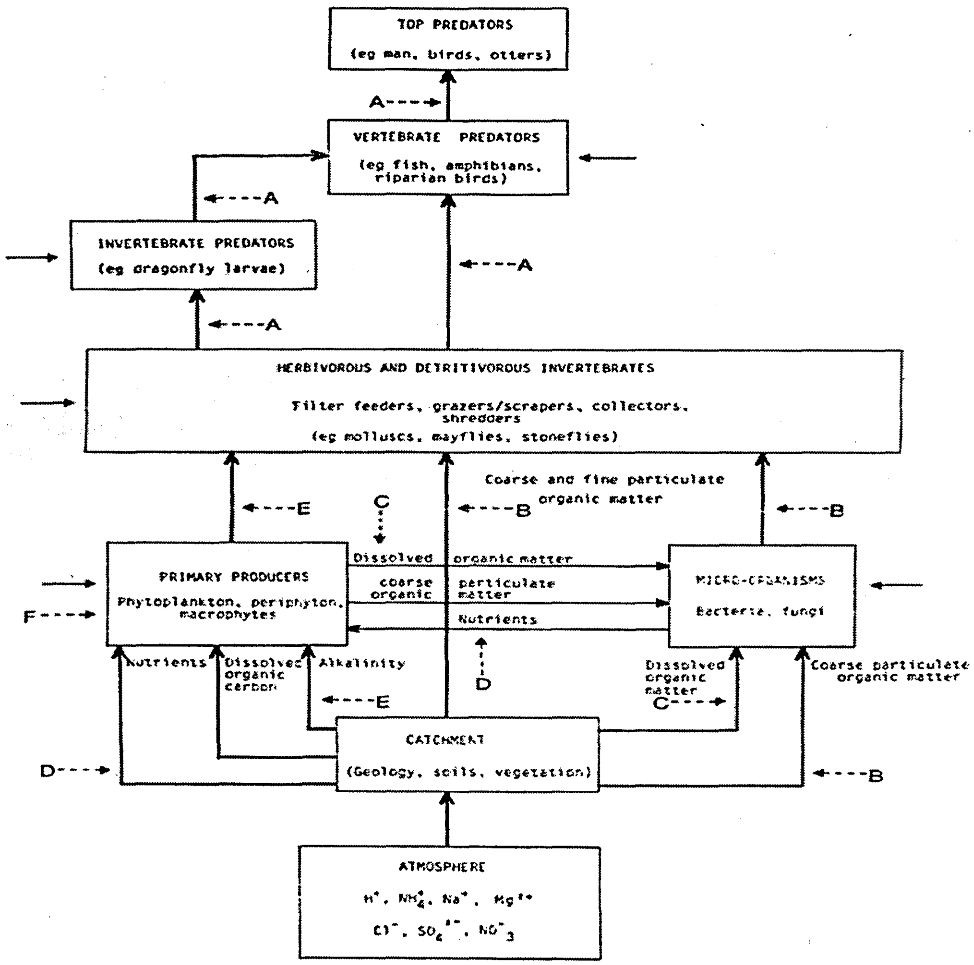
C3.3 Natural Terrestrial Environment

A general discussion of some of the elements of the natural terrestrial environment is presented below. Note that many of these elements exchange materials with other major subsystems and/or have similar components, such as soils, pests, etc. However, the discussion here is focussed on natural terrestrial environments and is intended to indicate the state of knowledge in key areas of the subsystem. Clearly, specific types of subsystem exists, for example, defined by particular types of dominant vegetation, such as trees (forest), grasses (savannah/grasslands), and upland species (bog/moorlands/mosslands). Here, one generic vegetation type is single out for illustrative purposes, that of natural forests, and a review if the state of knowledge of this type is presented in Section C3.3.2.

C3.3.1 General Review

Numerous models have been developed to predict acidification processes in natural ecosystems, including RAINS (and latterly, Smart RAINS), SAFE, MAGIC, ILWAS, RESAMDAS and MIDAS (UN ECE, 1993). In addition, many flowcharts, systems diagrams and schematic representations have been produced, some more detailed that others, to describe movements and effects of acidifying pollutants such as sulphur through natural ecosystems. An example is reproduced below and although it clearly cannot assist in quantifying pathways in its present form, it illustrates the points that there are numerous variables involved, that sources of effects often arise indirectly from sulphur pathways as well as directly involving sulphur contact and, furthermore, that these variables can be identified and modelled effectively.

The passage of sulphur through soils and the changes which occur as a result of its passage involves numerous and complex chemical processes. Ecosystem processes likely to be affected by acidification are illustrated in Figure C7. The vast majority of research in the area has been generated in the study of acidification of soils and related environments. Numerous reviews of the area have been undertaken (for example, Legge and Krupa, 1990, Reuss and Johnson, 1986 and Tabatabai, 1985) which illustrate the rapid and ongoing increase in knowledge of soils systems and the role of sulphur in them.



Ecosystem processes likely to be affected by acidification. The arrowed boxes represent trophic levels where acidification has direct toxic actions, whilst broken arrows represent possible indirect effects. These indirect effects include: A. Loss of acid-sensitive fish and invertebrates, resulting in the proliferation of some other invertebrates, and hence alteration in the quality of food available for top predators. B. Reduced decomposition of coarse particulate organic matter, affecting the quality of food available for detritivores. C. Complexation of dissolved organic matter by metals, leading to loss of substrate for microbial action. D. Aluminium and phosphorus may complex, reducing nutrient availability for plants. E. Loss of alkalinity may reduce the availability of inorganic carbon, changing the quality of primary producers. F. Changes in the invertebrate fauna may alter the quality of grazing pressure.

Figure C7. Ecosystem Processes Likely to be Affected by Acidification (Warren et al, 1988, cited in Edwards et al, 1990)

The soils of Europe can be roughly divided into 3 categories. First, old residual soils, mainly in Mediterranean countries, which have been strongly weathered residues that have been thoroughly leached. Second, the eiolian soils of central Europe, formed during the last glaciation, and containing high concentrations of calcium. Third, are the fine-grained sediment soils and coarse-textured tills of post-glacial northern Europe, which were deposited either by water or above water levels, and which are relatively young. Their age means that they are relatively immature, and so retain strong characteristics derived from parent rock material, which ranges from acidic granites to calcareous sediments. Thus, wide local variations in base-acid characteristics exist It is those soils which are acid-derived, young and subject to prevailing winds which bring polluted SO_2-laden air masses where the soils are most susceptible to damage from sulphur emissions. The buffering capacity of the soil, or alternatively, the proton budget (Gorham, 1998), in other words, its ability to counter/control acidification, is dependent upon micro-chemical conditions, including soil structure, presence of bases, water, etc. Although there are many variables, then, the nature of the parent rock type is clearly a major factor in soil buffering capacity, particularly across broadly similar climatic zones. A general indication of the rock type/buffer relationship is given in Table C2.

So, a central element to soil resistance to acidity is nutrient deficiency, since acids increase the hydrogen ion exchange rates with cations such as potassium, magnesium and calcium in the soil. The cations are necessary for plant growth. As the available stocks of such cations are used up by these reactions, they are not available for other nutrient cycling activities, and they are increasingly depleted within the surface layers of the soil as they are leached out in acid solutions. Thus, at an inorganic level, it is now well established that levels of these cations in soils will decline in response to acid deposition (for example, see Walna et al, 1998 and Abrahamsen et al, 1983), and the dose-response relationship can now be reasonably well established, despite the existence of complicating factors to this basic process. As these base cations are depleted, aluminium, which is normally bound up in silicate minerals such as clays, is increasingly mobilised to replace the cation function of buffering acid. This occurs when H+ levels have accumulated until pH is around 4.2. However, aluminium in moderate quantities has a toxic influence on many life forms, including plant and microbial communities. It is now well-established that soil acidification leads to loss of nutrients, aluminium mobilisation and root damage. Indeed, validated models have been developed which predict the development of levels of different aluminium species in acid soils, for example, SPECIAL (Goenaga and Williams, 1990). In some regions, cumulative deposition of SO_x and other acidifying compounds has leached all the available magnesium, potassium and calcium, leaving aluminium as the major base cation (for example, Hovmand, 1999).

Group	Acid neutralizing ability	Rock type
A	None - low	Granite, syenite, granite-gneisses, quartz sandstones (and their metamorphic equivalents) and other siliceous (acidic) rocks, grits, orthoquartz, decalcified sandstones, some quaternary sands/drifts
B	Low - medium	Sandstones, shales, conglomerates, high grade metamorphic felsic to intermediate igneous, calcsilicate gneisses with no free carbonates, metasediments free of carbonates, coal measures
C	Medium - high	Slightly calcareous rocks, low-grade intermediate to volcanic ultramafic, glassy volcanic, basic and ultrabasic rocks, calcareous sandstones, most drift and beach deposits, mudstones, marlstones
D	"Infinite"	Highly fossiliferous sediment (or metamorphic equivalent), limestones, dolostones

Source: Norton (1980); Kinniburgh and Edmunds (1986); Lucas and Cowell (1984).

Table C2. Typical rock type/buffer relationships (Hettelingh et al, 1991)

The linkage between pH (hydrogen ion availability) and buffer system is well-established. At pH levels of 8-6.2, calcium and bicarbonate generally dominate, while in the range pH 6.2-5, silicate-bicarbonate exchange operates. At pH 5-4.2, cation exchange systems such as involving magnesium, calcium and ammonia generally operate, while, as stated above, aluminium exchange dominates in the range pH 4.2-2.8. Below 3.8, iron exchange begins to provide the main buffer system. This is a gross simplification since many reactions can take place. However, they are governed by known variables and so can be theoretically successfully predicted.

Some research has shown that acidification changes biological activity in soils, reducing the rate of litter decomposition and therefore restricting nutrient release. It also affects mycorrhizal relationships between fungi and tree roots which facilitate tree nutrition. It can reduce respiration by organisms in soils (including bacteria), increase levels of ammonia by reducing mobilisation of nutrients previously released in decomposition, and decrease soil nitrate levels as a result of ammonification (Park, 1987). Increased soil acidity is also often associated with increased concentrations of heavy metals in addition to aluminium, including cadmium, zinc, lead, mercury, manganese, and iron (Tolba, 1983, cited in Park, 1987). Once mobilised by acids, these metals can travel in solution and be taken up by root osmosis into plants, and so enter the food chain. They may also enter the potable water supply (and so travel through two further pathways; water environment; human body). Indeed, this pathway is also potentially followed by the excess acid solutions, causing acidification of water bodies along the way.

It should be noted that meteorological factors such as rainfall greatly affect soil acidity, not just through loading adjustments (which is already built in to dispersion modelling) but also through affecting soil reaction rates, water content of soils, etc. Another major factor on soil acidity is land use, which is why it is essential that the pathway analysis method takes a subsystem based approach, and that each physical subsystem has a similar basic land use type.

Upland soils tend to be more susceptible to acidification, since rocks weather slowly and soils are thin, with low buffering capacity. Hence, softwater upland environments have few natural defence mechanisms, and high (acid) rainfall brings regular doses of acidity. The linkage between upland soils and acidity has been well established in Europe and north America (for example, Edwards et al, 1990). Thus, although not perfectly understood or proven, a range of effects within soils can be attributed to acid deposition, including atmospheric sulphur compounds. A fairly comprehensive summary of these is contained in Table C3.

The rate of reaction and resultant changes in pH and resultant Al-cation availability are clearly important in determining when environmental changes can be expected, and these can be estimated from basic data about quantities or reactants and soil conditions. For example, for one region, it was calculated that a reduction in pH from 5.3 to 5 could take 6-105 years depending upon loading scenarios, while for other areas, similar ranges were 50-280 years (for drop from pH 4.8 to 4), 5-15 years (for drop from pH 4.2 to 4), and 20-160 years (for drop from pH 4.2 to 3), the last two examples being for the same site (Legge and Krupa, 1990).

In addition to observing acidification, numerous experiments have been conducted to artificially induce phenomena for observation under relatively controlled conditions. For example, artificial irrigation experiments have shown that soil microbial communities may not be particularly badly affected by heavy metal deposition when combined with the pollutant effects of sulphur deposition (Pennanen et al, 1998). Observations of critical relationships between total cation concentrations for soil solution composition and the saturation of exchange sites are ongoing (for example, Matschonat and Vogt, 1998). With continued improvements in understanding of spoil/sulphur relationships, modelling continues to improve. Recent validation testing of SMART2, a revised version of the dynamic soil acidification model SMART, showed that desired accuracy in calibration was achievable; the test site was a forested catchment in Finland (Ahonen and Rankinen, 1999).

Process or Property	Hypothetical Impact of Acidic Deposition
I. Soil Exchange Complex	
Exchange Capacity	Decrease in CEC resulting from the influence of clay aluminum
	Increase in CEC of soils with oxy-hydroxides due to sulphate adsorption
Exchangeable Acidity	Increase
Base Saturation	Decrease
Clay Mineral Morphology	Increased formation of hydroxy-Al interlayers and acid weathering
Aluminum	Increased mobilization and leaching
	Increased availability and toxicity
II. Organic Matter	
Organic Matter Turnover	Decreased rate of C mineralization due to acidification and/or associated trace metal toxicity
	Decreased CO_2 flux from land to atmosphere
	Increased retention of organic matter
Microbial Community Dynamics	Shift from bacteria to more acid-tolerant fungi
Organo-Mineral Associations	Reduced organo-clay interaction due to disruption of cation bridge linkages
Root Uptake	Trace metal toxicity due to acidification
III. Plant Nutrients	
Nitrogen	Decreased ammonification
	Decreased nitrification
	Changes in products of denitrification
	Increase in leaching
	Enhanced cation leaching due to NO_3 inputs
	Reduced plant availability
Sulphur	Increased SO_4^{2-} reduction in low S, anoxic systems
	Increased reduced-S flux and reduced CH_4 flux to atmosphere
Sulphur (Continued).	Decreased leaching of S
	Decreased leaching of cations in sesquioxidic soils; increased leaching in others
	Reduced plant availability
Phosphorus	Decreased leaching and $AlPO_4$ precipitation in soil with high Al
	Increased PO_4^{3-} solubilization, plant availability and leaching in calcareous soils
	Reduced availability with pH reduction
Fe, Mn, Zn, Cu, Co	Increased availability
	Increased leaching
Mo, B	Reduced availability
Ca, Mg, K	Reduced availability
	Increased leaching
Toxic Elements	Some micronutrients may reach toxic levels due to increased solubility
	Increased concentrations, toxicity, and leaching of heavy metals
	Increased Al toxicity
IV. Weathering	
Carbonates	Increased dissolution
Primary Minerals	Increased dissolution
Clay Minerals	Increased influence of aluminum (formation of Al interlayers)
	Reduced surface charge

Table C3. Summary of the Potential Impact of Acidic Deposition on Soils (Turchenek et al, 1987, cited in Legge and Krupa, 1990)

Acid surges are a phenomenon that can cause particularly high levels of damage in short time periods. Sulphates as undissolved acids gradually filter down from the atmosphere at a rate determined by their size, aerodynamics and atmospheric conditions, and are deposited as dry deposition. During periods of drought or low precipitation, dry deposits accumulate as dust particles and become highly concentrated 'acid powders' that are washed out by subsequent rain and cause acid surges into an otherwise equilibrium environment. A similar form of build up can occur during prolonged icy conditions, when build-ups of dust in snow and ice are subsequently released with meltwater in the spring. Fog drip or occult deposition can also generate a large surge of sulphur input. Thus, acid deposition can be highly episodic, and a few isolated peaks of high acidity can cause much ecological damage (Park, 1987), a phenomenon which has been established both theoretically and experimentally (for example, Quist, 1998).

For surges to occur, there must be a build-up of the causative components earlier in the system. The early view that acid precipitation sulphate was directly and quickly released into soil drainage and streams has been substantially modified, as early mass balance assessments have often proved to be very approximate. It is recognised that sulphur deposited or drained to wetlands or reducing groundwaters can be retained as reduced sulphide. In wetlands, the amount of retention has been experimentally measured (by using ^{35}S tracing) as up to 70% of sulphur input (Brown, 1985, cited in Howells, 1990). Seasonal surges of sulphates are also common, demonstrating that residence time of sulphur in terrestrial environments can be considerable. Clearly, in its reducing phase, sulphur will be deposited, and only in oxidising dry phases may it then be released. On longer timescales of hundreds of years, the vast majority of sulphur deposited on natural systems will find its way into solution in large saline water bodies such as seas and oceans where, once again, it will eventually precipitate out into benthic muds as reducing conditions as critical concentrations of sulphur-bearing salts are reached.

Vegetation is also directly affected by SO_2 in the air. Where SO_2 concentration exceeds 0.01 -0.02ppm during extended periods, or 0.04-0.05ppm for a few days, coniferous trees in urban areas are directly damaged (anon, 1971). Different plants vary widely in their resistance to SO_2 sensitivity, for example, some lichens are particularly sensitive. Acid damage can potentially occur in two places; above ground direct air pollution contact with foliage, or through roots affected by transfer of acid and acid effects through soil/substrate pathways, for example nutrient deficiency or aluminium poisoning. Acid attack leads to a burning effect on the photosynthetic mechanism, and this is especially apparent in non-deciduous perennial plant species where leaf surfaces are exposed typically for several years, and so are subject to cumulative acid attack. Research in recent decades has tended to concentrate on the effects of acidification at realistic concentrations on stomatal responses, biochemical reactions, photosynthesis and assimilation.

Thus, plant physiology is affected in various ways by air pollutants, including; inhibition of photosynthesis, stomatal opening which increases pollutant uptake), decreased translocation of nutrients to fruiting bodies, more carbohydrates retained in leaves for repair and maintenance, and decreased root growth, restricting uptake of water and nutrients. Numerous publications and reviews have been published dealing with stomatal responses to SO_2 as well as many other atmospheric pollutants.

Sulphur may act as a nutrient when absorbed through leaves or taken up by roots. However, the overall effect may not be beneficial, as the stimulated plant growth may also be accompanied by decreased tolerance to frost, drought, and increased palatability of foliage to leaf-eating insects. Fungal attack may also increase, particularly where root growth and plant vigour is reduced.

Lichens and mosses have been suggested as potentially particularly sensitive to acid deposition. Ferguson et al (1978) demonstrated that Sphagnum growth is impaired by fumigation by SO_2 and treatment by sulphuric acid and concentrations found across northern Britain. This species has undergone major decline over the last century. As bryophytes, they lack a cuticle and absorb water rapidly after rain, thus they are exposed more directly to wet deposition than vascular plants. Mosses may therefore provide useful indicator species for early acid deposition effects Mäkipää, 1998). Lichens are flat plants which occur on a wide range of surfaces, and can obtain nutrients directly from the breakdown of rock materials. They have been shown to be sensitive to acidity levels in air. Although there is considerable variation between species, many die if subject to long periods of exposure to 0.01-0.02ppm SO_2. Since different species have varying tolerance to acidity and lichens can otherwise survive in a wide range of environments, they are a group of plants which have particular potential for providing indicator species. However, there remain considerable gaps in knowledge about lichens with regard to deposited anthropogenic sulphur dose-response relationships, although work is ongoing (for example, Mahoney et al, 1995).

One technique which is widespread use by environmental scientists to overcome the problem of the shear number of species and amount of biomass to be studied within any given system, is to identify a particular species or set of species within an ecosystem type, which is/are sensitive to particular changes, in this case, for example, acidity or sulphur. Within the complexity of an ecosystem with many species, this gives a rough and ready means of identifying the likely impact of such changes by undertaking relatively few measurements concerning only these indicator species. This approach is limited, and is problematic, since it does not take a materials flow approach and therefore does not allow mass balance to be applied, nor pathways to be established.

Some fungal diseases are increased by pollutants, while the growth rates of certain plant pests such as aphids have been shown to be stimulated by pollutants (TERG, 1988). However, some plant pathogens are inhibited by the presence of current anthropogenic levels of SO_2. Increases of background levels in rural areas have been associated with a decrease in incidence of pathogens, while decreases in SO_2 levels in urban areas have been associated with the re-appearance of diseases such as black spot, which was formerly controlled by the presence of SO_2.

C3.3.2 Forests

Arguably, effects of acidification on trees are more difficult to detect than those on annual crops (see Section C3.5) due to the greater biological complexity of forest ecosystems and longer response times to acid stress. Also, two types of forest stand can be differentiated. A natural type is a dynamic unit comprising a more or less continuous successional sequence, with natural tree death occurring for a variety of causes, and various stages of young growth replacing it. A climax stand is established through successional development. However, in planted and managed forests with even-aged stands, the succession concept has little utility. This review concentrates on the former, more complex and diverse problem of establishing sulphur pathways through natural forest systems.

Clearly, natural forests are affected by acidification by two routes; direct air pollution depositing on foliage and uptake of acids and acid-related effects in roots from the soil. Regarding the latter, the discussion and review of processes in soils (Section C3.3.1) clearly of direct relevance here and no further discussion of soil processes specifically is required here, apart from to note that forest damage research invariably involves examining key soil characteristics. Indeed, criteria for relating risk of forest damage to chemical characteristics of soils has been developed, for example, as shown in Table C4. Calcium levels in topsoils are related to productivity of forest land. Since sulphuric acid leads to leaching of calcium and other bases, it reduces plant productivity. An annual rate of reduction of about 0.3% has been postulated (anon, 1971).

Soil Property Risk	Increasing Risk	High Risk	Very High
$pH(H_2O)$	≥ 4.2	4.0 to 4.2	< 4.0
Base Saturation	≥ 0.05	< 0.05	0.0
Al^{3+} ($\mu mol\ L^{-1}$)	< 25	25 to 40	> 40
Ca/Al Ratio	> 0.4	0.1 to 0.4	< 0.1

Table C4. Linkage Between Risk of Forest Damage and Soil Chemistry (Ulrich et al, 1984, cited in Legge and Krupa, 1990)

Regarding direct atmospheric pollution, particles can become trapped on tree foliage and subsequently washed from leaves and stems by rain. Acid rain can also potentially interfere with and damage foliage and foliage processes. However, the simple comparison of SO_2 deposition patterns (mainly around urban areas) and areas of known forest dieback (mainly in upland remote areas) demonstrates that SO_2 damages can be eliminated as an overall cause for this dieback. Furthermore, needle analysis studies generally indicate that S-levels rarely exceed normal background concentrations. Also, some trees also tend to acidify soils without pollutant presence, and it is important to separate such natural processes from the effects of 'anthropogenic' sulphur. However, it has been suggested that acidifying pollutants could cause chronic damages and act as part of a more complex damage cycle. It has also been noted that SO_2, NO_x and other gaseous pollutants are "much more phytotoxic" when they occur in combination than as single agents (Krause, 1987, cited in MacKenzie and El-Ashry, 1989).

Thus, there has been some difficulty in establishing a clear materials link and dose-response between SO_2 deposition and so-called "forest dieback". The latter has been observed across north America and Europe, although it is not a universal phenomenon and apparently affects different tree species in different ways. A classification system of damage has been developed in Europe and used widely in forest assessment studies, based primarily on needle/leaf loss. However, it is now clear that numerous causes exist for forest dieback, and SO_2 deposition is only one of them, and may not be dominant or even directly causal. For example, some known causes of damage to Norway Spruce which can be confused with recent forest decline are presented in Table C5.

Agent	Symptom
Abiotic	
Low temperatures or 'winter chlorosis'	Needles of all ages slightly yellow in early spring
Drought or salting of nearby roads	Red coloration of older needles on branches in exposed localities
Pathogenic	
Spruce needle cast fungus (*Lophodermium piceae*)	Older needles red-brown with black spots or bands in spring
Spruce needle scorch (*Lophodermium macrosporum*)	Black spots along central rib of needle
Rizosphaera needle scorch (*Rhizosphaera kalkhoffii*)	Like the above symptoms but with much smaller black spots
Spruce needle rust (*Chrysomyya abietis*)	Orange banding of needles
Grey mould (*Botrytis cinerea*)	New needles go brown and hang down; isolated patches of infection; not to be confused with frost damage
Sawfly (*Paristiphora abietina*)	Young needles eaten on one side only, rest turns red
Spruce bell moth (*Epinotia tedella*)	Leaf bases eaten and needles turn red-brown
Spruce bark beetle (*Ips typographus*)	Bark peeling and resin droplets on trunk, trees ultimately become red-brown and die

Table C5. Causes of Damage to Norway Spruce which can be Confused with 'Recent Forest Decline' (after Wellburn, 1988)

Various hypotheses have been postulated to explain forest decline, including poor forest management practises, ozone, ammonium/excess nitrogen deposition, halo-carbon initiated ultraviolet damage, and trace metal accumulation, which are outside the scope of this Working Paper apart from to note the potential uncertainty in deriving the necessary acid dose-response link. To the extent that SO_2 is causal, the main function of acid deposition in forest decline must be the effects via soils, in other words, soil leaching of essential minerals and increases in soil toxicity such as aluminium to root and mycorrhizal associations. Various reviews of this phenomenon exist in relation to forest dieback.

The integration of research effort is a logical response to complex pathway analysis such as that required to establish dose-response relationships between SO_2 and natural forests. Several large-scale studies have been undertaken, including the IIASA forest study, which involved a Europe-wide timber assessment and development of a consistent and formalised dynamic model, the PEMU/AIR model, incorporating a cumulative dose-response approach to modelling needle loss in pine stands under sulphur and nitrogen deposition in Germany. As with several other modelling approaches, this has two components, the pollutant transport and deposition model (PEMU) and the pine stand decline model (PSD). Notwithstanding the problems of models which incorporate a deterministic approach with many variables, the results from the model demonstrate that it is possible to model basic effects within reasonable validity limits. A pollutant-pathogen-dieback link has also been sought and the roles of viruses, fungi and pests in various dieback episodes and sites have been studied. However, this area is invariably complex. For example, factors affecting the occurrence of outbreaks of the Pine bark bug Aradus in Finland have been established (see Figure C8).

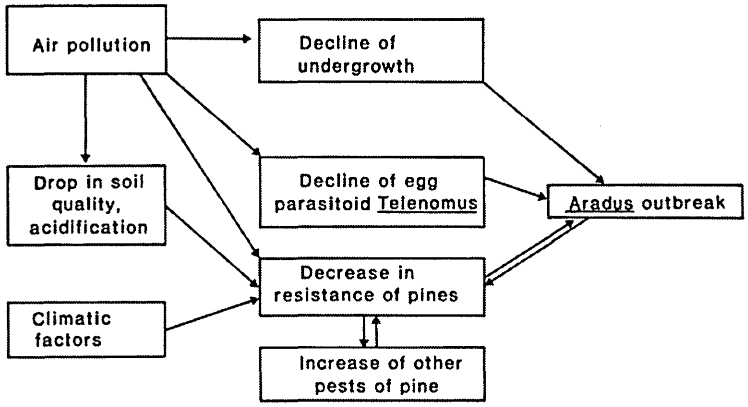


Figure C8. Factors Affecting the Occurrence of Outbreaks of the Pine Bark Bug Aradus in Finland (Heliovaara and Vaisanen, 1986, cited in Innes, 1987)

With specific regard to sulphur, it has been established that SO_2 treatment increases the susceptibility of Norway spruce to woolly aphid infestations (Keller, 1983, cited in Innes, 1987). Field observations have been combined with various pollutant exclusion and controlled exposure experiments under varying degrees of controlled conditions in order to try to establish dose-response relationships and inform model development. In one long-term study over 10 years, artificial additional loading of up to 20 g m^{-1} sulphur led to no significant difference in leaf biochemistry of mountain birch (Suomela et al, 1998). In another study, acid mist precipitation led to degradation of cuticular waxes and subsequent spring frosts brought an observed decline in needle water potential (Esch and Mengel, 1998). Broad-leaved tree species' leaves have also been shown to be susceptible to simulated acid rain, with Beech showing more predisposition to developing macroscopic necritic lesions than Holm oak (Paoletti, 1998). Observation studies of defoliation using satellite images and monitoring plots, compared to SO_2 concentrations, showed significant correlations, particularly during winter months (Šrámek, 1998).

Regarding root development, one study showed that Pinus pinaster seedlings develop root length markedly more slowly under increasingly acid conditions between pH 6.5

and 3.5. However, the most acid conditions also produced the highest root biomass due to thickening. It also confirmed that more acid conditions generated higher P, Fe and Al concentrations in roots (Arduini et al, 1998). Other work has demonstrated a link between soil acidity, base metal accumulation and needle strength in Scots pine (Sogn and Abrahamsen, 1998). As a result of such studies, continuous development is taking place in various aspects of forest decline modelling, providing better understanding of the variables involved and the interplay of dose-response relationships through integrated soil-plant models (for example, Augustin et al, 1998, Schall et al, 1998). Undoubtedly, gaps remain, particularly in the understanding of regulation of processes of nutrient uptake. However, model performance is good and improving. In general, this greater understanding has led to a recognition that sulphur-induced acid precipitation is only one factor in causing forest ecosystem stress, and indeed, that ozone is likely to be a more significant factor in many cases (for example, Miller et al, 1998).

A very large number of studies have been undertaken on the effects of sulphur on forest growth and a small sample of some of the results is presented in Table C6. Clearly, not all experiments give useful or accurate dose-response information, and gaps remain in dose-response relationships for various tree species and for various pollution scenarios. However, the vast majority of studies show that at relatively high concentrations above typically 50ppb, reductions in tree growth (for various demonstrable reasons) typically occur.

C3.4 Water Environment

The water environment includes all surface waters, groundwaters, and estuarine, sea and ocean waters. Estuarine and marine waters are the destination of most fluvial sulphur. However, they will not be considered in this review, except to note that sea spray and sea salt rain episodes can cause highly acidic flushes locally in terrestrial environments, and that SO_2 gaseous absorption is a major mechanism for direct transfer of sulphur from atmosphere to oceans. Groundwater, which is also a transport pathway for natural sinks and sources in rocks and pore spaces, is similarly not included in this review, which instead concentrates primarily on streams and lakes, and the ecosystems they support.

	Trigger/exposure level	Effect	Study/ref.
General Swedish coniferous forest	annual mean 0.002ppmSO_2	few percent average drop in growth rate	cited in anon, 1971.
General Ruhr Pine species	>0.07-0.08ppmSO_2	death	cited in anon, 1971.
General Ruhr Pine species	0.01ppmSO_2	slight damage	cited in anon, 1971.
General Ruhr Pine species	0.04ppmSO_2	medium-severe damage	cited in anon, 1971.
Tree species; Pinus Strobus (U.S.)	0.05 for one hour	some measurable damage	Dochinger and Seliskar, 1970
Cyanobacterial lichens	10$\mu g SO_2 m^{-3}$ (annual mean)	critical load	UN ECE, 1993
Forest ecosystems	20$\mu g SO_2 m^{-3}$ (annual mean)	critical load	UN ECE, 1993
Natural vegetation	20$\mu g SO_2 m^{-3}$ (annual mean)	critical load	UN ECE, 1993
Agricultural crops	30$\mu g SO_2 m^{-3}$ (annual mean)	critical load	UN ECE, 1993
General plant function	200-400ppb	"effects threshold" on photosynthesis	Howells, 1990
General plant function	>100ppb	"effects threshold" on stomatal function	Howells, 1990
General plant function	>40ppb	"effects threshold" on assimilation	Howells, 1990
General plant function	35-380ppb	"effects threshold" on respiration	Howells, 1990
Picea abies, Laix, in Germany	238ppm S measured concentration	Mg/Ca/K/Zn fertilisation showed Mg deficiency was occurring	cited in Legge and Krupa, 1990
Red spruce, USA	S content of poor foliage 0.12%, good foliage 0.10%	foliage condition related to S-level	cited in Legge and Krupa, 1990
Green ash, Paper birch, red pine, USA	SO_2 greenhouse fumigation experiment; various levels	Green ash growth not affected, Pine root growth lowered, varied with temperature.	cited in Legge and Krupa, 1990
Yellow poplar seedlings	SO_2, O_3, simulated acid rain in greenhouse experiment	Changes in growth rate and dynamics; O_3 more active at lower pH: SO_2 at higher pH.	cited in Legge and Krupa, 1990
Picea abies	10-week fumigation 25ppb SO_2	15% reduction in annual ring width	cited in TERG, 1988
Picea abies	10-week fumigation 50ppb SO_2	18-25% reduction in annual ring width	cited in TERG, 1988

Table C6. Summary of findings of some studies of acid precipitation on forest species

Note: am = annual mean

C3.4.1 Lakes and Streams

Hydrological systems, and specifically, surface waters, are of great importance in tracing pathways of environmental changes arising from sulphur. As a whole, hydrological systems provide the primary source of dynamism in acid circulation within the terrestrial environment. Since they provide various means and rates of transfer of acidifying pollutants, and they contain various elements, such as streams, lakes, groundwater flows at various levels, estuarine and marine water bodies, they also involve a complex and wide range of physical, chemical and biological conditions. Given this, it is not surprising that the earliest region-scale effects of transboundary air pollution were found in aquatic ecosystems, and that major programmes of study have been targeted at developing knowledge of cause-effect relationships, such as in the Surface Waters Acidification Programme (SWAP), which provided a major research input into the development of UK policy leading to the Large Plant Directive agreement and its associated wide programme of Flue Gas desulphurisation retrofitting in 1987. The SWAP programme established that progressive lake acidification over a century is common in northern Europe, and that there is convincing evidence of a clear causal link between anthropogenic atmospheric acidifying pollution (including SO_2) and acidification of surface waters (Mason, ed, 1990). The general picture is that acidification leads to lower biotic mass and diversity of both plant and animal species, although important variations to this exist. Several reviews of the evidence have been undertaken (for example, Legge and Krupa, 1990). Some contain useful indications of gaps in research and knowledge as well as what has been established and what is currently under investigation (for example, Howells, 1990).

Since the 1940s, fish population decline in areas receiving acid deposition have been documented and in some areas, over half of lakes with pH less than 5 are fish less. A 5000-lake survey in southern Norway showed that 1750 have lost their fish populations and a further 900 are seriously affected. In southern and central Sweden, fisheries damage is evident in 2500 lakes (UN ECE 1984). Crayfish were found to suffer more parasitism and impaired reproduction in an experimentally acidified lake (Gorham, 1998). Detailed accounts of biological effects of lake acidification have been presented (for example, Muniz, 1991), albeit without precise materials flows and dose-response calculations. However, there is "substantial agreement" in responses of both biogeochemical and the effects on biota of lower trophic levels (Gorham, 1998).

Invariably, water passes through soils and streams before reaching lakes. Stream acidity is a function of soil and base material metabolism and net atmospheric and water input, although the variables involved in these determinants are numerous. A conceptual model of hydrological factors influencing stream water acidity and aluminium levels in storm runoff events in upland Wales is shown in Figure C9. There are strong interrelationships between topography, soil type and vegetation, and runoff hydrochemistry. Therefore, acidification is related to catchment hydrology and geomorphology as well as acid inputs to the system, both in terms of timing and pathways taken by runoff waters. It is also related strongly to vegetation type; those streams draining conifer forests are often more acidic and contain higher aluminium levels than those draining equivalent moorland areas.

There are basically two main ways in which acidification of water can occur. One is by direct deposition of acidifying pollutants, the other is by water passing through an acidified environment before entering the affected hydrological system. For example, if water passes through an acidified soil it may become acidified and so transfer this acidity elsewhere in the hydrological system. Whether or not soil acidification leads to freshwater acidification depends upon the pathway the drainage water follows through or over the soil. If it penetrates below the near-surface acidified zone into higher pH layers beneath, then acidification is improbable, whereas if overland or near-surface flow over/through acidified soils occurs, water acidification will occur. Such hydrological conditions may typically only occur for short periods during storm events in upland areas, particularly with thin or impermeable soils and steep slopes.

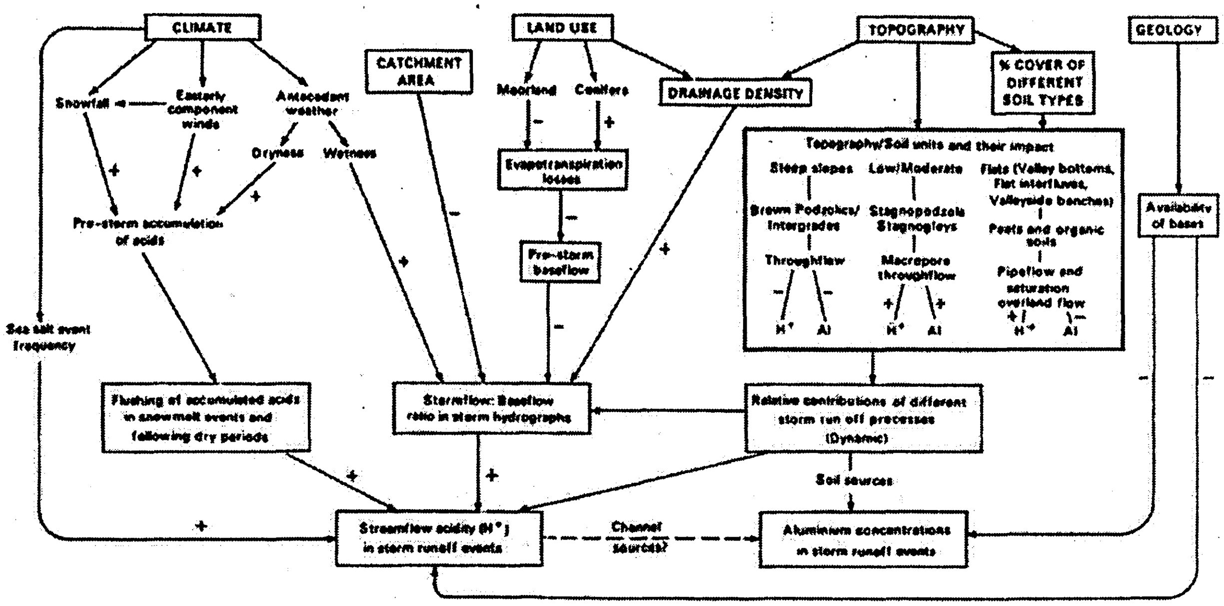


Figure C9. <u>Hydrological Factors Influencing Streamwater Acidity and Aluminium Levels in Storm Runoff Events in Upland Wales (Edwards et al, 1990)</u>

Note. + = direct relationship, - = inverse relationship).

The capacity of freshwater bodies to neutralise acidity (buffering capacity) is largely determined by bicarbonate content. High concentrations occur in hard waters, but below pH 5.4 bicarbonate is completely absent. There is a close relationship between various types of hardness and pH, such that three basic types of freshwater can be distinguished; acid, soft and hard. The very soft zone, with pH between 5.7-5.2 marks a transition, since below this, permanent acidity is marked by the elimination of many freshwater species. The link between acid deposition and acidification of poorly-buffered freshwaters is compelling, both theoretically and in experimental and observational reality. It has already been reported that sulphur can be temporarily precipitated in reducing conditions. A study of ombrotrophic bogs showed that there is a close relation between surface water pH and rainwater pH, and that buffering occurs due to the cation-exchange properties of the surrounding peat (Proctor and Maltby, 1998).

Freshwater primary producers such as algae have been observed to apparently benefit from acidification, possibly due to reduction in invertebrate grazing and/or reduced decomposition rates. Macrophyte biomass can be large in acid waters, though is usually variable. In general, floral composition shows a marked negative impact associated with declining pH, including, for example, declining phytoplankton and periphyton diversity. Acid tolerant taxa include Sphagnum and Juncas bulbosus among the macrophytes, peridinium and Gymnodinium among the phytoplanktons, and filamentous chlorophytes such as Mougentia or blue-green algae amongst the periphyton. Clearly, even alterations of biotic communities towards these acid-tolerant species could have consequences for animals which preferentially graze certain plants or are dependent upon them for habitat (Edwards et al, 1990).

Bacteria and fungi communities are also affected by acid effects, both directly and indirectly. Inorganic matter in solution or suspension is supplemented by large amounts of organic matter in the materials transfer from terrestrial to freshwater environments. Leaves and woody debris can be an important source of energy in upland streams, and are quickly colonised by freshwater fungi, which produce the necessary enzymes to degrade the material. Several studies have shown that plant litter decomposition is retarded in acid streams, and that this can be associated with reduced microbial populations.

Elsewhere in the food chain, acidification triggers biological changes at all levels in the aquatic ecosystem. In general large fish and plant species decline, while some mosses and filamentous algae species increase, occasionally forming impenetrable mats on lake beds. When pH reaches 5.5, negative effects upon life can be observed, while a pH of 5 is regarded as a critical level for the survival of most fish species (anon, 1971). The rate of fall of pH in a water body is a function of 3 factors; supply of acidifiers, presence of buffering actions, and residence time of water. The latter is critical, since a long residence time will inevitably lead to a higher cumulative effect of acid fall-out, particularly if the rate of acid supply exceeds the rate of buffering.

Residence time of water in a group of lakes used in a study in Sweden was found to be in the range 3-50 years (anon, 1971).

The primary cause of death to most fish is poisoning by toxic aluminium, which is mobilised in soils by acid fallout and them mobilised in solution. Aluminium affects gill function and causes a build-up of mucus which effectively leads to suffocation. Some fish species, such as brown Trout have high resistance to acidity, possibly because they have secure food source since they can feed on insects entering the stream from the surrounding terrestrial habitats.

Most species of mayflies, caddis flies, freshwater shrimps, limpets, snails and beetle larvae are absent from acidic water systems with pH below 5.4. Aluminium and heavy metals released by acidification of runoff waters are concentrated by many invertebrates and so enter food chains. Birds and other fish/invertebrate feeders such as otters therefore receive secondary doses of these. Bird numbers are reduced in many areas because of imperfect egg calcification which fail to produce live chicks (Wellburn, 1988). Prey rich in calcium such as fish and molluscs are rare below pH 5.7-6. The Dipper has been observed to require a larger territory (and therefore occurs at lower populations) as pH decreases. Food chain effects also apply to all higher lifeforms, so that toxicity build-up of acid-related releases of heavy metals can appear in fish, birds, amphibians and mammals. The latter two groups are potentially critical, and amphibians should not be overlooked in this regard. In the Hubbard Brook forest of Northeast USA, amphibian biomass was found to be twice that of birds and equal that of small mammals. Otters are an important large mammal in remote streams and their presence appears to be primarily controlled in terms of acidification by food supply reductions. They were found to be absent from acidic headwaters of the Severn, but present along circumneutral adjacent streams (Mason and Macdonald, 1987).

Recent research on acidification within hydrological systems demonstrates that model-based approaches are becoming increasingly sophisticated, valid and accurate in predicting a wider range of variables. For example, in one study, four watershed acidification models were applied to compute hydrogen ion, alkalinity and sulphate concentrations in Turkey Lakes watershed, Canada (Bobba et al, 2000). While errors in data and imperfect knowledge of variables controlling water quality parameters are acknowledged, further analytical techniques are demonstrated to allow assessment of uncertainty in the models. In the same paper, detailed discussion is included of problems of complexity and error in modelling and data collection, and recommendations are made as to how improved databases can be established.

The water cycle is clearly a critical element in driving the sulphur cycle. Modelling of water acidification is much advanced over the past two decades. For example, one study used the Model of Acidification of groundwater in catchments (MAGIC) along with dispersion predictions from the Hull Acid Rain Model (HARM) to predict expected recovery of acid waters, arising from the implementation of the Second Sulphur Protocol (Jenkins et al, 1998). Incidentally, there is some evidence that, following the removal of excessive acid deposition, river catchments are able to recover quickly (ETSU, 1998). However, other evidence does not support this. Dynamic models have been successfully applied to several catchments and model types have been reviewed (Whitehead et al, 1990). Remote sensing and GIS systems are also increasingly being envisaged to allow greater validity to be attained in modelling this complex issue (Li and Tang, 1998). In best practice, where “gaps” exist, risk assessment is applied to assess worst case and likely case scenarios, rather than not considering unknowns at all.

C3.5 Managed Terrestrial Environment

Land use has a potentially major impact on the acidity of drainage waters. Managed terrestrial environments therefore provide an opportunity to regulate water acidity, and knowledge of the cause effect relationship is important in informing the optimum land use and land management methods. To give two stark examples, the application of elemental sulphur to crops to increase yield, or the use of irrigation water containing sulphur, will clearly effectively increase the effect of acid deposition. Thus, the managed terrestrial environment is differentiated from the natural terrestrial environment by the level of human interference in land use/vegetation coverage. If vegetation is essentially cultivated, then the environment is managed. Therefore, this subsystem theoretically includes commercial planted forests and a wide range of perennial and annual crops, indeed, all managed land uses which occur within the zone of deposition. The review presented below concentrates on crops, meaning annual cultivated crops and grasslands.

C3.5.1 Crops

The health of a crop is governed by its interaction with various physical, chemical and biological climatology, which itself is influenced by management practices. It is therefore difficult to isolate the effects of atmospheric deposition of sulphur compounds arising from anthropocentric emissions. Various studies have been undertaken, including fumigation experiments both indoors and outdoors, filtration experiments to remove air particles, and field trials with point source emitters, in an attempt to establish dose-response relationships, with considerable success. Most studies indicate acid deposition does no visible damage to crop foliage at present rural deposition rates (Heck, 1989). However, early studies indicated that some crop plants such as peas show a progressive yield decrease along a transect from rural to urban areas, and this effect was tentatively ascribed to O_3, SO_2 and NO_x. Typically, measurable damage occurs in the range 0.1-0.5ppm SO_2.

Agricultural soils are often not at high risk of acidification because agricultural practices such as liming are established and designed to counteract it. However, liming is not always undertaken, and has been declining as a practice in recent decades, particularly in the UK. A range of crops including peas, beans and spinach have been found to decrease in the presence of long term elevated levels of SO_2, ozone and oxides of nitrogen in the air. Experiments have shown that yields are reduced without visible damage by exposure to SO_2 alone at concentrations above 40ppb (TERG, 1988). Other crops, such as broccoli and wheat were unaffected. However, cereals exposed to 40ppb SO_2 have indicated reduced growth of barley. Table C7 contains a summary of findings of some studies of acid precipitation on crop species.

	Trigger/exposure level	Effect	Study/ref.
General	0.1-0.1ppmSO_2	"measurable damage"	various
Peas, beans, spinach	40ppb	invisible damage, yield reduction	TERG, 1988
Agricultural crops	30$\mu g SO_2 m^{-3}$ (annual mean)	critical load	UN ECE, 1993
"Sensitive" crops	0.l9ppm (8h exposure)	foliar injury threshold	cited in Legge and Krupa, 1990
Intermediate crops	0.24ppm (8h exposure)	foliar injury threshold	cited in Legge and Krupa, 1990
Resistant crops	0.49pprn (8h exposure)	foliar injury threshold	cited in Legge and Krupa, 1990
Soybean	2.8h x 10 exposures of 300ppb	6% yield loss	Irving and Miller, 1984 cited in NAPAP, 1987
Winter wheat/potato	190ppb for 10-13% of growing season	10% yield loss	cited in Legge and Krupa, 1990
Snap bean	0.15ppmSO_2 4h, 3x per week for 4 weeks	9% yield loss against control	cited in Legge and Krupa, 1990

Table C7. Summary of Findings of Some Studies of Acid Precipitation on Crop Species

Note: this is a small sample: many studies have been conducted and reviews assessed for example, anon, 1987.

Wheat is particularly important because it is a major UK arable crop, particularly in the eastern counties, most likely to be affected by point source sulphur-bearing stack emissions. The dose-response effects have been well studied and a sample of resultant effect/responses on wheat growth or production has been compiled and is reproduced in Table C8.

Effect/ Response	Average SO_2 Conc. (ppb)	Effect Parameter
threshold	4.5	Winter wheat yield reduction
threshold	7.7	Spring wheat yield reduction
- 1%	10	Spring wheat yield reduction
- 1.4%	20	Spring wheat yield reduction
-11.7%	51	Spring wheat yield reduction
-26.6%	83	Spring wheat yield reduction
-16%	100	Winter wheat seed wt. reduct.
-33%	130	Winter wheat seed wt. reduct.
-36%	141	Spring wheat yield reduct.
-30%	200	Yield reduction
-1.00%	241	Threshold for leaf destruct. over entire season (1008 h)
-26.60%	440	Spring wheat yield effect
-30.30%	600	Thatcher(cv.)Hard Red Spring wheat dry wt. loss with 100 hour exposure
-20.00%	1315	Wheat yield decrease
-4.00%	1470	Percent foliar injury
-6.00%	1470	Percent reduction seed wt.

Table C8. <u>Responses of Wheat Growth/Production to SO_2 (Legge and Krupa, 1990)</u>

A more recent study using a fumigation experiment also showed increasing yield loss with SO_2 dose, although this did not occur at typical UK SO_2 atmospheric levels, and furthermore, no consistent dose-response curve or linear response was established, and it was noted that other variables such as weather and pathogen occurrence were important (McLeod et al, 1991).

The productivity of grass species in pasture lands is also affected by SO_2 (for example, Bell, 1985). However, there is little evidence that grasslands (in the UK at least) are directly damaged by current levels of acid precipitation. Indeed, some areas with low sulphur deposition (for example, northern Scotland and west Wales) are sulphur-deficient and so may benefit from fertilisation effects. Furthermore, there is evidence that grasses develop SO_2 tolerance rapidly, for example, for concentrations of 37-56ppb this has developed over as little as 4 years, while a decline in SO_2 deposition to

20-30ppb causes disappearance of this tolerance (TERG, 1988). Despite this, controlled exposure experiments have suggested that above 30ppb, SO_2 may adversely affect grass growth without visible damage. Dose-response relationships have been established experimentally, although they are not always replicable in the field, as other variables appear to operate in many cases. Cumulative and synergistic effects have been indicated.

C3.6 Built environment

Strictly, the built environment is a sub-system within the managed terrestrial environment. The bulk of work in the area has been conducted on structural building materials such as stone, concrete, and metals. Acid attack on other finishes, such as acrylic-melamine coatings, have been studied (for example, Mori et al, 1999), although these materials are not considered further in this review. Instead, the following concentrates primarily on calcareous stone and metals commonly used in buildings.

C3.6.1 Buildings and Materials

Structural building materials are attacked by S, particularly if they are based on carbonate formulations, as in limestone, concrete, and calcite-cemented sandstones. However, metals and painted surfaces are also attacked by S-deposition. The proficiency of many paints is reduced. Variables include atmospheric concentrations, strength and direction of wind and rain intensity, degree of exposure to wind and rain, time-of-wetness of the surface and the natural structure and reactivity of different materials.

A study carried out in the 1980s found that currently available damage functions did not perform well in predicting materials damage under various pollution scenarios. However, subsequent work has involved historical studies of pollution and damage rates, laboratory experiments, and field experiments, and these together have produced some success in modelling and deriving dose-response relationships for calcareous stone (Lipfert, 1989), while other work has identified dose-response functions for various building materials (Butlin et al, 1994). This work has formed part of the National Materials Exposure Programme in the UK, which has set out to establish damage functions (which involves measuring does-response relationships) for atmospheric acid on building materials. Although the primary aim of this work involves the setting of action levels (which are based on NOEL - no observed effect level) and environmental assessment levels (EALs - the upper level of tolerable damage), for policy purposes, logically, these two levels could be used directly in environmental change and human consequence assessment (see Figure C10). Dose-response functions are established for the effect of SO_2 on calcareous stone, steel and zinc, and it is noted that observed synergistic effects between SO_2 and NO_x are restricted to laboratory experiments.

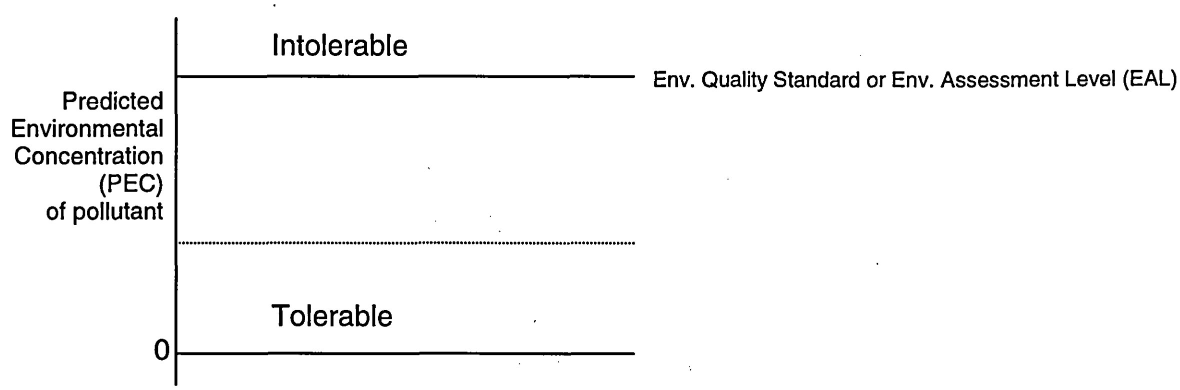


Figure C10. Tolerability of releases leading to building damage (adapted from Butlin et al, 1994)

Note: NOEL describes the lower boundary below which no changes are likely: EAL describes a point at which two distinct categories of human consequence exist.

With building materials, unlike many ecosystems, degradation is irreversible and is a natural process. It is not generally catastrophic, but rates of damage are aggravated by acid attack. Therefore, the measurement of environmental change is most appropriately seen as the change in damage rate. However, the process is not simple for SO_2 since several processes may be operating. Indeed, carbon dioxide undoubtedly plays a major role in natural weathering of calcareous surfaces. SO_2 is not necessarily even the dominant factor; one study estimates that halving SO_2 air concentrations would reduce dissolution rates by "only 15%" (Cooke and Gibbs, 1994).

Calcareous stone such as limestone is an important local structural material, and has been for many hundreds of years, making it a common material in old buildings. Sandstones and other siliceous-based sedimentary rocks (and metamorphic, such as marble) are also susceptible to carbonate weathering, where they have a calcite cement. Rainfall dissolves calcite at rates depending on the dissolved CO_2 content, with a modest increase in dissolution arising from rain acidification at typical urban levels. However, wet surfaces are good absorbers of SO_2, until they become saturated with gypsum (calcium sulphate), although direct rain washing can also remove gypsum and renew surfaces for further wet acid attack. Stone porosity is important in salt penetration and freeze-thaw, and SO_2 deposits can influence mechanical damage as well as dissolution.

Sulphate crusts are a fairly common feature of old limestone, and their formation is a complex process, involving particulates such as soot. They form in rain-sheltered zones, are rich in hydrated gypsum, and are formed by dry deposition of sulphur dioxide into pores of moist stone. Also, configuration effects are important on buildings or sculptures with complex patterns or relief, since they may suffer increased material loss due to larger surface area and varying SO_2 deposition velocities. Organic species such as algae, lichen, and bacteria may also have a role in stone erosion.

While local variations in weathering rates have been recorded, there is generally considerable similarity between rates predicted by a range of functions for atmospheric

SO_2 concentrations currently found in the UK (Butlin et al, 1994). The differences are typically due to the fact that different materials weather at different rates and regional variations in, for example, crystalline density will affect rates locally.
Detailed reviews of the effects of acid attack on metallic surfaces can be found elsewhere (for example, BERG, 1989, Butlin et al, 1994), although this issue warrants brief consideration here. Steel and zinc surfaces are perhaps the most important metallic materials for consideration, since they have been shown to corrode up to ten times faster in urban air than outside urban areas, although this clearly does not allow separation of SO_2 effects from other urban pollutants. However, these materials are also the most widespread metallic surfaces. Ferrous metals are potentially particularly significant in their corrosion response to SO_2 at current anthropogenic levels. However, CO_2 rather than SO_2 is the main actor in corrosion of reinforced concrete. Indeed, the former is thought not to play a significant role (Dunderdale, ed, 1990).

Atmospheric corrosion typically occurs due to oxidation during dry conditions in the presence of a corrosion stimulating agent, followed by removal of the oxidised layer during wet conditions. Rainwater contaminants such as sulphur contributes to the formation of corrosive solutions. Relative humidity is important, as is the total time for which the material is wetted. Highest corrosion rates occur when the water layer is thinnest, immediately before drying. Although less work has been done on establishing dose-response relationships for metals compared to that for calcareous stone, these have been reliably produced, for example, for a range of commonly used metals (see Table C9).

The establishment of an inventory of buildings is difficult, although clearly not as complex as an inventory of the natural environment. Materials vary locally and change over time. However, to estimate damage, the total area of exposed materials must be known. This can be done by an inventory/census approach where each building is examined, or by a probability distribution approach where areas of exposed materials are aggregated to provide a probability function of the exposed area per unit area. The former method would be particularly applicable for unique buildings such as heritage sites. A detailed review has been undertaken elsewhere, and dose-response functions have been devised for a wide range of building materials, including calcareous stone and metals (Butlin et al, 1994). The knowledge is therefore established to allow a QLOS approach to be successfully taken with regard to buildings attack by SO_2 emissions from a point source stack emission.

Metal	Predicted corrosion rate in SO_2 free atmosphere (μm)	SO_2 concentration predicted to double corrosion rate (μg m^{-3})
Ferrous metals	25	40-100
Copper	1	40-50
Zinc and galvanised steel	1	100
Aluminium	0.03	20-50

Table C9. Dose-Response Relationships for a Range of Commonly Used Metals (Butlin et al, 1994)

C3.7 Human Body

The transfer of sulphur through the human body may occur in a number of ways. Direct respiration of acidified air may provide a pathway into the respiratory systems, from which it may enter other parts of the body. Ingestion may occur, for example, through drinking of acidified sulphur-bearing water. However, potential effects are direct and indirect, and the latter includes food chain effects, such as heavy metal build-up in food/water due to increased leaching of them from rocks/soils, and assimilation into water supplies, crops, animals, which are then ingested, worn, used, etc., by humans. Clearly, these pathways involve passage through the natural environment. The following review is largely limited to direct respiration of sulphur, although some consideration is also given to indirect effects associated with ingestion of heavy metals through sulphur-based acidification of water supplies.

In general, rather than dealing with acid rain, where no direct effects on humans are known, the most likely source of direct atmospheric acid impact is in the case of episodic sulphur-bearing mists, fogs, aerosol or smogs. Knowledge of health hazards for normal urban exposures is incomplete. However, regarding aerosols, particle/droplet size is important, and sizes below 0.8 micrometres across are considered particularly dangerous. Thus smaller amounts of sulphur in smaller sized droplets may be more damaging to lung function than larger ones with larger total amounts of sulphur in the air. This is because the smaller the particle, the greater the chance of it reaching the most sensitive parts of the respiratory system (see Figure C11). Particles above PM_{10} are unlikely to penetrate beyond the trachea, while 90% of particles greater than 2μm are caught by the mucus layer and expelled through the action of the cilia. Some are swallowed, which transfers the sulphur to the respiratory system, with potential further health effects. Small particles below 2μm can penetrate deep into the lung, where some will be deposited. Acid particles will then reduce pH of airway fluids as they dissolve. Particles of all sizes can be removed by macrophages (scavenger cells).

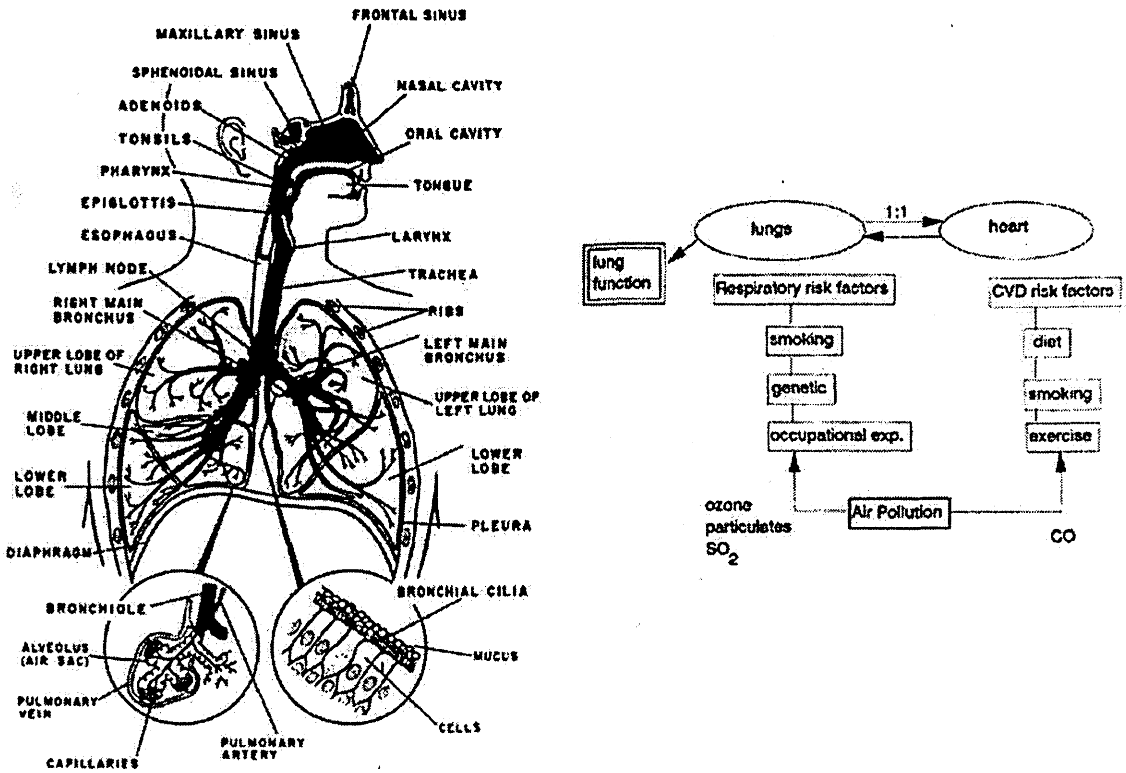


Figure C11. Human Respiratory System And (Right) Interrelationships in the Cardio-Respiratory System (after Lipfert, 1994)

The actual potential effects of small sulphur-bearing particles on the respiratory system include alterations to respiratory mechanics, reductions in the supply of oxygen (although this is likely to arise not from sulphur but from CO and O_3), reduced resistance to infection (although effects on macrophages and cilary action are not proven), ageing and chronic disease of the lung (although direct effects have not been proven) and lung cancer. Regarding the latter, no causal link is established, while it is for asbestos, smoking and light hydrocarbons such as benzo(alpha)pyrene. Lung cancer death rates were higher in urban than rural areas, but most investigators discount the role of urban air pollution at present community levels. Contributory effects in aggravating disease are certainly likely, but the effects of smoking are so strong that it has proven difficult in the past to determine relative contributions based on epidemiological data. However, there is enough evidence that air pollution can cause loss of lung function, and that loss of lung function can cause increased mortality, to indicate that air pollution can cause increased mortality. The mechanism for the linkage may well be the lung's role as the primary organ of defence against airborne toxins; a compromised lung implies compromised defences (Lipfert, 1994).

Standards have been set on the basis of avoiding short term increases in mortality and illness, although short-term acute health effects are possible during unusual weather conditions. The best known case of increased mortality in a short-term acute episode is in connection with a high level of pollution in London in December 1952. A maximum daily concentration of 1.3ppm was recorded, and over a two week period mortality rose by 70%, corresponding to 4000 extra deaths.

By the 1970s, a link had been established between respiratory disease and particulate and/or SO_2 air pollution, but there remained disagreement as to the level of pollution that would "significantly" affect human health. One review (Holland et al, 1979), concluded that relatively high levels of pollutants cause known hazards to human health, but that effects of lower (more "normal") anthropogenic pollution levels were inconclusive. Other research has suggested that human health may be adversely affected by particulate pollution, even at relatively low concentrations (Shy, 1979, Ware et al, 1981). By the late 1980s, a clearer picture of dose-response was occurring, with production of information such as that shown in Table C10.

Concentration (ppmv)	Period	Effect
0.03–0.5	continuous	Condition of bronchitic patients worsened
0.3–1	20 sec	Brain activity changed
0.5–1.4	1 min	Odour perceived
0.3–1.5	15 min	Increased eye sensitivity
1–5	30 min	Increased lung airway resistance, sense of smell lost
1.6–5	less than 6 hours	Constriction of nasal and lung passage
5–20	more than 6 hours	Lung damage reversible if exposure ceases
20 upwards	more than 6 hours	Water-logging of lung passageways and tissues, eventually leading to paralysis and/or death

Table C10. Doses and Effects on Human Health of SO_2 (Wellburn, 1988)

Note: concentrations are lower if aerosols, particulates or other pollutants are present.

More recent epidemiological studies have tended to show that morbidity and mortality effects are associated with low levels of particulate pollution too. An example is the major "Clara County" study (Fairley, 1990). A review has established coherence of effects across a range of health outcomes and a consistency of effects across independent studies with different investigators in different settings (Dockery and Pope III, 1994). This study also recognises that data and results of studies which can contribute to understanding of dose-response relationships should be reported transparently and in a form which allows them to be readily compared with other, previous and future investigations.

Nevertheless, although the evidence is now compelling, the most patchy area of sulphur effects on human health remains the effect of long exposures at low concentrations. In excess of 5ppm ($14mgSO_2/m^3$), irritation of air passages arises. However, cases vary, and susceptible people can be adversely affected by 1-2ppm, while others are apparently unaffected at concentrations of over 10ppm. While SO_2 in gaseous form is unlikely to be directly responsible, small droplets or particles containing sulphate or sulphuric acid are.

For longer term increased incidence of illness, annual average SO_2 concentrations of 0.03-0.04ppm has been suggested as sufficient to produce adverse health effects amongst vulnerable populations. However, the point has been made that many different indices are used to describe adverse health effects, with little inter-comparison between, for example, studies of long term effects and short term episodic effects, indicating the immaturity and complexity involved in this field, of study (Bates, 1992).

The UK Dept of Health Committee on the Medical Effects of Air Pollutants reported in 1995 (DOH, 1995), that non-biological particles can lead to changes in lung function, days in restricted activity, increased hospital admissions and mortality. However, it suggested policy should be based on PM_{10} measurements rather than particular particles within the respirable range, due to lack of clear evidence on the relative effects of aerosols derived from SO_2 and NO_x and particles from other sources. It also stopped short of identifying a causal link between particulate and health effects, but suggested the association should be prudently treated as causal. However, it also states that, although increased morbidity and mortality occurs amongst elderly and infirm, there is no evidence that healthy individuals are likely to experience acute health effects as a result of exposure to concentrations of particles found in ambient air in the UK. This report was based on particles of a size able to remain in suspension in air for hours or days, implying an aerodynamic diameter of 10-15 micrometres, which corresponds closely with the upper limit of material capable of entering the respiratory tract.

One of the ongoing problems of establishing respiratory dose-response to SO_2 bearing particulates is the inevitable cocktail of airborne particulate matter and the complexity of the respiratory system, which provides possibilities for various types of materials and damage to operate cumulatively or synergistically within the system. One review of epidemiological evidence concludes that SO_2 is considerably less linked to respiratory symptoms than particulate material, particularly very small particles much smaller than the 10 micrometer standard for respirable particulates. This does not mean that SO_2 is not linked to respiratory symptoms, for in many studies, it has been unequivocally associated. For example, in one review of 16 studies on acute effects of particulate pollution as measured by health care service use and other restricted activity indicators, 6 showed an association with ambient SO_2 or sulphate levels (Pope et al, 1995). However, considerably less correlation was found between these pollutants and respiratory morbidity and mortality, where the indicated link with particulates is significantly stronger.

Another route for sulphur or SO_2-related materials and energy into the human body is by ingestion. Most work in this area has been concentrated on drinking water. Evidence of contaminated groundwater, which may be exacerbated by acid deposition, has been reported in southern Sweden, parts of Ontario, Canada, and the Adirondack region in New York (Tolba, 1983), and this clearly involves the release of toxic substances through the water environment and thence into potable drinking water supplies. The direct effect on health of pH of drinking water has been described as "impossible" to ascertain, because of effects of pH on other aspects of water quality, for example, the corrosion and dissolution of pipework, including lead. The continued use of pipework for potable water supplies based on lead, or using galvanised fittings or lead-bearing solder is a major concern, particularly in areas where low pH water is concerned, since concern over linkage between lead ingestion and child development has lead to steadily reducing critical levels. It is anticipated that a level of 10 microgrammes per decilitre may be necessary as a safe level in order to protect foetuses and young children. However, the most likely hazard from alterations to public water supply are increased levels of aluminium leached from soils. A rare bone-wasting disease is associated with high levels of aluminium in drinking water (above 1000ppm), but another major disease attributed to aluminium in natural waters is Alzheimer's.

New methods of assessing effects of acid pollution on health are being developed, for example, through the SAVIAH programme, which utilises GIS-based information in epidemiological investigations of linkages (for example, Pikhart et al, 1997). The APHEA project sought to link SO_2 and particulates with mortality (Katsouyanni et al, 1997). It shows that, despite the uncertainty of long term effects, today's relatively low levels of SO_2 and particulates still have detectable short-term effects on health. More generally, another recent study reports an association between air pollution and daily consultations for asthma and other lower respiratory disease in London (Hajat et al, 1999), with the most significantly associated receptor group being children, and the most important pollutants being NO_2, CO and SO_2.

A major review and data sourcebook of long term community health and air pollution provides a picture of current levels of air pollution (in the western world) as being contributory to health effects rather than primary causal factors, although there is some evidence for chronic effects at higher pollution levels of the past (Lipfert, 1994). Classical toxicology was for a long time unable to determine causal links because it was based on known toxicities established using animal experiments, and therefore, when episodic air pollution disasters occurred with toxicity levels well below known levels, other apparent solutions to the problem such as synergistic effects were sought. However, the emphasis on population toxicity failed to identify that a human population actually consists of a wide range of health states from healthy to extremely unhealthy, or to put it more accurately as far as contributory sulphur-induced effects are concerned, unsusceptible and susceptible to aggravation of current health (respiratory) problems. The appropriate QLOS approach to this is therefore to establish this contribution and the additional aggregate response attributable to the anthropogenic sulphur portion for the various human consequence response categories.

In conclusion, dose response curves have been established which, for mortality and hospitalisation, tend to be linear or log-linear over modest ranges of pollution levels, with no sign of thresholds that were central to classical toxicology. For SO_2, logarithmic transforms fit, suggesting that the response tapers off at higher concentration levels (Lipfert, 1994). This is counterintuitive to classical toxicology, and the explanation lies in the spread of population response states. Parallels are found here with studies of heat-wave mortality, where heat wave deaths have been shown to be much higher in the first of 4 consecutive heat wave summers, despite temperature profiles showing higher temperatures in subsequent years. The explanation is that the later years' populations have less susceptible people in them.

C4. Pathway Analysis Method Discussion

At this point, the human body sub-system is selected to illustrate the application of current knowledge to pathway Coding, and compilation of the Pathway and Environmental Change Inventories. The selection is made since it is a good example of a complex sub-system with a range of pathways and environmental changes. A simplified material flow diagram for the demonstration subsystem is given in Figure C12. As this shows, the vast majority of sulphur is likely to eventually find its way into the kidney system and be excreted in solution in urine, whereupon it will continue its path through the sulphur cycle, by being eventually discharged into marine waters.

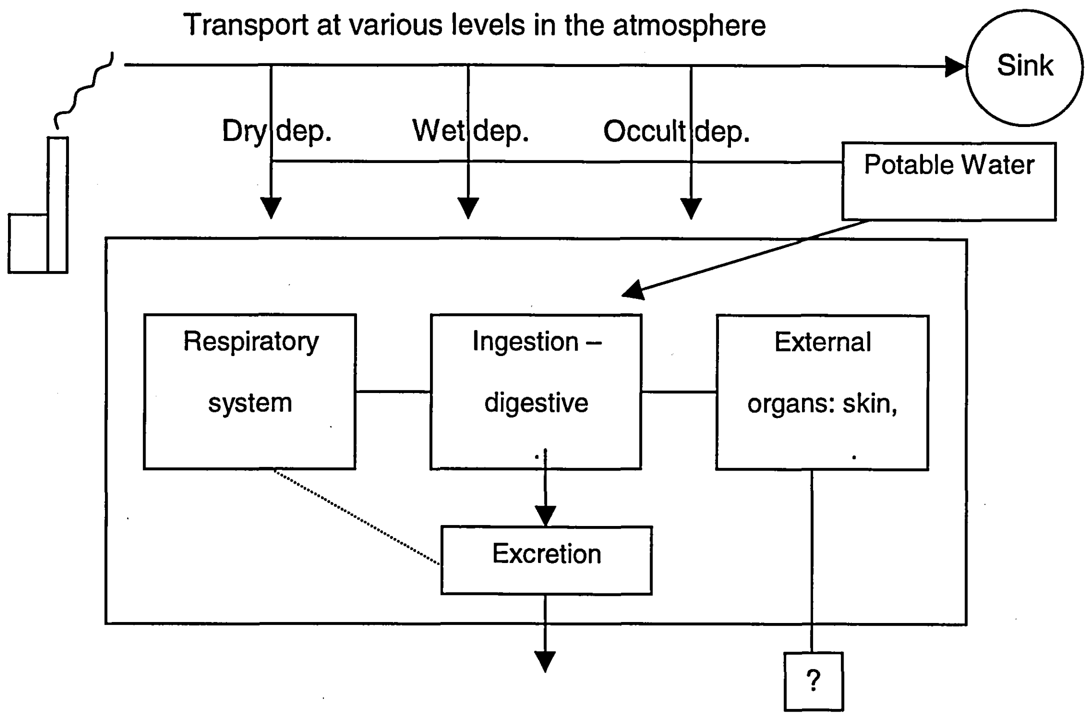


Figure C12. A Simplified Material Flow Diagram for the Human Body (Demonstration) Subsystem

Coding is undertaken systematically for each destination. For example, all first destination Codes may be common or similar to reflect the fact that all sulphur which ends up in the human body is first emitted to the atmosphere. Subsequent Codes may split according to which organs are involved (i.e. which subsystem within the human body). One pathway may therefore be Coded ATM/HB/RS/LU/KI/UR/BED/R/MARS, showing that the sulphur passes through the atmosphere (medium level) [ATM] directly to the human body [HB], where it enters the respiratory system [RS], and specifically, the lungs [LU]. Then, it is eventually mobilised in solution and reaches the kidney [KI], where it is incorporated into the urine stream and excreted [UR]. On leaving the human body subsystem, tracing must continue, and in this example it passes into the built environment drainage system [BED], into a river [R], and thence into a marine sink [MARS]. In this example, the Codes are generic, but in a real situation, it may be necessary to generate site specific Codes. For example, a particular river, or a particular stretch thereof. Any physical space may be assigned a Code, to suit the pathway analyst. However, all Codes must be assigned uniquely, although clearly the same specific Code can be used for several RIOs or pathways, provided that specific subsystem is being referred to. Thus, as pathway identification progresses, the cumulative Codes for each item become longer and more diverse. Each is unique, and contains detailed information about the route of the RIO, thus ensuring transparency and traceability by people other than the pathway analyst. A quantity will also be assigned to each pathway destination. Unknown quantities or apparent losses/gaps also need to be stated.

Every unique pathway-RIO combination forms a separate entry on the pathway inventory, which is the standard data format for the transfer of data from pathway analysis to environmental change assessment. The following rules must be adhered to in compiling the pathway inventory (after Horne, 2000a):

Pathway Inventory Rule 1. RIOs are not mixed (unless they are exactly alike, i.e. share the same Code);

Pathway Inventory Rule 2. Quantities are explicitly stated for each entry (i.e. for each pathway destination point);

Pathway Inventory Rule 3. Any unknowns, gaps or data losses (for example, as indicated by mass/energy balance accounting) are included as separate pathway inventory entries;

Pathway Inventory Rule 4. Any approximations of data are marked as a range or as an explicitly stated approximation with an indication of likely error range;

Pathway Inventory Rule 5. All pathway inventory entries for a given RIO are represented by a unique Code combination.

Thus, unknowns are entered as explicit quantities and at specific destinations, with unique Codes. For example, if there an apparent quantity loss is established between ATM/HB/RS/LU and subsequent destinations on the pathway then, in addition to all subsequent known destinations being entered on the pathway inventory, a Code ATM/HB/RS/LU/? is added, with the quantified apparent loss entered. Where estimations or uncertainties occur in pathway inventory quantity data, caveats must be attached to the appropriate entries as indicated in the rules above.

An environmental change is any shift in system equilibrium caused by the RIO energy/matter. This step is therefore dominated by the need to establish a direct or causal (indirect) link and an accurate dose-response relationship between the RIO and every possible environmental change at each pathway destination and pathway. There may be several environmental changes from any given pathway inventory item. In every case, it is essential that the following rules are adhered to (after Horne, 2000a):

Environmental Change Rule 1. Each environmental change is defined as a quantified shift in equilibrium of an inorganic system, or a quantified change in biomass, health or population of an organic system, as resulting from a given RIO or group of RIOs (see Rule 6);

Environmental Change Rule 2. Where there is some doubt as to the precise nature or existence of a dose-response relationship, the burden of proof must lie with disproving the link rather than proving it, in accordance with the precautionary principle. A reasonable level of likelihood is sufficient;

Environmental Change Rule 3. Pathway inventory entries are not summed (unless they act exactly alike, when the summed environmental change should be transparently attributed to the appropriate pathway inventory entries);

Environmental Change Rule 4. There are two principal types of environmental changes, those caused by direct physical contact with a pathway inventory item, and those which arise elsewhere as a result of a causal link - the latter require particular attention as they are more easily overlooked in environmental change analysis. Otherwise, indirect environmental changes are treated no differently from direct ones;

Environmental Change Rule 5. Any unknowns, gaps and approximations of data are included in all estimates (i.e. caveats must be applied and sensitivity tests undertaken, with ranges of uncertainty being provided along with a central estimate);

Environmental Change Rule 6. All environmental changes must be presented in site/type specific groups, and each examined and compared systematically to other group entries and to other groups for potential system-level, synergistic, cumulative, or threshold effects, including between different RIOs. Where present, the combined environmental changes should be adopted;

Environmental Change Rule 7. Every item on the environmental change inventory should retain its unique pathway inventory Code for transparency and traceability purposes (where inventory items have been summed to assess cumulative/synergistic effects, the Code should indicate the range of pathway inventory items it represents).

In order for summing of pathway inventory entries to occur, the environmental change concerned must be indistinguishable. For example, an alteration to kidney efficiency may be indifferent to whether the sulphur arrived at the kidney via the lungs or the digestive system. In such a case, the environmental change "loss of kidney function" could be summed for both pathway inventory entries: The need to search actively for cumulative changes is paramount, both within RIOs and across RIOs. Once the environmental change assessment exercise is complete, the data transfer to the human consequence assessment step can take place. The data required are the pathway inventory data (quantity and Code) and a short environmental change descriptor accompanied by quantity of change (there may be several quantities, for example, time, intensity, likelihood, etc.), whether direct or indirect, and whether indirect potential results of it have been checked for (since, in theory, all environmental changes can have knock-on effects within their respective systems).

Descriptors of environmental changes and human consequences provide clarity and will help lay-people viewing the inventory to immediately understand the entry (in accordance with transparency and accessibility requirements). Therefore, these are central to the data presentation tables, a sample of which is shown in Table C11.

Pathway Inventory for Pathway A1/B3/B3RS/	RIO quantity	Code	Causal links checked?	Causal change? (If no, it is a direct physical change)	Environmental Change Inventory	Quantities
Nasal Cavity		/B3NC	No	No	mucus generation	
Pharynx		/B3Ph	No	No	?? (no change identified)	
Oesophagus		/B3Oe	No	No	?? (no change identified)	
Trachea		/B3Tr	No	No	?? (no change identified)	
Bronchial and Lung activity		/B3Lu	No	No	mortality	
		/B3Lu	No	No	morbidity: long term bronchial illness/lung cancer	
		/B3Lu	No	No	episodic short term bronchial problems	
??? (gap = total A1/B3/RS/ - A1/B3/B3RS/B3NC + A1/B3/B3RS/B3Ph + A1/B3/B3RS/B3Oe + A1/B3/B3RS/B3Tr + A1/B3/B3RS/B3Lu)		/B3Lu/?	No	No	?? (gap and no change identified)	

Key:
?? Potential direct environmental change missing/unaccounted for
??? Possible pathway missing

Table C11. Pathway Inventory and Environmental Changes Data Presentation

Note: ATM/HB/RS/ is pathway Coding, denoting that the sulphur has passed into the atmosphere mixed layer, then directly to the human respiratorv system. Further Codes are added to this stem as shown in column 3.

Note: Causal links have not been investigated.

Note: Quantities columns are indicated for completeness: quantities would be entered in a full analysis. Note that where a pathway destination leads to more than one environmental change total RIO quantity should be stated for the inventory destination.

Table C11 should be seen as a small part of a much larger, multi-layered table containing information in separate steps and on many different subsystems and RIOs. The most appropriate means of portraying this information is in spreadsheet form, linked to spatial information on a GIS base, such as modelling and dose quantities, etc.

Integration in such a way minimises the errors and maximises the efficiency of the exercise.

Coding is of critical importance since each human consequence must be uniquely accounted for, transparent, and traceable to a given pathway destination and RIO quantity. The double question marks denote areas where pathway inventory production has identified that there are potential direct effects which have not been considered in the review. Therefore, this potential remains unknown, and should be Coded and identified, and, if possible, quantities of RIO applicable should be entered onto the appropriate part of the inventory. Single question marks show where indirect or direct data are missing. Quantities data would be required in a full analysis (or approximations, which would be marked as such). Thus, the presentation of data illustrates current gaps in data and knowledge. The presentation of data arising from the human consequence inventory is not shown.

It is essential that the pathway analysis method produces data which are complete, so that potential impacts are not overlooked, and that the data are sufficiently accurate for the purposes of valuation. The only way to check whether data are sufficiently accurate is to establish the range of likely actual values (by assumptions/knowledge of variance or margin of error). Then, endpoints (highest/lowest points) for each suspect/approximate datum can be established and used in addition to the calculated central point in all subsequent calculations up to and including valuation (referred to as the Test of Sensitivity, Horne, 2000a). To check all pathways are covered, once all environmental changes are listed, along with all human consequences, RIO quantities at each pathway destination can be summed to provide a mass balance check comparison with the known total RIO release. This is referred to as the Test of Completeness (Horne, 2000a). Where errors or losses of RIO are found, so that gap-entries have been entered on the pathway inventory, a Completeness Index should be produced (% accounted for under each consequence). This is then presented as another piece of data to the valuation process, where it may be used to provide potential ranges of additional value (for example, on a pro-rata basis). In every case, it must be presented to decision makers along with the results of the valuation. It must also be stressed that the QLOS approach will always lead to minimum values (unless factoring or adjustments for gaps are made). Thus, a QLOS value x should be stated as >x for the purposes of decision making, or x_a+/-y where y is the confidence interval and x is corrected by the Completeness Index-based adjustment.

Indeed, the Completeness Index is not the only piece of data in addition to the valuation that is likely to be of interest to the decision maker. The Coding system is simple but it makes the traceability and transparency of the method immediately apparent. The pathway analysis method data records are therefore potentially of interest to decision makers and other parties. Developers, for example, may find the information produced by the pathgway analysis method – and the output analysis and valuation methods - invaluable. By making design adjustments, they could investigate the effects on Total QLOS value as well as on other important full cost criteria. Another particular benefit of the QLOS approach and data format is the potential for updating and transfer of data. As knowledge and modelling techniques, etc. improve, the transparency, accessibility and stepped, sequential nature of the QLOS dataset enables simple and quick data updating to provide QLOS values based on the latest data techniques. Although data transfers carry all the problems of any site-specific dataset, there is clearly potential for many impacts to have similar consequences (or predictably different ones), thus providing efficiency in the data collection and environmental evaluation process.

C5. Conclusion: The State of Knowledge

Recent advances, such as the use of Geographic Information Systems (GIS) interfaces, second generation Gaussian dispersion algorithms, and ability and relative availability of necessary computational power ensure that the possible accuracy continues to improve. Local and regional models capable of necessary accuracy to predict relatively low concentrations of pollutants around buildings are now available. For example, UK-ADMS, which is a model developed specifically for UK point source applications, has been adapted to ADMS-Urban, a multi-point source model which has been used to successfully model levels of SO_2 in London using the Greater London Emissions Inventory (Owen et al, 1999).

Increasingly, many environmental scientists can (and do) approach modelling and measurement of sulphur through the environment at the site specific, detailed, atomic/ionic level. Information is being generated which could therefore be of direct use to applications of the pathway analysis method. This development has been made possible by improvements in knowledge, and accelerated by the development of research and computational tools, and the needs of the policy process. However, the latter is somewhat a “double-edged sword”. Policy-driven research may be expected to produce precise answers but, since it is invariably time and subject constrained, are these the right answers? The reality in the case of the demonstration RIO is that the developing political aspects to the sulphur control debate have generally focussed the research efforts towards understanding the role of anthropogenic emissions of sulphur in specific damage terms (as well as providing the research resources to do this). The results to date are of some use in moving the level of knowledge from the general to the specific in terms of quantified sulphur pathway tracing. The potential tools are available, and are currently being further refined, which will allow greater accuracy in data required in pathway analysis method applications. If the principles of the pathway analysis method are applied to future research, then suitable data will be more widely available and more widely transferable than currently.

During the peak of the “acid rain” debate of the mid to late 1980s, the idea that cause and effect had not been established unequivocally was seen by some researchers as an opportunity to do more research, and by some politicians as an opportunity to wait and see. Many genuine gaps in understanding do indeed persist (for example, see Howells, 1990). However, the adoption of a precautionary principle necessitates action to avoid unless a systematic review (for example, undertaken using the systematic framework approach) shows that impacts/effects will be both established and acceptable. There will inevitably be unknowns (denoted by apparent gaps in materials flows), and risk assessment is a tool which can be used to generate further information about actual risks, so that judgements about perceived values of these risks can be made during subsequent valuation. The burden of proof shifts under the precautionary principle, and the proof can only be found through application of a systematic approach such as the pathway analysis method.

The critical loads approach has great merit and potential for development for use within the pathway analysis method. It has been put to good use within the ExternE project, which remains the most relevant single work to date as far as the systematic framework is concerned (ExternE, 1995a). However, the critical loads approach focuses only on the final state of the soil, and improved, dynamic modelling approaches are being developed which provide temporal information, particularly in connection with a general policy shift towards optimal economic control modelling (Schmieman and Ierland, 1999). Furthermore, a combination of the development of finer resolution models and the integration of these with each other and with a wide range of GIS-based spatial datasets provides a major step forward (for example, Lowles et al, 1998). Also, novel approaches to modelling are being developed, in an effort to overcome the problem of validation of very large multi-variable dynamic time series based models. An example is the Data-Based Mechanistic modelling approach (Young, 1998). At the same time elsewhere in the literature, the concept of local and regional material balance as a means of producing more sophisticated and accurate assessments of changes arising at the local level is developing as a response to the rather broad-brush simplicity of the classical critical loads approach (Mayer, 1998). Provided models can produce data outputs which conform to the principles of the pathway analysis method and the wider systematic framework, then the current data problems can be substantially overcome.

Clearly, acidity arising from sulphur deposition is both complex and, potentially, a major problem simply because of the diversity of environmental changes it can produce. Equally clearly, sulphur is only one of several acidifying pollutants, and they can act cumulatively with each other, and possibly even synergistically (Guderian, 1985, Mackenzie and El-Ashry, 1989), although the evidence for the latter is extremely limited. Therefore, logically, many studies of acidity now seek to take a "total modelling" approach and include all acidifying compounds. Since the pathway analysis method operates at the single RIO level, this demonstration has not included an assessment of, for example, the combined effects of SO_2 and NO_x emissions. However, the approach remains valid, since it is essential that the pathway identification and Inventory is just that; compilation of mechanical transfer of materials through the environment. Once these data are established, then integrated modelling can be used to establish the extent of combination effects, for example, of pollutants, as suggested in many current studies (for example, Rasmussen, 1998).

Setting out to establish total acidity and involving a range of pollutants is valid if the aim is to assess acidity. However, the pathway analysis method seeks to establish the effects of sulphur. It is only after the compilation of the Pathway Inventory that cumulative and synergistic patterns should be investigated, by which time, a great deal of detailed *generic* information is in the public domain about how a particular material (sulphur) for a particular source (RIO) passes through particular environment types (subsystems). Only using this approach, could effects *other* than acidity from sulphur be expected to be detected, since an acidity study, which has a single impact as its starting point, does not seek to trace sulphur for this purpose, and even if other effects come to light incidentally, it must overlook them as they are outside its remit and aim. The policy and political process drives science and science is only one ingredient in decision making (Winstanley et al, 1998). The potential problem arises when the policy process exerts a dominant influence over science and determines the outcome of scientific inquiry by careful selection of research and researchers according to pre-existing policy goals. One source of dominance is control of resource available to scientific inquiry. However, the pathway analysis method provides some protection against such problems, since it is demonstrably objective and transparent in approach, and isolates subjectivity - where it is accepted that values may change according to policy decisions or various cultural-social-political factors.

Modelling has been repeatedly reported and advocated as a potential provider of knowledge about complex dose-responses and environmental changes for pathway analysis method applications. Software technology has advanced considerably over the past decade, and online model databases are now a reality. With existing levels of knowledge, useful and sufficiently accurate and inclusive assessments can be made of environmental changes due to SO_2 emissions from a power station stack point source. The prospect of "model federations", with data freely available and rapid development and validation of new models possible over networked systems (Rizzoli and Davis, 1999), would further assist the efficient practical application of the pathway analysis method. Indeed, if this is combined with the materials flow approach developed in life cycle based work (Ayres and Ayres, 1998), then the systematic framework would broadly be the result.

APPENDIX D. ISSUES IN QUALITY OF LIFE MEASUREMENT

There is inevitability in the need to devise a process for valuing environmental impacts in a common currency. There is no question that it cannot be done - it is being done implicitly at present. The basic economic necessity of being able to compare like with like is undeniable. The question is how it is best done. The approach taken in the current work is to devise a means of valuing environmental impacts in standard quality of life units.

D1. Measurement of Quality of Life

There are 3 basic options for establishing values; personal judgement, collecting existing values from the literature (established by other authors or inferred from data), or measuring values directly from a sample of people. The first is the quickest and easiest, and could be combined with sensitivity analysis to test robustness of values. The second is also potentially relatively simple. The key problems here are that, while there is a large and rapidly expanding literature on values and quality of life, there are very few actual values which have been established with any confidence. Also, where they are postulated, there may be variability between values from different authors, or the value established may not correspond to the values needed (or be in the form needed, for example. transferable ratio data).. In short, the literature is still somewhat "immature". However, much progress has been made and a review of values in the literature (and problems and methods of producing them) is essential. The third option could be expected to be the most accurate, and is necessary where values are either not available or suspect due to assumptions made, technique used or population measured.

In general, there are also 3 levels of measurement of quality of life. The first is the single global question, and this is rejected as it is well-established that this approach leads to unacceptably high loss of data and precision. The second is the development (by empirical means) of a very long list of descriptors of quality of life, containing, as a whole, all of the possible elements of the concept. This is often then compiled into a number of "domains" (usually between 3-16), within each of which questions/descriptors can be identified as being of the same group. The problem for then achieving a single index is that there are potentially millions of possible quality of life-states, since theoretically, every possible state within each domain may exist at the same time as every other state within every other domain. In practice, there are impossible combinations and many complex multi-dimensional health-related quality of life (HRQL) measures have thus been simplified into a number of common combination states. This brings the discussion to the third level of measurement, which is a combination of the simplicity of a global score and the accuracy of a multi-dimensional quality of life-status measure; the Index. While the term "index" is often used to mean different things, here, it is taken to mean a single ratio-scaled list of states measured in a single unit. Since a single global measure has been rejected, this index must contain several dimensions of quality of life, combined in different ways, to achieve the different points on it. It is this approach, the forerunners of which in HRQL measurement are the Rosser Index and the Index of Well-being (or Quality of Well-being Scale) discussed in Appendix E, which are envisaged for construction of the QLOS Index.

The question of how to value things raises the supplementary question of "whose values?". For quality of life, direct measurement of the person affected, a proxy, an expert or a decision maker/politician are some possibilities. In the current work, it is assumed that the former is the ideal choice. One important issue which arises immediately is whether the values obtained are rational, or are at least made accurately and in full possession and understanding of the relevant information. For example, asking a person to value something they have never experienced is likely to lead to a margin of error. Nevertheless, as Fox-Rushby (1994) states, hypothetical scenarios have usually been the preferred approach, because individuals invariably only experience a small number of conditions, and so use of scenarios allows any one individual to value a larger number of quality of life states.

D2. Problems in Measuring Quality Of Life

As iterated in Section D1, the fundamental aim of developing a standard framework for full (quality of life) valuation of environmental impacts in comparable terms is essential to improve on the current ad-hoc system of policy-led, opaque, interest-based decision making. Any other conclusion suggests a failure to accept that decisions based on value comparisons are being and have to be made and, therefore, that improvement in them is desirable, or an implied position that the current opaque inaccurate system is satisfactory. The field of health resourcing is similarly affected by a lack of acceptance of the need to establish life-based common value comparables. Thus, it is not surprising that there is a wide range of literature critical of the development of single scales of quality of life. The review of these problems associated with quality of life measurement, as presented here, is undertaken from the perspective of identifying weaknesses with current measurement techniques and developing solutions, rather than seeking to discount the approach to single scale measurement *per se* (since there is no alternative but the unsustainable present one, so such a conclusion would be invalid). The discussion is split into two parts; Section D2.1 deals with the general problems of quality of life scales, and Section D2.2 deals with the more specific problems which arise in combining scores and/or subscores, or other methods of producing a single scale of quality of life states. Each is presented in a similar format; summary of the problem and possible solutions. The conclusion summarises the key outstanding issues in Section D2.3.

D2.1 General Problems with Values, Measurement and Scaling

The general problems of scaling values can be grouped into 3 categories; appropriateness of the approach (or why the general approach is inappropriate), issues in measurement bias, and the problem of differentiation and distribution of quality of life. These are discussed below in Sections D2.1.1 (seven issues), D2.1.2 (seven issues) and D2.1.3 (five issues) respectively.

D2.1.1 Appropriateness of the Approach

D2.1.1.1 Does Average Quality of Life Exist?

HRQL self-rating scales like the General Health Questionnaire (see Appendix E) have been criticised as follows: "Their use implies that Quality of Life (sic) means the same to everybody and so can be defined in general terms. Even if the items they contain have been selected by studying an appropriate group of patients, the instrument may only strictly be applicable to a non-existent average individual" (Bech, 1994). However, since the QLOS approach is designed to measure total impact, individual average issues are not of concern; averaging is an acceptable device for aggregation. Nevertheless, the argument that quality of life is too personal/individual for standardised measure is compelling. People's perceptions of quality differ for every realm of evaluation imaginable and, furthermore, these differences are generally the result of legitimate individual preferences. There is also a large body of literature to support individual variation in the way various health states are rated. However, empirical data support the argument that, despite the differences, there is much common ground among people's relative evaluation of health states (Kaplan et al, 1993). As has been stated: "There is a strong consensus for the idea that a day spent with a runny nose is much closer to perfect health than is a day spent confined to bed.

Of course, there are exceptions, such as someone who would like a day away from work. To aid the many decisions which must be made for groups of individuals, there is a strong argument for using available information on areas of common ground in valuing health outcomes. This information is available in the form of well thought-out quality of life measures" (Walker and Rosser, 1993).

D2.1.1.2 Short Term Fluctuations

It has been established above that the most appropriate group to elicit values from is the receptors (those who receive impacts/experience quality of life states), but one of the problems with this is that work to date suggests that utility values obtained in this way tend to be prone to fluctuations over short periods of time. Are they therefore reliable? If this question is taken as one of an individual's preferences being transient, the solution is to capture a reasonable population sample - the aggregate scores will then reflect average preferences and will automatically smooth out any transience effects. Regarding longer term changes in social/value preferences, repeating valuation studies from time to time and from place to place will allow evidence of shifting values over time to be detected.

D2.1.1.3 Key Concepts are too Poorly Defined and Understood

As definition and concept of both utility and quality of life are poorly understood, to allow a few researchers to base major decisions about, say, building (or not) energy projects on such spuriously obtained results is ethically unsound. If this were to be the case, the only solution is to improve understanding by practice. However, the ideas that quality of life is not understood or well described by researchers seeking to measure it, the concept is not familiar enough to be meaningful to respondents, or that there is too much variation and not enough agreement as to what is in fact important to quality of life, are now outdated. In terms of respondent views, there is plenty of evidence that people are familiar with the term and that they relate relatively easily to it and, furthermore, that they also broadly agree on what the key constituents of quality of life are (for example, Farquhar, 1995). If there is a problem with the quality of life concept, it is in definition. Invariance of term usage has been suggested as the sole criterion for scientific language, and terms borrowed from the vernacular, particularly referring to emotions, cannot develop invariance of usage in common parlance (Horley, 1984). Thus, the use of such terms, which are widespread in HRQL and quality of life measures, is questionable, or at least, the achievement of invariance is not accomplished easily. This is a perennial problem for quality of life measurement, and one which has not received much more than the inevitable empirical attention. While terminological variance is a problem for meaning between groups of users and respondents, it is also a problem for theoretical development, which is much-needed in the entire quality of life measurement field. However, the solution is not to reject quality of life measurement, but to define it better.

D2.1.1.4 People are Irrational

One fundamental problem with any method of measuring people's values for quality of life is that it relies on the possibility that people can make unconstrained value decisions. In reality, people's decisions are affected by existing social values, and their judgements are therefore not purely existential in character. Thus utilitarianism is arguably flawed. However, there is no better way to elicit quality of life values than asking people or judging people's behaviour; the alternative, to base values on how people ought to value, is rejected, since a measurement method should be reflective (recording values) rather than prescriptive (determining values). While noting that people operate within their social world as well as their individual world, and that this affects their values, we must accept that the basic approach to measuring people's direct or inferred values is the right approach. There are specific problems that arise from apparent irrationality, for example, the "Preference Reversal Phenomenon" may affect studies where risk is involved. It occurs when an individual, asked to choose between risky prospects A and B, chooses A, but when asked to place separate certainty equivalent valuations on them, places a higher value on B. This is an outstanding theoretical issue within the field of risk valuation.

D2.1.1.5 Complexity

While the case has been strongly made that people should decide their own values, and that therefore scale points can only be determined by a proper sample of the general population, there are difficulties in actually securing these values. Quality of life is a complex concept, and its measurement often involves many and sometimes difficult questions. Therefore, notwithstanding the point in Section D2.1.1.3, one predictable criticism is that difficulty in collecting value data will lead to inaccuracy in results. The solution lies in the elicitation or measurement process; utilising several steps (so only one piece of information has to be evaluated at a time) and using simplified notes for guidance are two measures that can improve accuracy of results.

D2.1.1.6 Information Requirements and Error

The issue here is that there is not enough information available about the value or object of value, or it is patchy or erroneous. As with any new technique, the first data collection round is difficult because the data have not been collected in that form for that purpose before. Inaccuracies and gaps in information and knowledge undoubtedly exist. The effects of all impacts, diseases, etc. on people, or the response of the environment to many pollutants are not known. This does not invalidate the method set up to evaluate them, it merely identifies where information and knowledge is required.

D2.1.1.7 Risk and the Future

Firstly, it is important to make the distinction between probabilistic risk (the likelihood of an event occurring) and perceived risk (the subject view of value of a given risk). The former is not of concern here, since risk-weighting, in the sense of adjusting values to account for the actual likelihood of occurrences, is not a subjective exercise and can be carried out separately from measurement of quality of life outcome. However, where there is an extra risk element perceived by the subject, which affects their quality of life, then this (subjective) element of risk needs to be incorporated into the outcome measurement process.

Prior to experiencing the future, people understandably have a varying view of what it will be like, compared to that which they arrive at during and after experiencing it. In particular, risk in prospect may be expected to be valued more highly where there is less control or feeling of control. Some people are more frightened of flying (where they do not feel in control) than of driving (which is far more life-threatening but the driver feels in control and the process is more familiar). The solution is that, during the process of generating human consequences from environmental changes which have perceived risk elements, these elements must be assessed accounting for the specific attributes of the perceived risks involved, including whether they are current or future events (for example, by asking respondents to value their current and future risks).

D2.1.2 Issues in Measurement Bias

D2.1.2.1 Content or Scale Bias

There are a range of ways in which the researcher, by constructing response or questions in a particular way, biases the results obtained, causing untrue data to be achieved. The process of deciding what to include in utility (i.e. what questions to ask) means that researchers/experts who do this rather than getting respondents to determine the content/questions are inevitably biasing the process. Therefore, studies which are based on the former approach are suspect and should be avoided, pending verification by a process involving the latter approach. Examples where design of response categories lead to Scale Bias include the case of category and Likert scales, where experience suggests the use of middle values as possible response categories should be avoided, since subjects may prefer not to commit themselves and may therefore select this value as an easy option rather than the true measurement. In other cases, it may be deemed valid to measure neutrality, so neutral categories may be used. One author notes that, either way, "the researcher should be aware of the dangers of creating biased responses when forcing choices" (Bowling, 1995a). In reality, since both approaches could be viewed as potentially bias-inducing (and indeed both may be so, to a greater or lesser extent, among different subjects and for different measurements), it leaves the construction of any notion of a valid category scale in some problems. Fortunately for the current work, it is unlikely that scales with a central neutral point will be required, since, in the measurement of impact, it is assumed that only in one side of the attitude continuum is being considered - i.e. what is getting worse, or is negative. Thus, it is to be expected that the neutral value will form an anchor point at one end of the scale.

D2.1.2.2 Framing Effect

This is really a type of Scale Bias. Evidence has been presented that a pair of prospects presented to a sample population in terms of probabilities of gains gave different results when the (formally identical) pairs of prospects were presented in the negative terms of losses (Loomes and McKenzie, 1990). A parallel exists here with the contrast in results obtained between "Willingness to Pay" studies and those seeking the same values but couched in terms of "Willingness to Accept" (compensation). There is a difficulty in terms of choice between differing values (should they be available). However, since impacts are essentially about negative effects on quality of life, it would appear appropriate to couch questions in those terms, wherever there is some doubt over which approach to use.

D2.1.2.3 Respondent Bias

This is a general term used here to describe a range of ways in which the subject in a survey may deliberately choose not to give true answers. Principal among these are free riding situations or where the respondent feels that, by over/under emphasising, they can maximise their particular agenda. These can be dealt with by a combination of survey design (to avoid/detect), adjustment where applicable, and use of appropriate sample size.

D2.1.2.4 Method Bias

Several studies have set out to compare different valuation methods for the same parameters and population sample, and produced widely differing results. This is a problem familiar to environmental economics, and leads to method-based bias, where errors within methods are reflected in values. To give just one example, Llewelyn-Thomas et al, 1984 found that standard gamble and category rating approaches generated systematic differences in values from the same population sample for the same set of health states (examples also exist where similar results were obtained). In the absence of a gold standard against which to compare which method arrives at the most accurate values, the only means of differentiating between methods is how well they perform against the criteria set out in terms of predicted outcomes (validity), repeatability (reliability) and sound theoretical basis. The best criteria-fit indicates the most appropriate method.

D2.1.2.5 Design and Response Issues

The structure and format of questionnaire surveys is such that some people, who may have particularly chaotic lifestyles, or a lack of resources, or find questionnaires difficult or meaningless, etc., will not respond to quality of life surveys. In the Ventegodt (1996) survey, unemployed respondents had a quality of life 8-13% lower than employed respondents, so if, for example, they were to have a significantly lower response rate too, the overall population quality of life recorded will be artificially high. Thus, a net result of this known pattern is that there is a possibility that average quality of life is actually slightly marginally lower than that measured. However, statistical methods can be employed to partly compensate for this - for example, weighted linear regression.

D2.1.2.6 Hypothetical Preferences

These fall into 3 types; ignorance/lack of experience, differences between stated intention and action, and lack of reality. Reservations have been expressed about both the weakness of assumptions implicit in revealed preference approaches, and also the problems of direct questioning to elicit preferences. Regarding the latter, the point is made that; "unless people go about with hypothetical preference structures for choices they may never meet, then it is only by exploring choices that they can learn what trade-offs they are willing to make" (Brown and Green, 1981). Therefore, any direct method must be designed to allow respondents to first discuss and discover their preferences before stating them. The second problem is that intentions do not always mirror actions. For example, people may overstate or understate scores or response levels, depending on their perception of what they will attain from it (or not), or what they feel they should score. They will tend to attempt a response to any question, provided it is not obviously silly, so it is important to ensure that questions are within their experience (they have a frame of reference with which to answer, rather than a guess). Regarding the third problem, there is a tendency to not place great emphasis on thinking about what personal values actually are when it appears not to matter. The experience of being asked largely irrelevant questions in consumer surveys bears this out. Indeed, responses will often be biased towards those which the respondent thinks will shorten the survey, or which the interviewer wants (or doesn't want, if they wish to be awkward). Hypothetical bias affects a wide range of survey methods, and is well-known, and can be corrected once it is understood and measured for a particular survey type.

D2.1.2.7 Regression Bias

Regression Bias is the tendency to overestimate the intensity of weak stimuli and to underestimate strong stimuli. For example, in a QLOS index, this may mean an insufficient quantitative distance between serious illness or death, and more marginal impacts such as feeling uneasy. Regression Bias can apply to all response measures - both physical and social. A correction has been suggested for regression bias, for data sets which fit the following necessary requirements (Lodge, 1981):

- Two response modes for the same stimuli tests must be highly correlated (>95);

- Two response modes for the same stimuli tests must be linear on log-log graph, i.e. correspond with a power function;

- The bias direction for the two response modes must be the same on calibration and scaling tasks;

- The empirical exponent derived from the scaling should be approximately equal to the exponent obtained in calibration.

The ideal scale value, a, would become: $a + (NE^{1.0}\ LP^{1.0})^{1/2}$, since numerical estimation (NE) and line production (LP) have exponents of 1.0, and the use of 2 measures means they are raised to the power of 1/2. The corrected scale value is:

$a_c = (NE^{1/n1}\ LP^{1/n2})^{1/2}$, where n1 is the numerical exponent from calibration of NE to linelength matches, and n2 is the numerical exponent from calibration of LP to number stimuli. Note that this means that the empirical corrected and theoretical exponents are linearly related. Relative ratio relationships are the same, but their absolute value range is stretched.

D2.1.3 Differentiation and Distribution of Quality of Life

D2.1.3.1 Culture and International Consistency

Despite the number of quality of life instruments and their various applications in various parts of the world, as many researchers have pointed out (for example, Orley and Kuyken, eds, 1994), there are as yet no accepted, standardised and validated ways of assessing quality of life across cultural and national boundaries. Patrick et al (1994) cite the following 3 issues as central to this process;

- Questionnaire content and conceptual basis;
- Translation method;
- Testing and comparison of validity, reliability, responsiveness and effect size within each culture.

However, while accepting that this process is far from complete, the implication here is that cross-cultural and international consistency can be reached. Bullinger (1994) agrees that the process of developing such consistency is essentially iterative, but notes that this may lead to original measures being changed, with resultant loss of statistical data sets and proven track record. This observation is made in the light of the involvement of the SF-36 rating scale (see Appendix E) in the "Quality of Life" project, which is designed to internationalise this essentially US-based measuring instrument. A solution to this is to devise weights and scales simultaneously across the world. Others (for example, Kuyken and Orley, 1994) state that certain facets of quality of life are not universal, and advocate culture-specific questions. The pilot instrument for Thailand includes the question "How well are you able to rid yourself of negative feelings through meditation?", in recognition that "the vast majority of the population are Buddhists". This is illogical, since many people in Thailand are not Buddhists, and many Buddhists do not meditate. What is this question supposed to measure? Certainly, it would be more useful and universal (both inside and outside Thailand), if the last 2 words were omitted from the question; different people have different ways of "ridding themselves" of negative feelings and, presumably, it is the ridding not the means which is important to the researchers in this case. Of more substance is the observation (Kuyken and Orley, 1994) that self-esteem as a concept varies between the west, where "I" is more central, and east, where "we" is more central. However, such differing emphasis can be reflected in careful structuring of questions with self-anchoring aspects (i.e. of the "your social interaction compared to your perceived optimum" type).

D2.1.3.2 Transferability and Uniqueness

The issue of how transferable measurements are, or whether they are essentially specific to individual decisions, is somewhat problematic. The extent to which each value decision is individual or unique reflects the range of variables being considered (explicitly or, perhaps more importantly, implicitly) in the valuation. If one is asked to value a particular health state of identifiable symptoms, the experience (direct or indirect) and knowledge the individual has of the state will influence the value given. At a future point, further experience may alter the value. Such variations over time or life experience are subject to the same explanation as for the static question above. Another way of interpreting the question is how transferable values are between individuals. Unless there is very high variance, a sufficiently large population sample will ensure that aggregate values again reflect the population. Where there is high variance of values within a population, a determinant or indicator variable (such as age group or social class) needs to be established to allow transferability to unlike population profiles - for example, this would allow adjustment to be made to values for applying known weights to a different demographic population profile to the original, weighting population. By this means, any values of health/emotional states which were systematically related to age, sex, education, class, wealth, religion, health experiences, culture, etc. can be corrected prior to transfer of values, if the requisite information on both populations and preferences is known. In short, if sensitivity requirements are sufficiently high, then relatively sophisticated scales may eventually be needed, with the considerable data requirements that this would entail.

D2.1.3.3 Social Class

Maslow's Hierarchy of Needs (Maslow, 1962) has been used to generate a means of assessing people's needs in quality of life measurement (Gratton, 1980). Seventyfive separate needs were generated and placed within Maslow's (five) need levels. The ten needs that most consistently fitted each level were then selected, and these fifty needs were then scaled using Q-Sort methodology into five piles (most important at one end, least important at the other). Analysis showed that social class affects need priorities. While middle class respondents emphasised esteem and self-actualisation, the working class were esteem- and belonging-oriented, and lower class were physiology- and belonging-oriented. The same applies to this apparent problem as applies to culture;

care needs to be taken to ensure that any piloting and scaling exercises involve a representative sample of the population.

D2.1.3.4 Inequality of Utility Theory

Keeney (1982) doubts the validity of some utility approaches to valuation because of their essentially individual approach, in the context of societal risks. Instead, he advocates that organisational-based value judgements should be established in such cases, to ensure maximum utility is secured through minimising deaths, equitable distribution of risks, and catastrophe avoidance. Linnerooth (1982) also rejects the classic utility approach in the context of the equity-efficiency trade-off, for example, the inverse linkage between Willingness To Pay for increased survival and marginal utility of wealth (which is in turn very sensitive to survival probability itself). Utility theory is certainly not egalitarian, and there is a fundamental moral question over any theory which supports maximising utility gain by concentrating it in the hands of the minority. However, while such criticism of any method which implicitly adopts aspects of utility theory may appear valid, in fact, societal equality aspects of environmental impacts are not an issue here. This does not mean equality is to be ignored. It means that the place to deal with equality issues is not within the aggregated Scale approach to measuring impacts, but within the wider social policy framework. The QLOS Index is designed to measure environmental impacts, not equitably distribute them.

D2.1.3.5 Disenfranchised Populations

Questionnaires can only be responded to by those who are capable of understanding them fully. It is normally assumed that a quality of life questionnaire cannot be effectively used with young people under 18-years-old or mature, mentally incapacitated people. Therefore, corrections will need to be applied to response data to correct for disenfranchised groups whose values are not reflected in the survey, using the values of proxies, such as friends, relatives, etc. (This proposal is flawed if it is found to be possible to measure quality of life by using a method that is sensitive to the developmental repertoire of the young or incapacitated. Indeed, steps are being taken to develop methods for undertaking health-related quality of life measurements in children and adolescents (Drotar, 1998) and, when these are found to be adaptable to the measurement of impacts in quality of life terms, they should be adopted. Hence, the correction proposed is a stop-gap measure to prevent the quality of lives of these groups from being ignored, and ensure that, at least, some approximated representation is incorporated into total quality of life assessment).

D2.2 Specific Problems with Single Scales

The production of a single index brings further specific scaling and measuring problems, particularly related to the issue of aggregating or combining dimensions of quality of life, and the inevitable involvement of applying a time-related value to reflect the length of time the quality of life state lasts for.

D2.2.1 Aggregation of Multidimensional Constructs

The main argument which those who work with multi-dimensional quality of life measures use against summing scores is that, while it is possible to measure individual aspects, it is meaningless to sum these as they are fundamentally different units of quality of life. Various quantities can be measured for a moving object, such as mass, velocity and direction, but summing or otherwise combining these quantities (for example, into a measure of momentum), involves a loss of important information about each quantity. Thus, the argument is made that, while aspects of quality of life are measurable, the entire concept is either non-quantifiable in a single quantity, or its expression as such involves irretrievable or unacceptable loss of data. Accepting this view for a moment, multiplying dimension-values in quality of life measurement is equally invalid. Indeed, summing such items as mobility and aesthetic appreciation has been described as “at best meaningless and at worst redundant” (Williams, 1994). Many authors suggest that the fact that many different HRQL measures have different contents (and therefore measure different concepts of quality of life) lends weight to the view that items should be left unaggregated (for example, Bowling, 1995a). While accepting that this is inconvenient for those who need a global measure, the assertion is made that to sum scores would undermine the integrity of the instrument.

Thus, it is apparently problematic to weight and sum scores across dimensions. However, people do make judgements based on quality of life as a single entity, so it can be (and is) done every day. It may be as crude as devising a single number to represent a moving object and another for another, but proceeding heuristically will lead to improvements in understanding which can be expected to lead to a better quality of life measurement model. Indeed, at least one rating scale is already being iteratively altered to allow summing to take place, in recognition of the fact that summed scores are needed and are becoming more achievable as knowledge of the quality of life construct and its measurement develops (see Appendix E). So, the answer to the question of whether the different dimensions of health and emotional states can be valued independently and the results integrated, is a challenging one, and the short answer is “yes - provided the measure has been designed and validated appropriately”. Incidentally, one of the few ways of validating summed scores is to adopt a different, though theoretically similar approach to the same problem and compare results. For example, comparison of a method involving summing single scores with one involving a more holistic scenario approach (such as a more complex version of the single global index) may provide interesting results for comparison.

One of the problems with summing scores is that it requires more precision from scoring methods, since accuracy, theory and measurement method of various dimensions must be combined. The Standard Gamble, the theoretically most sound method, is difficult to execute, and may be contaminated with risk; Thurstone's "Paired Comparisons" method has been rejected as too involved for more than a few states, while direct estimates such as category scaling and magnitude estimation are difficult to transfer effectively from the laboratory to the large sample of the population. It has also been pointed out that in one early cost-benefit analysis experiment based on Quality of Well-Being (QWB) Scale scores (see Appendix E) was undertaken in Oregon, and produced results that were so counter-intuitive that informal procedures were used to re-order the resulting list (Jenkinson, 1995). While such examples of the extra burden of validity, reliability and precision required in stepping up from single dimension scoring to multidimensional global quality of life states do suggest there is some distance still to go in perfecting the latter, it is clear that major progress has been made and that workable models for such approaches do exist and will improve in the future.

D2.2.2 Aggregation Hides Detail

One of the common reasons stated for not summing multi-dimensional scale scores into a single index is that it leads to loss of information, as referred to in Section D2.2.1. For example, if twenty items are rated on a 1-5 Likert scale, the sum would range from 20-100. If these items covered several domains of health/quality of life, for example, physical mobility, social health and mental state, then the sum score of, say, 40 could be achieved by two very different individuals, one a physical invalid, and one with severe social problems. It is true that this approach therefore loses data and hides detail, but the stated goal of a sum score for measurement purposes is not affected by this problem. Loss of data at this point is only problematic where more information is required than just total size of impact (for example in HRQL, where a prognosis is required, which may rely on a specific pattern of dimension scores). A developer or policy maker may need to know which dimensions of quality of life are specifically affected by a particular environmental impact, for example, in order to build specific mitigation measures or legislative controls to reduce it.

What authors of HRQL measures mean when they state that domain scores should not be summed into a total score, is that they are confident the measure is useful for the use they intend (diagnosis or outcome in particular areas), but they are not confident that each domain is actually given due weight in the assessment of overall quality of life, or that the theory/model (if any), on which the measure is based, is designed to actually measure quality of life at all. Therefore, where measures do recommend that scores are not summed, they must be treated as potentially suspect for QLOS Index construction purposes.

D2.2.3 Constant Proportional Time Preference

A general criticism is that single index scales are invariably over-simplistic in assuming that days of illness/impact are of equal importance whenever they occur, proportionally however long they last for, and irrespective of following or preceding health state. If successive impacts occur in the same person, they may be expected to affect the total loss differently to if separate episodes (or impacts) affected different receptors. Other rivals to the constant proportional time preference assumption include the idea that quality of life might be better evenly spread across time (so there may be a willingness to accept less quality of life overall for a consistent level), or that it is preferable to have quality of life peaks - such as a good ending. Thus, the issues of accepting the assumption of constant proportional time preference are potentially problematic, given evidence suggesting it does not hold in reality.

D2.2.4 Constant Proportional Risk Attitude

Loomes and McKenzie (1989) also report that there is various evidence suggesting that people's risk taking fluctuates over time, and they may even practise systematic switches between risk aversion and risk taking behaviour. Therefore, any assumption of constant proportional risk attitude is also brought into question.

D2.2.5 Interaction, Cumulative and Synergistic Effects

Related to the issue of combining a number of quality of life domain/components to give a total quality of life measure, one problem with this is that quality of life may actually depend upon interactions between components. No evidence is presented for this phenomenon at this stage, but it is a possibility.

D2.2.6 Reliability

Since health and quality of life status are not yet well-measured concepts, measurement error is to be expected. However, the measure for this, the confidence interval, is often very high. The SF-36, the most validated health survey questionnaire, has a confidence interval of 12.8, even for its most reliable dimension, that of physical functioning. However, reliability is well-known and measurable and must improve over time, through iterative empirical processes.

D2.2.7 Sensitivity and Maturity

It has been suggested that, out of 5 different HRQL instruments compared, none consistently indicated greater sensitivity to change in all dimensions (Jenkinson, 1995). In other words, they indicated different sensitivity in different areas, indicating that there is considerable immaturity in the field. Indeed, the very fact that well in excess of 1000 health status measures have been developed over the past 3 decades is evidence itself of the immaturity of the field. Combining domain scores therefore will obscure inaccuracies in domains and items, any of which may not be sensitive or appropriate and may therefore affect the overall summed result. The solution to this problem is to develop a consistently accurate and reliable, sensitive measure, and to maintain transparency in all procedures.

D2.3 Conclusions on Quality of Life Measurement

As long ago as 1936, it was stated that a single index of health was not useful because of the loss of information involved in creating it (Hunt et al, 1986). To some extent, this view has continued to the present, with the notable exceptions (in health research) of health economists and others who input into resource-based decision making. Authors of scales often state that the domain scores should not be summed but rarely explain why. The real explanation is that the job of creating a single scale is rather complex and should be avoided if possible. Moreover, the absence of a comprehensive scientific theory which integrates quality of life concepts and aspects is the source of more severe problems than most scale authors recognise. This in turn has led to what has been called an “empiristic approach” (Rosenburg, 1995).

The most fundamental critique of a method for measuring quality of life would be one that challenges the process itself. The closest that criticism of the quality of life Index as a concept comes to this is in the issue of summing domain scores to achieve a single index. However, the problem is a practical one rather than a conceptual or theoretical one. In HRQL research, where the concept is best-developed, there is, after only a couple of decades of significant research effort, considerable and increasing consensus over the definition and description of quality of life and increasingly, over the issue of producing single quality of life scores. Given the numerous problems raised by researchers, the question of whether questionnaires/interview studies can provide enough accurate information to value different quality of life states relative to each other on a single scale is undoubtedly valid. However, as the vast majority of these researchers point out, these problems do not point us in another direction, as the only other direction would appear to be the current direction of policy and decision making, which is rejected as inefficient. Therefore, these problems are merely obstacles that must be overcome, and for the majority, they can be overcome satisfactorily, as demonstrated in the points raised above.

APPENDIX E. ASSESSMENT OF QUALITY OF LIFE

The last few decades have seen significant growth in interest in measuring quality of life. Here, attention is specifically placed on those approaches which may contribute to construction of a single index of quality of life outcome states, for use in valuing environmental impacts. Section E1 contains a discussion of the Quality Adjusted Life Year (QALY) concept, and Section E2 summarises the economic approaches to the problem. Section E3 reviews twenty measures of quality of life drawn from the health-related quality of life (HRQL) literature, and Section E4 provides resulting comparisons.

E1. QALY Concept.

The Quality Adjusted Life Year (QALY) concept encompasses a group of related approaches to producing values for quality of life states, and it has produced the most debate and results for single quality of life values. Its origins go back at least as far as 1980 (Weinstein et al, 1980). A QALY is a year of full life quality, unimpaired by poor health. In QALYs, length and quality of life are amalgamated into a single index. Each life-year is adjusted with a utility factor (1 = full health). The main purpose for developing the QALY concept was to enable different types of medical procedures and other interventions to be compared by calculations of the unit of cost per QALY gained. Hence, at least in theory, the QALY is a generic measure of health benefit, which is designed to allow comparison between any conceivable type of medical care for any conceivable condition. It is calculated as the product of the number of years gained from a particular treatment/intervention and the quality of life in each of those additional years. The apparent success of the QALY concept in practical operation has been demonstrated, with various studies compiling league tables of cost-effectiveness based on cost per QALY calculations for different conditions and medical interventions, a notable example being that of the Office of Health Economics (OHE, 1989).

E1.1 QALY Methods

QALYs can be measured in different ways. The Rosser Index of Disability (see Section E3), originally developed to indicate hospital performance (Rosser and Watts, 1972), has generally been adopted for use in providing measurements for use in QALY calculations in the British context. The Index of Well-being Scale (see Section E3) is a single-score scale designed to produce quality-adjusted life years by a different route. This is based on expected utility theory, developed by Neumann and Morgenstern in the 1940s, and adopts the standard gamble technique for eliciting preferences for different health states. Its attraction is that it can be expected to reflect global subjective value, including individual risk, and indeed, the method was designed to be used in resource allocation at an individual level. However, the method has been widely criticised, not least because of apparent internal inconsistencies in the standard gamble method, bias problems, and an apparent lack of validity. One author has proposed three approaches to the derivation of utility values which have varying relationship with the QALY approach; standard gamble, time trade-off and composite-index rating scale (Torrance, 1987).

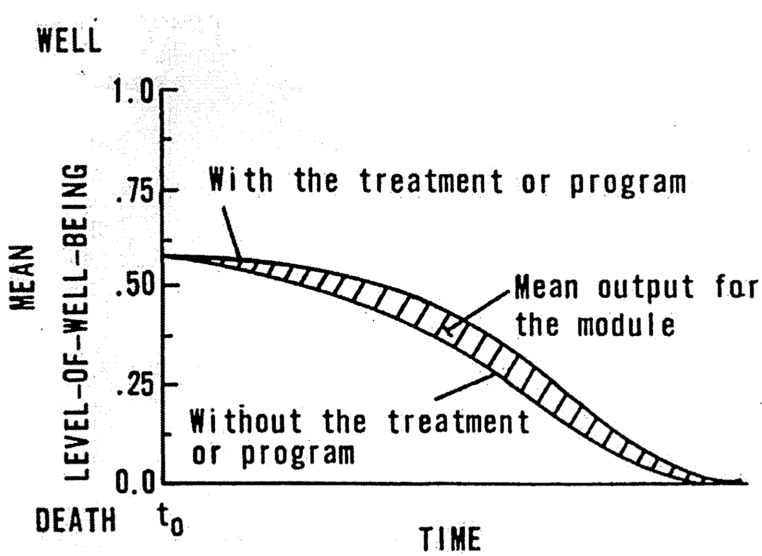


Figure E1. <u>QALY Concept (after Kaplan and Bush, 1982)</u>

The more UK-based QALY method has been developed for decision making at the aggregate or policy level. This is the one which receives most attention here, since it has more potential to contribute to the QLOS Index. It is based on production of a global index and measurement of health states against this common scale. Rosser's Classification of Illness States has been adopted for this purpose (see Section E3, Kind, Rosser and Williams, 1982).

The precise value of a QALY is lower the worse the quality of life of the unhealthy person (which is what the "quality adjusted" part is about). If being dead is worth zero, it is, in principle, possible for a QALY to be negative, i.e. for the quality of someone's life to be judged worse than being dead. The general idea is that a beneficial health care activity is one that generates a positive amount of QALYs, and that an efficient health care activity is one where the cost per QALY is as low as it can be."

Apart from the general logic of the approach, there are two things to note from this. Firstly, the assumption is implicit that time and quality of life can be rationally combined into a single measure. Secondly, it is clear that the QALY concept was developed to assist in making decisions about provision of health care services, and as such is a branch of health economics. The macroallocation application of the QALY concept requires a list of weightings for states of quality of life resulting from different health factors. Here, there is a potential contribution to the QLOS Index.

E1.2 Theory problems: UK Method

A key questionable assumption is that the approach assumes that time spent in a state is independent of the value of being in it - and of preceding or proceeding states. In reality the effect of health (or environmental impact) on quality of life may be dependent on preceding state, how one expects to remain in the state, and/or what the expected future states are. This leads directly onto a potential problem; the fact that the approach is risk-neutral - it ignores perceived risk and uncertainty. Rosser and Kind's Disability index is based on the *certainty* of being in different states. There is considerable evidence to suggest that risk and uncertainty do in fact have a substantial impact on valuation and choice - especially where the (even remote) possibility of immediate death is concerned (Loomes and McKenzie, 1990).

Since it is a framework rather than an explicit survey or scale, the QALY concept does not directly give rise to useful scale points. However, in order to operationalise the concept, scale points are needed, and one index which has been adopted to generate weightings of different quality of life states for QALY work in the UK is the Rosser Distress and Disability Index (see Section E3, Kind et al, 1982).

E1.3 Criticisms

The main criticism (of many) levelled at QALYs is that they are crude and unreal: "QALYs are not really measures of quality of life but measures of units of benefit from a medical intervention, combining life expectancy with an index of, for example, disability and distress. They are based on invalidated value judgements" (Bowling, 1995b). Indeed, lack of validation remains a key weakness, as does the lack of conceptual basis, particularly on the relationship between QALY and the real judgements faced by patients with the condition.

The assumption that the value given to any health state is independent of the time spent in it has been criticised, (for example, by Loomes and McKenzie, 1989), and the concept of the Healthy Year Equivalent (HYE) was developed partly to overcome this. HYEs do not assume inter-temporal additive separability whereas QALYs do.

The relationship between quality and quantity of life may seem relatively straightforward until the possibility of living states worse than death are contemplated. In this case, there is a conflict between the two. A solution has been suggested to this, by distinguishing between death as a state and as an event (Sintonen, 1981). Thus, the former would score zero on the typical scale, but the value of death as an event would depend on the individual circumstances - such that a negative value would be possible without conflicting with the state of death value, which bounds the state of being scale.

E1.4 QALYs and QLOSs

Firstly, a key similarity between QALYs and QLOSs is that both attempt to use a simple 2-dimensional model as a basis for measurement; quantity and time. The QALY seeks to measure the quantity defined as quality of life, and the QLOS, the (proportional or absolute) drop in quality of life caused by an environmental change. These are not dissimilar in concept, only in quantity, and may potentially share similarities in the approach to measurement. However, whereas the QALY measures time in terms of absolute well-years, the QLOS concept, as developed thus far, seeks to measure the proportion of life-time remaining. This is an important difference.

There are many points of comparison between the QALY approach and the QLOS approach adopted in the current work. Both are essentially frameworks, into which notions of intrinsic value (or philosophical concepts) must be inserted. For example, neither are necessarily strictly utilitarian in nature, depending on whether quality of life is defined in terms of happiness or not.

E1.5 Time

One problem identified with the QALY approach is the way in which time spent in an unhealthy/impacted state is included in calculation of value. In the QALY approach, time is measured in absolute terms, so, if one of two people had to be selected for the same lifesaving treatment, the youngest and/or fittest would be selected because of the expected QALY gain in the long term. This conclusion has been challenged as inequitable, not least because it takes as a basis for denying benefit, the fact that someone is already unfortunate (i.e. is unfit, say due to arthritis). As one author states: "people must value the metric (the variable against which health related quality of life is measured) at a constant rate. For instance, an additional year of life must be worth the same regardless of age when a time trade-off instrument is used to measure health status" (Holmes, 1995). An alternative which meets this need is to normalise time by expressing it in terms of "% of life remaining". In this example, both would score the same and a lottery would be the only means of choosing which is treated. Not only is this solution more equitable, it also reflects the way we view our life at any point (we view it much more in terms of what we can do with our remaining life, than what we are doing with our total life).

Another potentially unfortunate outcome of the QALY approach is the tendency for it to favour higher birth rates over prolonging the lives of those already living, since "making babies is doubtless a cheaper way of making QALYs than saving lives" (Broom, 1988). The proposal to normalise time in the valuation method partially overcomes this, since each human will have the same normalised unit of quality of life. Nevertheless, more babies would still lead to more total quality of life. While this is true, they also lead to more environmental impacts, and since the systematic framework is about measuring specifically NEGATIVE effect on quality of life state arising from environmental impacts, more people means more receptors and more impact-creators, and so the pressure is to reduce rather than increase the number of babies.

The proposal to normalise time in the (QLOS) valuation method overcomes the numerous accusations of ageism within the QALY approach. Each new life is an additional creator and receiver of environmental impacts. Had they not been born, the world would not contain the impacts they create, and they would not suffer the impacts which humanity makes. This may appear an overtly negative view of life, but the aim of the work is to measure environmental impacts in quality of life quantities, so inevitably, given this criteria only, each new birth leads to more impact, and old people are subject to generally the same scale of threats to their remaining life as young people are.

E2. Economic Approaches

This summary of the main economic techniques which have been applied to valuation of health states is included to inform quality judgements of potentially useful studies, as reviewed in Section E3. In contrast to multi-dimensional approaches, most of which are not *designed* to produce single global scores, economic approaches must all provide for this possibility, since invariably the end-point is to compare cost-effectiveness, cost-benefit analysis, or for other means of decisions making over resource efficiency.

So, the general aim is to measure utility (usually "quality of life" or "well-ness") in a single common currency measure, so that financial costs of particular medical services which add a known amount to quality of life can be compared on an equal basis of money invested and/or quality of life. The methods most commonly used are Standard

Gamble (SG) and Time Trade Off (TTO), although other indirect methods are used, invariably involving using data of actual actions to infer values. Good summaries of the economic approaches and their applications and shortcomings as they relate to health-related quality of life are found elsewhere (for example, Mooney, 1986).

E2.1 Productivity or Human Capital

This approach involves measuring the amount of lost productivity (for example, time off work) associated with a particular health state, as a proxy for the quality of life value. Since economists almost always use paid productivity here, it immediately ignores the 60% of the economy which is unpaid - housework, child care, voluntary work etc. Furthermore, even if it could be made to capture this, it revolves around productivity, which is not quality of life.

Thus the method is theoretically flawed, and while some authors have accepted that adjustment should be made to capture the remainder of quality of life value which is not productivity-based, the framework developed for the QLOS Index suggests that economic losses should be reflected in direct measurements of subjective quality of life changes, rather than vice versa. The only real case for considering human capital values is that they have status, being used in civil court cases as the basis for economic losses suffered in death and disability cases. However, this is no substitute for their theoretical weakness, and anyway, more recent trends in such cases suggest that court settlements now attempt to capture more than just economic value losses anyway. Therefore, human capital values are rejected as potential contributors to the QLOS Index, on weak theory grounds.

E2.2 Standard Gamble (SG)

Following Neumann-Morgenstern gamble theory, a scaling method has been developed where judges were first offered a choice of remaining health for a given time, or receiving a drug which gave them a stated probability of ensuring perfect health for a given time or death (Torrance, 1976a). The judges had to choose the probability point at which they would be indifferent to the two choices. Other scenarios were then presented along the same lines. It has been pointed out that this method is problematic because it is likely to measure risk aversion rather as well as health judgements (although other methods which do not include risk are often criticised for the opposite reason, and since future life is primarily about risk judgements, in such cases of future assessment, it is appropriate that risk aversion is included in values). However, 2 other problems arise; it is complex in use, and even when economists or other intelligent judges use it, they return scores with high variance. Nevertheless, scores should be of potential use for QLOS Index construction. This is a direct method - it asks for peoples values directly, not indirectly or inferred, like revealed methods, where the assumptions of perfect knowledge and choice are terminal problems.

The strength of the SG approach is its solid base in utility theory, and it is a classic method for recording preferences which explicitly account for uncertainty. The general method is as follows;

- Subjects choose between a guaranteed outcome of state A for T years or a chance (P) of being perfectly healthy for T years or dead;

- The chance (simple probability) is varied until the subject feels each option is equally desirable (the point of indifference).

However, the value obtained (probability of being healthy) is not a simple reflection of the severity of the original health condition. It also reflects risk aversity and gambling aversity.

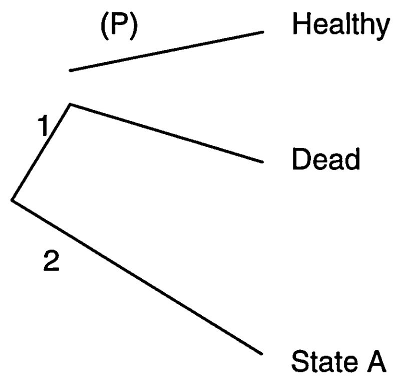


Figure E2. Illustration of Standard Gamble Approach to Measuring Quality of Life

E2.3 Time Trade Off (TTO)

TTO is a direct method - it asks for peoples values directly, not indirectly or inferred, like revealed methods. Subjects are required to consider a particular health state that is to last for a known length of time (Torrance et al, 1972, and Torrance, 1976b). Then, considering that a procedure will provide them with normal (perfect) health for a shorter period of time, followed by death or severe disablement, they are required to consider various time periods for this new normal health state, until they are indifferent about being in their current state or having the procedure. Another related approach is to require subjects to value other people's (societal) states, by describing state A and state B, and getting subjects to say how many people in state B would need to be helped (made better) in order to equal the good of treating one person in state A.

TTO is not derived from expected utility theory and requires the assumption that the subject's utility function for healthy years is linear with time - this assumption has been criticised. There is some evidence that people do not always make time-based decisions along these lines. This method has been described as a variant of magnitude estimation, and is simpler and more widely used than SG.

Using the SG method as a "gold standard" (due to its strong theoretical basis) Torrance found that studies using TTO and Category Scaling suggested the latter is less reliable than TTO. However, one author reports better reliability and validity in TTO when compared to SG (Bowling, 1995a). Time Trade Off is basically a variant of Willingness to Pay, except that time is used instead of money as the currency used to place the value. Since there are major problems with using money (not least, income bias, see Horne, 1996), and it could be expected that time is a more stable currency, it would appear to be a preferred method.

Potentially, useful scale points could be obtained, and some which may be of use have been. A general population sample was offered two choices for laryngeal cancer treatment and found that on average, individuals would trade off 14% of their life to avoid loss of speech (though there was a variance of between 17% for "executives" and 6% for fire-fighters; McNeil et al, 1981, cited in Spilker, 1990). However, Sintonen (1981) considered use of SG and TTO in health index construction to be inappropriate, due mainly to issues of complexity and (therefore) cost, and also that they would not be appropriate as methods to derive importance weights.

E2.4 Revealed Preference

To avoid the problems of hypothetical bias, for example in WTP, TTO and SG, some economists have attempted to directly value preferences by analysing data about actual actions, and using this to suggest what preference values are revealed by them. For example, if a population is found to choose an activity which exposes them to a 1 in

1000 higher probability of death but offers a £2,000 a year higher income, then, assuming linearity of risk, the population values life at 1000x2000 = £2 million.

Applying this to QLOS Index construction, revealed quality of life preferences could theoretically be used to generate values for different quality of life states. However, there are so many assumptions here associated with revealed preference approaches, not least that infinite choice is implied, along with usual free market conditions for zero inertia, perfect knowledge of risks and benefits, etc. Thus, it is difficult to see how such values could be justified on theoretical grounds.

Another approach to measuring implied values is to use current practice as evidence of values. For a simplified example, a decision to spend £400 million on a scheme to mitigate global warming-induced flooding to protect 400 people from certain death, implies a value of human life of £1 million. A number studies could be taken and the largest implied value would be that closest to the actual value. The key assumptions relate to the decision-making and policy process, namely, that it accurately reflect values, and that is efficiently translates those values. Therefore, the method is self-referencing, and if the decision-making and policy process were to fit this assumption then it would need to be already using some form of QLOS Index of values. This approach is therefore rejected on the grounds of weak theory.

E2.5 Comparison of Rating Scales and Economic approaches

Rating scales developed for economic valuation purposes provide a transition between HRQL scales and those economic methods involving direct money (WTP, CVM) or utility-based valuation. An example is where subjects place health state descriptions on a 0-100 scale (EuroQol, 1990). The major problems with this approach are that it has no theoretical basis, and it may be prone to bias such as response spreading.

In one study, the performance of an automated utility assessment instrument for measuring preferences for overall health was assessed (Nease et al, 1996). Using rating scale, time trade-off and standard gamble metrics, they assessed utilities for current health relative to perfect health and death. To validate the instrument, comparisons of utilities were made with the General Health subscale of the SF-36 Health Survey instrument (see Section E3), satisfaction with current health, and degree of bother due to current health. Utility for overall health was statistically significantly associated with the General Health subscale score and measures of satisfaction with current health and degree of bother. Other findings included substantial variation in utilities among patients with similarly severe overall health, and substantial overlap in utilities among subjects with different levels of overall health.

The importance of economic evaluation in the health sector has grown rapidly in the western world over the last decade. In the UK, decentralisation of National Health Service (NHS) budgeting and fundholding has led to intense interest in value for money aspects of health care provision. The commonwealth of Australia (1990) and in Canada, the Ontario Ministry of Health (1991) have issued guidelines to foster a standard approach to economic evaluation of health care (Fox-Rushby, 1994).

E2.6 WTP and CV

Contingent Valuation (CV) has received less attention in the HRQL field than in valuation of environmental losses, where it has been advanced by various neo-classical environmental economists and criticised by others (Horne, 1995). CV involves direct expression of preferences, and is theoretically well-developed, being centred on the estimation of consumer surplus as value of public goods. The two possibilities for eliciting such values from subject are willingness to pay for reduction in impact or willingness to accept for an increase in impact. Among the many problems are the fact that ability to pay affects valuations, and little is known about the psychometric properties of CV valuations.

E2.7 Economic Value Of Life

The body of literature which is devoted to economic valuation of life in money terms is of potential interest in the context of setting near-end points of the impact scale, for example, by placing a value on death (which is a major point close to the top end of the impact scale - only situations worse than death can score higher).

To illustrate the neo-classical economic approaches, Mooney (1986) documents a study by Culyer and Maynard (1981) where the value of life is used in valuing duodenal ulcer treatments. Three different figures were used for the value of the risk of death - the Department of the Environment's (then) value of life of £68,500, the expected productive output-derived value of life obtained using the human capital approach (£46,000), and the contingent valuation derived value that individuals were willing to pay to reduce their risk of death (£3 million). For one duodenal ulcer treatment option - the vagotomy - an operation with a case-fatality rate of 0.5% is required, giving a low estimate (including lowest value of life and treatment costs) of £1,180 and a high estimate (including the highest value of life and treatment costs) of £16,370. The alternative, cimetidine drug treatment, where there was no operation and therefore no risk of death, gave an equivalent cost range of £1,018-£1,239. It was noted that from the point of view of the NHS surgery seems the cheapest, whereas from the individual or community view, the drug (on this basis) would appear more attractive. However, the key issues for the purposes of developing the QLOS Index, are as follows;

- the human capital approach is flawed since it can only possibly capture a part of the total value of life;

- from an individual/community perspective, it would appear that, for any impact which results in a risk of death of the order of 0.5% or higher, this factor is likely to dominate the value of the impact - and probably means that impacts with any significant risk attached are clearly more highly valued than others.

E3. Existing Measures of Quality Of Life

The principal methods by which valuation of quality of life can be achieved are threefold; psychometric valuation by interview or questionnaire, analysis of behaviour, and utility measurement based on economic theory. The first is the most common method used in HRQL measurement literature. Single question or single-dimension measurements of quality of life are notoriously imprecise, and are rejected for use here on this basis, in favour of multi-dimensional measures. The scaling methods used include magnitude estimation, fractionation, equivalence and category scaling. The advantage over behavioural analysis is that the questions posed (the subject of valuation) can be specified precisely, and means to control the valuation experiment can be utilised, for example, sensitivity to known and applied variations can be measured. The classical economic approach is based on the Von Neumann-Morgenstern standard gamble, which can be extended to multi-attribute utility measurement. Multi-attribute utility theory can be used to reduce the theoretical number of combination health states suggested by a typical 5-6 quality of life-dimension set of data to a much shorter list of actual practical health states for inclusion in an index. For example, millions or billions of potential quality of life-state combinations can be reduced to 200 or so (Torrance, 1987). Some utility values which have been established by using economic techniques with the QALY approach have been summarised from the literature and these are reproduced in Table E1.

Some utilities for health states

Health state	Utility
Healthy (reference state)	1.00
Life with menopausal symptoms (judgment)	0.99
Side effects of hypertension treatment (judgment)	0.95–0.99
Mild angina (judgment)	0.90
Kidney transplant (TTO, Hamilton, patients with transplants)	0.84
Moderate angina (judgment)	0.70
Some physical and role limitation with occasional pain (TTO)	0.67
Hospital dialysis (TTO, Hamilton, dialysis patients)	0.59
Hospital dialysis (TTO, St John's, dialysis patients)	0.57
Hospital dialysis (TTO, general public)	0.56
Severe angina (judgment)	0.50
Anxious/depressed and lonely much of the time (TTO)	0.45
Being blind or deaf or dumb (TTO)	0.39
Hospital confinement (TTO)	0.33
Mechanical aids to walk and learning disabled (TTO)	0.31
Dead (reference state)	0.00
Quadriplegic, blind and depressed (TTO)	<0.00
Confined to bed with severe pain (ratio)	<0.00
Unconscious (ratio)	<0.00

Table E1. <u>Some Utility Values from the literature (Torrance, 1987)</u>

Note: TTO = derived using Time Trade-Off method, see Section E4 for further explanation

The remainder of this Section is structured around the measures under review, for ease of reference, and presented, generally in chronological order. For each of twenty general measures (selected on the basis of availability, general reputation in the literature, and intended application to general populations), a brief description is given, followed by an assessment of its suitability against key criteria. Any useful single scale points are also presented. Comparisons are presented in Section E4.

The 20 measures reviewed are;

1. Index of Activities of Daily Living (IADL)
2. McMaster Health Index Questionnaire
3. General Health Questionnaire (GHQ)
4. Sickness Impact Profile
5. Nottingham Health Profile
6. Rosser Classification of Illness States
7. Quality of Well-being Scale

8. SF-36
9. McGill Pain Questionnaire
10. Spitzer Quality of Life Index
11. Evaluation Ranking Scale
12. Healthy-Year Equivalent (HYE)
13. EuroQol
14. SEIQoL
15. Health Measurement Questionnaire (HMQ)
16. 15D
17. Health Utility Index
18. Index of Health-related Quality of Life (IHQL)
19. PCASEE
20. WHOQOL

E3.1 Index of Activities of Daily Living (IADL)

One of the oldest HRQL scales, this is not actually generic. It is a specific disability index developed by Katz et al (1963). The index is mainly concerned with basic self-care parameters such as bathing, toileting, dressing and feeding. Specific design means that the population range and the items of measurement are necessarily constricted, and could not be applied widely. On the basis of content (it could only detect relatively ill states) it is rejected.

E3.2 McMaster Health Index Questionnaire

This instrument goes back to 1970, but has been a long time in development, and has thus fallen behind more modern true indexes in terms of statistical reliability and validity and scale construction techniques. Weights have been derived for the measure but no significant difference was found between using the weights and not using them.

Physical, social and emotional scores are reported separately and no single index score is calculated, thus invalidating it for direct use in QLOS Index construction.

E3.3 General Health Questionnaire (GHQ)

The most commonly used rating scale technique is the Likert Scale, and the most widely-used Likert-type scale until the early 1990s was the GHQ, developed by Goldberg (1972). The scale was designed for detecting non-psychotic psychiatric disorder. This specific use and the fact that items are not weighted or intended for summing to a single score renders it unsuitable for use here.

E3.4 Sickness Impact Profile

Developed by Bergner et al (1976), this scale method is based on observations and behaviour rather than subjective feelings. In setting up the areas of response, a survey revealed 312 items, which were then grouped into 14 categories (one of which was emotions/feelings). After further piloting, the items were reduced to 146 in a short form survey. These were then scaled by intervals analysis of 15 point category ratings. Internal consistency is generally shown to be good. Also, the scale has been shown to discriminate for the purposes for which it was designed i.e. between sick and less sick people. It was shown in some early studies that it does not discriminate well between health status in the general population, and this led to the development of later scales for this purpose in the 1980s (for example, Nottingham Health Profile, SF-36).

The current SIP contains 136 items referring to dysfunction through illness in 12 areas or domains - work, recreation, emotion, affect, home life, sleep, rest, eating, ambulation, mobility, communication and social interaction. It has been found to be particularly useful in assessing illness impact in the chronically ill. The major problem is that it can only be used with people who are and believe themselves to be ill. The overall score is calculated by adding scale-weighted values and dividing the sum by the maximum possible dysfunction score and multiplying by 100 to give the SIP percentage. Two subscores can also be calculated, for the physical dimension (using ambulation, body care, movement and mobility) and the psychosocial dimension (using social interaction, alertness behaviour, emotional behaviour and communication). Used widely for approaching 25 years, the SIP has been adopted as a gold standard against which many more recent surveys have been compared. The UK equivalent, the Functional Limitations Profile (FLP) was adapted from the SIP where it was used in a study of disabled people in Lambeth. The modification consisted of rewording to enhance linguistic meaning and analysing them for conformity to usage in Britain.

Problems include the lack of measurement sensitivity. For example, questions are couched in can/can't terms, rather than do/don't terms, there is little attention to subjective aspects such as pain, and furthermore, many questions incorporate activities which require different levels of function, such as "I have difficulty doing housework, for example, turning faucets, kitchen gadgets, sewing, carpentry". Internal structure is also problematic in certain areas, for example, in the ambulation category, as one author states; "..the statement "I do not walk at all" logically precludes the statement "I walk more slowly". Consequently, it becomes logically impossible to affirm all statements, thus making the claim that scores range from 0-100% incorrect" (Williams, 1994).

One of the most problematic aspects of the SIP is the distinct possibility that logically inconsistent scores can result. For example, more disabled respondents may answer a smaller number of higher weighted mobility questions, while moderately disabled respondents answer many more moderately weighted questions, with the result that the latter appear more disabled in the final scores than the former. Furthermore, the item weights themselves suffer from similar problems to those for the NHP (see above), in that they are applied to closely related items, with little variation in significance of each (therefore, the magnitude of weights have low variance). Indeed, similar tests to those on the NHP, involving comparisons of weighted and non-weighted (or more accurately, equal-weighted) scores, suggested that the use of weighting in the SIP is of limited value. More significant than the item weights may be the implicit weighted embedded in the structure. Since categories contain different numbers of items and are summed separately prior to total summing to provide the index score, the more items there are in a category, the more items have to be selected in order to obtain the same weight as a category with less items. Last but not least, the SIP is about sickness, not quality of life, and thus has a content weakness for use in the QLOS Index.

E3.5 Nottingham Health Profile

The Nottingham Health Profile (NHP) was developed in the late 1970s, out of the need for an indicator of the typical effects of physical, social and emotional ill health on quality of life. The aim was to overcome problems with earlier scales such as the SIP, IADL and QWB. It is based on lay definitions of health, rather than relying on expert panels. It is relatively short, it is appropriate for both ill people and general populations, and it has undergone considerable testing for reliability and validity. Hunt, one of the original authors, was critical of existing measures because of length, ambiguity of statements, scoring and weighting methods which reflected physician values rather than respondents, narrowness of health concept, and theoretical flaws in indices which involved summing into a single score. Original interviews led to 2200 statements of ill health, which were then successively reduced by amalgamation, removal of overlapping terms and refining tests, to 38, grouped into 6 sections; physical mobility, pain, sleep, social isolation, emotional reactions, and energy level. Weighting of each item was undertaken using Thurstone's paired comparisons (a detailed account is given in Hunt et al, 1986). This suggests that the NHP is underpinned by good scaling theory. However, Jenkinson (1991) points out that it has been suggested that Thurstone's method was designed to measure attitude and is not appropriate for measuring factual statements, which the NHP contains (for example, "I am unable to walk at all"). Furthermore, statements do not cover the full range of the attitude continuum, and the authors themselves have noted that members of the normal population may affirm very few statements.

The major problems with the NHP appear with summing and weighting. Firstly, the authors stated that scores from different domains should not be summed into a single score, but later suggested that this could be done (Jenkinson, 1991). The main problem with weighting is the low variance of item weights. For example, the 3 items on energy have weights of 39.2, 36.8 and 24. The variance is so low that the full spread of states of energy cannot realistically be captured - Thurstone's original weights were of orders of magnitude difference, reflecting a far greater spread of attitude states. The NHP is therefore only capturing a narrow range of states in this respect. This in itself does not mean that NHP weights can not be used to indicate positions on a QLOS Index, only that one would expect these points to be bunched around one area of the scale.

Nevertheless, the NHP is short, easily administered and easily understood, and designed and proved to be of relevance to general populations. It only requires Yes/No answers to each of 38 Part 1 items (each is weighted to reflect aggregate value of the item). There is also a Part 2, which consists of a short list of social function questions. As the statements are negative (about ill-health, not health) some authors have suggested it is really an illness profile rather than a health profile.

The Nottingham Health Profile Distress Index is a 24-item shortened and adapted version of the original 38-item NHP. The index gives rise to a summed single figure, but for health-related distress, not quality of life. The original NHP covers a broader area of health related quality of life and incorporates a series of item weights, derived through application of Thurstone scaling. This is described as a two stage process; firstly, the determination of scale value for each statement, and secondly, derivation of weights from these (McKenna, Hunt and McEwen, 1981). A further example of the process is given elsewhere (Hunt et al, 1986).

However, not only has the use of Thurstone scaling been questioned here, since it was designed for application to psychological/subjective variables, not factual/objective series of questions as contained in the NHP, but the variance of item weights is also questionably low. Typically, weights within domains vary up to a factor of 1.5-2, and no weights reach a single order of magnitude in difference. One author shows that removal of weights and replacement with binary scores (i.e. equal weighting) makes very little difference to scores (Jenkinson, 1994).

NHP provides useful results in ill or unhealthy populations, but is not sufficiently sensitive at the “almost well” end of the spectrum, to distinguish between perfect health and slightly imperfect health, since it would record maximum scores for such respondents. Here lies an explanation for the low variance of weights; the NHP items are bunched at one end of the health-related quality of life impairment spectrum. Since all items are about severe impairments to health, they give rise to similar weights of severity under the application of Thurstone's method. Therefore, the conclusion is not that weights are useless for the NHP, as far as the current work is concerned. What is relevant is that the NHP is not sensitive enough to record a wide spectrum of health states among general populations. If it did, it could be expected to incorporate a much larger variance of weights, and thus make the weighting process more meaningful and significant.

E3.6 Rosser Classification of Illness States

First developed by Rosser and Kind (1978), this consists of a valuation matrix of 29 states of health derived from 8 degrees of disability and 4 levels of pain (see Table E2). The values were derived using a ratio-based method (i.e. questions of the type "how many times more ill is a person in state A than state B?"). For the original study, 6 widely dispersed illness states were selected (IC, IID, VC, VIB, VIIB, VIID), and each subject scored these 6 "marker states" relatively in pairs. The result was a ratio scale of marker states, which was used as a framework to rank all the remaining 23 states. After processing the results, a scale was produced, the Rosser Distress and Disability Index, with 0=death and 1=no impairment (see Figure below). A QALY score is generated by combining the expected time spent in each illness state, to produce a QALY score for the complete profile of progress through the matrix.

The Rosser classification and valuation matrix

		Distress rating			
	Disability rating	A. No distress	B. Mild	C. Moderate	D. Severe
I	No disability	1.000	0.995	0.990	0.967
II	Slight social disability	0.990	0.986	0.973	0.932
III	Severe social disability and/or slight physical impairment	0.980	0.972	0.956	0.912
IV	Physical ability severely limited (e.g. light housework only)	0.964	0.956	0.942	0.870
V	Unable to take paid employment or education, largely housebound	0.946	0.935	0.900	0.700
VI	Confined to chair or wheelchair	0.875	0.845	0.680	0.000
VII	Confined to bed	0.677	0.564	0.000	−1.486
VIII	Unconscious	−1.028	*	*	*

Scale defined so that: healthy = 1.0; dead = 0.0.
*Denotes invalid combination of disability and distress.

Table E2. Rosser classification and weights for illness states (From Kind et al, 1982)

It satisfies transparency and logic criteria, although questions have been raised about theoretical basis and reliability and validity of results. Also, although the states described are largely generic in nature, it may be argued that there are other aspects to quality of life-state than physical/mobility and pain/distress (indeed, criticism has been made that the spectrum and range of states, at only 29, is not sufficiently sensitive - however, given the need for simplicity and efficiency, it could be argued that such simple criteria subject to successful sensitivity testing would be acceptable).

Loomes and McKenzie (1989) point out that one assumption implicit in the method used to create the Rosser Index was that the quality of each state is independent of the time spent within it, and of the experience of any other states which may have preceded it or which may come after it. Evidence is cited suggesting that this is not the case. This also may challenge the logic of assuming that the approach here subscribes to the independence principle. In other words, the affect of environmental impact on quality of life is partly influenced by previous experience and the quality(s) of resultant state(s) is/are affected by preceding or proceeding states.

A further problem arises from the issue of risk and uncertainty. People were asked to compare the certainty of different states. Once values were derived, the scale authors then factored in risks (say, of success of a particular procedure) from objective data. Implicit here is either that people are assumed to be risk-neutral (i.e. there is no perceived element to risk which deviates from actual objective risk) or that attitudes to risk are disregarded.

As an update to the Rosser Classification, attempts have recently been made to expand the dimensions in the Index, to produce a more sensitive measure of global quality of life (Wilkin et al 1992, and IHQL above).

E3.7 Quality of Well-being Scale

The precursor to the QWB was the Index of Well-being (IWB), developed by Kaplan, Bush and Patrick, and consisting of function levels in 3 dimensions; mobility, physical activity and social activity. From these, 43 possible combination states were identified, and these are weighted to produce a single scale of weighted states of well-being. Notwithstanding that various criticisms have been made over apparent counter-intuitive weights possibly derived from unclear definitions of states or inaccurate scaling, the general approach would appear to be of potential use for QLOS Scale construction, as various health states are categorised on a single Index in terms of combinations of levels of mobility, physical activity and social activity. Public surveys were used to elicit weights for each level of mobility, and physical and social activity and a utility value was assigned to each function level. Questionnaire responses were then used to assign each subject to a given function state.

The resultant weights assigned to each function level within each area are presented below (for weighting method see also Kaplan et al, 1979). The index also incorporates "symptom/problem complexes", 25 of which are listed in what is purported to be an extensive list of perceived/subjective sources of loss of health status, which are complementary to the objective, measured function levels in the 3 areas (see Tables E3, E4 and E5).

Quality of Well-Being Scale elements and calculating formulas

Step no.	Step definition	Weight
	Mobility Scale (MOB)	
5	No limitations for health reasons	−.000
4	Did not drive a car, health related (younger than 16); did not ride in a car as usual for age, and/or did not use public transportation, health related; or had or would have used more help than usual for age to use public transportation, health related	−.062
2	In hospital, health related	−.090
	Physical Activity Scale (PAC)	
4	No limitations for health reasons	−.000
3	In wheelchair, moved or controlled movement of wheelchair without help from someone else; or had trouble or did not try to lift, stoop, bend over, or use stairs or inclines, health related, and/or limped, used a cane, crutches or walker, health related; and/or had any other physical limitation in walking, or did not try to walk as far or as fast as others the same age are able, health related	−.060
1	In wheelchair, did not move or control the movement of wheelchair without help from someone else, or in bed, chair, or couch for most or all of the day, health related	−.077
	Social Activity Scale (SAC)	
5	No limitations for health reasons	−.000
4	Limited in other role activity, health related	−.061
3	Limited in major (primary) role activity, health related	−.061
2	Performed no major role activity, health related, but did perform self-care activities	−.061
1	Performed no major role activity, health related, and did not perform or had more help than usual in performance of one or more self-care activities, health related	−.106

Table E3. Item weights for QWB (Kaplan and Bush, 1982)

List of Quality of Well-Being Scale symptom/problem complexes (CPX) with calculating weights for QWB scale, version 6B

CPX no.	CPX description	Weights
1	Death (not on respondent's card)	−.727
2	Loss of consciousness such as seizure (fits), fainting, or coma (out cold or knocked out)	−.407
3	Burn over large areas of face, body, arms, or legs	−.367
4	Pain, bleeding, itching, or discharge (drainage) from sexual organs—does not include normal menstrual (monthly) bleeding	−.349
5	Trouble learning, remembering, or thinking clearly	−.340
6	Any combination of one or more hands, feet, arms, or legs either missing, deformed (crooked), paralyzed (unable to move) or broken —includes wearing artificial limbs or braces	−.333
7	Pain, stiffness, weakness, numbness, or other discomfort in chest, stomach (including hernia or rupture), side, neck, back, hips, or any joints of hands, feet, arms or legs	−.299
8	Pain, burning, bleeding, itching, or other difficulty with rectum, bowel movements, or urination (passing water)	−.292
9	Sick or upset stomach, vomiting or loose bowel movements, with or without fever, chills, or aching all over	−.290
10	General tiredness, weakness, or weight loss	−.259
11	Cough, wheezing, or shortness of breath *with* or *without* fever, chills, or aching all over	−.257
12	Spells of feeling upset, being depressed, or of crying	−.257
13	Headache, or dizziness, or ringing in ears, or spells of feeling hot, or nervous, or shaky	−.244
14	Burning or itching rash on large areas of face, body, arms, or legs	−.240
15	Trouble talking, such as lisp, stuttering, hoarseness, or inability to speak	−.237
16	Pain or discomfort in one or both eyes (such as burning or itching) or any trouble seeing after correction	−.230
17	Overweight or underweight for age and height of skin defect of face, body, arms or legs, such as scars, pimples, warts, bruises, or changes in color	−.186
18	Pain in ear, tooth, jaw, throat, lips, tongue; missing or crooked permanent teeth—includes wearing bridges or false teeth; stuffy, runny nose; any trouble hearing—includes wearing a hearing aid	−.170
19	Taking medication or staying on a prescribed diet for health reasons	−.144
20	Wore eyeglasses or contact lenses	−.101
21	Breathing smog or unpleasant air	−.101
22	No symptoms or problem (not on respondent's card)	−.000
23	Standard symptom/problem (not on respondent's card)	−.257
24	Trouble sleeping	−.257
25	Intoxication	−.257
26	Problems with sexual interest or performance	−.257
27	Excessive worry or anxiety	−.257

* CPX 24-27 are assigned standard weights until empirical weights can be derived in new studies.

Table E4. Symptom-problem Complex Weights for QWB (Kaplan and Bush, 1982)

So, people are classified into various categories according to their responses, and each category effectively has a utility value (the weight). Once the process is complete, to calculate a well-being score, W:

$$W = 1 + (CPXwt) + (MOBwt) + (PACwt) + (SACwt)$$

where wt is the preference-weighted measure for each factor and CPX is the symptom/problem complex. The example below is cited in Kaplan and Anderson, 1990.

QWB Element	Description	Weight
CPX-11	Cough, wheezing, or shortness of breath, with or without fever, chill, or aching all over	−.257
MOB-5	No limitations	−.000
PAC-1	In bed, chair, or couch for most or all of the day, health related	−.077
SAC-2	Performed no major role activity, health related, but did perform self-care activities	−.061
	$W = 1 + (-.257) + (-.000) + (-.077) + (-.061) = .605$	

Table E5. Example Calculation of QOL-state (Kaplan and Bush, 1982)

Note that this calculation is for a particular day - the result can be multiplied by time to reflect the length of time the value is applicable.

The weights were derived from studies carried out in the San Diego area involving 400 subjects. Each subject rated each state on a scale of 1-15 equal appearing intervals, with death as zero and optimum function as 16. Judges tended to choose scores towards the centre of the scale. Equivalence scaling was also used, which has a better theoretical foundation in classical psychophysics, particularly for producing a ratio scale. Magnitude estimation was also used. In the event, the results were analysed and it was decided to adopt the category method to produce an interval scale, with statistical correction being applied to compensate for the data spread.

Median ratings for each case were taken as the dependent variable in regression analysis. Preferences were found to be stable over time, independent of prognoses, age or symptom/problem complexes and consistent across age/sex/status groups. Criticisms of the QWB weightings include vague case descriptions, inaccuracy in scaling method, and problems with model stability. These have been partly addressed, for example by Bush et al (1982). Indeed, validation was a prime consideration in the early stages of scale development, and this led to the development of validation procedures which remain classic to the present.

E3.8 SF-36

The SF-36 is the most researched and most used health status measure, and has the longest pedigree, having been developed out of the largest battery of health status measures ever reviewed (Brook et al, 1979). (Given the obvious link between valuation of health states and levels of insurance (and even actuarial studies) it is not surprising that there are considerable insurance company interests in the field of quality of life measurement. At its simplest, the task of working out how much compensation for loss of an arm versus a leg or an eye is, at its core, one of quality of life valuation.)

So far, the scale originators (Ware et al, 1993, 1994, 1995) have expressly argued against the possibility of summing dimension scores to form a single index score using the SF-36. However, the dimensions have now been reduced successfully to 2, in the amalgamation of the SF-36 into Physical component summary (PCS) and mental component summary (MCS) measures. According to Ware et al (1995), this makes it possible to reduce the number of statistical comparisons and thereby the role of chance in testing hypotheses about health outcomes. To test their usefulness relative to a profile of eight scores, results were compared across 16 tests involving patients (n=l,440) participating in the Medical Outcomes Study. Comparisons were made between groups known to differ at a point in time or to change over time in terms of age, diagnosis, severity of disease, co-morbid conditions, acute symptoms, self-reported changes in health, and recovery from clinical depression. The relative validity (RV) of each measure was estimated by a comparison of statistical results with those for the best scales in the same tests. Differences in RV, among the eight scales in the profile from the Medical Outcomes Study 36-Item Short-Form Health Survey (SF-36), were consistent with those in previous studies. One or both of the summary measures were significant for 14 or 15 differences detected in multivariate analyses of profiles and detected differences missed by the profile in one test. Relative validity coefficients ranged from .20 to .94 (median, .79) for PCS in tests involving physical criteria and from .93 to 1.45 (median, 1.02) for MCS in tests involving mental criteria. The MCS was superior to the best SF-36 scale in three of four tests involving mental health. Results suggest that the two summary measures may be useful in most studies and that their empirical validity, relative to the best SF-36 scale, will depend on the application. However, while the single summary score version of the SF-36 appears imminent, until it appears, unfortunately, this powerful health status measure cannot be used in QLOS Index construction.

E3.9 McGill Pain Questionnaire

Though not a full quality of life instrument, this is the most used Pain measurement questionnaire, and it does allow for total pain scores to be compiled through a system of weights. It has good reliability and validity and a potentially useful system of derivation of values. Restrictive content prevents use of values in the QLOS Index.

E3.10 Spitzer Quality of Life Index

Developed by Spitzer et al (1981), this is designed as a multinational, short, simple and widely applicable generic measure. However, it is specifically aimed at disease groups. Early pre-testing and correlation was primarily done with cancer patients in Australia, and it has been used primarily among cancer patients. It produces a single index score, but no weighting is applied and its disease-specific purpose and reports of weak theoretical validity render the instrument unusable for QLOS Index construction.

E3.11 Evaluation Ranking Scale

Content of this scale is inappropriate, as it is not a quality of life scale (Wilkin et al, 1992). However, its interest lies in the method; it has a 2-stage procedure, with patients first rating elements in order of importance and then attaching their own magnitude scaling weighting. This approach may be of use to quality of life measurement.

E3.12 Healthy-Year Equivalent (HYE)

This is a modified QALY method which overcomes its time-bias problems by fixing the quality of life measurement to life-span (Mehrez and Gafni, 1989). However, only limited data were available for this review. In general, it adopts the QALY approach, with modifications, as discussed in Section E1.

E3.13 EuroQol

The EuroQol is a relatively recent generic health-related quality of life scale designed to provide a single index value for each health state (EuroQol, 1990). However, early studies have suffered from low response rates, and have revealed low sensitivity, suggesting that it is rather a crude measure. It has a weak theoretical basis. The scale is based on 16 items chosen from 216 items assembled from a review of the GWB, SIP, NHP and Rosser Index. These cover 6 health domains; mobility, self care, main activity, social relationships, pain and mood. Psychometric comparisons with the UK

SF-36 showed highly skewed responses with over 97 of respondents recording ceiling scores for physical functioning, indicating a much stronger ceiling effect than the SF36. Therefore, it is unlikely to be applicable to healthy or near healthy population groups, and is therefore rejected on present information.

E3.14 SEIQoL

The Schedule for the Evaluation of Individual Quality of Life involves asking respondents what is important in their life (thereby giving content information). Then subjects rate their status against a vertical visual analogue scale (best to worst). They then rate their global quality of life on a similar horizontal version. Weights are then attached according to priority or frequency of responses for each item - results of a survey show that the method of weighting affects results (Bowling, 1995a). Although ratio scaled data was not provided, Table E6 gives some useful information into general priorities of respondents.

Items mentioned as most important in current life in priority order

	1st most important (n = 1968)		2nd most important (n = 1687)		3rd most important (n = 1113)		4th most important (n = 591)		5th most important (n = 243)	
	%	(No.)	%	(No.)	%	(No.)	%	(No.)	%	(No.)
Relationships with family/relatives	31	(602)	16	(264)	10	(112)	6	(38)	4	(11)
Relationships with other people	4	(69)	6	(103)	5	(54)	5	(31)	8	(18)
Own health	23	(460)	15	(246)	6	(71)	7	(41)	9	(22)
Health of someone close/responsible for	20	(397)	12	(197)	6	(62)	4	(23)	1	(3)
Finances/housing/standard of living	10	(192)	25	(430)	29	(322)	26	(156)	17	(41)
Environment (pollution, rubbish, noise, cleanliness, safety)	1	(16)	3	(53)	6	(62)	5	(28)	9	(21)
Conditions at work/job satisfaction	2	(36)	5	(81)	6	(69)	9	(52)	7	(21)
Availability of work/able to work	3	(59)	6	(101)	8	(95)	7	(42)	6	(14)
Social life/leisure activities	2	(43)	5	(88)	11	(119)	16	(92)	20	(48)
Religion/spiritual life	1	(21)	1	(19)	2	(21)	3	(15)	3	(8)
Education	1	(23)	2	(36)	3	(39)	3	(18)	4	(11)
Other*	2	(50)	4	(69)	8	(87)	9	(55)	12	(29)

*For example crime, politics/government, happiness/well-being, unspecified, etc.

Table E6. SEIQoL Weights (Bowling, 1995b)

E3.15 Health Measurement Questionnaire (HMQ)

This is a self-report survey for collecting data for processing into Rosser disability/distress categories (Kind and Gudex, 1991). As such, the scale points are provided by the Rosser Index, so no new scale points are produced (see below).

However, this classification method could potentially be adapted for use in the systematic framework. This potential will be examined in further work, where a classification/valuation method is developed for use with the QLOS Index.

E3.16 15D

Sintonen and Pekurinen (in Walker and Rosser, 1993) have developed this generic 15 dimensional measure of HRQL. With each dimension split into 4-5 levels and weights attached to each dimension and level, theoretically the measure now defines up to 10 billion states of quality of life on a single scale. The method used was a 2-stage approach, based on multi-attribute utility theory. Each respondent is required to ratio scale the dimensions and then ratio scale the levels within each dimension. Finally, a computer algorithm is used to combine social importance weights and level values to the levels checked by respondents and compute a 15D overall score. An additive aggregation rule is used to produce the final score from weighted dimension scores. Generic weights are not presented. While the approach is theoretically sound (notwithstanding a lack of weight of validity and reliability evidence), it is not envisaged for use in QLOS Index construction, on the grounds that it adopts a different (and possibly over-complex and sensitive) approach to measurement. However, if the QLOS Index were to prove insufficiently sensitive to value environmental impacts, this approach could be used.

E3.17 Health Utility Index

This encompasses 4 dimensions of HRQL; physical function, role function, social-emotional function and “specific health problems”. Based on a sample of the Canadian population, single attribute measurements were made (one dysfunction at a time), using category scaling. Multi-attribute measurements were made using time-trade off method. After data processing and transformation to make category scale values consistent with Time Trade-Off (TTO) values, a series of estimates of "disutility of dysfunction" were produced for different states (see Table E7). This has been used to estimate societal health of Canadians, and suggests that on average, Canadians would be willing to give up 6 years of life to completely eradicate themselves of the dysfunctional categories captured in the index (Holmes, 1995). This gives a general implication of the relative value of morbidity/non-fatal dysfunction against mortality (for example, total dysfunction is equal to about 5-10% of lifespan). The results were found to mirror similar values for the US population. The theory is based on QALY and economic TTO and so contains the weaknesses of these methods. In terms of Scale Points, the HUI does provide some scale points (sort out confusion of whether interval or ratio), and they are generic. Some evidence is presented of validity and reliability.

Dysfunctional category	Definition	Estimated regression parameter	Expected marginal disutility
Mobility	Trouble walking. climbing stairs. carrying objects or standing	0.402*	0.089
Severe Mobility	Completely unable to do at least one component of mobility category	0.077	0.107
Role	Activity limitation due to ill-health	0.685~	0.158
Severe Role	Permanently unable to work	0.260	0.226
Emotional	If unhappy or very unhappy	0.749*	0.174
Social	Less than the average number of contacts and visits with friends and family	0.061	0.013
Hearing	Trouble hearing normal conversation	0.144*	0.015
Sight	Trouble reading newsprint	0.l9l~	0.040
Severe perception	Completely unable to hear conversation or read newsprint	0.078	0.075
Short term incapacity	Any days in 2 week reporting period where normal activities had to be curtailed due to health	0.358*	0.078
Severe short term incapacity	Any days in 2 week reporting period spent bed-ridden due to health	0.259*	0.141
Agility	Trouble bending. grasping or reaching	0.104	0.022
Severe agility	Completely unable to do at least one component of agility category	1.08*	0.292
	*Denotes statistically significant at a 5% level of significance.	1.61*	

Table E7. Health Utility Index Estimates of Disutility of Dysfunction

E3.18 Index of Health-related Quality of Life (IHQL)

The IHQL is a relatively new instrument, which measures social, psychological and physical adjustment and combines these at 5 different levels of aggregation on a scale of utility values. The aggregation process provides single scores and it is claimed that accuracy and precision is preserved because of the multi-level scoring procedure. It also claims to combine the advantages of conventional HRQL rating scales and an index. Although the method appears to be based on Rosser's Index (see below) and is thus of interest, it has not been developed further and thus has not been widely tested for reliability and validity. Rosser's original 2-dimensional system is expanded into a 3-D system by splitting distress into "physical" and "emotional". A total of 175 quality of life states are produced as a result, weighted by using Standard Gamble methods (see Table E8).

Three-dimensional classification system: composite state valuations (0–1 scale of values)

		E1	E2	E3	E4	E5
P1	D1	1.000	0.970	0.894	0.791	0.643
	D2	0.990	0.960	0.884	0.781	0.632
	D3	0.971	0.940	0.864	0.762	0.614
	D4	0.946	0.917	0.840	0.738	0.590
	D5	0.917	0.887	0.811	0.710	0.561
	D6	0.885	0.855	0.780	0.678	0.530
	D7	0.838	0.804	0.729	0.628	0.481
P2	D1	0.944	0.915	0.838	0.736	0.588
	D2	0.934	0.904	0.828	0.726	0.578
	D3	0.915	0.885	0.810	0.708	0.559
	D4	0.891	0.861	0.785	0.684	0.537
	D5	0.861	0.831	0.756	0.654	0.508
	D6	0.829	0.799	0.724	0.623	0.477
	D7	0.779	0.750	0.675	0.574	0.427
P3	D1	0.867	0.837	0.761	0.660	0.513
	D2	0.857	0.827	0.751	0.650	0.503
	D3	0.837	0.808	0.732	0.631	0.485
	D4	0.814	0.784	0.709	0.608	0.461
	D5	0.785	0.755	0.680	0.579	0.433
	D6	0.753	0.723	0.648	0.548	0.402
	D7	0.702	0.674	0.598	0.498	0.353
P4	D1	0.714	0.685	0.610	0.510	0.365
	D2	0.703	0.675	0.599	0.499	0.354
	D3	0.685	0.656	0.581	0.481	0.337
	D4	0.661	0.632	0.557	0.458	0.313
	D5	0.632	0.604	0.528	0.429	0.285
	D6	0.601	0.572	0.497	0.399	0.254
	D7	0.551	0.522	0.449	0.350	0.207
P5	D1	0.468	0.439	0.365	0.267	0.125
	D2	0.457	0.428	0.355	0.257	0.114
	D3	0.439	0.410	0.337	0.239	0.097
	D4	0.416	0.387	0.314	0.216	0.074
	D5	0.387	0.358	0.285	0.188	0.047
	D6	0.356	0.327	0.255	0.159	0.017
	D7	0.308	0.279	0.207	0.111	-0.030

Table E8. <u>IHQL Quality of Life State Values (after Rosser et al, in Wilkin et al, 1992)</u>

Note: E1-5 - increasing emotional distress; P1-5 - increasing physical distress; D1-7 - increasing disability.

E3.19 PCASEE

Bech (1994) stresses that the PCASEE model aims to overcome the difficulties presented by the disparity between dimensions of functioning and quality of life in the majority of HRQL measures. The subjective PCASEE indicators are compared to the objective needs as presented in Maslow's hierarchy of human needs (Maslow, 1962). Drawing on this content model, and the self-anchoring based approach to quality of life, Bech (1993) developed a questionnaire for the PCASEE Model. Unfortunately, no data is available for this reason regarding Index states. The comparison of points on Maslow's hierarchy and the PCASEE Model is shown in Table E9.

Maslow's hierarchy of needs	PCASEE Model of Quality of Life
(objective needs)	(subjective indicators)
Physiological	Physical
Safety	Economic-social
Belongingness	Social
Self-esteem	Ego
Self-actualisation	Cognitive-emotional

Table E9. Comparison of Maslow's needs hierarchy and the PCASEE Model (adapted from Bech, 1994)

E3.20 WHOQOL

This is the World Health Organisation Quality of Life instrument (Kuyken and Orley, 1994). It has benefited form the experiences of earlier scales, and contains six domains (physical, psychological, independence, social relations, environment and beliefs/spirituality) and was developed from a theoretical framework of quality of life developed by the WHOQOL Group through lay focus groups. Questions were generated from eleven international field centres, pooled, reduced, and then ranked in each field centre. The results were compared and 276 relevant questions selected.

Anchor points were established for 4 Likert type response scales (evaluation, intensity, capacity, and frequency), and calibration of the adjectival descriptors was achieved using VAS trials in each field centre. The pilot instrument, containing 276 questions in 30 facets of quality of life, was trailed in 1994, with the aim of reducing the questions and facets and retaining the accuracy and sensitivity of the measure. Unfortunately, as no information is available about any attempts to combine scores or to otherwise produce single Index scores, it is thus rejected.

E4. Comparison

Three main summary Tables are presented. Table E10 contains summaries of the results of criteria analysis of the scales reviewed, Table E11 contains content descriptions of these scales, and Table E12 contains content descriptions of a sample of disease-specific scales (not described) for comparison purposes. It should be noted that the content analysis does not take the weight of questions into account but is designed to give a quick summary of the spread and focus of the subject area covered within the scale/questionnaire. This demonstrates that, while there are a wide range of subject areas and question types covered in the surveys sampled, all surveys contain similar core elements in their assessment of quality of life.

The criteria analysis shows that there is some theoretical weakness in many scales. Taken together with the evidence of similar subject coverage noted in the content analysis, this may indicate that content bias may be a factor in quality of life studies, having been all developed in a short space of time from a select group of researchers working with the same few theoretical texts. This underlines the need for theory. In Hays et al (1993) the authors discuss the trade-off between using preference weighted quality-of-life measures and quality-of-life instruments that offer precisely-measured states. Instruments that are weighted, such as EuroQoL, produce an overall score that is potentially useful in cost utility evaluations, the current process of weighting (for example, standard gamble, time trade-off, multi-attribute utility) limits the comprehensiveness with which an instrument can measure health states, because it requires that the number of health states assessed be restricted. The SF-36 (or the 15D), on the other hand, assess multiple health/quality of life domains with several possible levels in each, making the number of possible health states assessed extremely large. This precision in measuring health states defies application of standard health state preference weighting procedures. Efforts are underway to derive scores using empirical, rather than preference weighting, methods. The authors suggest that combining precise domain measurement with satisfactory methods of obtaining a summary score will advance the state-of-the-art.

	3.1 EuroQol	3.2 Evaluation Ranking Scale	3.3 General Health Questionnaire (GHQ)	3.4 Health Utility Index	3.5 Healthy-Year Equivalent	3.6 Health Measurement Questionnaire	3.7 Index of Activities of Daily Living	3.8 Index of Health-related Quality of Life (IHQL)	3.9 McGill Pain Questionnaire	3.10 McMaster Health Index Questionnaire	3.11 Nottingham Health Profile	3.12 PCASEE	3.13 Quality of Well-being Scale	3.14 Rosser Classification of Illness States	3.15 SEIQoL	3.16 SF36	3.17 Sickness Impact Profile	3.18 Spitzer Quality of Life Index	3.19 WHOQOL	3.20 15D
CRITERION 1																				
1. Is it (intended to produce) a single scale?						✓		✓						✓						
2. Does it produce single ratio scale points?				✓	✓						✓		✓	✓			✓	✓		
3. Does it produce ranked/interval scale points?				✓							✓				✓		✓			
4. Does it produce outputs that may contribute?																✓				
IF NO TO 1-4 - REJECT AT THIS POINT	✓	✓	✓				✓		✓	✓		?						✓	?	✓
CRITERION 2																				
What scaling theory does it draw on?				TTO/QALY	Mod. QALY	Rosser I.		Rosser/SG			Thurstone		Psych Theory	SG		Psych Theory	Psych Theory			
Transparent?						✓		✓			✓		✓	✓	✓	✓	✓			
Logical?						✓		✓			✓		✓	✓		✓				
Rational?						✓		✓			✓		✓	✓		✓				
Scaling Theory recognised				✓				✓			--		✓	✓		✓				
IS THE LEVEL OF THEORY ACCEPTABLE?				✓	✓			✓			--		✓	✓	?	✓	?			
CRITERION 3																				
Are all scale points/outputs generic?				✓		✓		✓			?		✓	✓	?	?	?			
CRITERION 4																				
Reliability (consistency)?				--				?			--		?	?	?	✓	?			
Validity/outcomes fit with theory?				--				?			--		?	?		✓	?			
Sensitivity of outputs?								✓			?		✓	?		✓	?			
Efficiency - are time/resource needs acceptable?								✓								?				
IS THE INTERNAL STRUCTURE ACCEPTABLE?				?				?								✓				
FURTHER INFO NEEDED?					✓	✓		✓				✓							✓	✓

Key: --= some, ?=not proven/unknown, ✓= yes

Table E10. Criteria Analysis of a Range of Health Related Quality of Life Survey Methods

	3.1 EuroQol	3.2 Evaluation Ranking Scale	3.3 General Health Questionnaire (GHQ)	3.4 Health Utility Index	3.5 Healthy-Year Equivalent	3.6 Health Measurement Questionnaire	3.7 Index of Activities of Daily Living	3.8 Index of Health-related Quality of Life (IHQL)	3.9 McGill Pain Questionnaire	3.10 McMaster Health Index Questionnaire	3.11 Nottingham Health Profile	3.12 PCASEE	3.13 Quality of Well-being Scale	3.14 Rosser Classification of Illness States	3.15 SEIQoL	3.16 SF36	3.17 Sickness Impact Profile	3.18 Spitzer Quality of Life Index	3.19 WHOQOL	3.20 15D
PHYSICAL FUNCTIONING:	✓	n/a		✓		✓	✓	✓		✓	✓	✓	✓	✓	n/a	✓	✓		✓	✓
Mobility	✓					✓	✓	✓		✓	✓		✓	✓			✓			✓
Ambulation/Body movement/phys dis						✓		✓					✓	✓			✓			
Eating								✓		✓			✓				✓			✓
Self Care	✓						✓	✓												
Vitality/Energy Level/Alertness/Fatigue								✓			✓		✓			✓	✓	✓		
Sleep								✓			✓						✓			✓
Senses								✓		✓			✓				✓			✓
Sex Life								✓			✓						✓			
Mental Capacity								✓		✓		✓	✓				✓		✓	✓
EMOTIONAL FUNCTIONING:	✓			✓		✓		✓	✓		✓	✓	✓	✓		✓	✓		✓	✓
Distress						✓		✓			✓		✓	✓						✓
Pain	✓					✓		✓	✓		✓			✓		✓				✓
Happiness								✓												
Well-being								✓				✓				✓				
Mood/Manic	✓							✓												
Anxiety			✓					✓			✓									
Depression			✓					✓		✓	✓		✓				✓			✓
Demeanour								✓										✓		
SOCIAL FUNCTIONING:	✓			✓		✓		✓			✓	✓	✓	✓		✓	✓		✓	✓
Social interaction, friends, rels, isolation	✓							✓		✓	✓		✓				✓	✓	✓	✓
Social Life								✓		✓	✓						✓			
Home life/family/daily living								✓		✓	✓							✓	✓	
Home Maintenance (and non-paid work)						✓		✓						✓			✓	✓		
Holidays											✓							✓		
Work (paid or non-specific)						✓		✓			✓		✓	✓			✓			✓
Interests and hobbies/recreation/leisure								✓		✓							✓			
Main activity/Role Functioning	✓			✓				✓		✓	✓		✓				✓		✓	
Social support/resources										✓	✓		✓				✓			
Communication								✓		✓	✓		✓				✓			✓
Behaviour (problems)/antisocial/social						✓		✓						✓			✓			
GLOBAL HEALTH/WELL-BEING/Quality of Life	✓															✓				
OTHERS;																				
Change in health	✓			✓				✓												
Perceived financial impact of illness																				
health perception								✓		✓	✓						✓	✓		✓
Financial difficulties												✓								
Expectations/perceived prognosis of								✓		✓							✓	✓		

Table E11. Content of a Range of 20 Generic Health Related Quality of Life Survey Methods

Note that clearly specific items listed under "additional aspects" (such as frequency of vomiting for cancer patients) is excluded as the purpose of the content analysis is to define the range of quality of life.

	Functional Living Index-Cancer	QLQ-C30	QLQ-C36	Spitzer QLOS	PULSES Profile	Barthel Index	Functional Status Rating System	Rapid Disability Rating Scale-2	Functional Status Index	Patient Evaluation Conf System	OECD Long-term Disability Q	Lambeth Disability Screening Q
Original designed application or target group	Cancer patients	Cancer patients	Cancer patients	Cancer patients	Disability (ADL)	Disability (ADL)	Disability (ADL)	Disability (IADL)	Disability (IADL)	Disability (IADL)	Disability (IADL)	Disability (IADL)
PHYSICAL FUNCTIONING:	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓
Mobility		✓			✓	✓	✓	✓	✓	✓	✓	✓
Ambulation/Body movement					✓		✓	✓	✓	✓	✓	✓
Eating												
Self Care		✓			✓	✓	✓	✓	✓	✓	✓	✓
Vitality/Energy Level/Alertness/Fatigue		✓										
Sleep			✓									
Senses					✓			✓		✓	✓	✓
Sex Life												
Mental Capacity					✓		✓	✓		✓		
EMOTIONAL FUNCTIONING:	✓	✓	✓									
Distress												
Pain	✓	✓	✓									
happiness												
well-being												
mood												
Anxiety	✓	✓	✓									
Depression	✓	✓	✓									
Demeanour				✓								
SOCIAL FUNCTIONING:	✓	✓	✓					✓	✓	✓	✓	✓
Social interaction, friends, rels, isolation	✓							✓			✓	
Social Life		✓										
Home life/family/daily living		✓		✓								
Home Maintenance (and non-paid work)	✓	✓		✓					✓	✓	✓	
Holidays												
Work (paid or non-specific)	✓	✓		✓						✓	✓	
Interests and hobbies/recreation/leisure	✓									✓	✓	
Main activity/Role Functioning	✓	✓	✓									
Social support/resources				✓							✓	
Communication								✓	✓		✓	✓
Behaviour (problems)									✓			
GLOBAL HEALTH/WELL-BEING/Quality of Life	✓	✓										
OTHERS;												
Change in health												
Perceived financial impact of illness			✓									
health perception				✓								
Financial difficulties		✓										

Table E12. Content of a Range of disease-specific Health Related Quality of Life Survey Methods

Note that clearly specific items listed under “additional aspects” (such as frequency of vomiting for cancer patients) is excluded as the purpose of the content analysis is to define the range of quality of life.

APPENDIX F. DEMONSTRATION OF THE VALUATION METHOD

The valuation method of the systematic framework has been developed as a means for valuing environmental impacts in terms of the quality of life outcome states they cause. here, this valuation method is demonstrated. Sections F1 and F2 introduce data requirements and the selection of trial impacts, respectively. Sections F3 and F4 present valuations and results. Comparisons are drawn with an alternative approach from the literature, and these are discussed in Section F5, before the demonstration is concluded in Section F6.

F1. Data Requirements

The identification of all RIOs, (Horne, 1996a, 1997) requires detailed knowledge of all stages in the fuel cycle and measurements of quantities and nature of all incidental outputs from the process at each stage. The pathway analysis method requires detailed knowledge of all receiving environments, how these receiving environments are changed by the introduction of each incidental output, and how combinations of incidental outputs act synergistically in the environment. The output of the pathway analysis method is an exhaustive list of all the environmental changes that occur as a result of the fuel cycle under study. These environmental changes are then assessed to produce a list of human consequences, and it is these human consequences which are then subject to valuation using the valuation method (Horne, 1999b).

For the purposes of this demonstration, data will be gathered on the basis of the list of environmental changes already having been established. This is because it is important that the environmental change and human impact data used in QLOS valuation is the same as in the comparable valuation. If changes were made or different data used to reflect the wider QLOS approach to establishing environmental change data, then this would affect the control aspects of the comparison between the two valuations (QLOS and comparable), thus invalidating any simple comparison between results.

F2. Selection of Trial Impacts

As stated above, it is important that a wide range of impact types are included in the trial, in order that the comparison of values can be made across this wide range. If only similar impacts are used in the trial, then the comparison would be limited. It is also important that necessary data are available to allow calculations using the QLOS valuation method, and furthermore, that data on valuations undertaken using other techniques (for comparables) are available. Therefore, a set of suitability criteria against which to compare potential trial impacts, based on these requirements, forms the basis for the selection exercise.

As discussed elsewhere, the European Commission ExternE project is the most comprehensive example to date of an external costing study for fuel cycles, and it adopts a methodology with some resemblance to that which is proposed here for identifying RIOs, pathways and environmental changes. Given this similarity, it is appropriate to use this as the reference study for this demonstration for the purposes of selecting trial impacts, data source and valuation comparisons. A large number of impacts were each assessed for suitability by establishing whether data were available, and what the relative magnitude values were, for each of the key QLOS variables; Index number (I), Risk Value (R), Time Value (T), Receptors Value (n). As a result, 10 impacts were selected, and the result of the assessment exercise for these is presented in Table F1. This shows that data are available for the selected impacts, and that they cover a wide range of impact types, as indicated by the range of magnitude combinations of the key variables. Thus, the objective of selecting a range of parameter magnitudes and combinations was achieved.

The number of valuations for the trial (10), was decided as a sample large enough to provide meaningful ranking-level comparison, and sufficient range of value types to allow examination of a range of issues within calculations, thus complying with a key aim of the demonstration. This number will therefore ensure a sufficiently broad basis for discussion of potential underlying reasons for similarities and/or differences in values. Also, it was decided to choose impacts from 3 different fuel cycles. Although this complicates the trial since it involves collecting data on three different projects, this provides a wider comparison than that which could be achieved within a single fuel cycle situation. As a result, three levels of comparison are attempted; between similar impact types in the same fuel cycle, between different impact types in the same fuel cycle, and between different fuel cycles.

Impact No.	1	2	3	4	5	6	7	8a	8b	9
Human Consequence (impact descriptor) / Criteria	Power Station related traffic; public accident deaths	Power Station related traffic; public accident major injuries	Power Station related traffic; public accident minor injuries	Power Station SO_2 emissions; agricultural crop losses; loss of livelihood	Power station: Global Warming; deaths associated with crop loss and starvation	Coalworkers' occupational health; mortality from lung cancer/radon exposure	Coalworkers' occupational health; death from Progressive Massive Fibrosis (PMF)	Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; fatal cancers	Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; non-fatal cancers	Amenity loss and annoyance caused by noise from wind turbine operation
IMPACT TYPE										
Index number (I)	H	M	M	M	H	L	H	H	H	L
Risk Value (R)	L	L	M	H	H	H	L	LL	LL	H
Time Value (T)	H	H	M	L	H	H	H	H	H	H
Receptors Value (n)	L	L-M	M	M	H	L	L	H	H	L
QLOS DATA AVAILABILITY										
Index number (I)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Risk Value (R)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time Value (T)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Receptors Value (n)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Table F1. Suitability Criteria and Trial Impacts Selected from the Reference Study (ExternE, 1995a to c)

Note 1: Impact type: H - High, L - Low, M – Medium; estimates of magnitude of parameter (see text).

Note 2: ✓ - data are available to allow an estimate calculation of this QLOS parameter.

F3. QLOS valuations

F3.1 The Valuation Method and Formula

The valuation method is discussed in detail elsewhere (Horne, 1999b). It is based on the existence of four key variables. The first is the intensity of the impact, expressed in terms of the resultant Quality of Life outcome state, and quantified by reference to the QLOS Index. The second is the probabilistic risk of the impact occurring. For example, if it is expected to occur, this value is 1, but if there is only an even chance of occurrence, the value is 0.5. Thus, the resultant QLOS Value will be risk-weighted. The third key variable is duration – the time for which the impact will occur. This is expressed in terms of the proportion of conscious life remaining, expressed mathematically as a fraction. An impact which is expected to be constantly experienced throughout the remainder of conscious life would therefore score 1, whereas, if it were only to affect, say one hour per day, every day, for the remainder of life, it would score 1/16, i.e. 0.0625, assuming that the receptor is asleep for an average of 8 hours per day and is conscious the rest of the time. The fourth key variable is the number of people (receptors) affected.

The QLOS Impact Value for any given Human Consequence is calculated by taking the product of the 4 key variables, as follows;

$$\text{Impact Value (Qu)} = I \times R \times T \times n$$

where; I = QLOS Index Classification, R = Probabilistic Risk,

T = Human Consequence duration/expected life-remaining, n = number of receptors,

Qu = QLOS units.

F3.2 Data

Reference study data are used in the valuations in this trial, in order to ensure that the results are not biased by variations in data on environmental changes, rather than values. As stated above, it is important that valuations and comparisons are derived using similar environmental change data, to ensure that it is the difference in method rather than data which is being measured.

The derivations for each specific datum are presented in the QLOS valuation calculation notes in Sections F3.3, F3.4 and F3.5 below. Note that Impact Numbers 1 to 7 are from the coal fuel cycle (Impact Numbers 1 to 5 being from traffic and emissions to air from the coal power station, and 6 and 7 from coal mining and coal handling, see Section F3.3), while Impact Numbers 8a and 8b are from the nuclear fuel cycle (see Section F3.4), and Impact Number 9 is from the wind power generation cycle (see Section F3.5).

In general, data are standardised in the reference study and can be used directly in QLOS valuation. However, some data requires interpretation or approximation, and this is explained wherever it is required, in the subsequent calculations.

F3.3 Coal Fuel Cycle: QLOS Calculations

The general data from the reference study (ExternE, 1995a, p.140) for the power plant are; 40 years design life, 1710 MW output, and 76% load factor, and, therefore, total output of 455.37 TWh (or 455.37×10^6 MWh).

The categories of accidental injury used by the reference study, from which much of the morbidity data are drawn, conform to the UK Reporting of Injuries Diseases and Dangerous Occurrences Regulations 1985 (RIDDOR), under which UK occupational accidents have been reported since April 1986. Three categories are distinguished as follows:

- fatalities,

- major injuries, defined to include major fractures, amputation, serious eye injuries, some causes of loss of consciousness and any injury requiring hospital treatment for more than 24 hours, and,

- minor injuries, defined to include other accidents responsible for the loss of more than three working days.

For transport accidents, the categories are similar. In this case, major injuries are generally those requiring a person to be detained in hospital, whereas minor injuries are those requiring other medical attention. The reporting is done by the police on the basis of observation at the scene of the accident, rather than by medical examination. For the purposes of categorising these on the QLOS Index (Horne, 1999a, 1999b), major injury could range from "severe distress, confined to wheelchair" (100, same as death) to "no distress, limited physical ability" (3.6). It is acknowledged that without further disaggregation of injury data, QLOS valuation will therefore necessarily be approximate, but, for the purposes of this exercise, it as assumed that "moderate distress, largely housebound" (10) is a reasonable mid-point score. For minor injuries, the range is less broad, and a mid-point of "mild distress, severe social/slight physical impairment" (2.8) is selected.

Impact Number 1: Power Station related traffic; public accident deaths

QLOS Index Number:

Death is rated as 100 on the QLOS Index.

Probabilistic Risk:

The reference study (ExternE, 1995a, p. 126-128) quotes UK Department of Transport figures on road safety. 5,217 deaths in a total of 401,307 million km implies an accident rate of 0.013 deaths/M vehicle km. Reference study estimates of 1,070 traffic movements per day, averaging 24km, gives 9.373 Mkm per year arising from the power station. Average expected deaths per year arising from the power station are therefore;

9.373 x 0.013 = 0.122 deaths per year, or (x40) 4.874 deaths in the lifetime of the plant. Since the number of receptors at risk is not given in the reference study data, this risk-weighted figure is effectively a composite of R and n. Probabilistic Risk is therefore entered as 1, for the purposes of this exercise.

Conscious Life Remaining:

Death affects all of remaining life, therefore T is 1.

Number of receptors:

As stated in probabilistic risk above, the risk-weighted receptor value is 4.874. Although this is not the true number of receptors of the risk, since this figure is not available, the value for number of receptors killed is used, 4.874.

Valuation:

Impact Value (Qu) = I x R x T x n

100 x 1 x 1 x 4.874 = 487.4 Qu

To normalise this value:

Total output = 1710 x 8760 (hours per year) x 40 (lifetime) x 0.76 (load factor) = 455.4 TWh (or 455.4×10^6 MWh)

Therefore, normalised value;

487/455.4 = 1.07 μQu/MWh

Impact Number 2: Power Station related traffic; public accident major injuries

QLOS Index Number:

As stated above, it is assumed that major injuries can be assigned an average QLOS Index number of 10.

Probabilistic Risk:

Using the same set of UK Department of Transport statistics as cited in Impact Number 3 above, but for major injuries (total 60,441 per year), implies an accident rate of 0.147 major injuries/M vehicle km.

Average expected major injuries per year arising from the power station are therefore; 9.373 (total vehicle Mkm from the power station per year, see above) x 0.147 = 1.378 major injuries per year, or (x40) 55.113 major injuries in the lifetime of the plant. Since the number of receptors at risk is not given in the reference study data, this risk-weighted figure is effectively a composite of R and n. Probabilistic Risk is therefore entered as 1, for the purposes of this exercise.

Conscious Life Remaining:

The definition of major injury is "major fractures, amputation, serious eye injuries, some causes of loss of consciousness and any injury requiring hospital treatment for more than 24 hours". This may not last for all of remaining conscious life. However, no data are available on this. An assumed average figure of 55% is therefore adopted, based on the notion that half of the injuries are permanent and half affect only 10% of remaining life (for example, provided by 3-4 years of illness or hospitalisation spread over remaining life). T is therefore 0.55.

Number of receptors:

As stated in probabilistic risk above, the risk-weighted receptor value is 55.113. Although this is not the true number of receptors of the risk, since this figure is not available, the value for number of receptors with major injuries arising is used, 55.113.

Valuation:

Impact Value (Qu) = I x R x T x n

10 x 1 x 0.55 x 55.113 = 303.12 Qu

Giving a normalised value;

303.12/455.4 = 0.666 μQu/MWh

Impact Number 3: Power Station related traffic; public accident minor injuries

QLOS Index Number:

As stated above, it is assumed that minor injuries can be assigned an average QLOS Index number of 2.8.

Probabilistic Risk:

Using the same set of UK Department of Transport statistics as cited in Impact Number 3 above, but for 275,483 minor injuries per year, implies an accident rate of 0.669 minor injuries/M vehicle km.

Average expected minor injuries per year arising from the power station are therefore; 9.373 (total vehicle Mkm from the power station per year, see above) x 0.669 = 6.271 minor injuries per year, or (x40) 250.821 minor injuries in the lifetime of the plant. Since the number of receptors at risk is not given in the reference study data, this risk-weighted figure is effectively a composite of R and n. Probabilistic Risk is therefore entered as 1, for the purposes of this exercise.

Conscious Life Remaining:

The definition of minor injury is "other accidents (other than major injuries described above) responsible for the loss of more than three working days". This may not last for all of remaining conscious life, however, no data are available on this. An assumed average figure of 2.5% is therefore adopted (i.e. for an average person with 40 years expected life remaining, the minor injury will affect 1 year of remaining life). T is therefore 0.025.

Number of receptors:

As stated in probabilistic risk above, the risk-weighted receptor value is 250.821. Although this is not the true number of receptors of the risk, since this figure is not available, the value for number of receptors with minor injuries arising is used, 250.821.

Valuation:

Impact Value (Qu) = I x R x T x n

2.8 x 1 x 0.025 x 250.821 = 17.56 Qu

Giving a normalised value;

17.56/455.4 = 0.039 μQu/MWh

Impact Number 4: Power Station SO_2 emissions; agricultural crop losses; loss of livelihood

For crop damage, the reference study adopts a classic damage cost approach, seeking to establish the economic loss associated with lower crop yields. Therefore, the impact is not characterised as "loss of food supply" but "loss of money". This is reasonable, since there is no undersupply of wheat and barley at present. For a QLOS valuation, however, it is necessary to establish the <u>human</u> impact, in this case, arising from the economic loss.

Reference study figures (ExternE, 1995a, p.166-171) are based on many assumptions, since there is a dearth of pathway data in this area. The QLOS valuation, as with others in this trial study, incorporate these assumptions unless specifically mentioned, in order to minimise the effect of differential approximation and assumption in the values calculated.

The general scenario is that the environmental change is the yield response of wheat and barley plants to increased SO_2. This is modelled using crop distribution data, information from pertinent indoor and outdoor trials where SO_2 concentrations are known, data for SO_x and NO_x emissions based on the power station described above, fitted with 90% efficient FGD (13,000 and 25,000 t/year respectively).

Two models were used, one for the immediate 100km grid square and another for other distant 100km grid squares across the UK (effects outside the UK were not assessed). Best estimate results were losses of wheat and barley 0.49% and 0.5% respectively in the immediate grid square and 0.13% and 0.06% for total UK national yields. Using international prices (which are less biased by subsidies) best estimates of economic losses are calculated from this as 162,000ECU for wheat and 87,000ECU for barley.

For the QLOS valuation, actual impact on UK farmers is required, and therefore the figures for domestic prices are used, giving total economic losses of 233,000ECU for wheat and 197,000ECU for barley (ExternE, 1995a, p.171). If the data and modelling had been undertaken for the QLOS valuation, then the number of farmers affected and extent (effect of given loss of turnover/profit on QLOS) would be established as part of the pathway mapping and human consequence data collection exercise. However, these data are not available. Therefore, the following assumptions have been made here to calculate the QLOS of the economic loss involved:

- Total bankruptcy/loss of livelihood would result in “severe distress, slight social disability” (I = 8.8);

- The dispersed effect of incremental crop loss across the UK is equivalent in overall impact to a concentrated effect of total crop loss of the same total loss;

- An assumed average arable farm of size 500ha would employ 4FTE staff, and would have gross output and income as in Table F2, and;

- For farm workers displaced by crop loss, it would take 5 years for them to retrain, readjust and find alternative economic subsistence such that quality of life outcome returns to previous level.

	Gross production (no SO_2 damage)* (t)	UK crop prices (ECU/t, 1990)	Gross income (ECU, 1,000s)	Total SO_2 damage loss (t)	Total SO_2 damage economic loss (,000ECU)	Equiv. ha loss (damage loss/av.yield)	Equiv. FTE staff loss**
100% Wheat	3320	158.4	526	1471	233	221.5	1.77
100% Barley	2545	156.8	399	1258	197	247.2	1.98

Table F2. Output and Income Losses Due to Wheat and Barley Crop Losses

Notes:

*based on average yields of 6.64 t/ha wheat and 5.09 t/ha barley

**based on equiv. ha loss/total ha. x total staff. For example, for wheat: 221.5/500 x 4 = 1.77

QLOS Index Number:

As stated above, the assumed QLOS Index category is "severe distress, slight social disability" (I = 8.8).

Probabilistic Risk:

The risk is 1, since the impact is expected. It should be noted that, based on the dispersion models for SO_2 pollution, the recipients can be predicted. It should be noted that this differs from the Probabilistic Risk value for Impact Numbers 1 to 3 above, where 1 is entered because n is risk-weighted and the data cannot be disaggregated to provide a true value for R. For Impact Number 4, R is 1 because the particular crops and extent of damage involved can be predicted with high confidence using the dispersion model.

Conscious Life Remaining:

T is based on how long it would take to retrain and gain alternative employment, etc. such that the impact (severe distress) subsides. It is estimated here that, for an average worker age of 40, this would take 5 years, or 5/45 = 0.11 T.

Number of receptors:

Based in the general assumptions and calculations presented in Table F2 above, the combined livelihood loss suggested by this scenario is 3.75 FTE staff.

Valuation:

Impact Value (Qu) = I x R x T x n

8.8 x 1 x 0.11 x 3.75 = 3.63 Qu

Giving a normalised value;

3.63/455.4 = 0.008 μQu/MWh

Impact Number 5: Power station: Global Warming; deaths associated with crop loss and starvation

Firstly, it should be noted that understanding and data on environmental pathways leading to global climate change are patchy, and that even the most conservative tentative calculations in the literature suggest that this is the single dominant externality associated with coal fuel cycles. Unfortunately, many studies do not attempt to value lives lost due to this externality, although some attempts have been made to measure economic losses using crude versions of the "Willingness to Pay" approach, where, invariably, developing world populations cannot "pay as much to avoid" and therefore deaths there are valued an order of magnitude lower than in the developed world (such calculations are usually undertaken by those in the developed world).

The reference study uses two models to attempt to establish environmental changes, and cites several valuation studies which have attempted to value either deaths or economic damage resulting. One study (Hohmeyer and Gärtner, 1992) suggests values much higher than the others (5030mECU/kWh assuming a 0% Discount Rate compared to Cline's (1992b) equivalent central estimate of 14.9 mECU/kWh). However, this study is based on equitable value of life for all humans and therefore is closest to the QLOS approach, and is the one used for comparison here. It is based on 45 million crop failure-related deaths.

QLOS Index Number:

Death is 100 on QLOS Index.

Probabilistic Risk:

The risk is 1, since the impact is expected.

Conscious Life Remaining:

Death affects all of remaining life, therefore T is 1.

Number of receptors:

Due to lack of transparency in the reference study calculations, the contribution of the power station as a proportion of total (2.5K) global temperature rise is not stated, but can be derived from data quoted based on Hohmeyer and Gärtners' (1992) estimate (ExternE, 1995a, p.284) as follows:

Total economic loss = 5030 mECU/kWh x 455.4 TWh (total output) = 2.3×10^{12} ECU

The reference study (ExternE, 1995a) uses a standard Value Of Statistical Life (VOSL – the economic value attributed to each person) of 2.6MECU, so this implies the total life-equivalent loss of the power station is $2.3 \times 10^{12}/2.6 \times 10^{6}$ = 884,615 lives.

Valuation:

Impact Value (Qu) = I x R x T x n

100 x 1 x 1 x 884,615 = 88.462×10^{6} Qu

Giving a normalised value;

88,460/455.4 = 194,250 μQu/MWh

Impact Number 6: Coalworkers' occupational health; mortality from lung cancer/radon exposure

QLOS Index Number:

Death is rated as 100 on the QLOS Index.

Probabilistic Risk:

Risk is stated as 0.002 (based on US estimates).

Conscious Life Remaining:

A T value is not used in the reference study valuation, so a figure is derived here. T is not 100, since death is delayed, so is based on the following estimate:

- average age of victim 40;
- average life exp. of pop. 80;
- av. time lost due to lung cancer (amongst those affected) 15 years.

The reference study calculates an average 0.03 years lost per individual, which translates into 120 years lost in total amongst the 4,000 workers involved. Since the risk-weighted number of people affected is 8 (R x n), this means that each of these individuals will lose 15 years of life on average (120/8).

Therefore; T =15/40 = 0.375.

Number of receptors:

An estimated 4,000 coal workers will be exposed to the risk to provide the coal.

Valuation:

Impact Value (Qu) = I x R x T x n

Impact Value = (100 x 0.002 x 0.375 x 4,000 = 300 Qu

Normalised value = 300/455.4 = 0.6588 μQu/MWh.

Impact Number 7: Coalworkers' occupational health; death from Progressive Massive Fibrosis (PMF)

The reference study (ExternE, 1995a, p. 141-2) assumes a coalmining population of 4,000 men, entering the industry at age 18 years, and working for 40 years exposed throughout to a dust concentration of 2 mg/m^3 respirable coalmine dust, and not consistently exposed to unusually high quartz concentrations. Central estimates of the numbers of men showing PMF or CWSP sufficiently severe to obtain compensation in Britain, i.e. Category 2 or 3 CWSP, are drawn from work involving coal of 86.2% carbon; i.e. slightly higher than the 84.8% carbon content of the coal planned for use in the trial study project. Assuming a diverse mining population, so that the occurrences are spread evenly throughout the 40 year period of the power plant, 1.2 miners per year are expected to contract PMF.

The reference study (ExternE, 1995a, p.141) also notes that, while, in principle, further occurrences of advanced CWSP or of PMF might be expected in this population after retirement, these estimates as they stand are already unrealistically high when compared with the prevalence data published in recent British Coal Medical Service annual reports. The estimates refer to prevalences at the end of a 40-year working life, and men who meet this criterion are only a small proportion of miners.

To derive an estimate of death impact (it is noted that neither the QLOS nor the reference study estimate include health endpoints other than death), work on long-term coal miner mortality is quoted reporting 21% of men having pneumoconiosis registered as the cause of death.

QLOS Index Number:

Death is rated as 100 on the QLOS Index. (Note: the long phase of disease leading up to death is not included or quantified here in order to retain comparability with the reference study).

Probabilistic Risk:

Since 21% of miners having pneumoconiosis have this registered as the cause of death, probabilistic risk for those with PMF is 0.21.

Conscious Life Remaining:

T is not 100, since death is delayed, so is based on the premise that deaths are distributed evenly from the point of diagnosis to the point of death, giving an average T of 0.5. It is accepted that this is a rough, assumed estimate, and that the true picture is likely to be skewed towards death later rather than earlier. However, since no deaths are considered after retirement age, which provides an underestimate to the valuation, this expected overestimate effect is accepted.

Number of receptors:

This is restricted to those who contract pneumoconiosis, and is expected to be 1.2 per year, or 48 over the lifetime of the project.

Valuation:

Impact Value (Qu) = I x R x T x n

Impact Value = 100 x 0.21 x 0.5 x 48 = 504 Qu

To normalise this value:

504/455.4 = 1.107 μQu/MWh.

F3.4 Nuclear Fuel Cycle: QLOS Calculations

The reference study from which the data in this Section are drawn is based on a nuclear fuel cycle incorporating the Tricastin 900MWe PWR located in Pierrelatte, France (ExternE, 1995b). Two impacts are calculated here, although they have been combined since the reference study value is expressed as a combined value, so for comparison purposes, the QLOS valuations are also combined into the single Impact Number 8. Acceptability of combining QLOS values as a general practice for disparate impacts is not implied by this case.

Impact Number 8a: Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; fatal cancers

QLOS Index Number:

Death is 100 on QLOS Index. (Note: the long phase of disease leading up to death is omitted here in order to retain comparability with the reference study).

Probabilistic Risk:

Two elements to probabilistic risk are used in the reference study; annual core melt probability ($5x10^{-5}$) and conditional probability of this event leading to massive containment release (0.19), giving a combined annual probability of $9.5x10^{-6}$. Over a reactor life of 30 years, this would accumulate to a total probability of $2.85x10^{-4}$.

Conscious Life Remaining:

Death affects all of remaining life, therefore T is 1.

Number of receptors:

Reference study estimates are a total 291,200 man.Sv collective dose, giving rise to a total of 14,537 fatal cancers across the locality and region.

Valuation:

Impact Value (Qu) = I x R x T x n

$100 \times 2.85x10^{-4} \times 1 \times 14{,}537 = 414.3045$ Qu

To normalise this, since these particular data are derived from a 1200MWe reactor, total output is given by The reference study (ExternE, 1995b) as 7.6 TWh/reactor.year, and reactor life is 30 years, this gives a total lifetime output of 228 TWh. This gives a normalised value;

$414.3045/228 = 1.817\ \mu Qu/MWh$

Impact Number 8b: Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; non-fatal cancers

QLOS Index Number:

Non-fatal cancers covers a potentially broad range of QLOS categories, from, say, "slight physical impairment, mild distress" (2.8) to "severe distress, confined to bed" (248, and worse than death). Therefore, without further information on the type of cancers and their effects on distress and disability, even an estimate can only be suggested with serious doubts over accuracy. With this in mind, a tentative central estimate of 10 ("unable to work/largely housebound, moderate distress") is adopted.

Probabilistic Risk:

Risk of occurrence of accident is as for Impact Number 8, i.e. 2.85×10^{-4}.

Conscious Life Remaining:

Since radiation cancers are incurable, T is 1.

Number of receptors:

The reference study estimates are a total 291,200 man.Sv collective dose, giving rise to a total of 34,889 non-fatal cancers across the locality and region.

Valuation:

Impact Value (Qu) = I x R x T x n

$10 \times 2.85 \times 10^{-4} \times 1 \times 34{,}889 = 99.434$ Qu

Giving a normalised value;

$99.434/228 = 0.436\ \mu Qu/MWh$

F3.5 Wind Power Generation Cycle: QLOS Calculation

The reference site selected from those dealt with in the reference study (ExternE, 1995c) is at Delabole. The wind farm is an array of 10 x 400kW turbines with an estimated load factor of 29%, giving 4,000 x 0.29 x 24 x 365 = 10.16 x 10^6 kWh/yr. For the purposes of QLOS valuation, the assumed life-span is 40 years, giving a total output of 406,464 MWh (0.406 TWh).

Impact Number 9: Amenity loss and annoyance caused by noise from wind turbine operation

An immediate problem for the QLOS approach is lack of data for assessing the QLOS value for different levels of noise increase, type and trace patterns. However, a dual approach is adopted to valuation of this impact in the reference study, and one of the approaches may assist in providing general estimates of some value. The first approach is a classic hedonic pricing valuation, involving comparison of noise profiles and annuitised house prices. This technique is used to establish noise values by using property prices as surrogates. The second approach is more direct, using "annoyance costing", where a formula is used to predict probabilities of being "highly annoyed" at various noise levels. Questions are raised about this approach and the formula, which, given the low-density of rural populations around the reference sites, predicts very low numbers of "highly annoyed". The reference study then adopts the hedonic method, on the grounds that it is the method most used elsewhere. It does not incorporate the low annoyance probabilities; it is based on the difference between background and new noise levels (so low rural background levels can give rise to relatively high impacts), and because it gives rise to higher externality values, which are a safer upper limit.

In fact, notwithstanding the potential flaws of the formula for predicting probability of noise annoyance, the more direct approach (rather than using surrogate hedonic pricing) is more similar to QLOS in concept. The approach to the QLOS valuation does, however, combine elements of both approaches.

QLOS Index Number:

The problems for the QLOS estimate here are lack of transparency in the reference study calculations and lack of data on exposure, level of distress, etc. As the maximum increase for the Delabole reference site is 5dB(A) (1 property) while most are below 1 dB(A), the assumed mean QLOS value is "mild disturbance" (0.1, an estimate QLOS value suggested in Horne, 1999).

Probabilistic Risk:

Probability of perception is high and close to 1 (probably only excluding the deaf), for those who are within the elevated noise level area. The value assumed here is 0.99.

Conscious Life Remaining:

Duration is calculated on the assumption that residents on average are exposed to the noise and are disturbed by it for 10% of their conscious time. This includes periods of non-operation, periods at work/elsewhere or indoors away from the noise, etc. Since assumed period of operation is 40 years, working on an average age of 40 and life expectancy of 85, $T = 0.1 \times 40/45 = 0.089$.

Number of receptors:

The reference study states that "fewer that 20 houses experience an increment in excess of 1 dB(A). At the nearby centres of population, Delabole and Camelford, the noise increment is typically only 0.1dB(A)" (ExternE, 1995c, p.47). On this basis, it is estimated that 40 people are affected.

Valuation:

Impact Value (Qu) = I x R x T x n

0.1 x 0.99 x 0.089 x 40 = 0.3524 Qu

Giving a normalised value;

0.3524/0.406 = 0.868 μQu/MWh

F4. Results

The results of the QLOS valuation calculations, along with comparables from the reference study (ExternE, 1995a to 1995c) are presented in Table F3. The valuations were carried out in 4 Stages. First a group of similar type impacts (Impact Numbers 1 to 3) were valued and compared. Then, this was extended to other, different impact types (Impact Numbers 4 and 5) within the same part of the fuel cycle under study (the coal-fired power station). In Stage 3, this was extended to 2 further impacts, from a different part of the fuel cycle – the coal mining process. Finally, in Stage 4, valuations were undertaken of impacts from entirely separate projects involving Nuclear and Wind fuel cycles (Impact Numbers 8 and 9). Explanation of the results is therefore presented for each of these Stages of the work, below, with more general discussion presented in Section F5.

	Stage 1			Stage 2		Stage 3		Stage 4		
Impact Number	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8a**	**8b**	**9**
Human Consequence / Criteria	Power Station related traffic; public accident deaths	Power Station related traffic; public accident major injuries	Power Station related traffic; public accident minor injuries	Power Station SO_2 emissions; agricultural crop losses; loss of livelihood	Power station: Global Warming; deaths associated with crop loss and starvation	Coalworkers' occupational health; mortality from lung cancer/radon exposure	Coalworkers' occupational health; death from Progressive Massive Fibrosis (PMF)	Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; fatal cancers	Nuclear fuel cycle; ST2 core melt accident with massive containment release scenario; non-fatal cancers	Amenity loss and annoyance caused by noise from wind turbine operation
QLOS DATA										
Index number (I)	100	10	2.8	8.8	100	100	100	100	10	0.1
Risk Value (R)	1*	1*	1*	1	1*	0.002	0.21	2.85×10^{-4}	2.85×10^{-4}	0.99
Time Value (T)	1	0.55	0.025	0.11	1	0.375	0.5	1	1	0.089
Receptors Value (n)	4.874	55.113	250.821	3.75	884,615	4000	48	14,357	34,889	40
TOTAL (Qu)	487.4	303.12	17.56	3.63	88.46×10^{6}	300	504	414.3	99.43	0.3524
QLOS VALUE (μQu/MWh)	1.07	0.666	0.039	0.008	194,250	0.659	1.107	1.817	0.436	0.868
COMPARABLE VALUE (mECU/KWh)	0.029	3.4×10^{-3}- 26.4×10^{-3}	0.23×10^{-3}- 1.87×10^{-3}	0.022	5030	0.047	0.057	0.068 combined value		1.1

Table F3. QLOS valuation Results and Comparables

Notes:

*Not the true value for risk, but due to data shortcomings, only a risk-weighted value for n is available, so the value for R is embedded in n.

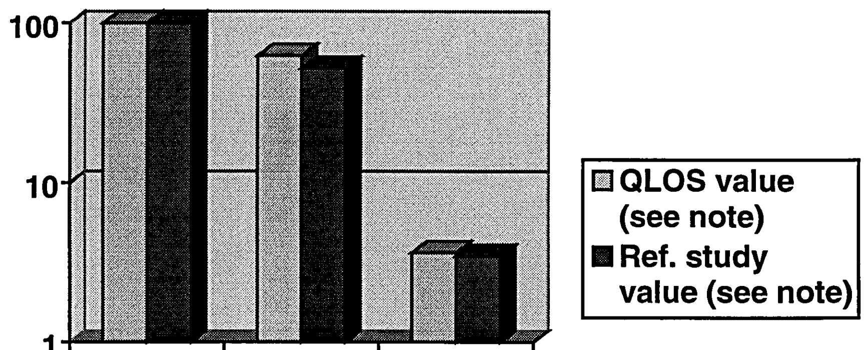
QLOS Value is normalised to output and expressed in μQu/MWh (Qu is QLOS unit)

The Comparable Value for Impact Number 5 is the upper value from the literature, and is approximately 1.5 orders of magnitude above the median value stated in the reference study (ExternE, 1995a).

The Comparable Value for Impact Number 8 is the combined value of Impact Numbers 8a and 8b. It also includes minor additional elements for hereditary effects, early diseases and deaths (see ExternE, 1995b, p.204-5) although the combined proportion of these elements is low.

F4.1 Stage 1: Coal-fired Power Station; traffic accidents

Impact Numbers 1 to 3 show a similar rank for both the QLOS and reference study values. Seeking closer comparison beyond simple ranking becomes increasingly problematic, since reference study and QLOS units use different scales which are not calibrated together. However, since each scale is ratio type, it is possible to normalise both sets of values, for example, on a 0-100 scale. Thus, Figure F1 shows Impact Numbers 2 and 3 normalised to "Impact 1 = 100" for both reference study and QLOS values. This shows that there is close proportional correlation between major injury values and deaths, and between minor injury values and deaths, for the two valuation methods.



	Impact 1	Impact 2	Impact 3
QLOS value (see note)	100	62.24	3.64
Ref. study value (see note)	100	51.72	3.45

Figure F1. Impact Numbers 1 to 3

Note: Log Scale. Reference study and QLOS units are normalised to Impact Number 1 value = 100, to show relative proportional values for Impact Numbers 2 and 3 (see text). For Impact Numbers 2 and 3, central values within the range presented by the reference study (ExternE, 1995a) are used in the normalisation.

F4.2 Stage 2: Coal-fired Power Station; power station impacts

Extending the trial to two further impacts (Impact Numbers 4 and 5), reveals a deviation from the close similarity seen in Impact Numbers 1 to 3. However, ranking order is still closely similar, with a maximum change of two places, which is that of Impact Number 4, ranked 5th for QLOS and 3rd for the reference study (for QLOS values, the rank order is 5-1-2-3-4, and for the reference study values, it is 5-1-4-2-3). Figure F2 shows Impact Numbers 1 to 5 plotted on a normalised scale, where for each valuation method respectively, the value for Impact Number 5 is 100. This shows that Impact Number 4 has a relatively higher reference study value when compared against the QLOS value.

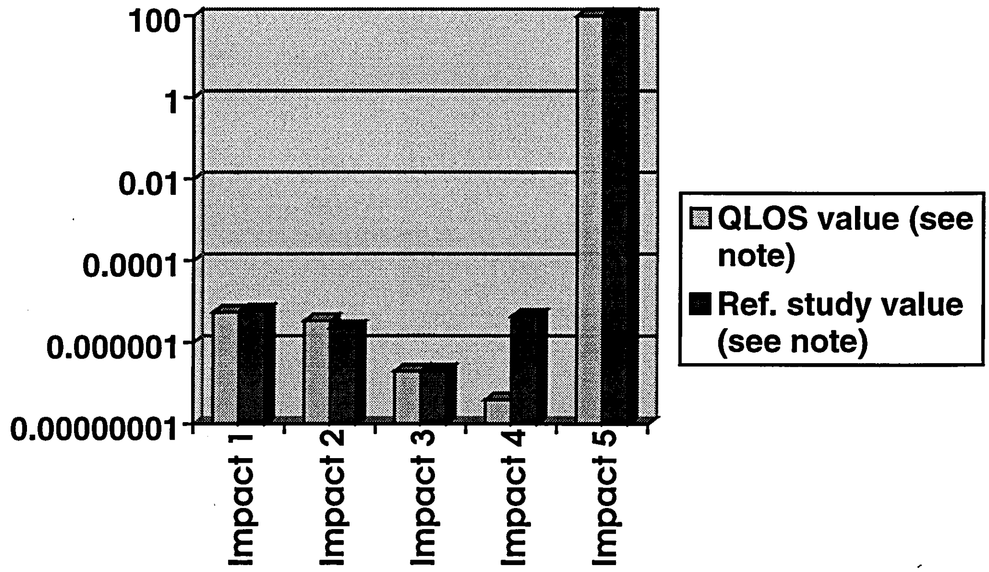


Figure F2. <u>Impact Numbers 1 to 5</u>

Note: Log Scale. Reference study and QLOS units are normalised to Impact Number 5 value = 100, to show relative proportional values for Impact Numbers 2 and 3 (see text). For Impact Numbers 2 and 3, central values within the range presented by the reference study (ExternE, 1995a) are used in the normalisation.

F4.3 Stage 3: Coal Fuel Cycle; Power Station and Coal Mining Impacts

Extending the trial to two further impacts (Impact Numbers 6 and 7) reveals further deviation from the close similarity seen in Impact Numbers 1 to 3. Ranking order also shows more deviation from that seen in Stage 2 (for QLOS values, it is 5-7-1-2-6-3-4, and for reference study values, it is 5-7-6-1-4-2-3), although the maximum change is still two rank places.

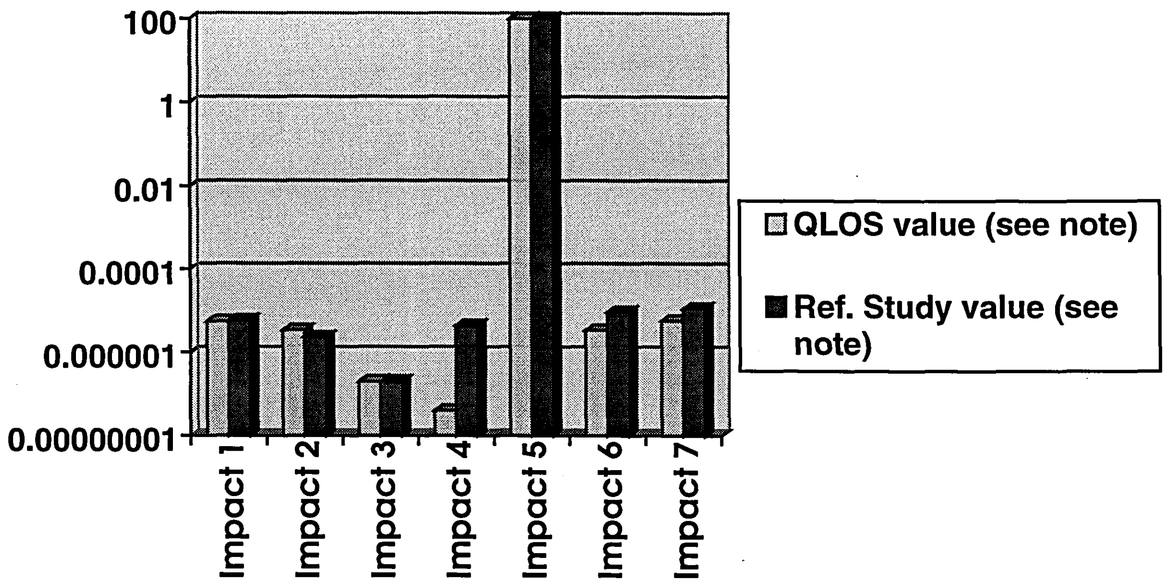


Figure F3. Impact Numbers 1 to 7

Note: Log Scale. Reference study and QLOS units are normalised to Impact Number 5 value = 100, to show relative proportional values for Impact Numbers 2 and 3 (see text). For Impact Numbers 2 and 3, central values within the range presented by the reference study (ExternE, 1995a) are used in the normalisation.

F4.4 Stage 4: Impacts across three Fuel Cycles

Stage 4 is the final Stage of the comparison study in this demonstration. This involves the addition of 2 further impacts, one from the nuclear fuel cycle, and one from the wind fuel cycle. It should be noted that comparison across fuel cycles is inherently more tenuous than within a single stage of a single fuel cycle, primarily because data variations and inconsistencies are more likely to affect results. However, notwithstanding this qualification, the trial results for Impact Numbers 8 and 9 show further interesting deviations from a simple parallel set of QLOS and reference study values. For example, for Impact Number 9, under the reference study, this is ranked 2nd, whereas under QLOS, it is ranked 5th, showing the largest rank difference in the trial study. Discussion and rank and graphical representations of the full set of trial results are presented in Section F5.

F5. Discussion

F5.1 Comparison of Method

The subject of the comparison is the valuation method, and the studies from which the data are drawn are concerned primarily with collating environmental change data and deriving human consequences for them. Valuation is invariably undertaken by converting impact data into money terms. The methods used vary and a description of these methods and of potential problems with them is presented elsewhere (Horne, 1995). Since every attempt has been made to use similar data for environmental change and human consequence scenarios in the QLOS valuations and the reference study valuations, the principal comparison should be between the method of adopting surrogate environmental economics valuations and these QLOS valuations.

The QLOS valuation method is a direct method, in the sense that it seeks to establish the human consequences on Quality of Life of a given environmental change directly, albeit in this exercise, by proxy, since the QLOS Index number was selected not by the receptor group but by this author. Methodologically, therefore, it is held that the QLOS valuations are more likely to reflect the true impact values, since they are derived from direct measures of impact. The caveat here is that the QLOS Index reflects the impact values of the receptor population.

F5.2 Comparison of Results

Note that only ranking (and some tentative interval) comparisons are possible, since no means has yet been developed for converting QLOS values into money terms, or for converting environmental economics-derived values into QOL-units. However, some discussion of calibration of the two types of units is discussed in Section F5.2.2 below. Also, Section F5.2.3, contains a discussion regarding the potential morbidity elements of impacts where the eventual result is death by illness.

F5.2.1 Rank Comparisons

The rank order of reference study values and QLOS values is presented in Figure F4. This shows that there is a variation in rank, with, for example, up to 3 rank place variation (Impact Number 9) between the ranks. The following discussion highlights some of the main rank differences and considers why they might occur.

Rank	Reference study	QLOS
1st	5	5
2nd	9	8
3rd	8	7
4th	7	1
5th	6	9
6th	1	2
7th	4	6
8th	2	3
9th	3	4

Figure F4. Rank Order of Impact Values for Reference Study and QLOS

Firstly, in general terms, three impact values are ranked lower under QLOS compared to the reference study; Impact Number 4, economic impact of agricultural crop loss, Impact Number 6, mortality from cancer, and Impact Number 9, amenity loss and annoyance caused by noise. Because of the relatively large drop in relative ranking (2, 2 and 3 places respectively), most other impacts are relatively higher ranked, Impact Numbers 1 and 2 showing the most marked trend in this regard (2 places). Only 4 impacts, Impact Number 5 (global warming), Impact Number 8 (cancers from nuclear accident), Impact Number 7 (occupational mortality from Progressive Massive Fibrosis) and Impact Number 3 (traffic-related accident minor injuries) show a rank position change of one place or less.

F5.2.2 Proportional Comparisons

Normalised or proportional comparisons for Stage 1, 2 and 3 of the trial have already been presented graphically in Section F4, in Figures 1 to 3 respectively. Figure F5 shows a similar representation, this time including all 9 impact values from all 4 stages, and Figure F6 shows the same data plotted on an ordinal scale designed specifically to illuminate the lower order impacts.

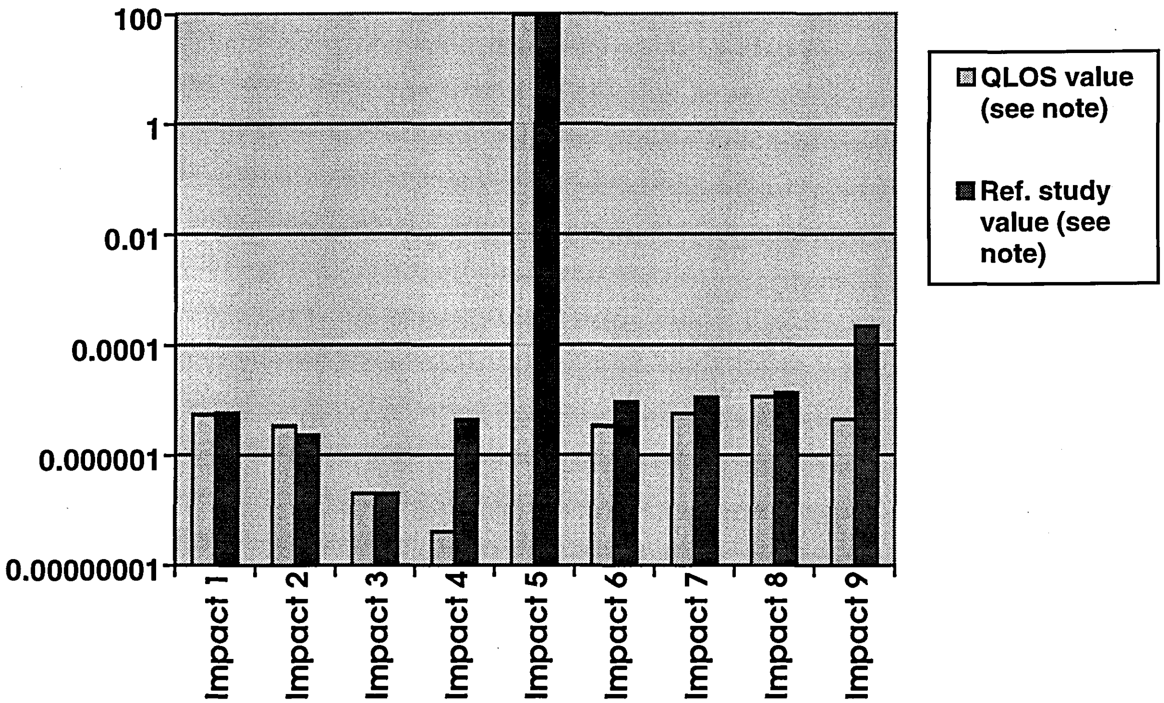


Figure F5. Graphical representation of QLOS and Reference Study Values (log scale)

Note: Reference study and QLOS units are normalised to Impact Number 5 value = 100, to show relative proportional values for Impact Numbers 2 and 3 (see text). For Impact Numbers 2 and 3, central values within the range presented by the reference study (ExternE, 1995a) are used in the normalisation. Impact Number 5 exceeds the scale.

In Stage 1, impacts were normalised to Impact Number 1 values – deaths from traffic accidents. In all other cases, they are normalised to deaths from global climate change (Impact Number 5). So, in all cases, the impact by which others are measured as a proportion is death. It is suggested that this is an unequivocal standard of impact, since it is simple to define and understand. If any calibration of the QLOS and reference study impacts is possible, it is therefore suggested that, using zero and death as effective scale points, a more direct and detailed comparison can be made than simply ranking.

For values which are relatively higher under QLOS, the most significant in ranking terms are Impact Numbers 1 and 2, deaths and major injuries from increased power station related traffic, although as a proportion of respective Impact Number 5 Values, Figure F5 shows that only Impact Number 2 has a significantly proportionally higher QLOS value than reference study value. Here, it is worth bringing Impact Numbers 1 and 3 into the discussion, since selection of these three impacts was intended to provide a trio of impacts which are very similar in all but seriousness of outcome on the quality of life of receptors. In other words, this trio of impacts provides an initial, clear test of the QLOS Index against an environmental economics based approach (hence, it was undertaken in Stage 1 of the trial). The outcome is clear; the rank order of Impact Numbers 1 to 3 overall is retained.

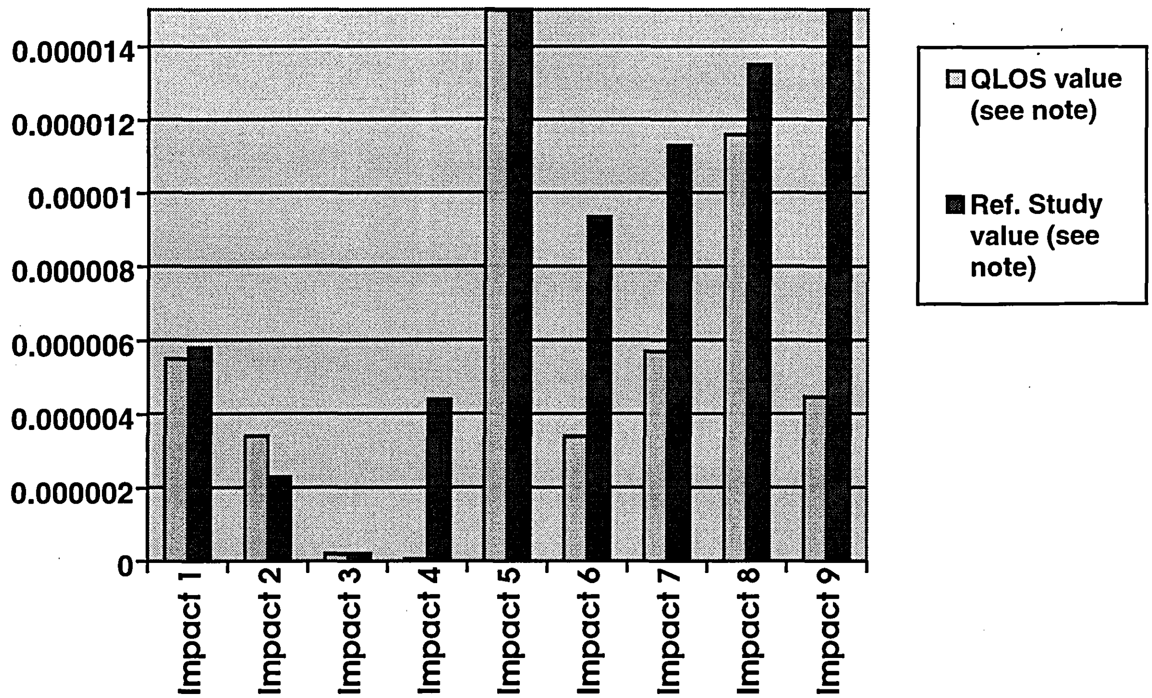


Figure F6. Graphical Representation of QLOS and Reference Study Values (ordinal scale)

Note: Reference Study and QLOS units are normalised to Impact Number 5 value = 100, to show relative proportional values for Impact Numbers 2 and 3 (see text). For Impact Numbers 2 and 3, central values within the range presented by the reference study (ExternE, 1995a) are used in the normalisation. Impact Number 5 exceeds the scale, as does the reference study value for Impact Number 9 (the purpose is to illuminate value comparisons of Impact Numbers 1 to 4 and 6 to 8).

Perhaps the single most striking observation from the comparisons shown in Figures F5 and F6 is that non-injurious impacts (Impact Numbers 4 and 9) provide lower relative values under QLOS. Impact Numbers 4 and 9 are essentially "economic" and "annoyance" in type respectively, rather than physically injurious (or potentially so) as in the other impacts in the trial.

In explaining why impact values for Impact Numbers 4 and 9 are relatively lower under QLOS, it is first important to re-examine the QLOS valuations. Both Impact Numbers 4 and 9 involved making assumptions beyond the proxy QLOS value selection, and these may be a source of error. The QLOS agricultural crop loss valuations include economic losses per farmer based on unsubstantiated assumptions about income per capita and (therefore) total per capita-equivalent loss of economic livelihood. Therefore, there is reason to suspect this value, however, basic sensitivity analysis (doubling and halving the assumed value for n) demonstrates that this would not affect the rank. Similarly, for Impact Number 9, the value for n is assumed, based on only patchy information provided in the reference study (ExternE, 1995c). Here, there is also the additional potential problem of some uncertainty over the average QLOS value. This time, basic sensitivity analysis, doubling and halving both n and I, shows there is the potential to affect rank. Indeed, a worst case (doubled I and n) would give a QLOS value of 3.472 μQu/MWh, ranking second overall, equal to the reference study rank. Therefore, it must be concluded that there is less confidence over the difference in rank between the reference study and QLOS for Impact Number 9. However, the conclusion remains that, with regard to Impact Number 4, this economic-based impact has a relatively lesser negative outcome on Quality of Life than they do on utility, as suggested by the reference study results, using a utility-based environmental economics approach.

The reference study approach, which draws on neo-classical economics, uses values for Impacts 4 and 9 which are derived using different methodological approaches than the other impacts. The QLOS values, on the other hand, are all derived from the same QLOS Index, which seeks to value the intensity of a given impact in terms of disability and distress, the level of which is the outcome on Quality of Life for any given impact. Thus, it is held that methodologically, the QLOS Index does actually measure the economic, annoyance and physical health (including death) impacts on the same scale, whereas the reference study and general neo-classical approach does not (it may in theory, but it may not in practice). It should be noted that the fundamental concept underlying the QLOS Index is the Von Neumann-Morgenstern standard gamble, which is itself a classical economic approach with a basis in utility theory.

F5.2.3 Global Climate Change

Under both methods, Impact Number 5, deaths caused by global climate change-related crop losses and starvation, ranks at the top. Both the reference study and QLOS values are derived using 0% Discount Rates (applying even 10% Discount Rates would still leave this impact orders of magnitude higher than all other impacts in the trial). However, it must be noted that no single reference study value is suggested, and that, on the basis of the results presented in the reference study (ExternE, 1995a), the reference study value quoted in Table F2 is approximately 1.5 orders of magnitude above the norm or median environmental economics-derived value quoted in this source. The value given in this trial is adopted because it is derived from a single human value for all humans (rather than differentiating for "Willingness to Pay" between developed world and developing world deaths), and this is therefore the most similar methodologically to the QLOS approach (where there is no "Willingness to Pay" effect and all human lives are valued equally). However, even if a more norm-based value is used, the reference study value would be around 10 mECU/kWh, and would still rank top, with approximately an order of magnitude difference between this and the next-ranking reference study value in the trial. It would, however, affect the relative interval difference between the reference study and QLOS values. Figure F5, where the reference study median value is marked, shows the effect on relative values that this would have.

F5.2.4 Death and Injury Comparisons

Examining proportional comparisons more closely, particularly for death and injury, it is possible to tentatively extend the conclusions which can be drawn from the results for

Impact Numbers 1 to 3 presented in Section F4.1 above. If the assumption is made that death is a suitable point of calibration between the QLOS Index and the monetary-derived estimates presented in the reference study, then a comparison can be drawn between values for levels of injury under QLOS and the reference study. Thus, under this assumption, Figure F1 suggests that the QLOS valuation method gives rise higher values for major injury than the reference study (20.3% higher) but only slightly higher values for minor injury (5.5% higher).

So, there is a significant difference between the economic values quoted for major injuries as used in the reference study, and the values suggested by the QLOS valuation. The question which arises here is to what extent this could be the result of the proxy system used in this trial for assigning QLOS Index categories to different impact states. In order for the proportional difference in Impact Number 2 to be accounted for by the proxy QLOS value alone, this would need to be varied considerably, as the following calculations show.

Impact Number 2 normalised value (Qu) = I x R x T x n/455.4

I x 1 x 0.55 x 55.113/455.4

I_a = 455.4/30.3121 x z

where z = 51.72% (see Figure F1) of Impact Number 1 and I_a is the QLOS Index number required to produce a proportionally similar value to the reference study result for Impact Number 2.

Now, z = 0.5534

Therefore;

I_a = 8.314

Therefore, a QLOS Index number of 8.3 would be required to return a value proportionally equivalent to the reference study value for Impact Number 2. The value actually used, 10, is equivalent to the state "moderate distress, largely housebound", whereas 8.3 would equate, for example, to slightly above the categories "moderate distress, physical ability severely limited" (6) and "severe distress, slight social disability" (7) and below "severe distress, severe social disability or slight physical impairment" (9). The problem here has already been mentioned in Section F3.3 above; without further disaggregation of injury data, and survey data on QLOS Index scores for major injury supplied by receptor groups (rather than the proxy scores used here) QLOS Index number choice is necessarily approximate. Therefore, the conclusion is that, while it appears the QLOS valuation method gives rise higher values for major injury than the reference study, further data would be required regarding QLOS Index categories for this apparent trend to be substantiated.

F5.2.5 Morbidity Elements of Mortal Illness

It should be noted that the QLOS valuation for Impact Numbers 6, 7 and 8 is only for the death phase of the impact. However, death from lung cancer, for example, is often preceded by a lengthy period of morbidity. Whether the reference study estimate could be said to represent the full externality value for these impacts is debatable, since it concentrates only on valuing death and not preceding illness (hence, to make the comparison valid, the QLOS valuations have only included the death phase).

However, illness phases prior to death would be important in any full QLOS valuation, particularly when death is delayed, because a value for illness during the period after suffering starts and before death occurs is integral to the value of the disease. Without this early phase being included in valuation, it should be noted that only lost years of proportional life remaining are being measured. It is suggested that illness preceding death may provide a significant additional element to the overall value of the mortality impact in some cases. In order to investigate the potential effect of this morbidity phase in death impacts, the example of Impact Number 6 is taken, and the calculations below demonstrate the potential contribution of this element.

QLOS Index Number:

A morbidity value for pre-death phase of the disease is estimated as "moderate distress, unable to work/housebound" (10 on the QLOS Index).

Probabilistic Risk:

Risk is stated as 0.002 (based on US estimates).

Conscious Life Remaining:

The death figure approximates the T for the death phase of the disease; the morbidity phase is therefore the remainder of life after diagnosis for all workers,

i.e. 1 - 0.375 = 0.625.

Number of receptors:

An estimated 4,000 coal workers will be exposed to the risk to provide the coal.

Valuation:

Impact Value (Qu) = I x R x T x n

Impact Value = 10 x 0.002 x 0.625 x 4,000 = 50 Qu (for morbidity phase)

Normalised value = 50/455.4 = 0.1098 μQu/MWh.

The Total Normalised value (death phase calculated in Impact Number 6 + morbidity phase) = 0.6588 + 0.1098 = 0.7686 μQu/MWh (see Table F4). The effect of this additional phase is enough to elevate the rank above Impact Number 2, major traffic-induced injury (where no morbidity phase would apply), demonstrating that this factor makes a material difference to relative as well as absolute impact value.

Human Consequence	Coalworkers' occupational health; mortality from lung cancer/radon exposure	
Criteria	Impact Number 6 (death phase only)	Morbidity Phase
VALUATION DATA		
Index number (I)	100	10
Risk Value (R)	0.002	0.002
Time Value (T)	0.375	0.625
Receptors Value (n)	4000	4000
TOTAL (Qu)	350	50
QLOS VALUE (μQu/MWh)	0.06588	0.1098
	Total: 0.7686	
COMPARABLE VALUE (mECU/KWh)	0.047	?

Table F4. Morbidity Elements of Mortal Illness

F6. Conclusions of the Valuation Demonstration

The aims of this demonstration have been achieved, since trial impacts have been calculated in QLOS value terms using substantially real data, and comparisons have been made with real comparables. Overall, the work demonstrates the practical applicability of the QLOS valuation method, along with an indication of its potential. Also, even with a small trial of sample impacts, there are grounds for justifying differences in values derived by QLOS and the reference study methods.

In terms of methodological differences, it is concluded that overall, the QLOS valuation method is more direct in the sense that it seeks to directly value changes in quality of life arising from changes, rather than by indicators or surrogates, as in many environmental economics approaches. The main conclusions in terms of differences in values for the trial impacts are that QLOS values tend to give more weight relatively to physical illness and disability arising from environmental changes, and hence less so to economic or emotional (for example, noise annoyance) effects. Also, there is some evidence that deaths and physical injuries tend to be valued relatively more highly in

QLOS values compared to the reference study, or, conversely, that enonomic and possibly annoyance type values are relatively lower using the QLOS valuation method.

APPENDIX G. LIFE CYCLE ANALYSIS AND THE SYSTEMATIC FRAMEWORK

Life Cycle Analysis is a technique for identifying and assessing the impacts arising from a given human activity. Thus, it has affinities with the approach of the systematic framework and, indeed, shares similarities in methodology. The purpose here, is to provide background information on LCA, followed by comparison between it and the systematic framework.

There is clear linkage between LCA, energy analysis and mass balance analysis. Energy analysis, based on the principles of systems analysis, was developed in the early 1970s in its modern form. With the emergence of this method for tracing energy through complex systems, there followed the development of a means for tracing materials through complex systems – the origins of mass-balance analysis. Although LCA has apparently developed rapidly as a new technique only relatively recently, during the 1990s, it actually has an established history since it is based on the same fundamental systems approach and analytical principles as these two earlier methods.

LCA (also known as life cycle assessment, ecobalance and product line analysis) is a technique for assessing the environmental impacts of a given industrial process, by identifying each stage of the process (and the process inputs) and quantifying the impacts arising. Therefore, in both concept and general approach, LCA shares many similarities with the output analysis method, including a systems approach, sequential methodology and impact assessment aims. However, the aims of many so-called LCA studies are very specific, such as the comparison of particular products, materials, or production processes for a range of purposes, including marketing, labelling, environmental management, or the supply of product information in various forms. The most common application of LCA techniques to date has been for comparative purposes. Typically, a company wants to compare their products or packaging to those of competitors, or to other options, or another body wishes to examine some products in a comparative context. Furthermore, despite common generic methodologies (for example, PIRA, 1993, SETAC, 1993, ISO, 1997) there are differences in approach across LCA studies.

It has been noted that the LCA methodology has four components: "goal definition and scoping, life-cycle inventory (LCI), impact assessment, and improvement assessment" (Curran, 1996). Goal definition involves defining the purpose of the study, the boundary conditions and the assumptions. Scoping involves reducing the scope or range of considerations to an appropriate size for the study. The LCI quantifies the resource use, energy use and environmental releases associated with the system being evaluated. Practitioners generally agree upon a common systems approach for performing LCI studies and this is, therefore, a less contentious area of the methodology. These first two components broadly correspond to the output analysis method.

Scoping is a tool for reducing the number of considerations in a study from an unmanageable number to a manageable number; in this case, the number of possible impacts. The aim is to identify those impacts which have no potential to be significant and discount them from further consideration, thereby reducing the number of possible impacts which must be examined in detail. Such scoping has been discussed elsewhere (for example, Horne, 1996b, 1997). While it has potential for assisting in meeting the requirements of efficiency in the valuation process, there are three problems with it. Firstly, given the uniqueness of environmental impacts, it is difficult to establish unequivocally that a life cycle output cannot lead to a significant impact until detailed examination has been undertaken. The tendency is therefore to create a detailed and involved scoping method, which then invalidates the benefits of scoping as an approach. Secondly, there is the problem of significance. Objectivity should be maximised and subjectivity concentrated within valuation. The output analysis method must therefore be concerned with magnitudes, not significance. The latter is an inherently subjective measure and requires judgement to be made about the potential outcomes of varying types of impact. Therefore, scoping involves practitioners exercising value judgement. Decisions should not be made about relative importance of RIOs since they are expressed in different units from each other and therefore cannot be compared. Thirdly, scoping out sub-significant impacts can lead to a lack of transparency since, often, these are simply left out of the valuation process. This can lead to uncertainty over whether scoped out impacts are "gaps" which are unaccounted for, or "gaps" which have been deemed not worthy of further consideration.

The solution to the problem of scoping lies in clarity of definition and preservation of transparency and objectivity. Scoping has two possible meanings. First, it can refer to picking out the impacts to be looked at such as, for example, the selection of carbon dioxide-induced global climate change and the rejection of visual impact from further consideration. This could be termed "scoping of impacts" and it is clearly subjective and inappropriate at this stage in the valuation process, since critical information is not yet known about the impacts and the RIOs are in dissimilar units. In the second meaning, it can refer to comparing magnitudes of similar RIOs and rejecting some on the grounds of complexity related to their small magnitude. For example, the carbon dioxide used in constructing coal hoppers for storing coal at a power station may be left out, whereas the carbon dioxide produced in coal combustion during power generation left in. This could be termed "scoping of like magnitudes" and is less likely to cause major inaccuracies or uncertainties in results of subsequent valuation. However, it does not necessarily follow that magnitude (size) and significance (importance of impact) are directly related. A large power station structure in an urban landscape may be less significant than pylons across a rural landscape. Likewise, local pollutants can have varying effects depending upon the release environment. In summary, scoping does not necessarily invalidate the results, but it cannot be subsequently maintained that all potential impacts have been examined or that all impacts have been included. In the worst case, using scoping of impacts techniques 20 years ago in valuing impacts of a power station, it is likely that what is currently recognised as the principal impact, carbon dioxide-induced global climate change, would have been "scoped out" and subsequently not considered.

The production of the LCI is broadly comparable to the production of the RIO Inventory. However, there are two main differences between these methods. Firstly, the LCI typically also lists natural resource use, that is, quantities of natural resource inputs. However, these are not listed on the RIO Inventory, since the systematic framework is only concerned with outputs. Quantities of natural resource inputs are required in constructing the RIO Inventory, in mass and energy balance checks to ensure that total process inputs are equal to total process outputs (RIOs plus products). However, they are not explicitly stated on the RIO Inventory. Secondly, the LCI may contain pathway information, whereas the RIO Inventory is limited to a list of materials and energy as they cross the boundary out of the human activity (defined as the point at which they enter the surrounding environment). For example, it has been stated (Curran, 1996) that typical quantities on a LCI may include;

- Energy - both embedded and that used in the processes;

- Emissions to air - typically 30-40 types, including carbon dioxide (both fossil and non-fossil) and pollutants, such as particulates, oxides of nitrogen and sulphur, and carbon monoxide;

- Emissions to water - typically more than 20, including pollutants, also Biological Oxygen Demand, Chemical Oxygen Demand, suspended solids, dissolved solids, iron, chromium, acid, ammonia, phosphates, etc.;

- Solid wastes - everything solid which leaves the system (notwithstanding that the system may include recycling).

While many of these are true RIOs, others may be referred to more properly as "burdens" on the environment, rather than outputs of the process. In contrast, RIOs are strictly defined as materials and energy leaving the process. When they enter the environment, they are already in a pathway, and the pathway analysis method is required in order to account for them, which is subsequent to the output analysis method.

Despite the differences, there are many similarities between the LCI and the RIO Inventory. Both are likely to be lengthy lists, typically involving various production stages and/or materials and energy. Neither are intended to determine the relative impact of outputs on the environment or on human health. Both may be followed by impact assessment. The LCI may be followed by the latter stages of the LCA. However, many studies have terminated at this point, with conclusions and improvement analysis being limited to seeking less resources use, less energy use, and lower levels of emissions to the environment. However, for impact assessment to be achieved, the production of the LCI must be followed by full LCA.

According to one commentator; "conceptually, impact assessment consists of three stages; classification, characterisation, and valuation" (Curran, 1996). These involve testing outputs for both the comparable and different impacts they cause. Where impacts are similar, factors can be produced, and where impacts are dissimilar, they can be classified as a prelude to magnitude estimation and valuation. Thus, the "classification" and "characterisation" exercises are the LCA equivalent to the pathway analysis method.

Classification involves aggregating and assigning LCI inputs and outputs to impact groupings, or categories. One widely accepted conceptual framework for life-cycle impact assessment (SETAC, 1993) lists four major categories; environment/ecosystem quality, quality of human life, natural resource use, and social welfare. For example, the use of fossil fuels may be assigned to a pre-determined impact group such as "depletion of finite resources". Characterisation is the process of developing conversion models to translate data to "impact descriptors". An example here would be the conversion of quantified LCI outputs of carbon dioxide and methane into units of global warming potential. Hence, the magnitudes of potential impacts on the chosen categories are evaluated. Some of the techniques and approaches used to accomplish this exercise include;

- Loading - assessing inventory data alone on the assumption that less quantity produces proportionally less impact;

- Equivalency - combining inventory data with derived equivalency factors to aggregate inventory data, thereby incorporating the assumption that the equivalency factors are correct;

- Inherent chemical properties - pooling inventory data based on chemical properties, for example on toxicity, persistence, bioaccumulation, etc.;

- Generic exposure and effects - estimating impacts based on generic environmental and human health information;

- Site-specific exposure and effects - determining actual impacts based on site-specific impact information.

Various documents have been produced which are designed to assist and standardise the use of the LCA methodology. A succinct and well-accessible example is the International Standard ISO14040, which seeks to establish principles for LCA (ISO, 1997). Although the series of ISO LCA standards is still incomplete, this holds out the best possibility to date that a full standard methodology will eventually be forthcoming. However, as yet, there is no such standard LCA impact assessment methodology. In one study, 36 different methods for doing the characterisation and/or valuation steps of impact assessment have been summarised (USEPA, 1994), none of which have been widely adopted by the scientific community. Additionally, a number of generic analysis models have been developed (for example, SETAC, 1993) and specific studies undertaken, although, being often semi-commercial, these are not always freely accessible and remain in the grey literature. In contrast, there is a relative plethora of information available discussing theory, principles and conceptual approaches to LCA.

In 1969, in the very first LCA study performed, the desirability of a single unit of impact assessment was noted, and a set of subjective factors was developed and applied. It was intended that diverse impact categories, such as energy resource depletion and toxic discharges, would be combined into a single quality of life category. However, because of the subjective nature of the analysis, the approach did not meet wide acceptance. There is still a high level of interest in expressing impacts in common units. However, while it is valuation for LCA theoretically involves assigning relative weights to different impacts to allow comparison across all impact categories, practically, no agreed method for doing this exists as yet.

REFERENCES

Abrahamsen, G, Stuanes, A O and Tveite, B. 1983. Effects of long range transported air pollutants in Scandinavia. Water Quality Bulletin, 8, pp.89-95.

Adger, N & Whitby, M. 1990. Appraisal and the Public Good: Environmental Assessment and the Cost Benefit Analysis. 29 pp. Countryside Change Working Paper Series, Working Paper 7, March 1990

Ahonen, J, Rankinen, K. 1999. Model application to study the effects of emission reduction scenarios on a forested catchment in Finland. Phys Chem Earth (B), 24, 7, pp.861-867.

Akai, M, Nomura, N, Waku, H and Inoue, M. 1997. Life Cycle Analysis of a Fossil Fuel Power Plant with CO_2 Recovery and a Sequestering System. Energy, 22, 2/3, pp.249-255.

Alastuey, A, Querol, X, Chaves, A, Ru iz, C R, Carratala, A and Lopez-Soler, A. 1999. Bulk deposition in a rural area located around a large coal-fired power station, northeast Spain. Environmental Pollution, 106, pp.359-367.

Alcamo, J, Shaw, R and Hordijk, L, eds. 1990. The RAINS Model of acidification. Science and strategies in Europe. Kluwer Academic publishers.

anon. 1971. Air pollution across national boundaries. The impact on the environment of sulfur in air and precipitation. Sweden's case study for the United Nations conference on the environment. Royal Ministry for Foreign Affairs and Royal Ministry of Agriculture. Stockholm, 1971.

anon. 1979. Pathways of pollutants through the atmosphere. The Royal Society, London.

anon. 1989. The electricity industry. UBS Phillips and Drew.

anon. 2000. Editorial: Introduction to a special issue on irreversibility. Resource and Energy Economics, 22, pp.189-196.

Arduini, I, Kettner, C, Godbold, D L, Omnis, A, Stefani, A. 1998. pH influence on root growth and nutrient uptake of *Pinus pinaster seedlings.* Chemosphere, 36, 4-5, pp. 733-738.

Augustin, S, Schall, P and Schmieden, U. 1998. Modelling aspects of forest decline in Germany: I. Theoretical aspects and cause-effect relationships. Chemosphere, 36, 4-5, pp.965-970.

Ayres, R U and Ayres, L W. 1998. Accounting for resources, 1. Economy-wide applications of mass-balance principles to materials and waste. Edward Elgar.

Bailey, R. 1991. Energy Policy in Confusion. National Westminster Bank Quarterly Review.

Baker, S, Kousis, M, Richardson, D, Young, S. ed. 1997. The Politics of Sustainable Development. Theory, policy and practice within the Europe Union. Routledge.

Barde, J-P and Pearce, D W. (eds). 1991. Valuing the Environment. Six Case Studies. Earthscan, London.

Bates, D V.1992. Health indices of the adverse effects of air pollution: The question of coherence. Environmental Research, 59, pp.336-349.

Bech, P. 1993. Rating scales for psychopathology, health status and Quality of Life. A Compendium with correspondence to DSM-III-R and the WHO systems. Springer, Berlin.

Bech, P. 1994. The PCASEE Model: An approach to subjective well-being. In: Orley, J and Kuyken, W. (eds). 1994. Quality of Life Assessment: International Perspectives. Proceedings of the joint Meeting organised by the WHO and the Fondation IPSEN, Paris, July 1993. Springer-Verlag, Berlin.

Beesley, M E. 1992. Privatization, Regulation and De-Regulation. Routledge.

Beesley, M E. 1994. Regulating Utilities: The way forward. Institute of Economic Affairs.

Bell, J N B. 1985. SO_2 effects of the productivity of grass species. In: Winner W E, Mooney, H A and Goldstein, RA, eds. 1985. Sulfur dioxide and vegetation: Physiology, ecology and policy issues. Standford University Press, US.

BERG. 1989. The effects of acid deposition on buildings and building materials in the United Kingdom. Building Effects Review Group, HMSO.

Bergner, M, Pobbitt, R A, Dressel, S, Pollard, W E, Gibson, B S,and Morris, J. 1976. The Sickness Impact Profile: Conceptual Foundation and Methodology for the development of a health Status Measure. Int. J Health Services. 6. pp. 393-415.

Beverland, I J, Crowther, J M, Srinivas, M S N, Heal, M R. 1998. The influence of meteorology and atmospheric transport patterns on the chemical composition of rainfall in south-east England. Atmospheric Environment, 32, 6, pp.1039-1048.

Bishop, M, Kay, J and Mayer, C. 1995. The regulatory challenge. Oxford University Press.

Bobba, A G, Singh, V P and Bengtsson, L. 2000. Application of environmental models to different hydrological systems. Ecological modelling, 125, pp.1549.

Boulding, K. 1966. The Economics of the Coming Spaceship Earth. In Environmental Quality in a Growing Economy, Resources for the Future. Johns Hopkins Press.

Bowling, A. 1995a. Measuring Disease. A Review of disease-specific Quality of Life Scales. Open Univ. Press.

Bowling, A. 1995b. What Things Are Important in People's Lives? A Survey of the Public's Judgements to Inform Scales of Health-Related Quality of Life. Social Science and Medicine, 41, 10, pp.1447-1462.

Brook, R H, Ware, J E, Davies-Emery, A, Stewart, A L, Donald, C A, Rodgers, W H, Williams, K N, and Johnson, S A. 1979. Conceptualisation and Measurement of Health for Adults in the health Insurance Study: Vol 8, Overview. RAND Corporation, Santa Monica, Cal.

Brookshire, D S & Coursey, D L. 1987. Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures. The American Economic Review, September, pp. 554-566

Brookshire, D S et al. 1982. Valuing Public Goods: A Comparison of Survey and Hedonic Approaches. The American Economic Review, March, pp.165-177

Brookshire, D S, Randall, A and Stoll, A. 1980. Valuing Increments and Decrements in Natural Resource Service Flows. Am Jnl Ag Econ, 62. pp.478-88.

Broom, J. 1988. Good, Fairness and QALYs. In: Bell, J M, Mendus, S (eds). 1988. Philosophy and Medical Welfare. Cambridge UP.

Brown, K A. 1985. Acid deposition: effects of sulphuric acid pH3 on chemical and biochemical properties of bracken litter. Soil Biology and Biochemistry, 17, pp.31.

Brown, R A and Green, C H. 1981. Threats to Health or Safety: Perceived Risk and Willingness-to-Pay. Social Science and Medicine, 15C, pp.67-75.

Buchanan, S C. 1990. Estimating Environmental Costs of Energy Resources. Electricity Jnl, July. pp.36-41

Budnitz, R J and Holdren, J P. 1976. Social and Environmental Costs of Energy Systems. Ann. Rev. of Energy. pp.553-580

Bullinger, M. 1994. Ensuring International Equivalence of Quality of Life measures: problems and approaches to solutions. In: Orley, J and Kuyken, W. (eds). 1994. Quality of Life Assessment: International Perspectives. Proceedings of the joint Meeting organised by the WHO and the Fondation IPSEN, Paris, July 1993. Springer-Verlag, Berlin.

Burnett, J. 1990. Ecology, Economics and the Environment. The Royal Bank of Scotland Review, 167, pp.3-15

Bush, J W, Anderson, J P, Kaplan, R M and Blische, W R. 1982. "Counterintuitive" Preferences in Health-Related Quality of Life Measurement. Medical Care, 20, 5, pp.516-525.

Butlin, R, Yates, T, Murray, M and Paul, V.1994. Effects of pollutants on buildings. DoE Report No. DoE/HMIP/RR/94/030. HMIP.

Camarero, L, Catalan, J. 1998. A simple model of regional acidification for high mountain lakes: Application to the Pyrenean lakes (north-east Spain). Water Resources, 32, 4, pp.1126-1136.

Chapman, P F, Leach, G and Slesser, M. 1974. The Energy Cost of Fuels. Energy Policy, 2, 3. p231.

Chapman, P F. 1974. Fuel's Paradise: Energy Options for Britain.

Charron, A, Plaisance, H, Sauvage, S, Coddeville, P, Galbo, J C, Guillermo, R. 1998. Intercomparison between three receptor-oriented models applied to acidic species in precipitation. The Science of the Total Environment, 223, pp.53-63.

Clarke, A. 1994. Comparing the Impacts of Renewables. International Journal of Ambient Energy, 5, 4, pp.343-362.

Cline, W R. 1992a. The Economics of Climate Change. Cambridge University Press.

Cline, W R. 1992b. Global warming: The Economic Stakes. Institute for International Economics.

Coase, R. 1960. The Problem of Social Cost. Jnl of Law and Economics. 3.

Cooke, R U and Gibbs, G B. 1994. Crumbling heritage? Studies of stone weathering in polluted atmospheres. National Power plc.

Cosemans, G. 1996. Environmental software conference report. 4th Workshop on harmonisation within atmospheric dispersion modelling for regulatory purposes, Oostende, Belgium, 6-9 May 1996. Environmental Software, 11, 4, pp.277-284.

Crane, A J and Cocks, A T. 1987. The transport, transformation and deposition of airborne emissions from power stations. CEGB Research, 20, pp.3-15.

Cropper, M L and Oats, W E. 1992. Environmental Economics: A Survey. Jnl. Econ. Lit., 30, pp.675-740

Culyer, A J and Maynard, A K. 1981. Cost-effectiveness of duodenal ulcer treatment. Social Science and Medicine, 15c, pp. 3-11.

Cummings, R G, Brookshire, R S and Schulze, W A. eds. 1986. Valuing Environmental Goods: An assessment of the Contingent Valuation Method. Rowman and Allanheld, New Jersey.

Curran, M A. 1996. Environmental Life-cycle Assessment. McGraw-Hill.

Daly, H E. 1999. Ecological economics and the ecology of economics. Essays in criticism. 191pp. Edward Elgar.

Davis, R. 1963. Recreation Planning as an Economic Problem. Natural Resources Jnl. 3, 2, pp.239-249.

Desvouges, W, Johnson, F, Banzhaf, H et al. 1995. Assessing environmental externality costs for electricity generation, Final report revised. Triangle Economic research, Durham, NC. Cited in: Freeman, A M. 1996. Estimating the environmental costs of electricity: an overview and review of the issues. Resource and Energy Economics, 18, pp.347-362.

DETR, 1999. A better quality of life. A strategy for sustainable development for the United Kingdom. Cm 4345. The Stationery Office.

DETR, 2000. Local quality of life counts. A handbook for a menu of local indicators of sustainable development. DETR.

Dochinger, L S and Seliskar, C E. 1970. Air pollution and the chloritic dwarf disease of Eastern White Pine. Forest Science, 16,1, pp.46-55.

Dockery, D W and Pope III, C A. 1994. Acute respiratory effects of particulate air pollution. Annual Review of Public Health, 15, pp.107-132.

DOE. 1955. A programme of nuclear power. Department of Energy. Cmnd. 9389. HMSO.

DOE. 1988. Privatising electricity. Department of Energy. Cmnd. 322. HMSO.

DOH. 1995. Non-biological particles and health. Dept. of Health, Committee on the Medical Effects of Air Pollutants, HMSO.

Drewnowski, J. 1974. On measuring and planning the quality of life. Publications of the Institute of Social Studies. Volume XI. Mouton, The Hague.

Drotar, D. 1998. Measuring Health-Related Quality of Life in Children and Adolescents. Lawrence Erlbaum Associates.

DTI. 1961. The financial and economic obligations of the nationalised industries. Department of Trade and Industry. Cmnd.1337. HMSO.

DTI. 1992. Digest of UK Energy Statistics. Department of Trade and Industry. HMSO.

DTI. 1994. Thinking about Regulating. A Guide to Good Regulation. Department of Trade and Industry.

DTI. 2001. New and Renewable Energy. Prospects for the 21st Century. The Renewables Obligation Preliminary Consultation: Analysis of the Responses to the Consultation Paper. March. HMSO.

Dubin, J. A. and Rothwell, G. S. 1990. Subsidy to Nuclear Power through Price-Anderson Liability Limit. Contemporary Policy Issues, July. pp.73-9

Dunderdale, D. 1990. Energy and the environment. Special Publication No.81, Royal Society of Chemistry.

EC. 1994. Potential benefits of integration of environmental and economic policies. European Communities Environmental Policy Series. Graham and Trotman and Office for Official Publications of the European Communities.

Edwards, R W, Gee, A S and Stoner, J H, eds. 1990. Acid Waters in Wales. Kluwer Academic Publishers.

Eikeland, P O. 1998. Electricity market liberalisation and environmental performance: Norway and the UK. Energy Policy, 26, 12, pp.917-927.

Energy Committee. 1992. Consequences of Electricity Privatisation. Energy Committee Second Report. Volume 1. HMSO.

Energy Select Committee. 1992. 4th Report. HMSO, London.

Esch, A, Mengel, K. 1998. Combined effects of acid mist and frost on the water status of young spruce trees *(Picea abies).* Chemosphere, 36, 4-5, pp.645-650.

ETSU. 1998. Reversal of surface water acidification in the Upper Duddon catchment (English Lake District). DTI Energy and Environment Programme Report No. ETSU NIO1/00051/REP.

ETSU. 1999. Estimation and projection of NO_x and SO_x emissions from UK industrial sectors: Phase IV. DTI Energy and Environment Programme Report No. ETSU N/02/00055/REP.

EuroQol. 1990. EuroQol - A new Facility for the Measurement of Health-related Quality of Life. Health Policy. 16. pp.199-208.

ExternE, 1995a. Externalities of Energy. Volume 1. Summary. European Commission DG-X11. Office for publications of the European Communities.

ExternE, 1995b. Externalities of Energy. Volume 2. Methodology. European Commission DG-X11. Office for publications of the European Communities.

ExternE, 1995c. Externalities of Energy. Volume 3. Coal and Lignite. European Commission DG-X11. Office for publications of the European Communities.

ExternE, 1995d. Externalities of Energy. Volume 4. Oil and Gas. European Commission DG-X11. Office for publications of the European Communities.

ExternE, 1995e. Externalities of Energy. Volume 5. Nuclear. European Commission DG-X11. Office for publications of the European Communities.

ExternE, 1995f. Externalities of Energy. Volume 6. Wind and Hydro. European Commission DG-X11. Office for publications of the European Communities.

ExternE. 1998. Externalities of Energy. Volume 8. Global Warming. European Commission DG-X11. Office for publications of the European Communities.

ExternE. 1999. Externalities of Energy. Volume 10. National Implementation. European Commission DG-X11. Office for publications of the European Communities.

Eyre, N. 1993. Comparison of environmental impacts of electricity from different sources. Nuclear Energy, 5, pp.321-326.

Eyre, N. 1997. External costs: What do they mean for energy policy? Energy Policy, 25, 1, pp.85-95.

Eyre, N. 1998a. A golden age or false dawn? Energy efficiency in UK competitive energy markets. Energy Policy, 26, 12, pp.963-972.

Eyre, N. 1998b. Environmental Impacts of Energy. Royal Commission on Environmental Pollution study on energy and the environment: Paper prepared as background to the study.

Faaij, A, Meuleman, B, Turkenburg, W, Wijk, A van, Bauen, A, Rosillo-Calle, F, and Hall, D. 1998. Externalities of biomass based electricity production compared with power generation from coal in the Netherlands. Biomass and Bioenergy, 14, 2, pp.125-147.

Fairley, D. 1990. The relationship of daily mortality to suspended particulates in Santa Clara County, 1980-86. Environmental Health Perspectives, 89, pp.159-168.

Farquhar, M. 1995. Elderly people's Definitions of Quality of Life. Social Science and Medicine, 41, 10, pp.1439-1446.

Feenberg, D and Mills, S. 1980. Measuring the Benefits of Water Pollution Abatement. Academic Press, New York.

Fells, I. 2000. Can the energy market protect the environment? Interdisciplinary Science Reviews, 25, 1, pp.29-33.

Ferguson R A D. 1994a. Environmental cost assessment of conventional electricity generation. Proceedings of the Environmental Impacts of Energy Technology, Circular 62 of the Solar Energy Society, London, pp.28-35.

Ferguson, P, Lee, J A, and Bell, J N B. 1978. Effects of sulphur pollutants on the growth of Sphagnum species. Environmental Pollution, 16, pp.151-162.

Ferguson, R A D. 1993. Environmental Costs of UK Electricity Generating Technologies. Draft Final Report for CEETS.

Ferguson, R A D. 1994b. Environmental Costs of Power Generation. IMechE C479/007/94.

Fox-Rushby, J A. 1994. The relationship between health economics and helth-related quality of life. In: Orley, J and Kuyken, W. (eds). 1994. Quality of Life Assessment: International Perspectives. Proceedings of the joint Meeting organised by the WHO and the Fondation IPSEN, Paris, July 1993. Springer-Verlag, Berlin.

Freeman, A M. 1996. Estimating the environmental costs of electricity: an overview and review of the issues. Resource and Energy Economics, 18, pp.347-362.

Gee, D. 1994. Eco-nomic Tax Reform: Shifting the Burden from Economic Goods to Environmental Bads. European Environment, 4, 2, pp. 17-22.

Geller, H, DeCicco, J, Nadel, S. 1993. Structuring an energy tax so that energy bills do not increase. ACE^3.

Gilbert, R J and Kahn, E P, eds. 1996. International comparisons of electricity regulation. Cambridge University Press.

Gladstone, B and Dewhurst, D. 1988. Electricity, Linked Industries and Privatisation. Special Report Number 4. Coalfields Communities Campaign.

Goenaga, X and Williams, D J A. 1990. determination of aluminium speciation in acid waters. In: Edwards, R W, Gee, A S and Stoner, J H, eds. 1990. Acid Waters in Wales. Kluwer Academic Publishers.

Goldberg, D. 1972. The Detection of psychiatric illness by Questionnaire. Oxford UP.

Goldsmith et al. 1972. A Blueprint for Survival. The Ecologist. Special Issue.

Goodin, R. 1980. Nor Moral Nukes. Ethics, 90.

Gorham, E. 1998. Acid deposition and its ecological effects: A brief history of research. Environmental Science and Policy, 1, pp.153-166.

Graber, W K, Siegwolf, R T W, Nater, W and Leonardi, S. 1996. Mapping the impact of anthropogenic depositions on high elevated alpine forests. Environmental Software, 11, 1-3, pp.59-64.

Gratton, L C. 1980. Analysis of Maslow's Need Hierarchy with Three Social Class Groups. Social Indicators Research, 7, pp.463-476.

Green, C H, Tunstall, S M, N'Jai, A and Rogers, A. 1990. Economic Evaluation of Environmental Goods. Project Appraisal, 5, 2, June. pp.70-82

Guderian, R, ed. 1985. Air pollution by photochemical oxidants. Formation, transport, control and effects on plants. Ecological Studies Vol.52. Springer-Verlag.

Hajat, S, Haines, A, Gou bet, S A, Atkinson, R W, Anderson, H R. 1999. Association of air pollution with daily GP consultations for asthma and other lower respiratory conditions in London. Thorax, 54, 7, pp.597-605.

Hall D C. 1990. Preliminary Estimates of Cumulative Private and External Costs of Energy. Contemporary Policy Issues, 8, July, pp. 283-307.

Hanley, N D. 1988. Using Contingent Valuation to Value Environmental Improvements. Applied Economics, 20, pp. 541-549

Hanley, N D. 1989. Valuing Non-Market Goods Using Contingent Valuation. Journal of Economic Surveys 3 (3), pp. 235-252

Hanley, N D. 1990. Valuation of Environmental Effects. 64pp. Ind. Dep. for Scotland, SDA. ESU Research Paper 22.

Hannah, L. 1982. Engineers, Managers and Politicians. The First Fifteen Years of Nationalised Electricity in Britain.

Hardin, G. 1968. The Tragedy of the Commons. Science, 162.

Hays, R D, Sherbourne, C D, Marshall, G N. 1993. The 'states versus weights' dilemma in quality of life measurement. Qual Life Res, 2, pp.167-8.

Heck, W W. 1989. Assessment of crop losses from air pollutants in the United States. In: Mackenzie, J J and El-Ashry, M T, eds. 1989. Air pollution's toll on forests and crops. World Resources Institute, Yale University Press.

Heliovaara, K and Vaisanen, R. 1986. Industrial air pollution and the pine bark bug, Aradus cinnamonomeus Panz (het., Aradidae). Journal of Applied Entomology, 101, pp. 469-478.

Henney, A. 1987. Privatise Power; Restructuring the ESI. Policy Study No. 83. Centre for Policy Studies.

Herfindahl, O C and Kneese, A V. 1974. Economic Theory of Natural Resources. Merrill.

Hettelingh, J-P, Downing, R J and de Smet, P A M. 1991. Mapping critical loads for Europe. Co-ordination Center for Effects, Technical Report No.1. National Institute of Public health and Environmental Protection, Bilthoven, the Netherlands.

Hinman, G., Rosa, E., and Kleinhesselink, R. 1990. Public Attitudes Toward Risks Associated with Various Activites, Substances and Technologies. For Bonneville Power Administration, DE-AI79-89BP98129. Cited in; Buchanan, S. 1991. Contingent Valuation Study of The Environmental Costs of Electricity Generating Technologies, pp.159-167, in; Hohmeyer, O. and Ottinger, R. L. (eds). 1991. External Environmental Costs of Electric Power. Analysis and Internalisation. Springer-Verlag, Berlin. 497pp.

Hoen, J and Randall, A. 1987. A Satisfactory Benefit Cost Indicator from Contingent Valuation. Jnl Env Econ and Mmt, 14, 3. pp.226-47.

Hohmeyer, O. 1988. Social Costs of Energy Consumption. Springer-Verlag, Berlin.

Hohmeyer, O. 1990. Social Costs of Electricity Generation: Wind and Photovoltaic versus Fossil and Nuclear. Contemporary Policy Issues, 8, July, pp. 255-282

Hohmeyer, O and Gärtner, M. 1992. The Costs of Climate Change. Fraunhofer Institut fur Systemtechnik und Innovationsforschung.

Hohmeyer, O. 1992. Renewables and the Full Costs of Energy. Energy Policy, 20, 4. pp.

Holdren, J P. 1982. Energy Hazards: What to Measure, What to Compare. Technology Rev. April. pp.32-75

Holdren, J P. 1987. Global Environmental Issues Related to Energy Supply: The Environmental Case for Increased Efficiency of Energy Use. Energy, 12, 10-11, pp.975-992.

Holland, W W, Bennett, A E, Cameron, I R, Florey, C D V, Leeder, S R, Schilling, R S F et al. 1979. Health effects of particulate pollution: Reappraising the evidence. American journal of Epidemiology, 110, pp.525-659.

Holmes, A M. 1995. A QALY-based Societal health satistic for Canada, 1985. Social Science and Medicine, 41, 10, pp.1417-1427.

Holub, H W, Tappeiner, G and Tappeiner, U. 1999. Some remarks on the 'System of Integrated Environmental and Economic Accounting' of the United Nations. Ecological Economics, 29, pp.329-336.

Hordijk, L, Kroeze, C. 1997. Integrated assessment models for acid rain. European Journal of Operational Research, 102, pp.405-417.

Horley, J. 1984. Life Satisfaction, Happiness, and Morale: Two Problems with the Use of Subjective Well-Being Indicators. The Gerontologist, 24, 2, pp. 124-127.

Horne, R E. 1994. A review of the organisational structure of the electricity supply industry in the UK. Working Paper 5, Resources Research Unit, Sheffield Hallam University.

Horne, R E. 1995. The potential for applying environmental economics techniques to the electricity supply industry. Working Paper 6, Resources Research Unit, Sheffield Hallam University.

Horne, R E. 1996a. A framework for internalising environmental externalities in the electricity supply industry. Working Paper 9, Resources Research Unit, Sheffield Hallam University.

Horne, R. E. 1996b. Scoping environmental impacts of electricity generation: Development of a methodology. Working Paper 10, Resources Research Unit, Sheffield Hallam University.

Horne, R. E. 1997. A procedure for internalising environmental impacts in the Electricity Supply Industry. Working Paper 11, Resources Research Unit, Sheffield Hallam University.

Horne, R. E. 1999a. QLOS: A single index of quality of life outcome states. Working Paper 18, Resources Research Unit, Sheffield Hallam University, September.

Horne, R E. 1999b. Quality of life outcome states (QLOS) Valuation Method. Working Paper 19, Resources Research Unit, Sheffield Hallam University, October.

Horne, R. E. 1999c. Quality of life outcome states (QLOS) Valuation Method: Trial and Comparison of Results. Working Paper 20, Resources Research Unit, Sheffield Hallam University.

Horne, R E. 2000a. Development of the Quality of life outcome states (QLOS) Pathway Method. Working Paper 21, Resources Research Unit, Sheffield Hallam University, January.

Horne, R E. 2000b. Demonstration of the QLOS Pathway Method: Sulphur Dioxide Stack Emissions. Working Paper 22, Resources Research Unit, Sheffield Hallam University, April.

Horne, R E. 2000c. Incorporating Quality of Life Outcome States into the Regulatory Framework of the Electricity Supply Industry. Working Paper 23, Resources Research Unit, Sheffield Hallam University, July.

Hov, Ø. 1990. Acid deposition. In: Dunderdale, D. 1990. Energy and the environment. Special Publication No.81, Royal Society of Chemistry

Hovmand, M F. 1999. Cumulated deposition of strong acid and sulphur compounds to a spruce forest. Forest Ecology and Management, 114, pp.19-30.

Howells, G. 1990. Acid rain and acid waters. Ellis Horwood.

Hubbard, H M. 1991. The Real Cost of Energy. Scientific American, 264, 4. pp18-23

Hufschmidt, M M et al. 1983. Environment, natural Systems and Development. An Economic Valuation Guide. 338pp. Johns Hopkins Univ Pr.

Hunt, D. and Johnson, C. 1995. Environmental Management Systems. McGraw-Hill.

Hunt, S M, McEwen, J and McKenna, S P. 1986. Measuring Health Status. Croom Helm.

Hunt, S M, McEwen, J, McKenna, S P. 1981. The Nottingham Health Profile User's Manual. NHP.

Innes, J L. 1987. Air pollution and forestry. Forestry Commission Bulletin 70, HMSO.

ISO. 1997. International Standard ISO14040: Environmental management – Life Cycle Assessment – Principles and Framework.

Jackson T J and Marcs N. 1994. Measuring Sustainable Economic Welfare - A Pilot Index: 1950-1990. 46pp. Stockholm Environment Institute.

Jarret, H ed. 1966. Environmental quality in a growing economy. 165 pp. Resources for the Future, INC. The Johns Hopkins Press, Baltimore and London.

Jenkins, A, Helliwell, R C, Swingewood, P J, Sefton, C, Renshaw, M, Ferrier, R C. 1998. Will reduced sulphur emissions under the Second Sulphur Protocol lead to recovery of acid sensitive sites in UK? Environmental Pollution, 99, pp.309-318.

Jenkinson, C. 1991. Why are we Weighting? A Critical Examination of the Use of Item Weights in a Health Status Measure. Social Science and Medicine, 32, 12, pp.1413-1416.

Jenkinson, 1994. Weighting for Ill-health: The Nottingham Health Profile. In: Jenkinson, C. (ed). 1994. Measuring Health Status and Medical Outcomes. Social Research Today, series No. 3. UCL Press.

Jenkinson, C. 1995. Evaluating the Efficacy of Medical treatment: Possibilities and Limitations. Social Science and Medicine, 41, 10, pp.1395-1401.

Jennings, A A and Kuhlman, S J. 1997. An air pollution transport teaching module based on Gaussian Models 1.1. Environmental Modelling and Software, 12, 2-3, pp.151-160.

Johansson, P. 1990. Valuing Environmental Damage. Oxford Review of Economic Policy, 6, 1.

Johansson, P. 1991. An Introduction to Modern Welfare Economics. Cambridge Univ. Press. 176pp.

Jones P. 1990. Social Costs of Energy. Atom 403, May. pp.23-7

Jones-Lee, M. W. and Loomes, G. 1992. Discounting and Safety. Working Paper, University of Newcastle upon Tyne, October. Cited in: Brand M and Jochem E. 1990. The Internal Energy Market. The New Coalition Against Energy Efficiency and Environmental Concerns? Energy Policy, Oct. pp.694-701.

Joskow P. 1992. Weighing Environmental Externalities -lets do it right!

Kahn, E. 1988. Electric utility planning and regulation. American Council for an Energy-Efficient Economy.

Kaplan, R M and Bush, J W. 1982. Health-Related Quality of Life Measurement for Evaluation Research and Policy Analysis. Health Psychology, 1, 1 pp.61-80.

Kaplan, R M, Anderson, J P and Ganiats, T G. 1993. The Quality of Well-being Scale: rationale for a single quality of life index. In: Walker, S R and Rosser, R M. 1993. Quality of Life Assessment: Key Issues into the 1990s. Kluwer Acad Publ.

Kaplan, R M, Bush, J W and Berry, C C. 1979. Health Status Index: Category Rating versus Magnitude Estimation for Measuring Levels of Well-being. Medical Care, 17, 5, pp.501-523.

Katsouyanni, K, Touloumi, G, Spix, C, Schwartz, J, Balducci, F, Medina, S, Rossi, G, Wojtyniak, B, Sunyer, J, Bacharova, L, Schouten, J P, Ponka, A, Anderson, H R. 1997. Short-term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. Air Pollution and Health: a European Approach. British Medical Journal, 314, 7095, pp.1658-1663.

Katz, S, Ford, A B, Moskovitz, R W, Jacobson, B A and Jaffe, M W. 1963. The Index of ADL: A standardised measure of biological and psychological function, JAMA, 185, pp.914-919.

Kealy, M J and Bishop, R C. 1986. Theoretical and Empirical Specification Issues in Travel Cost Demand Studies. American Jnl of Agricultural Econ, 68, pp.660-7.

Keeney, R L. 1982. Evaluating Mortality Risks from an Organisational Perspective. In: Jones-Lee, M W. (ed) 1982. The Value of Life and Safety. Proceedings of a Conference held by the 'Geneva Association'. North-Holland.

Keller, T. 1983. Die auswirkungen von immisionen auf waldbäume. In: Die Erhaltung des waldes- eine nationale und intemationale aufgabe. Grünes Forum Alpbach, Innsbruck.

Keoleian, G A and Lewis, G McD. 1997. Application of Life Cycle Energy Analysis to Photovoltaic Module Design. Progress in Photovoltaics, 5, 4, pp.287-300.

Kind, P and Gudex, C. 1991. The HMQ: measuring health status in the community. York: University of York, Centre for Health Economics.

Kind, P, Rosser, R and Williams, A. 1982. Valuation of Quality of Life: Some Psychometric Evidence. In: Jones-Lee, M W (ed). 1982. Value of Life and Safety. Proceedings of a Conference held by the 'Geneva Association'. North-Holland.

Kneese, A V. 1984. Measuring the Benefits of Clean Air and Water. Resources for the Future. Washington DC.

Krause, G H M. 1987. Forest decline and the role of air pollutants. In: Perry, R, Harrison, R M, Bell, J N B and Lester, J N (eds). Acid rain: Scientific and technical advances.

Krewitt, W, Heck, T, Trukenmüller, A and Friedrich, R. 1999. Environemntal damage costs from electricity generation in germany and Europe. Energy Policy, 27, pp.173-183.

Krupnick, A J and Burtraw, D. 1996. The social costs of electricity: Do the numbers add up? Resource and Energy Economics, 18, pp.423-466.

Krutilla, J and Fisher, A. 1985. The Economics of Natural Environments. 2nd ed. Resources for the Future, Wash. DC.

Kuyken, W and Orley, J. 1994. The Development of the World Health Oragnisation Quality of Life Assessment Instrument (the WHOQOL). In: Orley, J and Kuyken, W. (eds). 1994. Quality of Life Assessment: International Perspectives. Proceedings of the joint Meeting organised by the WHO and the Fondation IPSEN, Paris, July 1993. Springer-Verlag, Berlin.

Lancaster, K. 1966. A New Approach to Consumer Theory. Jnl of Political Econ., 74. pp.132-57.

Lee, J C, Fthenakis, V M, Morris, S C, Goldstein, G A, and Moskovitz, P D. 1997. Projected photovoltaic energy impacts on US CO_2 emissions: An integrated energy-environmental-economic analysis. Progress in Photovoltaics: Research and Applications, 5, pp.277-285.

Lee, R, Krupnick, A, Burtraw, D et al. 1994. Estimating Externbalities of electric fuel cycles: Analytical methods and issues, Estimating Externalities of Fuel Cycles. McGraw-Hill/Utility Data Institute, Washington, DC. Cited in: Freeman, A M. 1996. Estimating the environmental costs of electricity: an overview and review of the issues. Resource and Energy Economics, 18, pp.347-362.

Lefohn, A S, Husar, J D, Husar, R B. 1999. Estimating historical anthropogenic global sulfur emission patterns for the period 1850-1990. Atmospheric Environment, 33, pp.3435-3444.

Legge, A H and Krupa, S V.1990. Acidic deposition: Sulphur and nitrogen oxides. Lewis Publishers.

Leopold, L. et al. 1971. A Procedure for Evaluating Environmental Impact. US Geological Survey, Circular 645. Washington DC: US Geological Survey.

Levy, J I, Hammitt, J K, Yanagisawa, Y and Spengler, J D. 1999. Development of a new damage function model for power plants: Methodology and applications. Environmental Science and Technology, 33, 24, pp. 4364-4372.

Li, J, Tang, H. 1998. A theoretical calculation model for the acidification capacity of natural waters. The Science of the Total Environment, 212, pp.163-172.

Linnerooth, J. 1982. Murdering Statistical Lives...? In: Jones-Lee, M W. (ed) 1982. The Value of Life and Safety. Proceedings of a Conference held by the 'Geneva Association'. North-Holland.

Lipfert, F W. 1994. Air pollution and community health. A critical review and data sourcebook. Van Nostrand Reinhold.

Lipfert, F W. 1989. Atmospheric damage to calcareous stones: Comparison and reconciliation of recent experimental findings. Atmospheric Environment, 23, 2, pp.415-429.

Little, I M D. 1953. The Price of Fuel. Oxford.

Llewellyn-Thomas, H et al. 1984. Describing Health States: Methodological issues in Obtaining Values for Health States. Medical Care, 22, 6. pp.543-552.

Lodge, M. 1981. Magnitude Scaling: Quantitative Measurement of Opinions. A Sage University Paper, No. 07-025. Sage Publications.

Loomes, G and McKenzie, L. 1989. The use of QALYs in Health Care Decision Making. Social Science and Medicine, 28, 4, pp.299-308.

Loomes, G and McKenzie, L. 1990. The Scope and Limitations of QALY measures. In: Baldwin, S, Godfrey, C and Propper, C. 1990. Quality of Life: Perspectives and Policies. Routledge.

Loomis, J, Creel, M & Park, T. 1991. Comparing Benefit Estimates from Travel Cost and Contingent Valuation Using Confidence Intervals for Hicksian Welfare Measures. Applied Economics, 23, 1725-1731

Lowles, I, ApSimon, H, Juda-Rezler, K, Abert, K, Brechler, J, Holpuch, J, Grossinho, A. 1998. Integrated assessment models - tools for developing emission abatement strategies for the *Black Triangle* region. Journal of Hazardous Materials, 61, pp.229-237.

MacKenzie, J J and El-Ashry, M T, eds. 1989. Air pollution's toll on forests and crops. World Resources Institute, Yale University Press.

Madariaga, B & Mc Connell, K E. 1987. Exploring Existence Value. Water Resources Research 2 (5), May, pp. 936-942

Mahaney, W C, Wilson, E, Boyer, M G, Hancock, R G V. 1995. Marginal bleaching of Thalli of *Rhizocarpon* as evidence for acid rain in the Norra Storfjället, Sweden. Environmental Pollution, 87, pp.71-75.

Markyanda, A and Richardson, J (eds). 1992. Earthscan Reader in Environmental Economics. Earthscan, London.

Marwell, G and Ames, R E. 1981. Economists Free Ride, Does Anybody Else?. Experiments on the Provision of Public Goods IV. Jnl Econ Lit, 21. pp. 281-317.

Maslow, A H. 1962. Toward a Psychology of Being. Van Nostrand.

Mason, B J, ed. 1990. The surface waters acidification programme. Cambridge University Press.

Mason, C F and Macdonald, S M. 1987. Acidification and otter Lutra lutra distribution on a British river. Mammalia, 51, pp.82-87.

Matschonat, G, Vogt, R. 1998. Significance of the total cation concentration in acid forest soils for the solution composition and the saturation of exchange sites. Geoderma, 84, pp.289-307.

Mayer, R. 1998. Soil acidification and cycling of metal elements: Cause-effect relationships with regard to forestry practices and climatic changes. Agriculture, Ecosystems and Environment, 67, pp.145-152.

McKenna, S P, Hunt, AS and McEwen. 1981. Weighting the Seriousness of Perceived Health Problems Using Thurstone's Method of Paired Comparisons. Int Journal of Epidemiology, 10, 1, pp.93-97.

McLaren, D. 1996. Enduring Decisions: Evaluating environmental impacts for sustainability planning. Local Environment, 1, 1, February. pp. 51-62.

McLeod, A R, Roberts, T M, Alexander, K and Cribb, D M. 1991. The yield of winter cereals exposed to sulphur dioxide under field conditions. Agriculture, Ecosystems and Environment, 33, pp.193-213.

McNeil, B J, Weichselbaum, R and Pauker, S G. 1981. Speech and Survival: Trade-offs between quality and quantity of life in laryngeal cancer. New England Journal of Medicine, 305. pp.982-987.

Meadows, D H et al. 1974. The Limits to Growth. Pan, London.

Mehrez, A, Gafni, A. 1989. Quality-Adjusted Life Years, Utility Theory, and Healthy-years Equivalents. Medical Decision Making, 9, pp.142-149.

Midttun, A and Thomas, S. 1998. Theoretical ambiguity and the weight of historical heritage: A comparative study of the British and Norwegian electricity liberalisation. Energy Policy, 26, 3, pp.179-197.

Miller, P, Bytnerowicz, A, Fenn, M, Poth, M, Temple, P, Schilling, S, Jones, D, Johnson, D, Chow, J, Watson, J. 1998. Multidisciplinary study of ozone, acidic deposition and climate effects on a mixed conifer forest in California, USA. Chemosphere, 36, 4-5, pp.1001-1006.

Mirasgedis, S and Diakoulaki, D. 1997. Multicriteria analysis vs. externalities assessment for the comparative evaluation of electricity generation systems. European Journal of Operational Research, 102, pp.364-379.

Mooney, G H. 1986. Choices for Health Care: A Practical Introduction to the economics of health provision. Macmillan.

Mori, K, Tachi, K, Muramatsu, M, Torita, K. 1999. Mechanism of acid rain etching of acrylic-melamine coatings. Progress in Organic Coatings, 36, pp.34-38.

Muniz, I P. 1991. Freshwater acidification: its effects on species and communities of freshwater microbes, plants and animals. In: Last, F T, and Watling, R (eds). Acidic deposition: its nature and impacts. Royal Society of Edinburgh, Scotland.

Mylona, S N. 1989. EMEP MSC-W Report 1/89. Norwegian Meteorological Institute, Oslo.

NAPAP. 1987. The National Acid Precipitation Programme. NAPAP, Washington DC.

Navrud, S. 1989. Estimating Social Benefits of Environmental Improvements from Reduced Acid Depositions: A Contingent Valuation Survey. In: Valuation Methods and Policy Making in Environmental Economics. Folmer, H and Vanlerland, E. Elsevier.

Nease, R F, Tsai, R, Hynes, L M, Littenberg, B. 1996. Automated utility assessment of global health. Qual Life Res, 5, 1, pp.175-82.

Nicolaisen, J, Dean, A and Hoeller, P. 1991. Economics and the Environment: A Survey of Issues and Policy Options. OECD Econ Studies 16, Spring. pp.7-43

Nilsson, J and Grennfelt, P (eds). 1988. Critical loads for sulphur and nitrogen. Milrapport 1988:15, Nordic Council of Ministers, Copenhagen.

Nordhaus, W. 1991. The Costs of Slowing Climate Change: A Survey. The Energy Journal, 12, 1. pp.37-65

O'Riordan, T, and Jager, J. 1996. Politics of climate change: A European perspective. Routledge.

Oates, W E. ed. 1992. The Economics of the Environment. E. Elgar, Vermont.

O'Brien, K J. 1992. Environmental Factors Affecting Generation Planning. Power Engineering Journal, January pp. 32-41

OECD. 1992. Climate Change. Designing a Tradeable Permit System. OECD.

OECD. 1994. Managing the Environment. The Role of Economic Instruments. OECD.

OECD. 1999. National Climate Policies and the Kyoto Protocol. OECD.

OECD. 2000. Voluntary Approaches for Environmental Policy: An Assessment. OECD.

OFFER. 1991. Report on Pool Price Inquiry. Office of Electricity Regulation (OFFER), December.

OHE. 1989. Measurement and Management in the NHS. Office of Health Economics, London.

OPEC. 1993. Energy in the 21st Century; New Challenges and Goals. OPEC Review, Summer. pp.127-149.

Openshaw Taylor, E and Boal, G A, eds. 1969. Power System Economics. Edward Arnold.

O'Riordan, T, Cameron, J. ed. 1994. Interpreting the Precautionary Principle. Earthscan.

Orley, J and Kuyken, W. (eds). 1994. Quality of Life Assessment: International Perspectives. Proceedings of the joint Meeting organised by the WHO and the Fondation IPSEN, Paris, July 1993. Springer-Verlag, Berlin.

Ottinger, R et al. 1991. Environmental Costs of Electricity. Pace University Center for Environmental Legal Studies. Oceana Publications Inc., New York.

Owen, B, Edmunds, H A, Carruthers, D J and Raper, D W. 1999. Use of a new generation urban scale dispersion model to estimate the concentration of oxides of nitrogen and sulphur dioxide in a large urban area. The Science of the Total Environment, 235, pp.277-291.

Pal, D. undated. A numerical model of acidification of raindrops by oxidation of SO_2. International Journal of Environmental Studies.

Paoletti, E. 1998. UV-B and acid rain effects on beech *(Fa gus sylvatica L.)* and holm oak *(Quercus ilex L.)* leaves. Chemosphere, 36, 4-5, pp.835-840.

Papadopoulos, R. 1981. Growth and Overcapacity in the UK Electricity Industry. Energy Policy. 9, 2.

Park, C C. 1987. Acid rain: Rhetoric and reality. Methuen.

Patrick, D L, Wild, D J, Johnson, E S, Wagner, T H and Martin, M A. 1994. Cross-cultural validation of Quality of Life measures. In: Orley, J and Kuyken, W. (eds). 1994. Quality of Life Assessment: International Perspectives. Proceedings of the joint Meeting organised by the WHO and the Fondation IPSEN, Paris, July 1993. Springer-Verlag, Berlin.

Patterson, W. 1999. Transforming electricity. Royal Institute of International Affairs and Earthscan Publications.

Pearce, D W and Turner, K. 1992. Benefits Estimates and Environmental Decision Making. 59pp. OECD-OCDE

Pearce, D W et al. 1991. Blueprint 2. Earthscan, London.

Pearce, D W et al. 1993. Blueprint 3. Earthscan, London.

Pearce, D W et al. 1994. Blueprint 4. Earthscan, London.

Pearce, D W, and Markyanda, A. 1989. Environmental Policy Benefits: Monetary Valuation. OECD, Paris. 83pp.

Pearce, D W, Bann, C and Georgiou, S. 1992. The Social Cost of Fuel Cycles. Report to the UK Dept of Trade and Industry. CSERGE. HMSO.

Pearce, D W, Markyanda, A, and Barbier, E B. 1989. Blueprint for a Green Economy. Earthscan Publications, London.

Pennanen T, Perkiömäki, J, Kiikkilä, O, Vanhala, P, Neuvonen, S, Fritze, H. 1998. Prolonged, simulated acid rain and heavy metal deposition: separated and combined effects on forest soil microbial community structure. FEMS Microbiology Ecology, 27, pp.291-300.

Pezzey, J. 1989. Economic Analysis of Sustainable Growth and Sustainable Development. Env Dept Working Paper No 15. World Bank, Washington DC.

Pigou, A. 1920. Economics of Welfare. MacMillan, London.

Pikhart, H, Príkazský, V, Bobák, M, Kríz, B, Celko, M, Danová, J, Pyrl, K, Pretel, J. 1997. Association between ambient air concentrations of nitrogen dioxide and respiratory symptoms in children in Prague, Czech republic. Preliminary results from the Czech part of the SAVIAH Study. Small Area Variation in Air Pollution and Health. Central European Journal of Public Health, 5, 2, pp.82-85.

PIRA. 1993. PEMS Life Cycle Inventory Analysis Model. PIRA International.

Pirog, R L and Stamos, S C. 1987. Energy Economics. Theory and Policy. Prentice-Hall. 319pp.

Pope III, C A, Bates, D V and Raizenne, M E. 1995. Health effects of particulate air pollution: Time for reassessment? Environmental Health Perspectives, 103, 5, pp.472-480.

Portney, P R and Weyant, J P (eds). 1999. Discounting and intergenerational equity. 186pp. Resources for the Future, Washington, DC.

Proctor, M C F, Maltby, E. 1998. Relations between acid atmospheric deposition and the surface pH of some ombrotrophic bogs in Britain. Journal of Ecology, 86, 2, pp.329-340.

Quist, M E. 1998. Acid episodes decisive for plant performance. Chemosphere, 36, 4-5, pp.627-632.

Rasmussen, L. 1998. Effects of afforestation and deforestation on the deposition, cycling and leaching of elements. Agriculture, Ecosystems and Environment, 67, pp.153-159.

RCEP. 2000. Energy - The Changing Climate. Royal Commission on Environmental Pollution. 22nd Report. CM4749. The Stationery Office.

Reich, S L, Gomez, D R, Dawidowski, L E. 1999. Artificial neural network for the identification of unknown air pollution sources. Atmospheric Environment, 33, pp.3045-3052.

Reuss, J O and Johnson, D W. 1986. Acid deposition and the acidification of soils and waters. Springer Verlag, New York.

Reynolds, B, Lowe, J A H, Smith, R I, Norris, D A, Fowler, D, Bell, S A, Stevens, P A, Ormerod, S J. 1999. Acid deposition in Wales: The results of the 1995 Welsh Acid Waters Survey. Environmental Pollution, 105, pp.251-266.

Rizzoli, A E, Davis, J R. 1999. Integration and re-use of environmental models. Environmental Modelling & Software, 14, pp.493-494.

Roberts, J, Elliott, D and Houghton, T. 1991. Privatising Electricity. The Politics of Power. Belhaven.

Roberts, T M. 1984. Long term effects of sulphur dioxide on crops: an analysis of dose-response relations. Phil. Transcripts of the Royal Society of London, B 305, pp.299-316.

Rosenburg, R. 1995. Health-Related Quality of Life Between Naturalism and Hermeneutics. Social Science and Medicine, 41, 10, pp.1411-1415.

Rosser, R and Kind, P. 1978. A scale of valuations of states of illness. Is there a social consensus? International Journal of Epidemiology, 7, pp. 347-358.

Rosser, R M and Watts, V C. 1972. The measurement of hospital output. International Journal of Epidemiology, 1, pp. 361-8.

Rothman, D S. 2000. Measuring environmental values and environmental impacts: Going from the local to the global. Climatic Change, 44, pp.351-376.

Rowe W D and Oterson P. 1983. Energy: Assessment of Comparative and Non-comparative Factors in Alternate Energy Systems. 138pp. CEC, Luxemburg

Rowe, R D, Lang, C M, Chestnut, L G et al. 1995. The New York electricity externalities study. Oceana Publications, New York. Cited in: Freeman, A M. 1996. Estimating the environmental costs of electricity: an overview and review of the issues. Resource and Energy Economics, 18, pp.347-362.

Rowe, R D, Lang, C M and Chestnut, L G. 1996. Critical factors in computing externalities for electricity resources. Resource and Energy Economics, 18, pp.363-394.

Sáez, R M, Linares, P and Leal, J. 1998. Assessment of the externalities of biomass energy, and a comparison of its full costs with coal. Biomass and Bioenergy, 14, 5/6, pp.469-478.

Samuelson, P A. 1954. Pure Theory of Public Expenditure. Rev of Economics and Statistics, 36. pp.387-9.

Sandness, H and Styve, H. 1992. Calculated budgets for airborne acidifying components in Europe 1985, 1987, 1988, 1989, 1990 and 1991. DET Norske Meteorologiske Institutt Technical Report No 97. EMEPIMSC-W Report 1/92.

Schall, P, Augustin, S and Schmieden, U. 1998. Modelling aspects of forest decline in Germany: II. Application and validation of an integrated Soil-Plant-Model. Chemosphere, 36, 4-5, pp.971-978.

Schmieman, E C, van Ierland, E C. 1999. Dynamics of soil acidification: An economic analysis. Ecological Economics, 31, pp.449-462.

Schöpp, W, Amann, M, Cofala, J, Heyes, C and Klimont, Z. 1999. Integrated assessment of European air pollution emission control strategies. Environmental Modelling and Software, 14, pp.1-9.

Schulz, H, Popp, P, Huhn, G, Stärk, H-J, Schüürmann, G. 1999. Biomonitoring of airbone inorganic and organic pollutants by means of pine tree barks. Temporal and spatial variations. The Science of the Total Environment, 232, pp.49-58.

Schulze, W D, d'Arge, R C and Brookshire, D S. 1981. Valuing environmental commodities: Some recent experiments. Land Economics, 57, 2, pp.151-172.

Seland, Ø, Iversen, T. 1999. A scheme for black carbon and sulphate aerosols tested in a hemispheric scale, Eulerian dispersion model. Atmospheric Environment, 33, pp.2853-2879.

Seller, C, Stoll, J R and Canvas, J-P. 1985. Validation of Empirical Measures of Welfare Change: A Comparison of Nonmarket Techniques. Land Economics, 61, 2, pp.156-75.

SETAC. 1993. SETAC – Europe: Guidelines for Life Cycle Assessment: A Code of Practice. SETAC Workshop, Sesimbra, Portugal, March 31.

Shilberg, G M, Hahigan, J A, and Marcus, W B. 1989. Valuing Reductions in Air Emissions and Incorporation into Electric Resource Planning: Theoretical and Quantitative Aspects. JBS Energy Inc, CA.

Shy, C M. 1979. Epidemiologic evidence and the United States air quality standards. American Journal of Epidemiology, 110, pp.661-671.

Siebert, H. 1981. Economics of the Environment. 231pp. Lexington Books, Mass., Toronto.

Sintonen, H. 1981. An Approach to Measuring and Valuing Health States. Social Science and Medicine, 15C, pp.55-65.

Smith, F B and Hunt, R D. 1979. The dispersion of sulphur pollutants over western Europe. In: anon. 1979. Pathways of pollutants through the atmosphere. The Royal Society, London.

Smith, R A. 1852. On the air and rain of Manchester. Memoirs of the Literary and Philosophical Society of Manchester, Series 2 10, pp.207-217.

Smith, R A. 1872. Air and rain: The beginnings of chemical climatology. Longmans Green, London.

Smith, V K, Desvousges, W H and Fisher, A. 1986. A Comparison of Direct and Indirect Methods for Estimating Environmental Benefits. American Agricultural Economics Association, May, pp. 280-290

Smith, V K. 2000. JEEM and non-market valuation: 1974-1998. Journal of Environmental Economics and Management, 39, pp.351-374.

Söderholm, P and Sundqvist, T. 2000. Ethical limitations of social cost pricing: An application to power generation externalities. Journal of Economic Issues, 34, 2, pp.453-462.

Sogn, TA, Abrahamsen, G. 1998. Effects of N and S deposiUon on leaching from an acid forest soil and growth of Scots pine *(Pinus sylvestris L.)* after 5 years of treatment. Forest Ecology and Management, 103, pp.177-190.

Spedding, L S. 1999. Environmental Law Guide. Old Bailey Press, Law in Practice Series.

Spilker, B (ed). 1990. Quality of Life Assessments in Clinical Trials. Raven Press.

Spiro, P A, Jacob, D J and Logan, J A. 1992. Global inventory of sulfur emissions with 1° x 1° resolution. Journal of Geophysical Research, 97, D5, pp.6023-6036.

Spitzer, W O. et al. 1981. Measuring the Quality of Life of Cancer Patients. J Chronic Diseases. 34. pp.585-597.

Šrámek, V. 1998. SO_2 air pollution and forest health status in northwestern Czech Republic. Chemosphere, 36, 4-5, pp.1067-1072.

Stanord, A. 1998. Liberalisation of the UK energy market: An opportunity for green energy. Renewable Energy, 15, 1-4, pp.215-217.

Stern, J P. 1990. European Gas Markets: Challenge and Opportunity in the 1990's. Royal Institute of International Affairs. Dartmouth.

Stirling, A. 1992. Regulating the Electricity Supply Industry by Valuing Environmental Effects. How much is the emperor wearing? Futures, December, pp.1024-47.

Stirling, A. 1997. Limits to the value of external costs. Energy Policy, 25, 5, pp.517-540.

Stirling, A. 1998. Valuing the environmental impacts of electricity production: A critical review of some "first generation" studies. Energy Sources, 20, pp.267-300.

Suomela, J, Neuvonen, S, Ossipova, S, Ossipov, V, Pihlaja, K. 1998. A long-term study of the effects of simulated acid rain on birch leaf phenolics. Chemosphere, 36, 4-5, pp.639-644.

Tabatabai, M A. 1985. Effect of acid rain on soils. Critical Reviews in Environmental Control, 15, pp.144-171.

Tellus Institute. 1991. Valuation of Environmental Externalities: Sulfur Dioxide and Greenhouse Gases. Report to Massachusetts Division of Energy Resources, Boston, MA.

TERG. 1988. The effects of acid deposition on the terrestrial environment in the United Kingdom. United Kingdom Terrestrial Effects Review Group, 1st report, HMSO.

Thayer, M. 1991. Estimating the air quality impacts of alternative energy resources, Phase IV. California Energy Commission (Regional Economic Research).

Thurston, G D, Gorczynski Jr, J E, Currie, J H, He, D, Ito, K, Hipfner, J, Waldman, J, Lioy, P J and Lippmann, M. 1994. The nature and origins of summer haze air pollution in metropolitan Toronto, Ontario. Environmental Research, 65, pp.254-270.

Tietenberg, T. 1994. Environmental Economics and Policy. Harper Collins, London.

Tolba, M K. 1983. Acid rain - a growing concern of industrialised countries. Water Quality Bulletin, 8, pp.115-20.

Torrance, G W, Thomas, W H, and Sackett, D L. 1972. A utility maximization model for the evaluation opf health care programmes. Health Services Research, 7, pp. 118-33.

Torrance, G W. 1976a. Social preferences for health states: An empirical evaluation of three measurement techniques. Socio-economic Planning Science, 10, pp.129.

Torrance, G W. 1976b. Towards a Utility Theory Foundation of Health Status Index Models. Health Services Research, 11, pp.349-69.

Torrance, G W. 1987. Utility Approach to measuring health-related quality of life. Journal of Chronic Diseases, 40, pp.593-600.

Turchenek, L W, Abboud, S A, Tomas, C J, Fessenden, R J, and Holowaychuk, N. 1987. Effects of acid deposition on soils in Alberta. Preparation for the Alberta Government-Industry acid deposition research programme by the Alberta research Council, Edmonton. ADRP-B-05-87.

Turner R K and Bateman I J. 1990. A Critical Review of Monetary Assessment Methods and Techniques. Environmental Appraisal Group, Univ. East Anglia.

Turner, R K. 1993. Sustainable Environmental Economics and Management: Principles and Practice. Belhaven, London.

Ulrich, B, Miewes, K J, Konig, N and Khanna, P K. 1984. Untersuchungversfahren und Kriterion zur Bewertung der Versaurung und ihrer Folgen in Waldboden. Forst. u. Holzwirt, 39, pp.2378-286.

UN ECE. 1984. Airborne sulfur pollution: Effects and control. ECE, UN, Geneva, Switzerland.

UN ECE. 1993. Manual on Methodologies and criteria for mapping critical levels/loads and geographical areas where they are exceeded. UN ECE Task Force on Mapping, Federal Environmental Agency, Berlin.

UN. 1993. The Global Partnership for Environment and Development. A Guide to Agenda 21 Post Rio Edition. c.230pp. UN.

USEPA. 1994. Life-Cycle Impact Assessment. US Environmental Protection Agency (draft).

Van Ierland, E C. ed. 1994. International Environmental Economics. Theories, Models and Applications to Climate Change, International Trade and Acidification. Developments in Environmental Economics, 4.

Veljanovski, C. ed. 1991. Regulators and the Market. An assessment of the growth of regulation in the UK. Institute of Economic Affairs.

Ventegodt, S. 1996. Measuring the Quality of Life: Form Theory to Practice. Forskningscentrets Forlag, Copenhagen.

Walker, S R and Rosser, R M. 1993. Quality of Life Assessment: Key Issues into the 1990s. Kluwer Acad Publ.

Walna, B, Drzymala, S, Siepak, J. 1998. The impact of acid rain on calcium and magnesium status in typical soils of the Wielkopolski National Park. The Science of the Total Environment, 220, pp.115-120.

Ware, J E, Kosinski, M, Bayliss, M S, McHorney, C A, Rogers, W H, Raczek, A E. 1995. Comparison of methods for scoring and statistical analysis of SF-36 health profiles and summary measures: summary of results from the Medical Outcomes Study. Med Care; 33 (Suppl 4), pp.A5264-A5279.

Ware, J E, Kosinski, M, Keller, S D. 1994. SF-36 Physical and Mental Summary Scales: A User's Manual. Boston, MA. The Health Institute.

Ware, J E, Snow, K K, Kosinski, M, Gandek, B. 1993. SF-36 Health Survey: Manual and Interpretation Guide. Health Inst, New England Med Centre, Boston, Mass.

Ware, J H, Thibodeau, L A, Speizer, F E, Colome, S, and Ferris, B G Jr. 1981. Assessment of the health effects of atmospheric sulfur oxides and particulate matter: Evidence from observational studies. Environmental Health Perspectives, 41, pp.255-276.

Warren, S C, Alexander, G C, Bache, B W, Battarbee, R W, Crawshaw, D H, Edmunds, W M, Egglishaw, H J, Gee, A S, Hildrew, A G, Hornung, M, Howells, G D, Hunt, D T E, Pugh, K B, Wells, D E, Whitehead, P G, Cape, J N, Harriman, R, Neal, C and Ormerod, S J. 1988. United Kingdom Acid Waters Review Group Final Report. Department of the Environment, London.

WCED. 1987. Our Common Future. World Commission on Environment & Development Report. Oxford University Press.

Weinstein, M C, Fineberg, H V, Elstein, AS et al. 1980. Clinical Decision Analysis. Saunders, US.

Wellburn, A. 1988. Air pollution and acid rain. The biological impact. Longman Scientific and Technical.

Whitehead, P G, Musgrove, T J and Cosby, B J. 1990. Hydrochemical modelling of acidification in Wales. In: Edwards, R W, Gee, A S and Stoner, J H, eds. 1990. Acid Waters in Wales. Kluwer Academic Publishers.

Wiel, S. 1991. The New Environmental Accounting: A Status Report. The Electricity Journal. November.

Wilkin, D, Hallam, L and Doggett, M. 1992. Measures of Need and Outcome for Primary Health Care. Oxford.

Williams, S J. 1994. Measuring Functional Limitations and sickness Impact: A critical Review of the FLP and SIP. In: Jenkinson, C. (ed). 1994. Measuring Health Status and Medical Outcomes. Social research Today series No. 3. UCL Press.

Willig, R. 1976. Consumers Surplus Without Apology. Amer. Econ. Rev., 66. pp.589-97.

Willis, K G. 1990. Valuing Non-market Wildlife Commodities: An Evaluation and Comparison of Benefits and Costs. Applied Economics, 22. pp.13-30.

Wilson, E. 1993. Strategic Environmental Assessment: Evaluating the Impacts of European Policies, Plans & Programmes. European Environment vol 3 no.2 pp2-6

Winpenny, J T. 1991. Values for the Environment. A Guide to Economic Appraisal. HMSO, London. 269pp.

Winstanley, D, Lackey, R T, Wamick, W L and Malanchuk, J. 1998. Acid Rain: Science and policy making. Environmental Science and Policy, 1, pp.51-57.

Woolf, T. 1992. Accounting for the Environmental Externalities of Electricity Production: A Summary of US practice. Assoc for the Conservation of Energy, London.

YEG. 2000. Environmental Performance Report 1999. Yorkshire Electricity Group Plc.

Young, P. 1998. Data-based mechanistic modelling of environmental, ecological, economic and engineering systems. Environmental Modelling and Software, 13, pp.105-122.