Deducing Water Parameters in Rivers via Statistical modelling

Ahmed Moustafa

Myriad Vision Ltd
58 Regent Road
Leicester
LE1 6YJ
UK

Centre for Innovative and Collaborative Engineering
Department of Civil & Building Engineering
Loughborough University
Loughborough
Leicestershire, LE11 3TU

Loughborough University

Thesis Access Form

Copy No ……………………………… **Location**………………………………

Author Ahmed Moustafa

Title Deducing Water Parameters in Rivers via Statistical modelling

Status of access ~~OPEN~~ / RESTRICTED / ~~CONFIDENTIAL~~

Moratorium period: …0…-…5…years, ending …30……-…09… / 20…1…-…5……

Conditions of access proved by (CAPITALS): …………………………………………………

Director of Research (Signature) …………………………………………………………………

Department of Department of Civil and Building Engineering

Author's Declaration: I agree the following conditions:

OPEN access work shall be made available (in the University and externally) and reproduced as necessary at the discretion of the University Librarian or Head of Department. It may also be copied by the British Library in microfilm or other form for supply to requesting libraries or individuals, subject to an indication of intended use for non-publishing purposes in the following form, placed on the copy and on any covering document or label.

The statement itself shall apply to **ALL** copies:

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

Restricted/confidential work: All access and any photocopying shall be strictly subject to written permission from the University Head of Department and any external sponsor, if any.

Author's signature ……………………………… **Date** ………………………

Users declaration: for signature during any Moratorium period (Not Open work):
I undertake to uphold the above conditions:

Date	Name (CAPITALS)	Signature	Address

Loughborough University

Certificate of Originality

This is to certify that I am responsible for the work submitted in this thesis, that the original work is my own except as specified in acknowledgments or in footnotes, and that neither the thesis nor the original work contained therein has been submitted to this or any other institution for a higher degree.

Author's signature ....

Date ....

DEDUCING WATER PARAMETERS IN RIVERS VIA STATISTICAL MODELLING

By
Ahmed Moustafa

A dissertation thesis submitted in partial fulfilment of the requirements for the award of the degree Doctor of Engineering (EngD), at Loughborough University

[March 2011]

Myriad Vision Ltd
58 Regent Road
Leicester
LE1 6YJ
UK

Centre for Innovative and Collaborative Engineering
Department of Civil & Building Engineering
Loughborough University
Loughborough
Leicestershire, LE11 3TU

ACKNOWLEDGEMENTS

I would like to express my heartfelt thanks to a number of people who have supported me and made this research possible.

I would like to thank my academic supervisors: Dr. Ashraf El-Hamalawi and Prof. Andrew Wheatly. Ashraf's continual guidance, support, and encouragement throughout the duration of this project have been invaluable. Andrew's assistance and positive attitude have helped me stay on course. I consider it a privilege and an honour to have had the opportunity to work with them and share their valuable knowledge and expertise. Sincere thanks go to Andrew's secretary, Ms Christine Barton, for her support in booking Ashraf's and Andrew's meetings with me and for being there to talk to.

Special thanks and gratitude to CICE ex-director, Prof Chimay Anumba, for his guidance and encouragement during my early start of the research. I am also grateful to the CICE Centre Staff, in particular Jo, Colette, Sara, and Nadine for the administrative assistance.

Many thanks go to Myriad Vision Ltd and specially Dr Salman Ahmed for providing me with this research opportunity.

I owe my loving thanks to my wife and children, for giving me love and comfort to do my research and for helping me through the difficult times and reassuring me that I had done the right thing.

Special gratitude goes to my father and my mother for being there for me and for pushing me to finish this degree. Their unconditional love and support kept me going and spurred me on to greater heights. To simply put it, I would not, and could not, have done this degree without them.

I would like to heartily thank my sister and brother for helping me put this thesis together and for being always there for me.

Finally, warm thanks to Prof. Tarek Hassan and Dr Shabbir Ahsan for their friendship and endless support during tough times until the present day.

Dedicated to my father Ibrahim,
my mother Samia,
my loving wife Safa,
my daughter Noura,
my son Munir,
my sister Shaimaa,
and my brother Khaled,
For their love, inspiration and belief

ABSTRACT

Advanced monitoring of water quality in order to perform a real-time hazard analysis prior to Water Treatment Works (WTW) is more of a necessity nowadays, both to give warning of any contamination and also to avoid downtime of the WTW. Downtimes could be a major contributor to risk. Any serious accident will cause a significant loss in customer and investor confidence. This has challenged the industry to become more efficient, integrated and attractive, with benefits for its workforce and society as a whole.

The reality is that water companies are not yet prepared to invest heavily in trials, before another company announces its success in implementing a new monitoring strategy. This has slowed down the development of the water industry.

This research has taken the theoretical idea that the use of advanced online monitoring technique in the water industry would be beneficial and a step further; demonstrating by means of a state-of-the-art assessment, usability trials, case studies and demonstration that the barriers to mainstream adoption can be overcome. The findings of this work have been presented in four peer-reviewed papers.

The research undertaken has shown that Turbidity levels in rivers can be measured from the rivers' mean flow rate, using either Doppler Ultrasound device for real-time readings or based on past performance history. In both cases, the Turbidity level can also help estimate both the Colour and Conductivity levels of the subject river. Recalibration of the equations used is a prerequisite as each individual river has its own unique "finger print".

KEY WORDS

Acoustic-Doppler, Online, Drinking Water, Sensors, Turbidity, Colour, Conductivity, Flow, Rive

PREFACE

This thesis represents the research conducted between 2004 and 2008 to fulfil the requirements of an Engineering Doctorate (EngD) at the Centre for Innovative and Collaborative Engineering (CICE), Loughborough University, UK. The research was undertaken within an industrial context; supervised by CICE, funded by the Engineering Physical Sciences Research Council (EPSRC) and sponsored by Myriad Vision Ltd, a web technology and management consultancy firm.

The essence of the Engineering Doctorate is to solve one or more significant and challenging engineering problems within an industrial context that can be shown to be of benefit not only to the sponsoring company but also to the wider construction industry.

The EngD is examined on the basis of a thesis containing at least three (but not more than five) research publications and/or technical reports. This discourse is supported by one journal paper and three conference papers. The main body of the thesis allows the reader to gain an overview of the work undertaken, while more specific aspects of the research can be found in the papers presented in the appendices at the back of the thesis. These papers are an integral part of, and should be read when referenced in conjunction with, the thesis.

USED ACRONYMS / ABBREVIATIONS

ANN	Artificial Neural Networks
COD	Chemical Oxygen Demand
DO	Dissolved Oxygen
DOM	Dissolved Organic Matter
DUF	Doppler-Ultrasound Flow
EA	Environment Agency
EEC	European Economic Community
EngD	Engineering Doctorate
EPA	Environmental Protection Agency
FTU	Formazin Turbidity Units
GI	Gastro-Intestinal illness
GQA	General Quality Assessment
GTAP	Ground Troublesome Asset Project
IR	Infrared
JTU	Jackson Turbidity Units
MLR	Multivariate Linear Regression
MRA	Maximum Redundancy Analysis
NTU	Nephelometric Turbidity Units
Ofwat	The Water Services Regulation Authority
PAHs	Polycyclic Aromatic Hydrocarbons
PCA	Principal Component Analysis
PCR	Principal Components Regression
PLS	Partial Least Squares
SCADA	Surveillance Control and Data Acquisition
SS	Suspended Sediments
ST	Severn Trent
USGS	United States Geological Survey
WFD	Water Framework Directive
WTW	Water Treatment Works
WWT	Waste Water Treatment
…	…

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF PAPERS

The following papers, included in the appendices, have been produced in partial fulfilment of the award requirements of the Engineering Doctorate during the course of the research.

PAPER 1 (SEE APPENDIX A)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2008. The impact of bad sensors on the water industry and possible alternatives, ITcon, 13, Special Issue, Sensors in Construction and Infrastructure Management, pp. 166-178.

PAPER 2 (SEE APPENDIX B)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2007. Case studies on the need for monitoring of water quality in the UK. Eleventh International Water Technology Conference, IWTC11 2007 Sharm El-Sheikh, Egypt, pp. 389-994.

PAPER 3 (SEE APPENDIX C)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2007. Online Laboratory Investigation on the use of Acoustic Doppler in Turbidity Measurement. 8th UK's National Young Water Professionals Conference - Guildford, UK.

PAPER 4 (SEE APPENDIX D)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2008. Deciphering river flow data to determine river properties. Twelfth International Water Technology Conference, IWTC12 2008, Alexandria, Egypt, pp. 305-315.

1 BACKGROUND TO THE RESEARCH

1.1 INTRODUCTION

This chapter sets out the circumstances to the research undertaken to fulfil the requirements for the award of an Engineering Doctorate (EngD) at Loughborough University. It provides an opening to the general subject field, outlines the aim and objectives, validates the need for the research and sets it within an industrial perspective. The structure of the thesis is presented to provide clarity and direction to the reader and a synopsis is provided for each of the published papers that should be read in conjunction with the discourse (Appendices 1 - 4).

1.2 CONTEXT OF THE RESEARCH

All earth's water originates from the atmosphere as precipitation (rain, snow and hail). This is collected either above ground in rivers, natural lakes, man-made impounding reservoirs or below ground in aquifers. Water rapidly absorbs both natural and man-made substances, generally making the water unsuitable for drinking prior to some form of treatment.

The objective for water treatment is to produce an adequate and continuous supply of water that is chemically, bacteriologically and aesthetically pleasing. More specifically, water treatment must produce water that is (American Waterworks Association, 1990):

a) Palatable (i.e. no unpleasant taste);
b) Safe (i.e. does not contain pathogens or chemicals harmful to the consumer);
c) Clear (i.e. free from suspended solids and turbidity);
d) Colourless and odourless (i.e. aesthetic to drink);
e) Reasonably soft (i.e. allows consumers to wash clothes, dishes, or even themselves, without the use of excessive quantities of detergents or soap);

f) Non-corrosive (i.e. to protect pipework and prevent leaching of metals from tanks or pipes);

g) Of low organic content (high organic content results in unwanted biological growth in pipes and storage tanks that often affects quality).

Early European water legislation began in a "first wave", with standards for those of our rivers and lakes used for drinking water abstraction in 1975, and culminated in 1980 in setting binding quality targets for our drinking water (European Commission Environment website, 2006).

In 1988, the Frankfurt ministerial seminar on water reviewed the existing legislation and identified a number of improvements that could be made, and gaps that could be filled. This resulted in the second phase of water legislation; the first results of this were, in 1991, the adoption of (European Commission website, 2009):

- The Urban Waste Water Treatment Directive, providing for secondary (biological) waste water treatment, and even more stringent treatment, where necessary.
- The Nitrates Directive, addressing water pollution by nitrates from agriculture.

Other legislative results of these developments were the European Commission's proposals for action on:

- A new Drinking Water Directive, reviewing the quality standards and, where necessary, tightening them (adopted November 1998),
- A Directive for Integrated Pollution and Prevention Control (IPPC), addressing pollution from large industrial installations (adopted in 1996).

In 2000, the European Union introduced the Water Framework Directive (WFD) 2000/60/EC. The Water Framework Directive is the most substantial piece of water legislation ever produced by the European Commission, and will provide the major operator for achieving sustainable management of water in the UK and other member states. Some amendments have been introduced since, but the main directive remains the main doctrine.

In general, Water Treatment Works (WTW) must be able to produce a finished product of consistently high quality regardless of how great the demand might be. Like Waste Water Treatment (WWT), WTW consists of a range of unit processes, usually used in series, which provides some design and operational flexibility to achieve this.

The treatment required by water prior to being delivered to consumers will depend upon its initial quality, which is normally related to its source. In other words, the cleaner the raw water, the fewer treatment steps that are required, and hence the overall cost of water is less. The most expensive operations in conventional treatment are sedimentation and filtration, while water softening can also be very expensive. Groundwater is generally much cleaner than surface water, and thus does not require the same degree of treatment, apart from aeration and disinfection before supply. Naturally occurring substances that may need to be reduced or removed in groundwater include iron, hardness (if > 300 mg/l as $CaCO_3$) and Carbon dioxide. Substances originating from humans are becoming increasingly common in groundwater and those requiring treatment include nitrates, pathogens and trace organics such as pesticides. Surface water, on the other hand, requires more complex treatment due to its complex nature, although the quality of surface waters can be very high, e.g. upland reservoirs (Geldreich, 1996; Reasoner, 1992).

The selection of water sources for supply purposes depends on the ability of the resources to meet consumer demand throughout the year. Furthermore, this relies not only on the nature of the water, but on the cost of treating the water as well.

The principal quality characteristics of sources of potable (drinkable) water are:

1. Ground water – normally of a high bacteriological quantity and of a considerable salinity.
2. Surface water (upland sources) – of a relatively good quality, but may be corrosive and may be coloured.
3. Surface water (lowland sources) – may be appreciably polluted, requiring extensive and costly purification.
4. Rain water harvesting – of a high quality, but often with a low pH. Care should be taken with regard to the collecting surfaces and the storage containers' cleanliness.
5. Wastewater re-use.
6. The sea – desalination requires the use of high technology, costly processes such as distillation and reverse osmosis.

The need for a new advanced monitoring system was crucial after the Camelford water pollution incident in July 1988. The Camelford incident involved the accidental contamination of the drinking water supply to the town of Camelford, UK, with 20 tonnes of aluminium sulphate which was poured into the wrong tank at the Lowermoor water treatment works. As the aluminium sulphate broke down, it produced several tonnes of sulphuric acid which stripped a cocktail of chemicals from the pipe networks as well as lead and copper piping in people’s homes. Almost 20,000 people were poisoned.

According to many news articles (BBC, 2007; Metro, 2010) many people complained to the water authorities of the taste and the cloudiness of the tap water, but their complaints went unheeded. This disastrous and lethal incident could have been prevented through the use of smart and intelligent sensors.

The research area in this thesis was invoked by the findings of the literature review and an examination of several case studies. The case study and the operators' observations highlighted these shortcomings and inefficiencies in the UK water industry. In general, these problems are:

1. Leakages in pipes; with increased labour costs, it is very expensive to keep digging and replacing old pipes.
2. Supply and demand; climatic changes do have its toll on supply sources and demand at different times.
3. Quality improvement; the continuous development in water technologies will leave its mark on suppliers to improve the quality to customers.

In 2004, 375 sites were monitored by the Environment Agency for compliance with Surface Water Abstraction Directive (75/440/EEC) in England and Wales. Of these, 155 sites failed to comply with the directive (Environment Agency Website, 2007). This is more than 40% of the sites monitored.

The inefficiencies contribute massively to the rise in cost, inadequate water processing and increased processing times. It also affects the water quality produced by the water treatment plants.

Research has shown that increased use and continuing advancements of real-time remote monitoring and sensing technologies will become a progressively more important tool for evaluating water quality. But evidence to the contrary, the water industry is very much slow in picking up these new technologies and deploying them and, which has been proven by the EA findings.

In summary, this research will help lead to the following:

1. Increase reliability and efficiency of Water Treatment Works' processes.
2. Reduce operating costs by reducing unnecessary processes working time.
3. Flexibility for Surveillance Control and Data Acquisition (SCADA) system growth.
4. Improved operator interface and communication with WTW.
5. Increased data measured accuracy and improve predicted models.
6. Effective data handling and management through online content web designs.
7. Fully integrated information systems with data management tools.
8. Ultimately monitor and/or control operations at remote sites via the web.
9. Relational database server to provide and feed the SCADA system with data for improved system maintenance and better water modelling.

1.3 RESEARCH AIM AND OBJECTIVES

1.3.1 OVERALL AIM

The aim of this project is to enhance the Surveillance Control and Data Acquisition (SCADA) system operating Water Treatment Works (WTW) in the UK through the improvement of sensors' design and control software, enabling better control over changes in rivers, and consequently water quality. This research used mathematical modelling techniques to correlate the river flow with its constituents, correlate flow measurements in the Midlands region in the UK, and redeveloped an existing Ultrasound Flow meter to measure in real-time the Turbidity level, and thus help deduce the Conductivity and Colour levels.

1.3.2 OBJECTIVES

1. To conduct a literature review in order to investigate the different types of water treatments' processes, critical water parameters and valuable sensors to measure them.

2. To look at the inefficiencies existing in water treatment works by conducting trend analyses of archived critical parameter measures.

3. To analyse, via case studies, the sensors' readings and establish their correlative relationships with each other. This would reduce the number of sensors if readings can be abstracted from other sensors. The data would also be used to validate any working models.

4. To investigate a new approach to non-invasive sensors based on existing sensors with the application of artificial intelligence and multiplexing of sensors.

5. To examine the potential of Stochastic or IT-based modelling and improve archiving and presentation of data, and in turn use it as a benchmark for future water treatment works.

1.4 THE INDUSTRIAL SPONSOR

The EngD sponsoring company, Myriad Vision Ltd, is a UK based innovative software company, which specialises in the development of Content Management and E-learning solutions for the desktop and mobile/Hand-held arena. Myriad Vision was constantly aspiring towards expanding the possibilities of information sharing, and knowledge distribution, via novel tools and pioneering technology. This created the opportunity to expand into the water industry arena, and further enhance its products' range of solution management, via online web designs.

The overall aim of Myriad in undertaking this research project was to spring board the company range of product solutions, and how it can help solve technology related problems within the water industry via its online content management solutions. The research output was designed to be fed back into Myriad's range of web engines tools' to be accommodating to water industry requirements in terms of online control and management support.

1.5 PROBLEM DEFINITION

The problem in hand can be classified as a control theory and mathematical/experimental problem. The automation processes run by a SCADA system at specified WTW are insufficient to maintain a high level of drinking water quality. The automation processes are very much dependent on human input. This has been flagged up by the operator of the sites, Severn Trent Water, through the Groundwater Troublesome Assets Project (GTAP), to be of major concern. With this simple definition in mind, this research focused on finding alternative backup solutions to the SCADA sensory system by trying to eliminate human dependability during sampling procedures and data validation of the working sensors.

1.6 THESIS STRUCTURE

In addition to this introductory chapter, the structure of this EngD thesis comprises the following:

- **CHAPTER TWO**: discloses the findings of the literature review on the subject of non-invasive water constituent's detection techniques to acknowledge existing research work and the latest research milestones achieved within this field.

- **CHAPTER THREE**: presents the adopted research methodology before outlining the method and tools applied to the research project, along with justification for their use.

- **CHAPTER FOUR**: describes the research undertaken with reference to the set aims and objectives; and

- **CHAPTER FIVE**: discusses the findings of the research and their impacts and implications for the water industry as a whole. It also contains an evaluation of the new sensory techniques developed, and makes recommendations for further research and development.

1.7 CHAPTER SUMMARY

Chapter 1 has presented the reader with an introduction to this EngD research project and the structure for which this thesis is based on. It highlighted the need for the research undertaken, its aim and objectives, and what advantages the new and highly improved measuring technique can achieve. It also pointed to the need for the water industry to open up to more trials of ideas and research. Finally, the project's problem definition was provided, followed by the thesis breakdown structure.

2 LITERATURE REVIEW

2.1 INTRODUCTION

This literature review explores the four dominant themes of the research; 1) Flow as the focal element; 2) Turbidity as a measureable ingredient; 3) Sensors as the operating tool ; and 4) Correlation as the binding linkage.

The scope of this literature review is expanded to include research that examines the dominant themes of the research, regardless of the specific academic subject area. The literature review concludes with a review of what this research has contributed to the research community. Figure 2-1 illustrates this relationship.

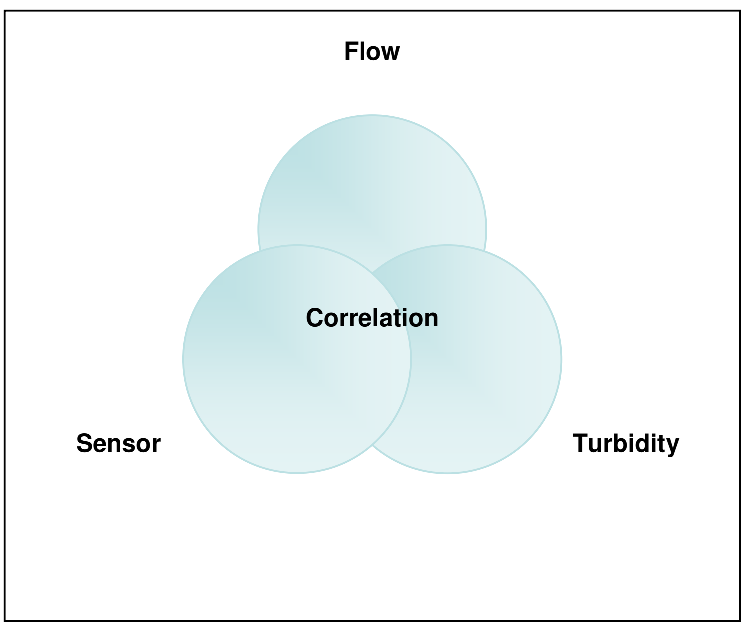


Figure 2-1 Illustration of the relation between the four dominant research themes

2.2 BACKGROUND

Generally, over the decades, water has been named differently; one of which is the "universal solvent" because of its ability to dissolve numerous substances. The term "water quality" relates to all the constituents of water, including both dissolved substances and any other substances carried by the water.

The quality of water in UK rivers is monitored by the Environment Agency in order to assess the overall health of the inland water ecosystem. Water quality is one of the 68 indicators of the Government's Sustainable Development Strategy.

River water quality can be influenced by a wide range of factors, either acting independently or collectively and interacting together. These factors include:

- River flow can have a significant effect on quality - lower than average river flows can result in poorer water quality as there is reduced dilution of pollutants. On the other hand, high river flows can improve water quality by increasing the dilution of pollutants.

- Temperatures can result in the growth of algae, which affects dissolved oxygen demand in rivers;

- High rainfall can adversely affect river quality by causing increased leaching of pollutants in the soil. Overflows from the sewerage system can also occur during periods of particularly intense rainfall; and

The way the Environment Agency measure water quality is changing. For twenty years, since 1990, a General Quality Assessment (GQA) scheme has been used to assess river water quality in terms of chemistry, biology and nutrients. GQA has helped drive environmental improvements by dealing with many of the major point sources of pollutants, such as discharges from sewage treatment works or other industry. A more sophisticated way of assessing the whole water environment is needed to help direct the action to where it is most needed.

The European Water Framework Directive (WFD) will provide the means to monitor by looking at over 30 measures, grouped into ecological status (this includes biology as well as 'elements', like phosphorus and pH) and chemical status ('priority substances'). The WFD covers estuaries, coastal waters, groundwater and lakes, as well as rivers. More information about the Water Framework Directive is available on the Environment Agency website.

2.2.1 RIVER FLOW

Many large rivers throughout the world are subjected to some form of water resource development, resulting in river regulation and altered flow regimes. The widespread concern about the environmental effects of river regulation is based on the logical assumption that the biota and ecological functioning of river systems depend on the volume and timing of flows, and the longitudinal, lateral and vertical connections they facilitate.

The primary purpose of freshwater flow quality monitoring is to generate sufficient and timely information to enable managers to make informed management decisions regarding the exposure health risk of the populations who are utilizing this resource. The secondary purpose of freshwater quality monitoring may include issues such as monitoring the quality of the environment for the essential needs of the habitat, identification of pollution sources, and thus potential polluters, immediate initiation of clean-up operations after an accidental or deliberate spill, concerns on potential terrorism events and the use of the data collected to identify stringent rules and regulations to avert the adverse effects that may be caused by the consequential environmental degradation.

This literature reviewed a subset of refereed and un-refereed literature to assess the evidence for geomorphological responses to flow modifications in rivers. The research undertaken produced overwhelming evidence that both river ecology and river geomorphology change in response to flow modification.

History

The design issues associated with freshwater flow quality monitoring have received wide attention since the 1970's (Beckers and Chamberlain, 1974). Various attempts were made in the 1980's to improve monitoring efficiency with regard to basic design criteria and the constraints of the problem (Skalski and McKenzie, 1982; Ilker et al., 2009). Groot and Schilperoort (1984) discussed the optimization analysis, whereas an emphasis on data collection was addressed in Whitfield (1988), and the interpretation of the monitoring outcome was chosen as the main theme in Ellis (1989). Following these preliminary studies, the literature that followed in the 1990's was concerned with the optimal location of water quality monitoring stations (Esterby, 1996; Loftis et al., 1991; Smith and McBride, 1990). Kwiatkowski (1991) identified the essential objectives that are common to many large scale water quality monitoring networks that are extensively utilized since then. Later studies used integer programming (Hudak et al., 1995), and multi-objective programming

(Cieniawski et al., 1995; Harmancioglu and Alpaslan, 1992), in which more complex issues were addressed. These methods were implemented ever since in various site specific applications. However, in this early literature, for most cases, steady state solutions were considered simply because of the complexity of the watersheds and the rivers that are considered in those studies. A comprehensive review of this literature can be found in Strobl and Robillard (2008). More recently, an online water quality management system is also proposed for the Liming River system in China (Yang et al., 2008).

Why Flow?

Flow is related to the incidences of natural hazards, such as flood and drought, which occur abruptly and may result in loss of human and animal lives and damage to human properties. Flood alert systems hold the highest possibility of reducing the damages from the floods. On the other hand, drought analysis also depends on appropriate forecasts of flow. Flow prediction, therefore, provides crucial information for adaptive water resources management. However, fluctuations of global climate change poses a challenge to scientists and engineers to estimate and forecast the magnitude and timing of stream discharges with higher accuracy. Recent scientific advances in remote sensing and artificial intelligence technologies empower such an effort of flow forecasting.

The importance of river flow is well recognized by ecologists and water resource managers. Nevertheless, water is a valuable commodity as well as a destructive force. Human society requires water for life, while seeking protection from floods and droughts. As a consequence, many rivers have been heavily modified to enable water managers to control flows to meet human needs (e.g., industry, agriculture, development of historical floodplains), while dampening or eliminating normal floods and droughts (Gleick, 2003). By trapping floods rather than conveying them downstream, many regulated rivers retain little of their original flow variability (Poff et al., 2007). Human control of river flows is now nearly ubiquitous (Lytle & Poff, 2004; Vörösmarty et al., 2004). Today, rivers are managed to meet multiple human demands (e.g., steady and dependable water supplies, flood control facilities to protect populated areas), but these factors severely constrain the flow variability that is required to meet ecological demands (Naiman et al., 2000; Postel & Richter, 2003).

River flow forecasting is required to provide basic information on a wide range of problems related to the design and operation of river systems. The availability of extended records of rainfall and other climatic data, which could be used to obtain stream flow data, initiated the practice of rainfall-runoff modelling (Dibike & Solomatine, 2001). While conceptual or physically-based models are of importance in the understanding of hydrological processes, there are many practical situations where the main concern is with making accurate predictions at specific locations. In such situations, it is preferred to implement a simple "black box" (data driven, or machine learning) model to identify a direct mapping between the inputs and outputs without detailed consideration of the internal structure of the physical process.

Flow affects turbidity and dissolved oxygen (D.O.) concentrations. Affects in general are;

- High-velocity streams are more erosive and suspend sediments for a longer time, leading to greater turbidity.
- Turbulent, fast-moving streams are better aerated and therefore have higher concentrations of dissolved oxygen.

- Greater volume maintains cooler temperatures.

Long-term surveys and monitoring programs of water quality are an adequate approach to a better knowledge of river hydrochemistry and pollution, but they produce large sets of data that are often difficult to interpret (Dixon and Chiswell, 1996). Most discussions on trend detection focus on analysing a single variable, while routine monitoring programs ordinarily measure several variables. The problem of data reduction and interpretation of multi-constituent chemical and physical measurements can be approached through the application of multivariate statistical methods and exploratory data analysis (Massart et al., 1988; Wenning and Erickson, 1994). The usefulness of multivariate statistical tools in the treatment of analytical and environmental data is reflected by the increasing number of papers cited in Analytical Chemistry Reviews (Brown et al., 1994, 1996).

Cluster analysis and Principal Component Analysis (PCA) have been widely used in the chemical engineering field, as they are unbiased multivariate methods which can indicate associations between samples and/or variables (Wenning and Erickson, 1994). These associations, based on similar magnitudes or variations in chemical and physical constituents, may indicate the presence of seasonal or man-made influences. Hierarchical agglomerative cluster analysis indicates groupings of samples by linking inter-sample similarities and illustrates the overall similarity of variables in the data set (Massart and Kaufman, 1983).

Reservoir impoundments are of great social and economic importance, as their primary functions are to provide a reliable water supply, flood management and means of generating renewable energy. However, regulated flow discharges from dams have had adverse ecological impacts in rivers throughout the world (e.g. Australia (Arthington and Pusey, 2003); Kazakhstan (Starodubstev et al., 2004); South America (de Merona et al., 2005), UK (Cowx and Gould, 1989); United States (Postel and Richter, 2003; Bell and DeLacy 1967 in Bizere 2000). Regulated rivers may have a significantly altered hydrological regime and/or water quality. For example, low flows may be increased, high flows may be reduced and/or occur at unnatural times of the year, flow variability may change, and hydrographs may be a different shape. Furthermore, water released from a reservoir may be cooler in summer/warmer in winter (Jensen, 2003), lower in dissolved oxygen and suspended sediment (often with complete absence of coarse sediment), and higher in nutrients (e.g. Collingwood, 1966) than downstream river water. Water discharge from high spillways may also be supersaturated with gases (Bell and DeLacy, 1967, in Bizere, 2000).

The EU Water Framework Directive (WFD) requires that water bodies, including rivers and lakes, reach “Good Ecological Status” by 2015. The quality elements for the classification of a river’s ecological status are presented in Annex V of the Directive (EEC, 2000). The biological elements include the composition and abundance of aquatic flora, and invertebrate and fish fauna. The hydromorphological elements supporting these biological elements include: hydrological regime, river continuity, and morphological conditions. The chemical and physico-chemical elements supporting the biological elements include thermal, oxygenation and nutrient conditions. All these elements may be influenced by the regulation of a river downstream of an impoundment.

In general, there are a number of elements that contribute to river flow levels;

1. *Climate*. Weather patterns have the greatest influence on flow. Areas with higher precipitation produce streams with greater average volume.

2. *Season*. River flow will vary throughout the year.

3. *Watershed*. If all other factors, such as precipitation, are the same, flow volume will increase as the size of the watershed increases.

4. *Groundwater*. Groundwater will contribute to flow if the channel is lower than the top of the groundwater.

To conclude, the author of this literature review found it very hard to find relevant research papers where direct connection between freshwater flows levels are associated with other hydrochemistry constituents. Most of the research carried out in this field was indicative of the influence of flow on these constituents, but not as a measuring tool to derive these constituents. This under-researched area will form a major component of this thesis.

2.2.2 TURBIDITY

As mentioned in chapter 1, 375 sites were monitored by the Environment Agency for compliance with the Surface Water Abstraction Directive (75/440/EEC) in England and Wales in 2004. Of these, 155 sites failed to comply with the Directive. Major pollutants are Suspended Sediments (SS), turbidity, pathogens, nutrients, metals, dissolved organic matter (DOM), pesticides, chlorophylls (algae, plants), temperature, and oils. However, over 90% of these 'failures' were due to insufficient sampling. Sampling shortfalls occur for a number of reasons, such as abstractions not being operated at the time of sampling, problems at the laboratory, and sampling error. The quality of abstracted water has generally improved since 1993. It was found that levels of coloration, nitrate and Polycyclic Aromatic Hydrocarbons

(PAHs) most commonly exceeded the Directive's standards in 2004 (Environment Agency, 2004).

What is Turbidity?

In simple terms, Turbidity is a measure of water's lack of clarity. Water with high turbidity is cloudy, while water with low turbidity is clear. The cloudiness is produced by light reflecting off particles in the water; therefore the more particles in the water, the higher the turbidity.

Many factors can contribute to the turbidity of water in rivers. An increase in river flow due to heavy rains or a decrease in stream-bank vegetation can speed up the process of soil erosion. This will add suspended particles, such as clay and silt, to the water. Sources of turbidity in rivers can generally be attributed to:

1. Soil erosion
    a. Silt
    b. Clay
2. Urban Runoff
    a. Road grime
    b. Rooftops
    c. Parking lots
3. Industrial Waste
    a. Sewage treatment effluent
    b. Particulates
4. Organics
    a. Microorganisms
    b. Decaying plants and animals

c. Gasoline or oil from roads

Measurement of turbidity

Turbidity is the measure of the light scattering properties of water. It is typically measured using in-situ equipment which record the attenuation (i.e. attenuance turbidimeters – measure the loss in intensity of a narrow parallel beam or dual beams) or scattering (i.e. nephelometric turbidimeters – measure light scattered at an angle to the beam), of a beam of radiation (Lewis, 1996). Nephelometric turbidimeters have been most widely used, recording turbidity data in nephelometric turbidity units (NTU) (Lewis, 1996).

In turbidimeters, the light is scattered in all directions off the particles in the water. A detector, consisting of a photodiode, is placed at a 90° angle to the light source. The amount of light being scattered directly into the detector is measured in 'Volts' and translated into turbidity units. This style of turbidity sensor is called a Nephelometer or Nephelometric turbidity sensor. A standard is used to calibrate the Turbidity Sensor in units of NTU, or Nephelometric Turbidity Units.

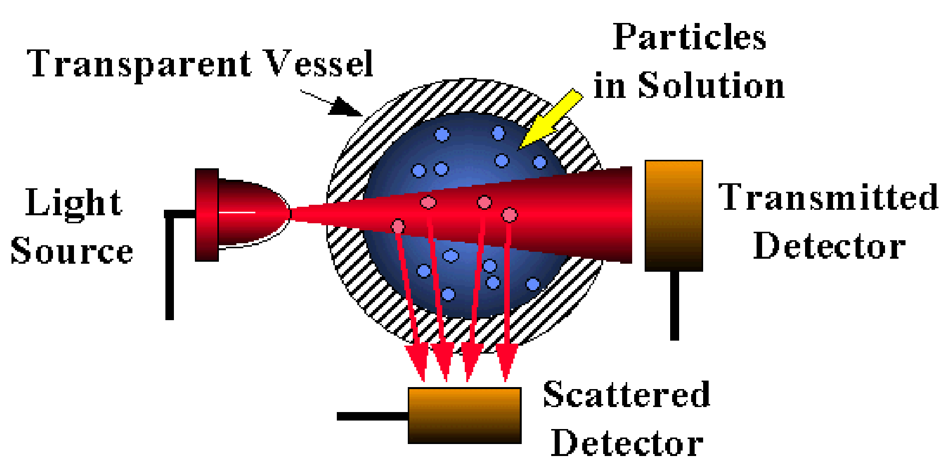


Figure 2-2 Turbidity Sensor diagram (Honeywell website, 2008)

Figure 2.3 illustrates light from an LED as it passes through the vessel, and is impressed on two photoreceptor diodes that monitor transmitted and scattered light. The ratio of these two signals provides a measure of turbidity (Honeywell Website, 2008).

Other units such as JTU (Jackson Turbidity Units), and FTU (Formazin Turbidity Units), have values similar to NTU, but are not exactly the same. The NTU units will be used through the course of the research. Water samples that are removed from the source and are measured with bench top meters are considered static, while online monitors are considered dynamic.

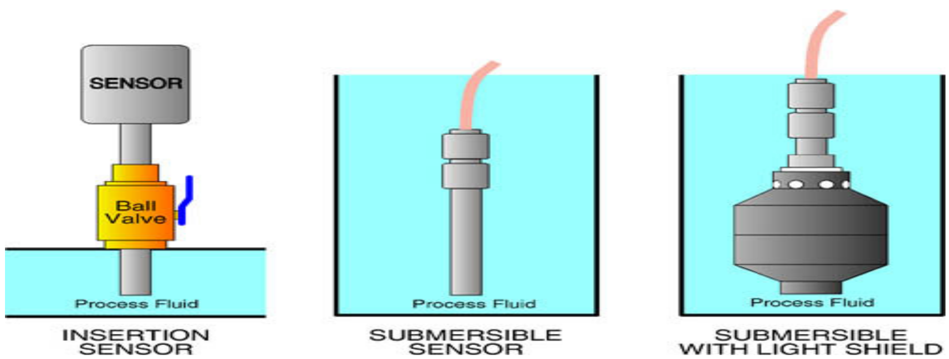


Figure 2-3 Three different Turbidity sensors usage techniques (Keyence website, 2007)

Turbidity measured by different nephelometers has been reported to produce different numerical NTU values (McGirr, 1974) because of differences in optical design of nephelometers even when the different instruments are all calibrated to Formazin. For this reason, the instrument used for turbidity measurements should always be specified. More recently, Davies-Colley and Smith (2001) reported turbidity measurements on 77 New Zealand rivers on two laboratory bench nephelometers: a 'traditional' Hach 2100A and its more modern replacement, a Hach 2100AN, to differ by 30% on average over a 3 order of magnitude range.

Advantages of current turbidity sensors are:

- Simple designs.
- Low cost.
- Not easily damaged.
- Clear indicator of Turbidity concentration.

Disadvantages of current Turbidity sensors are:

- Invasive.
- Requires maintenance frequently.
- Samples taken can sometimes be deceiving; as they could be unsystematic and unreflective of the water being sampled.
- Light based meters are susceptible to interference from other light sources in its vicinity.
- Laser-based meters consume a lot of power and require special maintenance.
- In-situ meters can be clouded repeatedly during stormy seasons.
- Sensors cannot be used as indicators for any other parameter in water.
- Particles sizes in water are wavelength-dependent.
- Measurements taken under static conditions compared to those taken under dynamic conditions differ primarily because static-measurement techniques do not completely account for particle settling, whereas dynamic-measurement techniques more accurately reflect the dynamic nature of particle movement within the water body.
- Temperature changes in the sample during transport from source water to the laboratory can cause differences between measurements taken on a static sample and measurements taken under dynamic conditions.

Why measure Turbidity?

Microbial or chemical contamination of drinking water resulting from inadequate treatment at the plant or poor control of the distribution system can cause acute Gastro-Intestinal (GI) illness (Eife et al., 1999; Knobloch et al., 1994). Although microbiological contamination is commonly accompanied by increases in turbidity, other factors, including silt and organic matter, also affect turbidity levels of water leaving the treatment plant (IRSEC, 2005). Limits of acceptable turbidity for water leaving the treatment plant vary between countries, but are generally below 1 or 2 Nephelometric Turbidity Units (NTU) (IESWTR, 1998; Rouse, 2001); most effluent turbidity readings are well below these limits (LeChevallier, 2004).

Outbreaks of GI illness have been linked to incidents in which turbidity exceeded acceptable limits (CDCP, 1993; Schuster, 2005). However, it is unclear whether endemic GI illness is associated with drinking water within acceptable turbidity levels ('normal' turbidity). Schwartz et al. (1997) reported an association between variations in normal drinking water turbidity and endemic GI illness in children in Philadelphia, USA. The Environmental Protection Agency (EPA) concluded that the study's results were invalid, citing flaws in turbidity measurement and analysis techniques (EPA, 1998; Sinclair, 2000). Despite this, subsequent studies in different settings have also suggested the existence of such an association (Morris et al., 1996; Aramini et al., 2000; Beaudeau et al., 1999; Gilbert et al., 2006; Morris et al., 1998; Schwartz et al.; 2000).

With the medical and scientific fields' interests in turbidity influence on humans in particular that make it through the water treatment works, the Drinking Water Inspectorate of England and Wales commissioned a study to determine what existing evidence there is for an association between drinking water turbidity and endemic GI illness in settings with public water supplies, similar to that in the United Kingdom (UK) (Mann et al., 2007). Mann study concluded that associations between drinking water turbidity and GI illness have been found in two settings. It is likely that studies observed different results because of differences in the mean turbidity level between settings. More specifically, the study called for alternative indicators of water quality technology that would give a particle count data to differentiate between bacterial and parasitic agents based on their size. The study does submit that such technology is expensive and not widely used, probably due to cost.

What's the difference between Suspended Sediments and Turbidity?

Suspended Sediment (SS) causes a range of environmental damage, including benthic smothering, irritation of fish gills, and transport of sorbed contaminants. Much of the impact, while sediment remains suspended, is related to its light attenuation, which reduces visual range in water and light availability for photosynthesis. Thus measurement of the optical attributes of suspended matter in many instances is more relevant than measurement of its mass concentration.

On the other hand, Nephelometric Turbidity is widely used as a simple, and cheap, instrumental surrogate for SS concentration. Turbidity may relate more directly than mass concentration to optical effects of suspended matter. However, turbidity is only a relative measure of light scattering (versus arbitrary standards); albeit a useful quantity where a relative index of water cloudiness is sufficient. Turbidity has no intrinsic environmental relevance until calibrated to a 'proper' scientific quantity (Davies-Colley and Smith, 2001).

Davies-Colley and Close (1990) showed that broad correlations between the three variables: turbidity, SS concentration, and visual clarity, exist. They showed significant interrelationships between these variables in a wide range of New Zealand rivers at baseflow (Davies-Colley and Close, 1990). All three variables were reasonably closely correlated, although the two optical variables (turbidity and visual clarity) are more closely related to each other than either is to SSC.

Generally speaking, it would be difficult to broadly correlate between turbidity and SS concentration based on one experiment. This thesis has highlighted that each river, or ecosystem, does have its own environmental fingerprint that will need deciphering in a more "individual" and "specified" method to eliminate errors and standardise the operation.

Bilotta and Brazier (2008) reported on the influence of Suspended Solids on water quality by calling for the turbidity to be treated with caution, and that it should be considered as a surrogate to SS measurements. They highlighted that turbidity readings respond to factors other than just concentrations of SS, as well as being influenced by the particle-size distribution and shape of SS particles. In addition, turbidity is a measure of only one of the many detrimental effects reviewed in the paper (Bilotta and Brazier, 2008). Despite their push for the use of SS measurement technology, the authors do concede that SS measurement is time-consuming and expensive, particularly if a large number of samples are to be collected and analysed. The authors, however, recommend that high resolution turbidity monitoring should be supplemented with direct measurements of SS (albeit at lower resolution due to resource issues). This would allow the turbidity record to be checked and calibrated against SS, and effectively building a rating-relationship between SS and turbidity, which would, in turn, provide a clearer picture of the exact magnitude of the SS problem. Figure 2-1 below illustrates the various components when measuring turbidity and suspended solids.

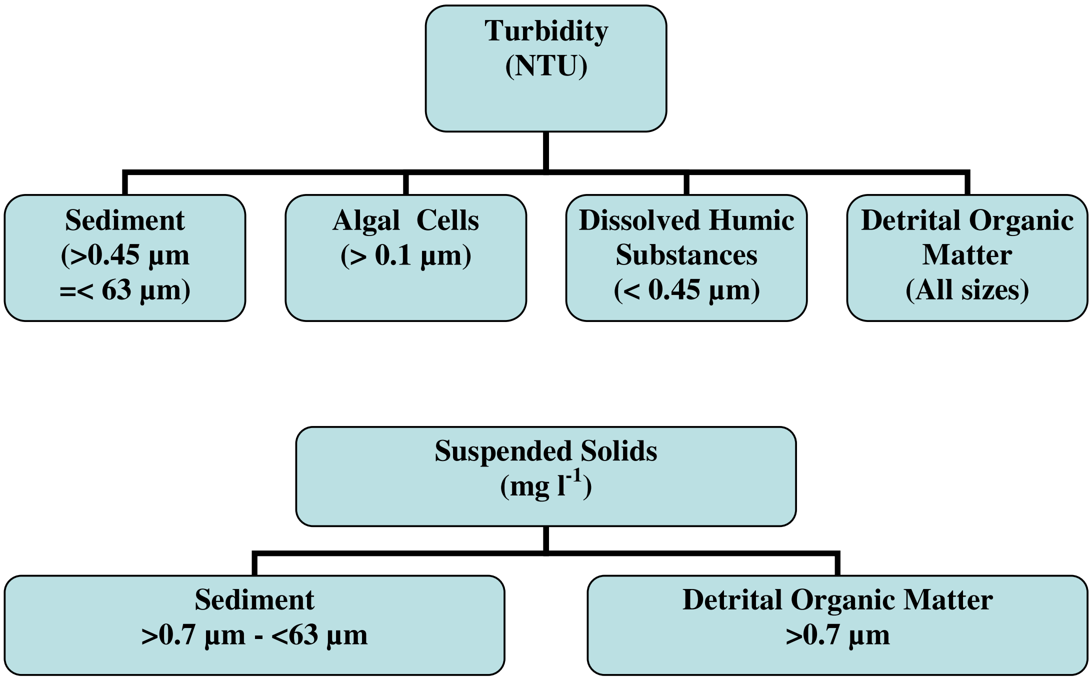


Figure 2-4 Schematic diagram illustrating the various components that are incorporated when measuring turbidity (top) and suspended solids via the conventional method (bottom) (Bilotta and Brazier, 2008).

Other turbidity effects include;

1. Aesthetically displeasing; reduced clarity also makes the water less aesthetically pleasing. While this may not be harmful directly, it is certainly undesirable for many water uses.
2. Decreases photosynthetic rate; high turbidity will decrease the amount of sunlight able to penetrate the water, thereby decreasing the photosynthetic rate.
3. Increases water temperature; because the SS particles in the water absorb the sunlight, they start to warm the surrounding water directly. This can lead to other problems associated with increased temperature levels.

2.2.3 CORRELATION

Quantifying relationships between freshwater flow rate and water quality components (e.g. turbidity) is a major goal of many water quality monitoring programs. This is a challenging task that is rarely achieved through simple analysis of raw data alone. Multiple regression analysis provides one approach, which, despite significant limitations, can be successful when very large data sets are available and only annual estimates are required.

In order for the WTW to meet its objectives and obligations, in terms of water quality and maintain production cost at minimum, it would need to rely heavily on automated processes where less human interaction would be needed, and artificial intelligence would dominate the processes for quick and less expensive responses. The automated processes would, in turn, rely heavily on the pre-calibration steps that the human operator would perform to optimize its operations and correlate its readings, as to avoid the need for human operator interference with it.

Current techniques for measuring water quality involve in-situ measurements and/or the collection of water samples for subsequent laboratory analyses. While these technologies provide accurate measurements for a point in time and space, they are expensive, and do not provide either the spatial or temporal view of water quality needed for monitoring, assessing, or managing water quality for an individual water body or for multiple water bodies across the landscape (Ritchie & Cooper, 2001).

Although parametric, statistical and deterministic models have been the traditional approaches for modeling the water quality, these require a vast amount of information on various hydrological sub-processes in order to arrive at the end results. In recent years, several researchers have conducted studies on water quality forecast models (Chen et al., 2003; Kurunc et al., 2005; Li, 2006). However, since a large number of factors affecting the water quality have a complicated non-linear relation with the variables; traditional data processing methods are no longer good enough for solving the problem (Wu et al., 2000; Xiang et al., 2006).

Artificial Neural Networks (ANN)

On the other hand, the Artificial Neural Networks (ANNs) are capable of imitating the basic characteristics of the human brain, such as self-adaptability, self-organization and error tolerant, have been widely adopted for model identification, analysis and forecast, system recognition and design optimization (Niu et al., 2006; Shu, 2006). Unlike many statistically-based water quality models, which assume a linear relationship between response and prediction variables and their normal distribution, ANNs are able to map the non-linear relationships that are characteristic of aquatic eco-systems (Lek et al., 1996). During the last two decades, ANNs have undergone an explosive development in application in almost all the areas of research (Rahim et al., 1993; Chu and Bose, 1998; Kung and Taur, 1995; Smits et al., 1992; Lerner et al., 1994; Lo et al., 1995; Messikh et al., 2007; Cabreta-Mercader and Staelin, 1995; Hanbay et al., 2008).

Artificial neural networks are flexible mathematical structures that are capable of identifying complex non-linear relationships between input and output data sets (Dibike and Solomatine, 2001). A neural network consists of a large number of simple processing elements called neurons, units, cells, or nodes. Each neuron is connected to other neurons by means of direct communication links, each with an associated weight that represents information being used by the net to solve a problem. The network usually has two or more layers of processing units where each processing unit in each layer is connected to all processing units in the adjacent layers. Neural networks have the ability to learn from experience in order to improve their performance and to adapt themselves to changes in the environment. In addition, they are able to deal with incomplete information or noisy data and can be very effective, especially in situations where it is not possible to define the rules or steps that lead to the solution of a problem.

The ANN approach has several advantages over traditional phenomenological or semi-empirical models, since they require known input data sets without any assumptions (Gardner & Dorling, 1998). The ANN develops a mapping of the input and output variables, which can subsequently be used to predict a desired output as a function of suitable inputs (Schalkoff, 1992).

In a nutshell, ANN can be considered as a black box that is able to predict an output pattern when it recognizes a given input pattern. Once trained, the neural network is able to recognize similarities when presented with a new input pattern, resulting in a predicted output pattern. In principle, ANNs can compute any computable function, i.e. they can do everything a normal digital computer can do, especially anything that can be represented as a mapping between vector spaces can be approximated to arbitrary precision by Neural Networks.

In practice, ANNs are especially useful for mapping problems which are tolerant of some errors and have lots of example data available, but to which hard and fast rules can not easily be applied. ANN models have been widely applied to water quality problems (Rogers & Dowla, 1994; Raman & Chandramouli, 1996; Wen & Lee, 1998; Lek & Guegan, 1999; Bowers & Shedrow, 2000; Kuo et al., 2004, 2007). The literature review conducted showed that the problem with ANN is the learning steps that it would need in order to be fully reliant. ANN will base its judgment on the input and expected output only, and would build the model to fit in between. The problem will then stem from environmental and global changes. Higher temperatures and higher sea levels would certainly have a major effect and influence on the inland habitat, including the freshwater ecosystem and its associated river quality parameters. In other words, in order for ANN to be fully reliant to undertake and anticipate any problems, it would essentially need updating with input and output data to "learn" and to adjust its model, which then re-require the same level of huge data sets, similar to parametric statistical and deterministic models. Another problem with employing ANN is that the binding linkage between Flow and Turbidity in our model would be considered as time consuming and inefficient, as the model has only one input (Flow) and one output (Turbidity). The author does submit that ANN would be able to deal with this correlation, but as the author has no prior knowledge with ANN modelling, a simpler and direct approach was adopted. Due to the nature of data on hand, the author opted to use a standard approach of correlation and regression to investigate the relationship between the Flow and Turbidity.

The aims of using correlation and regression are (Experiment Resources, 2011):

1. To test hypotheses about cause-and-effect relationships. In this case, the values of the Flow are varied to see whether this variation causes variation in Turbidity.
2. To see whether the variables to be correlated are associated, without necessarily inferring a cause-and-effect relationship. In this case, neither variable is predetermined; both are naturally variable.

Chemometrics

With the statistical based models and ANN both requiring data to optimize their models, this caused the interpretation of the data and the results obtained from some of the analytical methods to be questioned. With the increase in the amount of analyte information obtained from a single sample, and difficulties with the interpretation of the test results obtained from a complex sample, matrix has led to the development and application of statistical and mathematical methods from nearly a century ago (Wold, 1995).

These statistical and mathematical methods have resulted in new analytical applications, often by omitting the otherwise imperative sample pretreatment. Svante Wold (1974) was the first investigator to apply 'chemometrics' to organic chemistry applications. Nowadays, many of the statistical and mathematical methods developed for analytical chemistry applications, are referred to as 'chemometrics'. Chemometrics is concerned with the application of mathematical and statistical methods, as well as those methods based on mathematical logic, to extract useful information from chemical measurements (Brown et al., 1994). Chemometric methods may be subdivided into the following categories:

- Statistics (e.g. method validation, sampling strategies, detection limits, etc),
- Optimization (minimization or maximization of a function of one or more independent variables, e.g. mixture designs, liquid/liquid extractions, etc.),
- Signal processing (digital filtering, smoothing, background correction, domain transformations, trend analysis, image analysis, etc),
- Resolution (e.g. identification of peak patterns in unresolved regions),
- Parameter estimation (curve fitting and mathematical modelling of chemical properties of e.g. spectral band shapes),

- Structure-activity relationships (relation of the molecular structure to its chemical, physical or biological properties),
- Pattern recognition (classification of an unknown into one of a set of predetermined classes),
- Artificial intelligence (e.g. automated chemical workstations, including a scheduler for the initiation and monitoring of parallel experiments),
- Calibration (relating or modelling measured responses to the composition of a set of analytes),
- Exploring chemical data (for understanding/finding underlying phenomena), and
- Library searching (identification of unknowns and qualitative analysis of mixtures).

This research only makes use of calibration techniques (partial least squares regression and neural networks) and simple library search algorithms. These methods are described in the subsequent chapters. There are also a large number of articles concerning the development and application of chemometric methods applied to each of the above-mentioned categories, in a series of review articles (Brown et al., 1994; Brown et al., 1996; Levine, 1998; Levine, 2000).

2.2.4 SENSORS

Assessment of the quantity and quality of running water is important in hydro-environmental management. The physical and chemical properties of running water are driven by numerous environmental variables such as climate, storm water runoff, wastewater effluent, and tidal effect at the estuary. The plethora of processes known to interact at different scales that influence solute turnover in the soil, groundwater, and stream water lead to detailed physically-based models. However, the underlying interactions and dependencies of physical processes are only partially understood. Furthermore, the complex non-linear interrelations of variables are normally not well distributed. In addition, the imbalance between the required and available information (input parameters) at the scales of physically based models results in non-uniqueness of model solutions and hampers water quality parameter trend analysis.

Characterizing spatial and temporal variability in the fluxes, and stores of water and water-borne constituents, is important in understanding the mechanisms and flow paths that carry constituents to a stream and through a watershed (Montgomery et al., 2007; Wilkinson et al., 2009). Our ability to predict watershed response, which is becoming increasingly important as we work to manage growing pressures on limited water resources, is dependent upon our knowledge of watershed behaviour and the interacting processes that drive that response. In some watersheds, the time scale of many important hydrologic and water quality processes is of the order of minutes to hours, not weeks to months (Tomlinson & De Carlo, 2003), and understanding the process linkages between catchment hydrology and stream water chemistry, which is necessary for incorporating these processes into predictive models, requiring measurements on a time scale that is consistent with these processes (Kirchner et al., 2004). Indeed, the need for high frequency monitoring is well recognized (Kirchner et al., 2004; Kirchner, 2006; Hart & Martinez, 2006) and has motivated community initiatives towards the establishment of large-scale environmental observatories. The goal of these initiatives is to create a network of instrumented sites where data are collected with unprecedented spatial and temporal resolution, aiming at creating greater understanding of the earth's water and related biogeochemical cycles and enabling improved forecasting and management of water processes (Montgomery et al., 2007). Within observatories, environmental sensor networks have been proposed as part of the cyber infrastructure required to generate data of both high spatial and temporal frequency.

Estimating the flux, or mass flow rate, of a water quality constituent requires estimates of both the constituent concentration and the volumetric flow rate, or discharge of a stream. High frequency monitoring of stream discharge has long been practiced by organizations like the United States Geological Survey (USGS) due to the relatively simple methods and technology used to gage discharge and the established need for water quantity measurements in managing water resources (Nolan et al., 2005). However, high frequency discharge monitoring is done at a relatively small number of sites due to the costs associated with continuous monitoring.

Traditional water quality monitoring, on the other hand, is generally done at a much greater number of sites, but involves the collection and analysis of grab samples that are usually collected with a frequency too low to accurately characterize the temporal variability in concentrations of water quality constituents (Etchells et al., 2005; Scholefield et al., 2005). Even though both discharge and concentrations are required for estimating constituent fluxes, there is a spatial and temporal disconnect between the traditional methods of monitoring these variables.

High frequency in-situ monitoring can capture time periods and characterize trends that may be omitted or overlooked by periodic grab sampling (Kirchner et al., 2004; Tomlinson & De Carlo, 2003; Jordan et al., 2007). For many water quality constituents (e.g., sediment and phosphorus) though, sensor technology does not currently exist for making high frequency measurements of concentrations in-situ. For these constituents, it is impossible, or even impractical, to collect samples with high frequency for extended periods due to cost or logistical constraints, leaving us with a paucity of data available for testing models and fostering process understanding (Gascuel-Odoux et al., 2009; May & Sivakumar, 2009). Due to the limitations in existing sensor technology, many studies have examined the use of variables that can be measured in-situ with high frequency as surrogates for other water quality constituents that cannot economically be measured with high frequency.

This research found many papers that focused on the use of remote sensing as applied to various aspects of freshwater, estuarine and near-shore benthic ecosystems. The need for this applied research has now become more crucial by the day, as routine monitoring and mapping of the changes taking place in these ecosystems enable managers to focus their efforts in time and space, and to prioritize their responses to the most pressing issues. The most compelling aspect of much of the research carried out is that the best advances emerge from the combined use of technologies across the spectrum, and across a range of spatial and temporal scales. No single issue threatening the health of these ecosystems can be addressed using a single approach or with a single image acquired at a given point in time. Despite the additional resources and expenses required to make use of multi-sensor, multi-temporal and multi-scale image data, the utility of such data becomes evident across a diverse range of ecosystems and a wide range of conditions.

Remote sensing

Remote sensing can be defined as the "acquisition and recording of information about an object without being in direct contact with that object" (Gibson, 2000).

Monitoring and assessing the quality of waters in streams, reservoirs, lakes, and estuaries are critical to improve water quality. Current techniques for measuring water quality involve in-situ measurements and/or the collection of water samples for subsequent laboratory analyses. Remote sensing of indicators of water quality offers the potential of relatively inexpensive, frequent, and synoptic measurements using non-invasive sensors.

The overall process of remote sensing can be broken into 4 main components (US Army, 2003):

1. An energy source;
2. The interaction of this energy with the surface of the target sought;
3. Energy recorded by a sensor as data; and finally
4. Data displayed digitally for visual and numerical interpretation.

Remote sensing applications to determine water quality are limited to measuring those substances or conditions that influence and change optical and/or thermal characteristics of the surface water properties (Ritchie & Cooper, 2001). Suspended sediments, chlorophylls, DOM, temperature, and oil are water quality indicators that can change the spectral and thermal properties of surface waters, and are most readily measured by remote sensing techniques. Substances (i.e., nutrients, metals) that do not change the optical and/or thermal characteristics of surface waters can only be inferred by measuring surrogate properties (i.e., chlorophylls), which may have responded to an input of chemicals. These remote sensing techniques should improve our ability to monitor changes in the water topography and contents.

The research carried unearthed many technologies and theories that can be categorized as remote in-situ sensing techniques (e.g. Fluorescence xcitation (Ahmad & Reynolds, 1999; Ahmad, 1991); Laser (Fang & Ahmad, 2007; Kuwako et al., 2003); X-Ray spectrometry

(Moreira et al., 2008; Valentinuzzi et al., 2006) and Doppler-Ultrasound (Wass et al., 1997; Lynnworth & Liu, 2006)).

The sensors' literature review research concluded to use the Doppler-Ultrasound sensor method mainly because of its durability and availability in the entire surface water extraction sites, which would make it more appealing to the majority of the water companies. For example, despite fluorescence technology having a potential of being completely in-situ water quality monitor, it is not readily available to all water treatment sites due to its disadvantages in use. On the other hand, the Doppler-Ultrasound is almost available at all of the UK water treatment sites, as it is used as a flow-meter.

Doppler Flow-meter

The Doppler Effect, named after the 19th century Austrian scientist Christian Doppler, can be used to measure the flow in a pipe. The Doppler Effect is the frequency shift that occurs when a sound source (transmitter) is in relative motion with a receiver of that sound source. In the case of a Doppler flow-meter, we have two sensors mounted or strapped on the outside of a pipe. One of the sensors is the transmitter, and transmits a high frequency (ultrasonic) signal into the pipe. This signal is reflected off particulate matter or entrained gas bubbles in the fluid. The reflected signal is then picked up by the receiving signal and the frequency difference between the transmitted and reflected signals is measured and correlated into an instantaneous flow rate or flow total . Figure 2-5 below shows the principles of how Doppler ultrasound works.

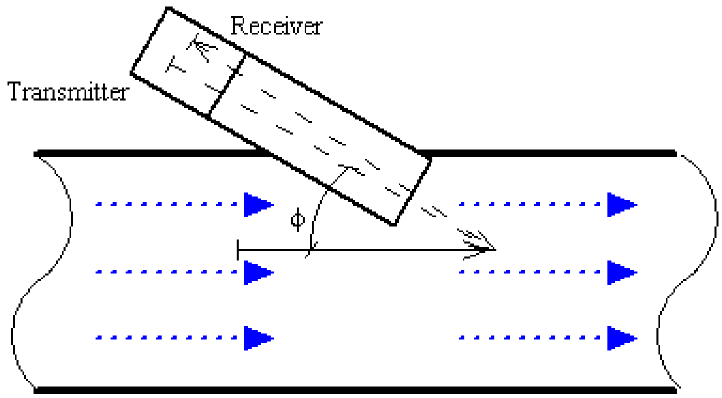


Figure 2-5 Illustrating the Doppler signal penetrating a pipe and then reflecting off the particulates in the stream. The signal phase shift is measured and correlated to a flow velocity (www.EngineeringToolBox.com).

The frequency is subject to two velocity changes; one upstream and the other downstream. Travelling upstream, the velocity of the wave is given as (V_s - V cos Φ) where V_s equals the velocity of sound in the fluid, V equals the average fluid velocity and Φ equals the angle of the ultrasonic beam to the fluid flow. Similarly, the downstream velocity is given as (V_s + V cos Φ). The Doppler relationship between the reflected and transmitted frequencies can now be expressed as:

$$f_r = f_t[(V_s + V \cos \Phi)/(V_s - V \cos \Phi)] \quad (2.1)$$

Here, f_r is the received frequency and f_t is the transmitted frequency. To further simplify this equation, one can assume that the velocity of the fluid in the pipe is much lower than the velocity of sound in the pipe; that is, $V << V_s$.

With this assumption, one can write:

$$f_r = f_t[(V_s + V \cos \Phi)/ V_s + (V \cos \Phi)]/ V_s = f_t[1 + (2V \cos \Phi)/V_s] \quad (2.2)$$

The frequency shift is given by

$\Delta f = f_r - f_t$ so that $\Delta f = [2(f_t) \cos \Phi / V_s]V = kV$ (2.3)

Where $k = 2(f_t) \cos \Phi / V_s$

This indicates that the fluid velocity in the pipe is directly proportional to the change in frequency between the transmitted and reflected ultrasonic signals for a given Φ, V_s and f_t. With knowledge of the pipe size, the electronics of the flow-meter will correlate the fluid velocity into a flow rate in the engineering unit of choice. Software corrections may have to be made for V_s, since the sound velocity through the medium will change with pressure and temperature fluctuations.

There are ultrasonic designs on the market that use a series of pulsed signals, as opposed to a continuous ultrasonic beam. The main advantage of the pulsed technology is that it can measure the vertical velocity profile within the pipe. Fluid flow will be faster along the middle of the pipe than along the pipe walls and the pulse-design allows one to obtain a better image of the flow profile within the pipe.

Another sensor design, which minimizes external noise, uses dual-frequency Doppler technology to send two independent signals into the pipe at different frequencies. Since both signals are subject to the same Doppler shift, yet the noise signals are random, the signals can be combined to calculate a flow velocity while subtracting out the noise.

Ultrasonic sensors can be used with a wide variety of pipe materials, but some will not allow the signal to pass through. Although pipe material recommendations will vary depending on the sensor design, one should not expect to have any problems with carbon steel, stainless steel, PVC, and copper. However, pipes made of concrete, fiberglass, iron, and plastic pipes with liners, could pose transmission problems. One should check with the particular manufacturer to ensure that the pipe material is suitable. Some Doppler designs utilize a section of pipe with built-in transducers that make direct contact with the fluid. This design, although no longer non-invasive, eliminates the problem of incompatible pipe materials.

The accuracy of the ultrasonic Doppler meter is typically around ±2% of full scale. The vast majority of Doppler meters are used for liquids (roughly 88%), while the rest are used for gas (11%) and steam (1%) applications.

Advantages

The main advantage of the Doppler ultrasonic meter is its non-intrusive design. An acoustic-coupling compound is used on the surface of the pipe, and the sensors are simply held in place to take a measurement or, for a more permanent installation, they are strapped around the pipe. Some manufacturers offer a special clamp-on probe that allows connection to smaller pipe sizes (down to 1/4-in. diameter). Other advantages include:

- Easy installation and removal—no process downtime during installation;
- No moving parts to wear out;
- Zero pressure drop in pipes being monitored;
- No process contamination of the measured substance;
- Works well with dirty or corrosive fluids;
- Works with pipe sizes ranging from 1/2" to 200";
- No leakage potential;

- Meters are available that work with laminar, turbulent, or transitional flow characteristics;
- Battery powered units are available for remote or field applications;
- Sensors are available for pulsating flows;
- Advanced software and data logging features available; and
- Insensitive to liquid temperature, viscosity, density or pressure variations.

Disadvantages

The main disadvantage of this technology is the fact that the liquid stream must have particulates, bubbles, or other types of solids in order to reflect the ultrasonic signal. This means that the Doppler meter is not a good choice for distilled water or very clean fluids. Although strides have been made with the Doppler technology so that it can work with smaller particulate sizes and smaller concentrations, one still needs to have some particulates present. A good rule of thumb is to have a bare minimum of 25 ppm at roughly 30 microns in order for the ultrasonic signal to be reflected efficiently. Some flowmeter designs may require a little more than this, so it is advisable to check the specifications of the meter one is considering.

An important point to note is that if the solids' content is too high (around 50% and higher by weight), the ultrasonic signal may attenuate beyond the limits of measurability. This possibility should also be checked with the manufacturer, referring to one's specific application. Another disadvantage is that the accuracy can depend on particle-size distribution and concentration, and also on any relative velocity that may exist between the particulates and the fluid. If there are not enough particulates available, the repeatability will also degrade.

Finally, the only other potential problem of this technology is that it can have trouble operating at very low flow velocities.

2.3 CHAPTER SUMMARY

This chapter provided an insight into the research that has been conducted within the primary areas of research. This serves to provide a knowledge foundation from which to learn, and to build upon, ensuring the research conducted for this thesis adds to, rather than duplicates, existing or other on-going work.

The literature review showed the potential use of river flow and possible usage of it by Water Treatment Works to depend its operational process mainly on the level of flow of rivers' water being extracted from. River flow forecasting was shown to be able to provide fundamental information on a wide range of problems related to the design and operation of river systems. For example, deciphering river flow can yield the exact amount of Turbidity, Colour, Conductivity and Oxygen levels present in the river.

Research into various correlation techniques resulted in the introduction to Chemometrics and its underlying methods of data extraction, analysis and interpretation. The statistical tool of Partial-Least Square was introduced and shown that, based on past history, more accurate forecasting is achievable when you have more than one variable present.

The review into different water sensors technologies led to more understanding of what the project requires to achieve its aim. The fundamentals of ultrasound review have given more robust theories of the Doppler-ultrasound and its potential usefulness to this project. The

Doppler-ultrasound works on water particulates that rebound the sound waves back. These particulates represent mainly the Turbidity level in the water sample and can therefore be quantified depending on the percentage frequency sent back of the particulates.

Other technological reviews have shown other potential solutions to the project, e.g. fluorescence technology can detect Colour levels in water, which, in turn, is proportional to the Turbidity levels, but by comparing the advantages of these technological reviews against each other, rather than a direct comparison between their individual disadvantages and advantages, it was decided upon to use the Doppler-Ultrasound method mainly because its durability and availability in all of the surface water extraction sites, which would make it more appealing to the majority of the water companies. For example, despite fluorescence technology having a potential of being completely in-situ water quality monitor, it is difficult to realise that it is not readily available to all water treatment sites. On the other hand, the Doppler-Ultrasound is almost available at all of the UK water treatment sites, as it is used as a flow-meter.

3 RESEARCH METHODOLOGY

3.1 INTRODUCTION

Research methodology identifies what the activity of research is, how to ensue and measure progress, and what constitutes success. This chapter briefly reviews research methodologies in general, and advises on the most suitable methods to adopt within this research project. The adopted methods are discussed, including the research issues applicable to this project. The use of each method is justified and the research philosophy is proposed.

3.2 RESEARCH METHODOGLOGY: *BACKGROUND*

It is possible to categorise research methodologies into quantitative, qualitative and triangulation research methods (Fellows & Liu, 1999). The choice of methodology is dependent upon the nature of the research project, the aim of the project, and the type of information available (Fellows & Liu, 1999). A brief description of the three methodologies is given with a view to justifying the methodology adopted within this research project.

i) *Quantitative research* is deductive, and it tests existing theories with the aim of proving or disproving them. It can test theories composed of variables, numerical measurements and can be analysed statistically in order to validate the theories. The methods deal well with large quantities of data and are fast at covering a wide scope of variables.

ii) *Qualitative research* is used to find facts about a concept, a question, or an attribute where factual evidence is collected and compared against a theory. Quantitative research uses methods such as structured surveys, experimentation, research of secondary data and numerical methods (Brannen, 1992). Qualitative research is subjective, and the methods aim to generate new theories and ideas. Information gathered can be classed as exploratory or attitudinal (Naoum, 2001).

iii) *Triangulation* can be defined as the application and combination of several research methodologies in the study of the same phenomenon, used to obtain confirmation of findings through convergence of different perspectives (Jakob, 2001), taken from both quantitative and qualitative methodologies. There are conflicting reports on the different number/types of triangulation, but the general consensus (including Mathison (1988), Begley (1996) & Guion (2002)) appears to indicate that five variants exist. These are:

a. Data – triangulation of different sources of data across time, space or persons.
b. Investigator – triangulation of work amongst several researchers.
c. Methodological – triangulation of multiple methods to study a single problem.
d. Theory – triangulation of two or more contrasting theoretical positions.
e. Analysis – triangulation via use of more than one analysis technique.

In addition to the above research methodologies, there are a number of other principal research styles, including ‘Case Studies’, ‘Experimental’ and ‘Correlational’. The most suitable forms for this type of research project were examined in more detail to enable further understanding of the techniques available, and to facilitate taking a decision on the most appropriate research strategy to facilitate completion of the overall research project.

Case Studies is an empirical inquiry that investigates an existing event within its real-life context; when the boundaries between phenomenon and context are not clearly evident; and in which multiple sources of evidence are used (Yin, 1984). According to Soy (1996), case study research excels at allowing an understanding of complex issues and can extend experience or add strength to what is already known through previous research. However, this approach can be time consuming and care must be taken, so not to draw generalised conclusions from limited cases to ensure academic rigour. Conversely, it can lead to new and creative insights, and have high approval amongst practitioners. Through triangulation with multiple means of data collection, the validity can be increased further (Voss et al, 2002).

Experimental research is an attempt by the researcher to maintain control over all factors that may affect the results of an experiment. In doing this, the researcher attempts to determine or predict what may occur (Key, 1997). This type of research is best suited to 'bounded' problems or issues, in which variables involved are known, or at least hypothesised, with some confidence (Fellows & Liu 2005).

Correlational research is a quantitative method of research in which a researcher has two or more quantitative variables from the same group of data, and is trying to determine if there is a relationship (or covariation) between the two or more variables. Theoretically, any two or more quantitative variables can be correlated as long as there are scores on these variables; however, it is probably a waste of time to collect and analyse data when there is little reason to think that these variables would be related to each other. In correlational research, it cannot be said that a variable causes something to happen, unless it is done through experimental research. Correlational Research Process involves:

- Variables to be studied are identified; identifying variables is a tedious process in its own right. This must be accomplished with great attention also paid to the impact weight of the variable on the overall process;

- Questions and/or hypotheses are stated; what drives the research process in the first place are theories that are born out of necessity for an engineering question;
- A sample is selected; sampling size would dictate how realistic the correlation is;
- Data is collected depending on sampling size; data should be collected to mirror this size;
- Correlations are calculated based on theories that were developed from either mathematical models or engineering concepts;
- Results are reported.

No precise method to broach engineering research problems exists, and the search for a method appropriate to this field is becoming a research field in its own right (Dobson, 2001; Glass et al., 2002; Gregg et al, 2001; Miles & Huberman, 1984; Myers, 1997). The solution of problems purely concerning engineering requires methods of a different kind, since in these cases it is directly possible to apply neither empirical methods nor methods which have to do with social and cultural component as the object of study does not yet exist (Wohlin, 2000). It would be necessary to study existing methodologies, reflecting on them to determine their advantages and disadvantages and proposing a new one, which, while retaining the advantages of the methodologies studied, would, as far as possible, lack their shortcomings (Lázaro & Marcos, 2005). Arriving at a better final proposition would largely depend on the creativity and common sense applied to the construction of the new method. This method is proposed by Klein & Hirschheim (2003) who stated that this method, when applied, consists of the formulation of experiences and the identification of the best practices.

3.3 ADOPTED RESEARCH METHODOLOGY

The adopted research methodology was broken into four distinct key processes. The processes included case-study, data analysis, numerical modelling and lab experimentation. These work tasks were in-line with the literature review. The literature review was carried out to facilitate comprehension of existing theories and work by others; to form a coherent argument for further research and to demonstrate a fundamental understanding of the concept. It also gave an insight into the advantages and disadvantages of different techniques that can be used to achieve the aim of the research. The selected methods were discussed in relation to research issues, and their selection is justified.

By means of a literature review, the weak spots of a model can be established and its respective technique of creation, followed by establishing a description of a new technique and new built model (Fig 3-1).

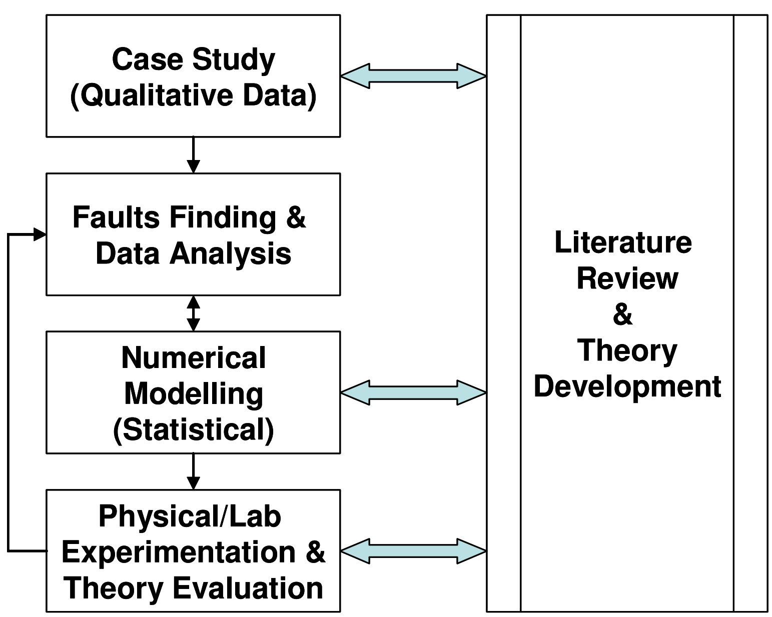


Figure 3-1 Diagram showing the working relation of key Research Methodology components for the undertaken research.

3.3.1 CASE STUDY

The Case Study process was aimed at investigating and customising the detailed trend of archived critical parameters on boreholes and river abstraction sites. Various sites were examined with varying output quantity and different operational procedure.

The origins of the case study were instigated by Severn Trent (ST) Ground Troublesome Asset Project (GTAP). The project was initiated by ST initially aimed at resolving problems that were hampering the continuous operation at Chequer House WTW. Chequer House processes are mainly computer controlled (un-manned site) and is a borehole WTW. The project objectives were to:

- Increase reliability and efficiency
- Reduce costs
- Introduce flexibility for system growth
- Improve operator interface and communication
- Increase data accuracy
- Introduce effective data management
- Fully integrated information systems
- Ultimately monitor and/or control operations at 400+ sites remotely
- Improve redundant servers, workstations and laptops for increased reliability
- Improve relational database server to provide the SCADA data to the Water Department Enterprise System for improved system maintenance and improved water modelling.

Since the objective of the project was to improve the operational running of the SCADA system of the WTW through the use of advanced sensor network and automated processes, it would be fitting to undertake a second case study of a manned WTW. This was to ensure that findings and obstacles faced with the first case study were very much real and un-noticed. Hence, a second case study looking at Little Eaton WTW (manned site) processes was initiated with the same project objectives as Chequer House. Little Eaton is a surface water (Derwent river) subtraction site.

3.3.2 FAULTS FINDING AND DATA ANALYSIS

The faults finding and data analysis process was essential to complete the case study research and to analyse the systems' performance of the case studies in relation to the sites' location and hold up time between the sensors readings and their coherent relation. Statistical modelling was introduced to research the application's intelligence effectiveness and try and build a workable formula between the sensors readings and other third party readings (i.e. flow measurements from the EA). The data obtained were analysed statistically in order to establish meaningful relationships and findings from the research. Raw data are processed for preliminary analysis so that comparisons between different tests in the laboratory and in the field can be compared and contrasted. Data analysis is a common method used within both quantitative and qualitative methodologies to interpret the information collected and manipulated to fit the research projects scope. This research had vast amounts of physical sensors' readings obtained from ST which required careful detailed analysis.

3.3.3 NUMERICAL MODELLING

Numerical modelling deals with quantitative data and aim to test a course of action, classify a design, or identify a best option from amongst a range of scenarios by numerical methods rather than physical methods. Modelling was used to compare and confirm the analytical data produced against actual physical test data and vice versa. Models can be derived from first principles (analytical modelling) or by using relationships derived from physical experimentation (empirical modelling). Some models are a combination of both analytical and empirical methods: these are termed mechanistic-empirical models (Ullidtz, 2002). An advantage of numerical modelling over physical experimentation is that it can permit the user to perform many more scenarios, potentially in less time, consuming fewer resources, whilst in a safer and more workable environment (Holt, 1998). This method can also be used to complement the physical test methods and support their findings.

The statistical modelling of the results was based on the use of Partial-Least Square (PLS) method. PLS is a statistical tool, part of Multivariate Data Analysis family of tools. That refers to any statistical technique used to analyse data that arises from more than one variable. This essentially models reality where each situation, product, or decision involves more than a single variable. The variables in this context are the elements being monitored by the SCADA operating system. Full description of the PLS technique will be presented in the next chapter.

3.3.4 EXPERIMENTAL EVALUATION

The final research process would enable to physically experiment the theories developed from the past processes and put them into practice by compiling a number of laboratory tests that can also be represented as an evaluation process of the theories devised. Physical experimentation represents models, or investigates physical phenomena. Experiments may be performed in the field (e.g. in-situ performance test) or in the laboratory (e.g. element test). Often research projects will utilise both field and laboratory techniques to provide a fuller picture of physical phenomenon. Whilst in-situ testing provides a realistic understanding of conditions in the field, it also provides much less control of key variables influencing the materials' behaviour. Laboratory testing provides a more controlled environment in which the experiment can isolate the significance of certain required variables, from minimising the effect of others (EAL, 1997).

This literature review highlighted the types of technologies used in the water industry and their limitations of use. After careful consideration, and based on the operating behaviours of all the devises researched, it would be logical to base the laboratory experiments using Doppler-Ultrasound Flow (DUF) device because of its availability and common use with the industry and, furthermore, due to the operational functionality that it works on the back of rebounded sound waves of particulates floating within the raw water.

3.4 CHAPTER SUMMARY

This chapter has outlined, and briefly reviewed, some of the main types of methodologies available, within the context of the subject area. It also provided justification for the application of the research methodology chosen to address the given problem. The research methodology was based to achieve the GTAP project aims and establish the work needed to achieve it. As the GTAP project aim was the instigation factor for this project, the literature review was the directional instigator, which based its conclusions on the collective work made by fellow researchers in the same working field.

Case studies have been looked into to provide history of the sites investigated and helped formulate theories of the best optimal operational processes that are needed to run a WTW in the most optimum way. Two case studies were investigated, Little Eaton and Chequer House. Little Eaton is considered to be a fully manned site where it operates on surface water (i.e. rivers); while Chequer House is considered to be un-manned as it operates on groundwater; boreholes.

Data Analysis examined the data readings collected from the case studies to establish a working pattern to the data readings and help identify the controlling variables of the sensors. Data analysis also helped formulate the theory on which the numerical modelling can be based upon.

The numerical modelling process initially looked at Principal Component Analysis (PCA) as the statistical tool to derive the required objectives, but due to its shortcomings, which are outside the scope of this research, PLS was employed to achieve the required targets. Both techniques are considered to be part of the Multivariate Data Analysis techniques.

The final process was based on laboratory experiments of the derived formulas from the previous processes. In other words, the experimental evaluation looked to investigate the DUF operational work and wither a different operational dimension can be added to it through signal analysis of the rebounded sound waves of the particulates floating in the water. The actual steps taken during the research, as part of the chosen methodology, is described in details in chapter 4.

4 THE RESEARCH UNDERTAKEN

4.1 INTRODUCTION

In order for the research to commence, a background field research was needed. This included being introduced to the general requirements needed by the water industry in terms of water quality produced, extraction rates, processes and procedures, etc. Another important step was the field reconnaissance of the actual and real life problems faced by the water companies. This included visits to different types of Water Treatment Works (WTW) with different extraction types (river/boreholes) and different operational procedures (manned/unmanned). These steps were necessary to draw out realistically what parameters need improving for control, and what constitutes an acceptable research objective.

4.2 PHASE 1: CASE STUDIES

The case studies investigated at the beginning of the research provided an adequate knowledge and overview into the performance needs of the UK water industry. Two water treatment sites were investigated; Little Eaton and Chequer House. The two sites were primarily chosen based on their operating behaviour. Chequer House was looked into first because it represented the main problem Severn Trent were facing with their un-manned borehole sites. The problems were primarily causing the site to shut down its processes due to the continuous number of alarms being raised by the SCADA sensors. In most cases, once an engineer is sent on site, s/he ends resetting the site processes. On the other hand, Little Eaton represented the fully manned river abstraction site on the Derwent River in the East Midlands.

There were originally a number of other borehole sites' data that were handed from Severn Trent (Blackstone, Bromsberrow, Diddlebury, Hob Hill & Meriden Shafts), but after thorough analyses of the data given, which took considerable time due to the size of the data given (about two full DVD's worth of text files) and the format it was provided, it was narrowed down to Chequer House. The main reason behind the vetting process and why it took so long was because a large part of the data given was irrelevant to the research and only useful for maintenance purposes. For example, some of the data listed the number of sites' failures and causes. However, the causes were almost always found to be not to the degree of seriousness to cause site shutdown such as intruder alarm shutdown, pumps failure, or fan shutdown. All these were accumulated into the data given by Severn Trent for a different number of other boreholes. Despite this data, there was no significant or relevant data that could be used to determine the performance of site.

The final deduction made out from the data analysis was that the data gathered from the two sites were insufficient in running the sites, let alone running any other water treatment works. A summary of the main findings is provided in the next chapter.

Another case study was looked at in the final year of research; Strensham WTW. Strensham is a river abstraction site, similar to Little Eaton, except it lies in the West Midlands on the River Severn between the English and Welsh borders. Strensham was added to the case studies' list to evaluate and cross-check the earlier findings that were drawn out from the Little Eaton case study data analysis.

Analysis of the history of flow readings from different WTW's in the Midlands region revealed that these flow readings are correlated and similar in pattern. In other words, flow readings recorded can be inter-correlated to each other and derived from each other.

4.3 PHASE 2: DATA COLLECTION AND ANALYSIS

With the case studies' data being available from the start of the phases, data identification was not difficult. The difficulty was in the collection of the data from both Severn Trent and the Environmental Agency. The data analysis process took about one and a half years of research and was based around the use of the correlations factor between the different measured parameters. The idea was to try and link the parameter readings measured at the sites together and to establish a common base for their fluctuations.

Data given for Chequer House did not result, in general, to any sort of indication for the readings being recorded by the sensors. Reasons for that were:

a. Insufficient data available for analysis; data recorded were inline with maintenance records rather than quality check operations.
b. Data is not continuous; i.e. data is missing a significant amount of readings. No apparent reason has been identified as the cause of this, but two significant causes were proposed:
    1. Some sort of communication failure with the sensors, which would explain why the system did not have a backup reading storage if the system itself was down. Since only the data was available, it was almost impossible to pinpoint the exact cause of this failure; and

2. Human intervention; this highlights the fact that unmanned sites are not fully independent. Human intervention could have been in the form of maintenance being carried on the sensor or system reset was initiated.

Data for Little Eaton were both based on lab measurements and sensor readings. The data given was comprised of Turbidity, Colour, Conductivity, Temperature, pH and Ammonia. Despite the fact that the data readings initially given were made every 15mins and for a period of two years, the data was still incomplete and missing a significant amount of entries. This was attributed to human operator error as most of the data were inputted by the operator of the site.

The correlation coefficient can be used as a measure of the extent to which two measurement variables “vary together.” Unlike the covariance, the correlation coefficient is scaled so that its value is independent of the units in which the two measurement variables are expressed. The value of any correlation coefficient must be between -1 and +1 inclusive.

Leading on from literature review findings and data analysis processes, it was found that two or more sensors’ readings can be deduced from other sensors readings on site or in the network or sensors associated with the site. A summary of the main findings is available in the next chapter.

4.4 PHASE 3: STATISTICAL MODELLING

As part of the aim and objectives and with the correlations between the measured parameters established, together with the backing from the literature review, it was necessary to try and provide these correlations in a more manageable way by establishing equations linking these parameters together.

Several multivariate and statistical tools were researched and examined. Tools, for example, like the Principal Component Analysis (PCA) were used to try and derive a smaller set of "synthetic" variables that could explain the original set of parameters. Months of the project time were spent trying to statistically decipher which multivariate tool would actually achieve the aim and objectives set. Research was done through attendance of lectures and workshops, mathematics' support work and extensive literature review through which Partial-Least Squares (PLS) method was identified as the possible solution.

4.4.1 PARTIAL-LEAST SQUARES METHOD

Partial Least Squares (PLS) is a method for constructing predictive models when the factors are many and highly collinear. Note that the emphasis is on predicting the responses and not necessarily on trying to understand the underlying relationship between the variables. For example, PLS is not usually appropriate for screening out factors that have a negligible effect on the response. However, when prediction is the goal, there is no practical need to limit the number of measured factors.

PLS was developed in the 1960's by Herman Wold as an econometric technique, but some of its most avid proponents are chemical engineers and chemometricians. PLS has been applied to monitoring and controlling industrial processes; a large process can easily have hundreds of controllable variables and dozens of outputs (Dijkstra 1983; Geladi & Kowalski, 1986; Stone & Brooks, 1990).

In principle, PLS is part of Multivariate Linear Regression (MLR) family of tools which can be used with very many factors. However, if the number of factors gets too large (for example, greater than the number of observations), you are likely to get a model that fits the sampled data perfectly, but that will fail to predict new data well. This phenomenon is called over-fitting. In such cases, although there are many manifest factors, there may be only a few underlying or latent factors that account for most of the variation in the response. The general idea of PLS is to try to extract these latent factors, accounting for as much of the manifest factor variation as possible while modelling the responses well. For this reason, the acronym PLS has also been taken to mean "projection to latent structure". It should be noted, however, that the term "latent" does not have the same technical meaning in the context of PLS as it does for other multivariate techniques. In particular, PLS does not yield consistent estimates of what are called "latent variables" in formal structural equation modelling (Dykstra 1983; Dijkstra, 1985).

The overall goal of PLS is for the factors to predict the responses. This is achieved indirectly by extracting latent variables T and U from sampled factors and responses, respectively. The extracted factors T (also referred to as X-scores) are used to predict U (Y-scores), and then the predicted Y-scores are used to construct predictions for the responses.

$$[Y] = b_0 + b_1[X_1] + b_2[X_2] + ... + b_p[X_p] + \zeta_i \tag{4.1}$$

In this equation, b_0 is the regression coefficient for the intercept and the b_i values are the regression coefficients (for variables 1 through p) computed from the data. To put it in perspective, the dependent variable would be turbidity level, and the predictor variable is the daily mean flow rate. ζ_i is the statistical error by which the observed value differs from its expected value.

PLS regression extends multiple linear regression without imposing the restrictions employed by discriminate analysis, principal components regression, and canonical correlation. In PLS regression, prediction functions are represented by factors extracted from the Y'XX'Y matrix. The number of such prediction functions that can be extracted typically will exceed the maximum of the number of Y and X variables. The procedure actually covers various techniques, depending on which source of variation is considered most crucial.

- Principal Components Regression (PCR): The X-scores are chosen to explain as much of the factor variation as possible. This approach yields informative directions in the factor space, but they may not be associated with the shape of the predicted surface.
- Maximum Redundancy Analysis (MRA) (van den Wollenberg 1977): The Y-scores are chosen to explain as much of the predicted Y variation as possible. This approach seeks directions in the factor space that are associated with the most variation in the responses, but the predictions may not be very accurate.
- Partial Least Squares: The X and Y-scores are chosen so that the relationship between successive pairs of scores is as strong as possible. In principle, this is like a robust form of redundancy analysis, seeking directions in the factor space that are associated with high variation in the responses, but biasing them toward directions that are accurately predicted.

An important feature of the method is that, usually, a great deal fewer factors are required. The precise number of extracted factors is usually chosen by some heuristic technique based on the amount of residual variation. Another approach is to construct the PLS model for a given number of factors on one set of data and then to test it on another, choosing the number of extracted factors for which the total prediction error is minimized. Alternatively, van der Voet (1994) suggests choosing the least number of extracted factors whose residuals are not significantly greater than those of the model with minimum error. If no convenient test set is available, then each observation can be used in turn as a test set; this is known as cross-validation.

Practical applications usually involve so many variables that it is not practical to seek a “hard” model explicitly relating to all of them. PLS is one solution for such problems, but there are others, including:

- Ridge Regression, a technique that originated within the field of statistics (Hoerl and Kennard 1970) as a method for handling collinearity in regression;
- Neural networks, which originated with attempts in computer science and biology to simulate the way animal brains recognize patterns (Haykin 1994, Sarle 1994).

Ridge regression and neural nets are probably the strongest competitors for PLS in terms of flexibility and robustness of the predictive models, but neither of them explicitly incorporates dimension reduction – that is, linearly extracting a relatively few latent factors that are most useful in modelling the response. For more discussion of the pros and cons of soft modelling alternatives, see Frank and Friedman (1993).

In any case, PLS has become an established tool in chemometric modelling, primarily because it is often possible to interpret the extracted factors in terms of the underlying physical system

– that is, to derive “hard” modelling information from the soft model. More work is needed on applying statistical methods to the selection of the model. The idea of van der Voet (1994) for randomization-based model comparison is a promising advance in this direction.

In short, PLS is the least restrictive of the various multivariate extensions of the multiple linear regression models. This flexibility allows it to be used in situations where the use of traditional multivariate methods is severely limited, such as when there are fewer observations than predictor variables. Furthermore, PLS regression can be used as an exploratory analysis tool to select suitable predictor variables and to identify outliers before classical linear regression.

4.5 PHASE 4: EXPERIMENTAL EVALUATION

The literature review has highlighted a number of methods that can be implemented as a solution to the targeted aim, such as the combine use of laser and infrared (IR). For example, theoretically, laser can be used to heat up and vibrate a water sample on molecular level, while the IR detector can read off the heat signature frequency difference between the agitated sample and the non-agitated one. This would not be practical as power usage; licensing and special containments units would need to be installed at every WTW. The one feasible solution that was flagged as an optimum solution was the use of Doppler-Ultrasound in Turbidity detection as well as flow measurement. In order to verify this method, it was necessary to try and evaluate the finding practically. Hence a lab setup was administered were a Doppler-Ultrasound was used to try and measure the predefined Turbidity level in running water, together with the flow of the water.

4.5.1 LAB SETUP

The experiment design was based on the use of a Vectrino Doppler probe placed on the top of a water flume like rig. The rig was designed to emulate natural running water in flume. The water is drawn from a tanker and sent back to the top of the flume. A controlled amount of clay (Bentonite) was put in the water to try and simulate natural Turbidity in rivers. The added amounts were controlled by adding the right dosage equivalent to 5 mg/l at each stage. The turbidity was increased from approximately 0 NTU (clear tap water) to approximately 30 NTU, which is double the average river turbidity reading measured.

The Ultrasonic Doppler, used during the experimental phase, was not a non-invasive one. To the authors' knowledge, the only non-invasive ultrasonic Doppler sensor available off the shelf in the market is a clamp-on sensor. Unfortunately due to their unavailability to the author, a fully invasive sensor was used instead to prove the theory.

The experimental procedures set were as follows;

1. The water flume is to be switched on for at least 3 minutes prior to start of readings; this ensured that any turbidity particles present will be evenly distributed;
2. Vectirno probe placed at the top of the flume and secured tightly as not to vibrate; vibration would cause the ultrasonic readings to fluctuate;
3. Vectrino probe reading time to be set for equal time interval (5mins) to avoid over heating the probe in accordance with manufacturer guidance;
4. Readings saved as text file on control station (laptop) and read into Excel workbook;

5. Use Matlab to run PLS code and derive equations variables; there are 17 variables returned by the code to derive the Y variable, Turbidity;
6. Increase Bentonite levels in the water flume tank and repeat experimental procedure.

The lab findings were very much in line with the literature review. Please see chapter 5 for a more detailed description of the findings.

4.6 CHAPTER SUMMARY

This chapter has given a breakdown description of the research undertaken towards achieving the EngD project aim and objectives. It describes the aim of each of the project phases and the work done during each phase. It also highlighted the hurdles that this project faced during each of the phases. The breakdown of the project structure was necessary to manage the project from the start and to have a set of clear milestone achievements that would indicate the progress of the overall research.

The research undertaken have gone through five main phases of work; case studies, data collection, literature review, statistical analysis and, finally, experimental evaluation. These phases were designed to help meet the aim and objectives of the project within the time duration of the research. The case studies identified were Chequer House and Little Eaton, which the main difference between them is that Cheque House is an un-manned WTW, while Little Eaton is a fully manned WTW. The data collection phase was mainly dependent on the information gathered from Severn Trent (ST) and the Environment Agency (EA), while the literature review helped look into the possible alternative solution to the Groundwater

Troublesome Asset Project GTAP project. Statistical modelling phase identified methods of analysis of the data gathered to try and interpret what the data meant. Experimental phase was setup to determine, in general, the validity of the research and solutions derived.

5 RESEARCH CONCLUSIONS AND IMPLICATIONS

5.1 INTRODUCTION

The purpose of this chapter is to present the derived practical conclusions of this EngD research project and to set out the implications of the project findings to the water industry. The research conclusions are highlighted in Section 5.2, followed by the implications in Section 5.3. A critical evaluation of the EngD project is provided in Section 5.4, and recommendations for further research are in Section 5.5. This chapter is a conclusion and a summary of the achievements of the project in Section 5.6.

5.2 RESEARCH CONCLUSIONS

In regards to the original project aim that was initiated by Severn Trent Groundwater Troublesome Asset Project (GTAP), together with preliminary observations of the water industry (Ahmed et al, 2007a), it revealed different key needs with regards to the management of WTW. These are as follows (Ahmed et al, 2007b):

1. Improvements on the operational automated methodologies used;
2. Enhancement to the existing SCADA operations;
3. The full use of potential sensory readings currently available;
4. Improved and updated modelling techniques; and
5. Improved data acquisition and storage

Every finding of this EngD project can be associated with addressing one or more of these needs. The findings also lie within the framework of the aim and objectives set at the start of the project. It remains to be proved that the solution can be rolled out to generic water utilities companies, but suggestions are offered for further research to extend and develop the framework and applications further to the wider engineering industry. The research findings can be summarised as;

1. A statistical correlation between flow readings and turbidity measurements were established;
2. A statistical correlation was found between Turbidity, Colour and Conductivity measurements in rivers. This established the theory that different rivers have different signature correlation;
3. Correlations of Flow readings for different rivers in the Midlands region were recognized;

Doppler-Ultrasonic flow sensors were found to be capable of determining the turbidity levels. Hence, Conductivity and Colour can also be deduced from Turbidity readings.

5.2.1 BACKBONE MONITORING NETWORK

With the correlation between the Flow rate and Turbidity levels established (Ahmed et al, 2007), and based on previous historical readings taken, it is possible to drive the WTW more efficiently without a drop in the performance targets (Ahmed et al, 2008). A mean flow rate driven site would yield a more robust operational work by the WTW process, as it would anticipate in advance the levels of Turbidity that the WTW would have to remove from the raw extracted water (Ahmed et al, 2008).

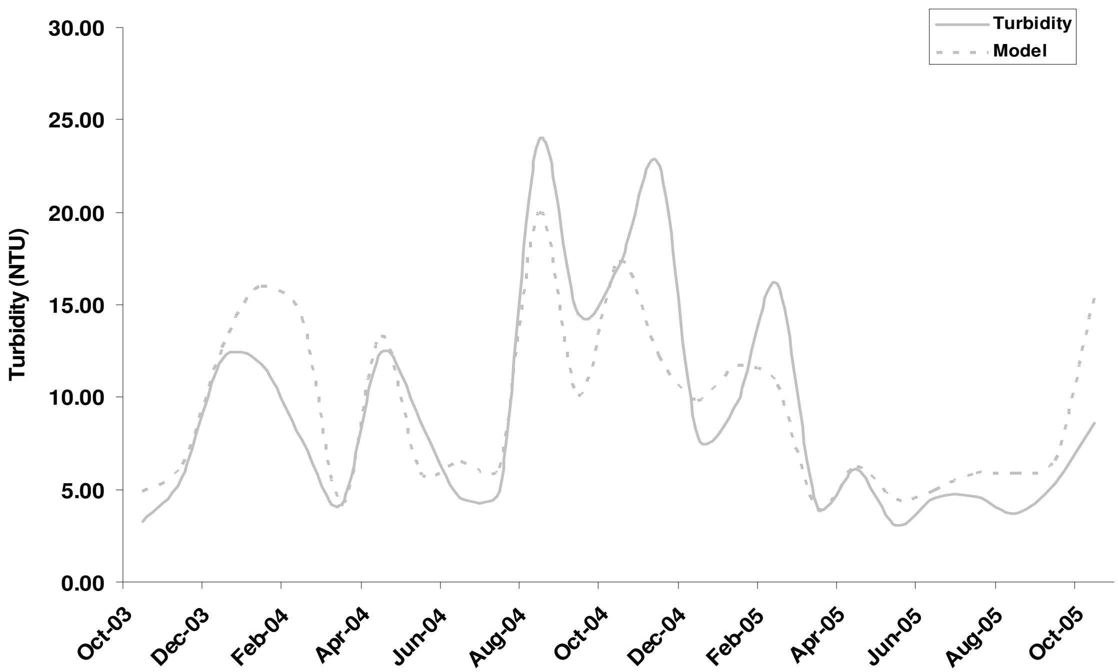


Figure 5-1 Measured Turbidity vs predicted Turbidity from flow rates at St Mary's Bridge.

Graph 5.1 illustrates the findings in deducing the trend shape of the Turbidity level from the flow rate at St Mary's Bridge. The turbidity estimation level was calculated using the PLS technique (described in chapter 4) and, also, through the use of the Flow rate of the River Derwent at St Mary's Bridge.

Where turbidity model is derived from the correlation and PLS regression analysis based on the exponential function of the derived relationship between the Flow and Turbidity:

$$\text{Model} = [0.4264 * R + 0.7 * F - 0.0591]^{e} \tag{5.1}$$

Where R = Ln (Rain)

F = Ln (Flow).

5.2.2 STATISTICALLY CORRELATED DRIVEN SYSTEM

A statistically correlated driven system is based on the idea that rivers signatures would hold up for at least one to two years ahead, based on previous data taken from the past twenty years (Ahmed et al., 2007). The correlation needs to be recalibrated to drive a more accurate representation of the system correlation based on latest performance. This method is very much a necessity, as changes in the climate weather can have an impact on the rivers' signature due to increased water temperature and adverse changes in weather patterns.

The correlation derived for Turbidity, Colour and Conductivity can help increase measurements' readings and also reduce human error factors. Acceptable marginal errors can alert the operator if sampling technique is faulty and/or expected levels are high/low such that processes need to be adjusted in line with this increase/decrease.

Table 5-1 Covariance coefficients between the parameters monitored

	Turbidity	*Colour*	*pH*	*Conductivity*	*Ammonia*
Turbidity					
Colour	**81.223**				
pH	-0.532	-0.483			
Conductivity	**-362.84**	**-388.41**	4.068		
Ammonia	0.334	0.105	-0.002	-0.004	
Temperature	-5.748	-10.514	0.007	17.486	-0.002

Table 5-2 Correlation coefficient between the parameters monitored

	Turbidity	Colour	pH	Conductivity	Ammonia
Turbidity	1.000				
Colour	0.570	1.000			
pH	-0.166	-0.416	1.000		
Conductivity	-0.622	-0.534	-0.012	1.000	
Ammonia	0.205	0.023	-0.426	0.067	1.000
Temperature	-0.145	-0.172	0.009	-0.028	-0.010

The correlation coefficients (table 5-2) calculated for the measured parameters at the river WTW have very small significance. This is an indication that all the parameters measured are not a realistic representative of the environment and cannot be interpolated from other types of sensors. The only explanation that can be noted from the correlation is that the relationship between Turbidity and Ammonia can be attributed to the rainfall effect on farmlands adjacent to the river. The rain is expected to wash some of the fertilizers into the river stream and causing the turbidity to rise at the same time.

On the other hand, table 5-1 shows that the most obvious positive link exists between turbidity and colour. This is acceptable as turbidity contributes largely to the colouring in the water as a result of sediments floating in the water. The higher the covariance coefficients between the parameters monitored, the stronger the correlation between these parameters. It is also clear that the temperature has some effect on the conductivity readings. This sounds logical, as it would be anticipated that temperature would affect solubility of contaminants, and therefore conductivity.

Having estimated the Turbidity levels from the Flow measurements, Colour and Conductivity levels can also be estimated using the same PLS technique. This method would prove invaluable as it would indicate to the operator the best performing processes levels that would eventually lead to cut operational holding time and cost. Graph 5.2 depicts the difference between the measured conductivity and the predicted (calculated) one based on calculated turbidity levels.

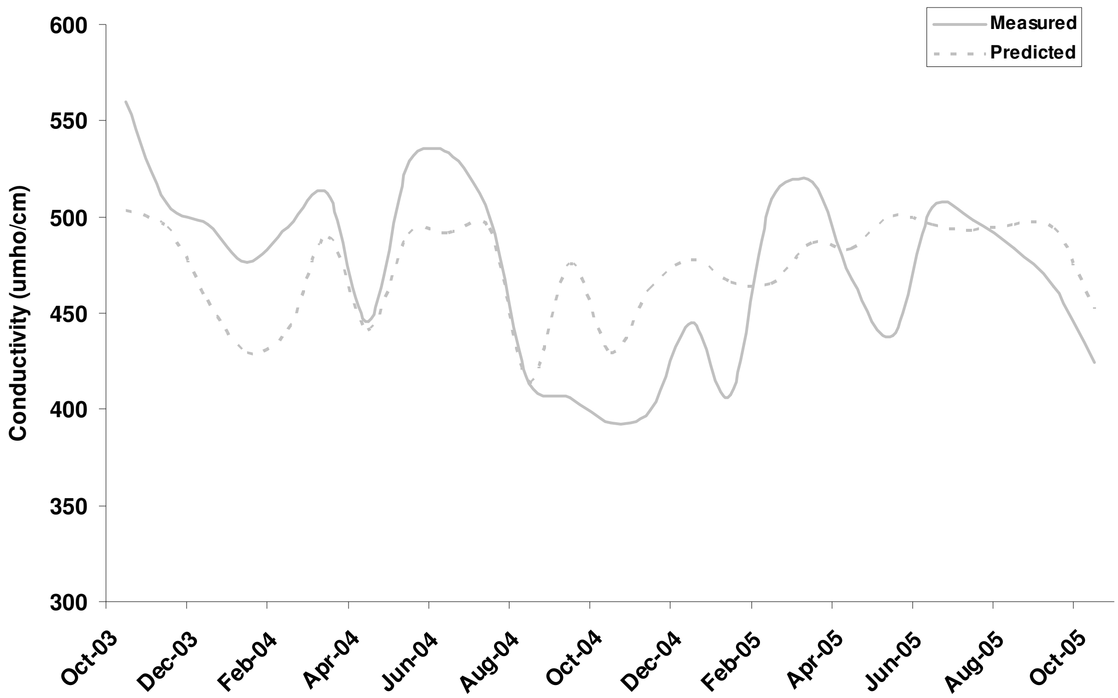


Figure 5-2 Measured Conductivity vs derived Conductivity from flow rates at St Mary's Bridge.

Having established the turbidity and conductivity estimation levels, the process can be repeated again to derive the expected colour level.

Where the Conductivity model was also calculated from the Flow rate at St. Mary's Bridge and the Rain level;

Conductivity Model = 531.474 – (12.495*R) – (2.4415*F) (5.2)

One possible explanation for why the predicted levels differ from the measured ones in graphs 5.1 and 5.2 is because the measured values were laboratory measurements, rather than online readings. This has numerous errors associated with it. For example, the point of where the sample was taken might not be appropriate, but rather where its more convenient and easy for the operator to take his/her sample; the time taken from collecting the sample to the time of laboratory inspection will cause most of the particulates floating to settle in the sample bottle; and human errors, which might include the wrong calibration of the machine used and sometimes by forgetting to record the readings taken as was found with the data in hand.

It would be, therefore, assumed that the calculated data for Turbidity, Conductivity and Colour are a more realistic representation of the real values in the River Derwent. However, it should be stressed that these relationships may not be appropriate for other WTW. These relationships were derived using the data supplied from one site, which means that it is more related to this site rather than being generalised.

Again, by having a third party data, e.g. Flow rate and/or Rain levels, would substantiate the estimated levels of the constituents being measured, i.e. Turbidity/Colour/Conductivity. The process leads on from the estimation of Turbidity levels from Flow rate by adding Rain levels to the correlation equation. It was noticed that the Rain level has insignificant, if any, effect of the estimated Turbidity levels.

This can be attributed to a number of reasons:

1. The amount of rain effect, in comparison to the river flow rate effect, is very minute;
2. The rain gauge is not optimally positioned to highlight its effect;
3. Not enough rain data to register its full potential; and
4. The intake pipe of the WTW is deep enough in the river that floating debris from the river banks is not sucked in.

Statistical Significance

Based on the sample of two years readings of Flow, Turbidity and Conductivity, the probability (p-value) was obtained to be of the order of 0.005 (0.5%). In other words, there is a chance of 0.5% that there exists no relationship between Flow & Turbidity and also between Flow & Conductivity (null hypothesis). Hence, the null hypothesis is clearly and evidentially rejected. The original research hypothesis states that there does exist a relationship between Flow and many of the water constituents found in rivers (e.g Turbidity).

Tests for statistical significance indicates what the probability is that a relationship found is due only to random chance. It shows what the probability in making an error if an assumption of a relationship exists.

5.2.3 REGIONAL CONFIGURATION

From the data analysis of the material available, it emerged that possible sources of the Derwent river flow rate would be the flow rate at St Mary's Bridge in Derby, which is considered to be the closest monitoring station to the manned site being investigated, and flow rate at Buttercrambe in York, which is approximately 95 miles up the river from St Mary's Bridge on the Derwent River. The statistical correlation between the two flow rates was very strong, 0.84. This has implied that there might be an undiscovered underground water waterway underneath the Derwent River running parallel to it. This was argued on the basis that there are a lot of obstructions to the river, which would cancel this correlation. Graph 5.3 shows the two flow rates measured at Buttercrambe and St Mary' Bridge (Ahmed et al., 2007).

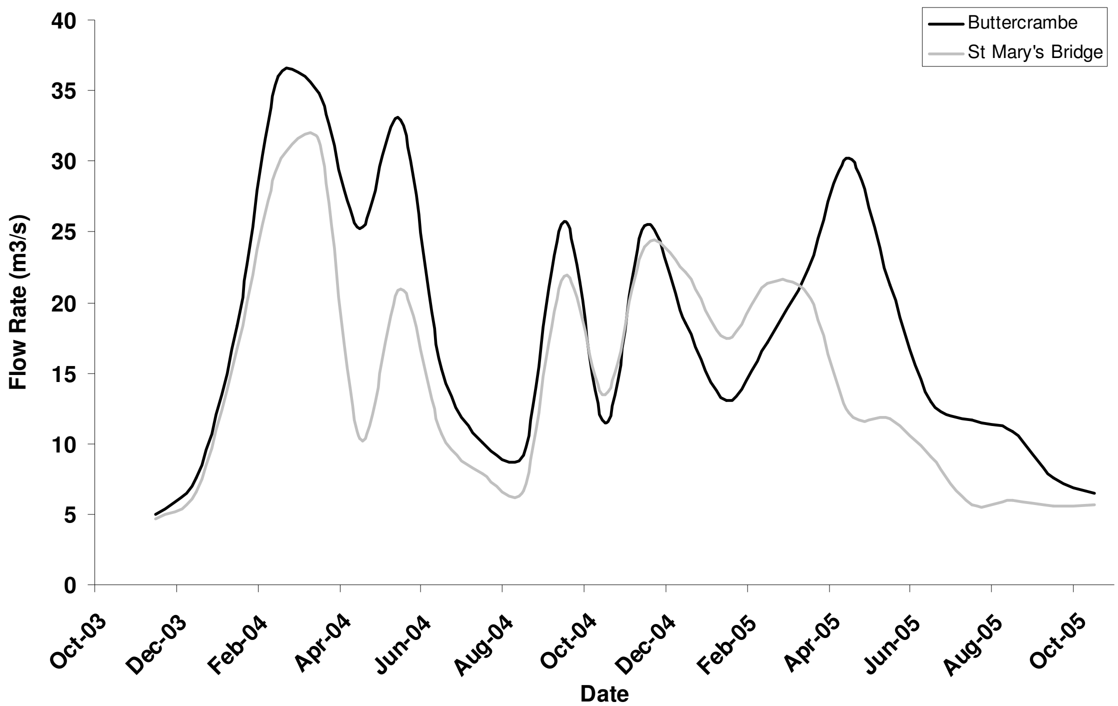


Figure 5-3 Comparison between the flow rates at Buttercrambe and St Mary's Bridge, both readings contribute to Little Eaton WTW.

From graph 5.3 above, it is clear that the two flow rates have similar trends. After graphical inspection of the graphs, the theory of underground waterway may well be a reality. The flow rate of the Buttercrambe readings is higher than those taken at St Mary's Bridge. This would be expected since there are numerous obstructions on the River Derwent, from which these readings are taken from.

The aim of this geological finding is that Virtual Flow Sensors can actually be placed on the system derived from other networked flow sensors. In other words, by knowing one or more flow rates of rivers in the Midland region, it is possible to derive an estimation of rivers within the region without having a flow sensor device in the river. This method can reduce maintenance cost and also give early flood warnings to areas lower than rivers.

5.2.4 REENGINEER DOPPLER-ULTRASONIC IN-SITU SENSOR

Using the fundamentals of Doppler-Ultrasound probes, and by redeveloping the coding operating the sensor, it has been possible to operate the sensor as a Turbidity sensor as well. This can have many implications for the operational work of the WTW, since the Doppler probe needs little maintenance and can achieve two different sets of constituent's readings at the same instant; Flow and Turbidity. Leading on from this, Colour and Conductivity can also be easily derived from the statistical correlation, giving an instant reading of the constituents at the probe location (Ahmed et al., 2007).

The probe can also be an in-situ version, clamped to the outside of a pipe. This version of the probe would take the maintenance work done on the probe a step further by eliminating maintenance and increasing the life expectancy of the probe. In fact, the probe lifetime is very much independent of anything apart from the lifetime of the pipe being attached to. Once the pipe reaches its end of working life, the probe can be taken out then re-clamped back again onto the new pipe.

The experimental setup was basically simple as a controlled amount of Bentonite clay powder being added to a water container, and connected to a pump and flume, which was holding approximately 100 l of water at any time during the experiment. The added amounts were controlled by adding the right dosage equivalent to 5 mg/l at each stage. The turbidity was increased from approximately 0 NTU (clear tap water) to approximately 30 NTU, which is double the average river turbidity reading measured. The ultrasonic Doppler sensor was inserted on top of the flume in the direction of the running water. Graph 5.4 illustrates the experimental results displayed in a graphical format.

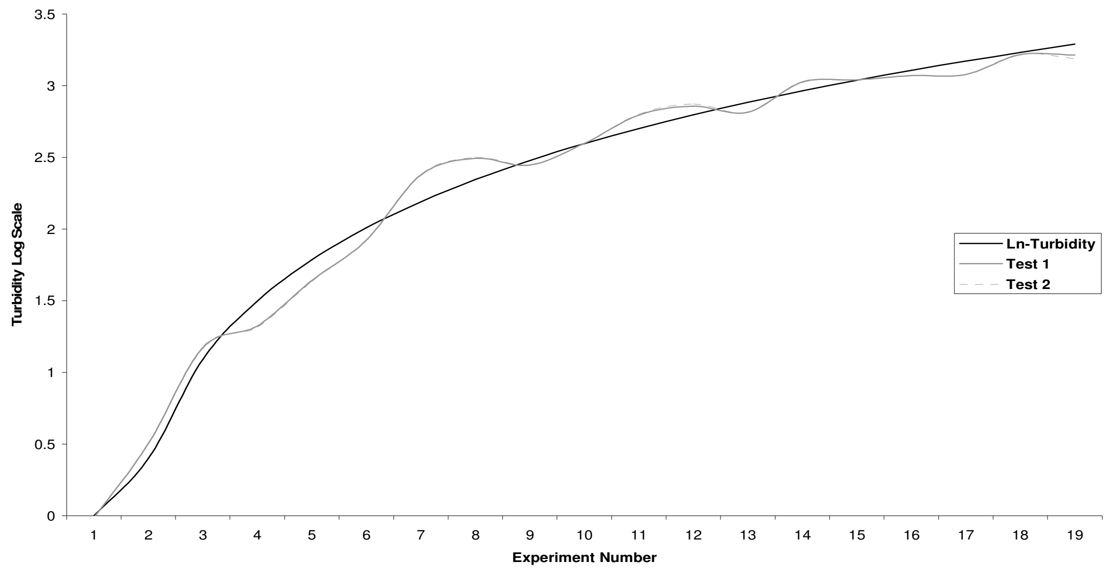


Figure 5-4 Comparison between the measured Turbidity levels versus readings taken by the Ultrasonic probe.

From graph 5.4, it is clearly visible that the probe had a high consistency and capable of detecting the amount of turbidity particulates flowing by it. But the major setback was that even after repeated trials, the test graphs could not be improved much further than this. This was accounted for the insufficient sensitivity of the probe. In other words, the probe was not designed to investigate this type of experiment, it was rather designed to merely measure the sounds coming off and back.

5.3 IMPLICATIONS AND IMPACT OF THE PROJECT

The implications and the consequent impact of the research within the water industry have already been methodically covered in the preceding chapters. Based upon a presentation of the research and recommendations (as well as previously completed work), Severn Trent would have;

1. Leased a new life for ageing systems; by working on similar methodologies to the one employed in this research, it can reproduce a similar effect using other available sensors.
2. Saved money by eliminating the dependency on wet lab results; online real-time readings would eliminate human intervention to a minimum and reduce operators' errors.
3. Improved WTW's efficiency; accurate water constituents readings would supersede the "play it safe" attitude within the operational procedures of the WTW's.
4. Improved quality.
5. Pushed the water industry closer to full autonomy of human intervention.

5.4 CRITICAL EVALUATION

The main limitations of the research project pertain to the enormity of the given subject matter; in-situ water monitoring covers a wide range of technologies and theories that were difficult to fully cover and understand during the research period. The research aim was driven by the need for a fully working solution to satisfy the sponsor and make a return on investment. The results do not, however, quantify the save on processes that can be achieved using the proposals given. Therefore, the confines of the work are considered to be:

- Ambiguous in terms of how much will the new proposals save Severn Trent in terms of cost and any other measures;
- The capital cost to implement such a proposal;
- Any laws affecting the installation of such systems; and
- The inability to compare the performance of the new Doppler Ultrasound to other devices.

5.5 RECOMMENDATIONS AND FUTURE WORK

Based upon the foregoing research findings and conclusions, the following recommendations are submitted for consideration by the water industry and, where applicable, for future work:

1. All utilities' companies and the Environmental Agency flow measuring devices are to be grid connected for flood warnings and quality checks;
2. Each river within England and Wales is to be investigated (using previous available data) and have its signature correlation noted;
3. Ofwat (The Water Services Regulation Authority) needs to put forward a new regulation aimed at utilities' companies to try and standardise the storage of data format;
4. A new in-situ monitoring operation needs to be forcefully introduced to the industry as standard operational monitoring procedure. The situation is far from simple to introduce this sort of monitoring procedure in reality. What tends to currently happen is that water utilities' companies would wait until another company takes the device on and test it before they can follow on;
5. More research funding needs to be set up collaboratively by the water companies, together with Environmental Agency, to try and aim for alternative (if not optimum) solutions for water monitoring; and
6. Cross industry panels are need (especially with the Oil industry) to exchange technologies and expertise.

The research undertaken could have given a whole new perspective of the data on hand had ANN been employed to analyse it. It is the recommendation of the author that further research would be carried out on the use of ANN together with Turbidity, Conductivity and Colour extraction levels from the Flow rate.

Temperature and sun index coefficients should also be incorporated in the ANN model as to be able to provide a complete picture of the environmental effect between Flow and Rain have on the Turbidity, Colour, Conductivity and Temperature parameters.

5.6 CHAPTER SUMMARY

This chapter introduced the last section of the research project and highlighted the achievements and shortfalls that this research has encountered. It also highlighted the recommendations that would be applicable to the engineering industry in general, and to the water industry in particular. New technological monitoring techniques have been discussed and experimental results findings shown.

The research conclusions were drawn out from the needs observed and from the analysis of the data gathered. These findings are:

a. A backup sensory network based on the flow measurements' readings taken up the intake pipe, which can ultimately give warning of the anticipated amount of Turbidity at the intake.
b. Using the correlation processes, a correlation between Turbidity, Colour and Conductivity were established based on particular river signatures. This can be simply termed as 'the more Turbidity in the water, the higher the Colour and Conductivity levels would tend to be', as Turbidity particles would have a direct influence on the charges in the water. It is commonly agreed that all rivers have their own signature in

terms of the correlation between the constituents dissolved within them. This is due to the surroundings and the topography of the land the river is running through.

c. The correlation leads from the first finding indicating that Colour and Conductivity levels, in addition to Turbidity, can actually be estimated prior to water reaching the intake using Flow readings. Further analysis found that another possible sensor, Flow, that was not initially looked at can also lead to the readings made by the later parameters mentioned. Correlation between the Flow and the other measured parameters were found to be strong. This meant that the higher the flow, the higher the Turbidity; which leads on to the last theory based from the correlation of the three parameters.

d. The Flow data was collected by the Environmental Agency. Hence, the flow data was not collected in the vicinity of the water treatment works. It was rather collected upstream from the intake of the WTW; which means that the flow was indicative of how much the flow of river is expected to reach at the intake.

e. Further analysis showed that the Flow sensor can be used as an indicative element of how much flow is expected at the intake, but this would not be the case if it started raining. Rain would most definitely have an impact on the flow; if not, then it would at least change the Turbidity in the water. The rain data was then collected and included in the calculation for the model

f. A complete drawn-out map of the Midlands region of river gauges can be used to estimate flows in other rivers. This was drawn out from Flow data supplied by the EA; the analysis of the data showed a strong cohesion between certain river data flows, indicating that rivers in the Midlands region all originated from one underground natural water reservoir.

g. A new alternative, non-invasive, Turbidity sensor device was reconfigured based on the working methodology of a Doppler-Ultrasonic Flow sensor.

Having highlighted the conclusions and findings for further research, it would be credible to state that this research has demonstrated the aims and objectives to a full extent, and this would be of benefit to the water industry as well as to the environmental agencies.

6 REFERENCES

- Ahmad, S.R, and Reynolds, D.M. 1999. Monitoring of water quality using fluorescence technique: prospect of on-line process control, Water Research, Volume 33, Issue 9, pp 2069-2074.
- American Water Works Association. 1990. Water Quality and Treatment: A Handbook of Community Water Supplies, McGraw-Hill, New York.
- APHA, 1992. Standard Methods for the Examination of Water and Waste-Water, 18th Edn, American Public Health Association, Washington, D.C.
- Aracil J. 1986. Máquinas, Sistemas y Modelos. Un Ensayo sobre Sistemática. TECNOS, S.A. Madrid.
- Aramini J., Wilson J., Allen B., Holt J., Sears W., McLean M. and Copes R., 2000. Drinking water quality and health care utilization for gastrointestinal illness in Greater Vancouver. Ottawa, Health Canada.
- Arthington, A.H. and Pusey, B.J., 2003. Flow restoration and protection in Australian rivers. River Research and Applications, volume 19, pp 1-19.
- Bazzani, M. and Cecchi, G. 1995. Algae and mucillagine monitoring by fluorescence LIDAR experiments in field. EARSeL Advances in Remote Sensing v. 3, pp. 90-101.
- Beaudeau P., Payment P., Bourderont D., Mansotte F., Boudhabay O., Laubies B. and Verdiere J. 1999. A time series study of anti-diarrhoeal drug sales and tap-water quality. Int J Env Health Res, volume 9, pp 293-312.
- Beckers, C.V., Chamberlain, S.G., and Grimsrud G.P. 1972. Quantitative methods for preliminary design of water quality surveillance systems. Socioeconomic Environmental Studies Series Report No. EPA-R5-72-001, US EPA, Washington, DC.

- Begley, C. M. 1996. Using triangulation in nursing research', Journal of Advanced Nursing, July, volume 24, pp 122-128.
- Berets, S.L, Milosevic, M. and Lucania, J.P. 2002. Diffuse Reflectance Sampling Methods. Harrick Scientific Products, Inc.
- Berger, P.S. 1992. Revised total coliform rule, In: Gilbert, C.E., and Calabrese, E.J., eds., Regulating Drinking Water, Lewis, Bocta Raton, FL, pp 161-6.
- Bilotta G.S. and Brazier R.E., 2008. Understanding the influence of suspended solids on water quality and aquatic biota. Water Research 42 (12), pp 2849-2861.
- Bolgrien, D.W., Granin, N.G., and Levin, L. 1995. Surface temperature dynamics of Lake Baikal observed from AVHRR images. Photogrammetric Engineering and remote Sensing v. 61, pp 211-216.
- Bowers, J.A., and Shedrow, C.B. 2000. Predicting stream water quality using artificial neural networks. WSRC-MS-2000-00112. http://www.osti.gov/bridge/.
- Brannen, J. 1992. Mixing Methods: Qualitative and Quantitative Research, Ashgate Publishing Company, England.
- Britton, G. 1994. Wastewater Microbiology, Wiley-Liss, New York.
- Brown SD, Sum ST, and Despagne F. 1996. Chemometrics. Anal Chem; volume 68, pp 21R-61R.
- Brown, S.D., Blank, T.B., Sum, S.T., and Weyer, L.G. 1994. Chemometrics. Anal. Chem. Volume 66, pp 315R–359R.
- Brown, S.D., Sum, S.T. and Despagne, F., 1996. Chemometrics. Anal. Chem. Volume 68, pp 21R–61R.
- Bryant, E.A., Fulton, G.P., and Budd, G.C. 1992. Disinfection Alternatives for Safe Drinking Water, Van Nostrand Reinhold, New York.

- Cabelli, V. 1978. New Standards for enteric bacteria. In: Mitchell, R., ed., Water Pollution Microbiology, Vol. 2, Wiley-Interscience, New York, pp 233-73.
- Cabreta-Mercader, C.R., and Staelin, D.H. 1995. Passive microwave relative humidity retrievals using feed forward neural network. IEEE Trans. Geo. Remote Sens. Volume 33, pp 842–852.
- Camel, V., and Bermond, A. 1998. The use of Ozone and associated oxidation processes in drinking water Treatment. Water Research, Volume 32, pp 3208-22.
- Campbell, J.B. 2002. Introduction to remote sensing, 3rd ed, The Guilford Press, New Yourk.
- Casey, T.J. 1997. Unit Treatment Processes in Water and Wate-Water Engineering, John Wiley and Sons, Chichester.
- Chalmers, J. M. and Griffiths, P. R. 2002, Handbook of Vibrational Spectroscopy, Vol. 2, John Wiley and Sons, Ltd.
- Chen, J.C., Chang, N.B., and Shieh, W.K. 2003. Assessing wastewater reclamation potential by neural network model. Eng. Appl. Artificial Intelligence, Volume 16, pp 149–157.
- Chilton, R. (ed.), 1977. Water Pipeline System: Leakage Management, Network optimisation and pipeline rehabilitation technology, BHR Group Conference Series 23, Mechanical Engineering Publications, London
- Chu, W.C., and Bose, N.K. 1998. Speech signal prediction using feed forward neural network. Electro. Lett., Volume 34, pp 999–1001.
- Cieniawski, S.E., Eheart, J.W., and Ranjithan, S. 1995. Using genetic algorithms to solve a multi-objective groundwater monitoring problem. Water Resources. Research, Volume 31, pp 399–409.
- Cole Parmer: Free Catalogues/Literature, 2008: http://www.coleparmer.co.uk/

- Collingwood, R.W. 1966. Water storage and reservoir management: Their effect on water quality. Association of River Authorities Year Book, pp 24-34.
- Cook, B. W., and Jones, K. 1992. A programmed introduction to infrared spectroscopy, Heyclen.
- Cowx, I.G. and Gould, R.A. 1989. Effects of stream regulation on Atlantic Salmon, Salmo salar L., and Brown trout, Salmo trutta L., in the Upper Severn catchment, U.K., Regulated Rivers: Research and Management, Volume 3, pp 235-245.
- Craun, G.F. 1986. Waterborne Diseases in the United States, CRC press, Boca Raton, FL.
- Cubitt, D.W. 1991. A review of the epidemiology and diagnosis of waterborne viral infections: II, Water, Science and Technology, Volume 24, pp 197-203.
- Curran, P.J. and Novo, E.M.M. 1988. The relationship between suspended sediment concentration and remotely sensed spectral radiance: A review. Journal of Coastal Research, Volume 4, pp 351-368.
- Davies-Colley R. J. and Close M. E., 1990. Water Colour and Clarity of New Zealand Rivers under Baseflow Conditions. New Zealand Journal of Marine and Freshwater Research, Volume 24. pp 357–365
- Davies-Colley R. J. and Smith D.G,, 2001. Turbidity, Suspended Sediment, and Water Clarity: A Review. Journal of the American Water Resources Association, Volume 37. pp 1085–1101.
- De Jong, S. 1993. SIMPLS: an alternative approach to partial least squares regression. Chemometrics and Intelligent Laboratory Systems, Volume 18. pp 251-263
- De Jong, S. and Kiers, H. 1992. Principal covariates regression. Chemometrics and Intelligent Laboratory Systems, Volume 14, pp 155-164.

- De Merona B., Vigouroux, R. and Tejerina-Garro, F.L. 2005. Alteration of fish diversity downstream from Petit-Saut Dam in French Guiana. Implication of ecological strategies of fish species, Hydrolbiologia, Volume 551. pp 33-47.
- Dekker, A.G., Malthus, T.J., and Hoogenboom, H.J. 1995. The remote sensing of inland water quality, pp. 123-142. In: Danson, F.M. and S.E. Plummer (Eds.), Advances in Remote Sensing. Chichester: John Wiley and Sons.
- Dibike, Y.B., and Solomatine, D.P. 2001. River flow forecasting using artificial neural networks, Journal of Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere, Volume 26, pp 1–8.
- Diem, M. 1993. Introduction to modern vibrational spectroscopy, Wiley, New York.
- Dijkstra, T. 1983. Some comments on maximum likelihood and partial least squares methods. Journal of Econometrics, Volume 22, pp 67-90.
- Dijkstra, T. 1985. Latent variables in linear stochastic models: Reflections on maximum likelihood and partial least squares methods. 2nd ed. Amsterdam, The Netherlands: Sociometric Research Foundation.
- Dixon, W., and Chiswell, B. 1996. Review of aquatic monitoring program design, Water Research, Volume 30, pp 1935–1948.
- Dobson, P. J. 2001. The Philosophy of Critical Realism-An Opportunity for Information Systems Research. Information Systems Frontiers, Volume 3, pp 199-210.
- Du Zuane, J. 1990. Handbook of Drinking Water Quality Standards and Controls, Van Nostrand, Reinhold, New York.
- Dykstra. R. L. 1983. An algorithm for restricted least squares, Journal of the American Statistical Association, Volume 78, pp 837-842.
- EAL, 1997. Validation of test methods – general principles and concepts. EAL-P11, European cooperation for Accreditation of Laboratories.

- Effler S.W., 1988. Secchi Disc Transparency and Turbidity. Journal of Environmental Engineering, ASCE, Volume 114, pp 1436–1447.
- Eife R., Weiss M., Barros V., Sigmund B., Goriup U., Komb D., Wolf W., Kittel J., Schramel P. and Reiter K. 1999. Chronic poisoning by copper in tap water: I. Copper intoxications with predominantly gastrointestinal symptoms. Eur J Med Res, Volume 4, pp 219-223.
- Elemental Analysis Inc. Website, 2008: http://www.elementanalysis.com/
- Ellis, J.C., 1989. Handbook on the Design and Interpretation of Monitoring Programs. Water Research Center, Medmenham.
- Environment Protection Agency, 1998. Review finds Philadelphia turbidity study seriously flawed. Health Stream, Volume 9.
- Esterby, S.R., 1996. Review of methods for the detection and estimation of trends with emphasis on water quality applications. Hydrol. Process., Volume 10, pp 127–149.
- Etchells, T., Tan, K.S., and Fox, D. 2005. Quantifying the uncertainty of nutrient load estimates in the Shepparton irrigation region. In: Proceedings MODSIM05 International Congress on Modelling and Simulation, Advances and Applications for Management and Decision Making. Available at:

http://www.mssanz.org.au/modsim05/papers/etchells_2.pdf.

- European Union Drink Water Directive, accessed 2006.
- Fellows, R. and Liu, A. 1999. Research methods for construction. London: Blackwell Science Ltd. ISBN 0 632 04244 3.
- Fellows, R. and Liu, A. 2005. Research Methods for Construction, 2nd Edition, Blackwell Publishing, UK.
- Frank, I. and Friedman, J. 1993. A statistical view of some chemometrics regression tools Technometrics, Volume 35, pp 109-135.

- From the Centers for Disease Control and Prevention, 1993. Assessment of inadequately filtered public drinking water. Washington, D.C., Volume 272, pp 1401-1402.
- Gardner, M.W., and Dorling, S.R. 1998. Artificial neural network: the multilayer perceptron: a review of applications in atmospheric sciences. Atmos. Environ., Volume 32, pp 2627–2636.
- Gascuel-Odoux, C., Aurousseau, P., Cordier, M., Durand, P., Garcia, F., Masson, V., Salmon-Monviola, J., Tortrat, F., and Trepos, R. 2009. A decision-oriented model to evaluate the effect of land use an agricultural management on herbicide contamination in stream water. Environmental Modelling & Software, Volume 24, pp 1433–1446.
- Geladi, P, and Kowalski, B. 1986. Partial least squares regression: A tutorial. Analytical Chimica Acta, Volume 185, pp 1-17.
- Geldreich, E.E. 1996. Microbial Quality of Water Supply in Distribution Systems, Lewis Publishers, Bocta Raton, Fl.
- Gibbons, D.E., Wukelic, G.E., Leighton, J.P., and Doyle, M.J. 1989. Application of Landsat Thematic Mapper data for coastal thermal plume analysis at Diablo Canyon. Photogrammetric Engineering and Remote Sensing Volume 55, pp 903-909.
- Gibson, P.J. 2000. Introductory Remote Sensing: Principles and Concepts, Routledge publishers.
- Gilbert M.L., Levallois P. and Rodriguez M.J., 2006. Use of a health information telephone line, Info-sante CLSC, for the surveillance of waterborne gastroenteritis. J Water Health, Volume 4, pp 225-232.
- Gitelson, A., Mayo, M., Yacobi, Y.Z., Paroarov, A., and Berman, T. 1994. The use of high spectral resolution radiometer data for detection of low chlorophyll concentrations in Lake Kinneret. Journal of Plankton Research, Volume 16, pp 993-1002.

- Glass, R.L., Vessey, I. and Ramesh, V. 2002. Research in Software Engineering: an analysis of the literature. Information and Software Technology, Volume 44, pp 491-506
- Gleeson, C., and Gray, N.F., 1997, The Coliform Index and Waterborne Disease: Problems of Microbial Drinking Water Assessment, E., and F.N. Spon, London.
- Gleick, 2003 H.P. Gleick, Water use, Annual Review of Environment and Resources, Volume 28, pp 275–314.
- Gray, N.F. 1992. Biology of Wastewater Treatment, Oxford University Press, Oxford.
- Gray, N.F. 1994. Drinking Water Quality: Problems and Solutions, John Wiley and Sons, Chichester.
- Gregg, D. G., Kulkarni, U. R. and Vinzé, A. S. 2001. Understanding the Philosophical Underpinnings of Software Engineering Research in Information Systems. Information Systems Frontiers, Volume 3, pp 169-183.
- Groot, S., and Schilperoort, T., 1984. Optimization of water-quality monitoring networks. Water Sci. Technol., Volume 16, pp 275–287.
- Guion, L. A. 2002. Triangulation: Establishing the Validity of Qualitative Studies, publication FCS6014, Institute of Food and Agricultural Sciences, University of Florida.
- Hanbay, D., Turkoglu, I., and Demir, Y., 2008. Prediction of wastewater treatment plant performance based on wavelet packet decomposition and neural networks. Expert Syst. Appl., Volume 34, pp 1038–1043.
- Harding, L.W., Itsweire, E.C., and Esaias, W.E. 1995. Algorithm development for recovering chlorophyll concentrations int Chesapeake Bay using aircraft remote sensing, 1989-91. Photogrammetric Engineering and Remote Sensing, Volume 61, pp 177-185.

- Harmancioglu, N.B., and Alpaslan, N., 1992. Water-quality monitoring network design – a problem of multiobjective decision-making. JAWRA Journal of the American Water Resources Association, Volume 28, pp 179–192.
- Hart, J.K., and Martinez, K., 2006. Environmental sensor networks: a revolution in earth system science? Earth-Science Reviews, Volume 78, pp 177–191.
- Haykin, S. 1994. Neural Networks, a Comprehensive Foundation. New York: Macmillan.
- Hoerl, A. and Kennard, R. 1970. Ridge regression: Biased estimation for non-orthogonal problems. Technometrics, Volume 12, pp 55-67.
- Holt, G. 1998. A guide to successful dissertation study for students of the built environment, Second Edition, Built Environment Research Unit, Wolverhampton, ISBN 1-902010-01-0.
- Honeywell Website, 2008: http://www.honeywell.com/.
- http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31998L0083:EN:HTML
- Hudak, P.F., Loaiciga, H.A., and Marino, M.A. 1995. Regional-scale ground-water quality monitoring via integer programming. J. Hydrol., Volume 164, pp 153–170.
- IESWTR (Interim Enhanced Surface Water Treatment Rule), 1998. Washington, D.C., United States Environmental Protection Agency.
- Ilker, T.T., Kijin, N., Jiabao, G., and Mustafa, M.A. 2009. Optimal water quality monitoring network design for river systems. J. Environ Management, Volume 90, pp 2987–2998.
- Jakob, A. 2001. The Triangulation of Quantitative and Qualitative data in Typological Social Research: Reflections on a Typology of Conceptualising 'Uncertainty' in the context of Employment Biographies, Forum Qualitative Social Research, Volume 2.

- Jensen, A.J. 2003. Atlantic Sakmon (Salmo Salar) in the regulated river Alta: Effects of altered water temperature on Parr growth. River Res. Applic., Volume 19, pp 733-747.
- Jones, K. 1994. Inside Science: 73, Waterborne Diseases, New Scientist, 9 July, 1-4.
- Jordan, P., Arnscheidt, A., McGrogan, H., and McCormick, S. 2007. Characterising phosphorus transfers in rural catchments using a continuous bank-side analyser. Hydrology and Earth System Sciences, Volume 11, pp 372–381.
- Key, J. P. 1997. Experimental Research and Design, Research Design in Occupational Education, Presentation.
- Kirchner, J.W., 2006. Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. Water Resources Research , Volume 42, W03S04.
- Kirchner, J.W., Feng, X., Neal, C., and Robson, A.J. 2004. The fine structure of water quality dynamics: the (high-frequency) wave of the future. Hydrological Processes Volume 18, pp 1353–1359.
- Klein H. and Hirschheim R. 2003. Crisis in the IS Field. A Critical Reflection on the State of the Discipline. Journal of AIS, Volume 4, Article 10.
- Knobeloch L., Ziarnik M., Howard J., Theis B., Farmer D., Anderson H. and Proctor M., 1994. Gastrointestinal upsets associated with ingestion of copper-contaminated water. Environ Health Perspect, Volume 102, pp 958-961.
- Kortum, G. 1969. Reflectance Spectrsoscopy, Springer, New York.
- Kubelka, P. and Munk, F.Z. 1931, Tech. Phys., Volume 12, pp 593.
- Kung, S.Y., and Taur, J.S. 1995. Decision-based neural networks with signal image classification applications. IEEE Trans. Neural Network., Volume 6, pp 170–181.

- Kuo, J., Hsieh, M., Lung, W., and She, N. 2007. Using artificial neural network for reservoir eutriphication prediction. Ecol. Model., Volume 200, pp 171–177.
- Kuo, Y., Liu, C., and Lin, K.H. 2004. Evaluation of the ability of an artificial neural network model to assess the variation of groundwater quality in an area of blackfoot disease in Taiwan. Water Res., Volume 38, pp 148–158.
- Kurunc, A., Yurekli, K., and Cevik, O. 2005. Performance of two stochastic approaches for forecasting water quality and streamflow data from Yesilirmak River, Turkey. Environ. Model. Soft., Volume 20, pp 1195–1200.
- Kwiatkowski, R.E., 1991. Statistical needs in national water-quality monitoring programs. Environ. Monit. Assess., Volume 17, pp 253–271.
- Latham, B. 1990. Water Distribution, Chattered Institution of Water and Environmental Management, London.
- Lázaro M. and Marcos E. 2005. Research in Software Engineering: Paradigms and Methods. The 17th Conference on Advanced Information Systems Engineering (CAiSE'05).
- LeChevallier M.W., Karim M., Aboytes R., Gullick R., Weihe J., Earnhardt B., Mohr J., Starcevich J., Case J., Rosen J.S., Sobrinho J., Clancy J.L., McCuin R.M., Funk J.E. and Wood D.J., 2004. Profiling Water Quality Parameters: From Source Water to the Household Tap. London, IWA Publishing, 230.
- Lek, S., Delacoste, M., Baran, P., Dimopoulos, I., Lauga, J., and Aulagnier, S. 1996. Application of neural networks to modeling nonlinear relationships in ecology. Ecol. Model., Volume 90, pp 39–52.
- Lek, S., and Guegan, J.F. 1999. Artificial neural networks as a tool in ecological modeling, an introduction. Ecol. Model., Volume 120, pp 65–73.

- Lerner, B., Levinstein, M., Rosenberg, B., Guterman, H., Dinsteun, I., and Romen, Y. 1994. Feature selection and chromosomes classification using a multilayer perceptron neural network. In: IEEE International Conference on Neural Networks, Orlando, Florida, pp 3540–3545.
- Levenson, M.D. 1982. Introduction to nonlinear laser spectroscopy, Academic Press, New York, London
- Levine BK. Chemometrics. Anal Chem 1998; Volume 70, pp 209R-28R.
- Levine BK. Chemometrics. Anal Chem 2000; Volume 72, pp 91R-7R.
- Lewis J., 1996. Turbidity-controlled suspended sediment sampling for runoff-event load estimation. Water Resources Res., Volume 32, pp 2299–2310.
- Li, R.Z., 2006. Advanced and trend analysis of theoretical methodology for water quality forecast. J. Hefei Univ. Technol. Volume 29, pp 26–30.
- Lo, J.Y., Baker, J.A., Kornguth, P.J., and Floyd, C.E. 1995. Application of artificial neural networks to the interpretation of mammograms on the basis of the radiologists impressions and optimized BI-RADS TM image features. Radiology, Volume 197, pp 242–1242.
- Loftis, J.C., McBride, G.B., and Ellis, J.C. 1991. Considerations of scale in water-quality monitoring and data-analysis. JAWRA Journal of the American Water Resources Association, Volume 27, pp 255–264
- Lytle, D.A., and Poff, N.L. 2004. Adaptation to natural flow regimes, Trends Ecol. Evol., Volume 19, pp 94–100
- Mann A.G., Tam C.C., Higgins C.D. and Rodrigues L.C. 2007. The association between drinking water turbidity and gastrointestinal illness: a systematic review. BMC Public Health, Volume 7, pp 256.

- Marcos, E. and Marcos, A. 1998. An Aristotelian Approach to the Methodological Research: a Method for Data Models Construction. Information Systems- The Next Generation. L. Brooks and C. Kimble (Eds.). Mc Graw-Hill, pp 532-543
- Massart DL, Vandeginste BGM, Deming SN, Michotte Y, and Kaufman L. Chemometrics: a text book. Amsterdam: Elsevier, 1988.
- Massart, D.L., and Kaufman, L. 1983. The Interpretation of Analytical Chemical Data by the use of Cluster Analysis, John Wiley & Sons, Inc., New York.
- Massart, D.L., Vandeginste, B.G.M., Deming, S.N., Michotte, Y. and Kaufman, L. 1988. Chemometrics: a Text Book, Elsevier Science Publishers B.V., Amsterdam, The Netherlands.
- Mathison, S. 1988. Why Triangulate?. Educational Researcher, Volume 17, pp 13-17.
- May, D.B., and Sivakumar, M. 2009. Prediction of urban stormwater quality using artificial neural networks. Environmental Modelling & Software, Volume 24, pp 296–302.
- McFeters, G.A. (ed.), 1990. Drinking Water Microbiology, Springer, New York.
- McGirr D. J., 1974. Turbidity and Filterable and Non-filterable Residue. Interlaboratory Quality Control Study No. 10. Canada Centre for Inland Waters Report Series No. 37. Burlington, Ontario. 9 pp
- Messikh, N., Samar, M.H., and Messikh, L. 2007. Neural network analysis of liquid–liquid extraction of phenol from wastewater using TBP solvent. Desalination, Volume 208, pp 42–48.
- Metcalf and Eddy. 1991. Waste-Water Engineering: Treatment, Disposal and Reuse, McGrew-Hill, New York.
- Miles M. B. and Huberman, A. M. 1984. Quality Data Analysis: A sourcebook of New Methods. SAGE. NewBury Park-CA.

- Montgomery, J.L., Harmon, T., Kaiser, W., Sanderson, A., Haas, C.N., Hooper, R., Minsker, B., Schnoor, J., Clesceri, N.L., Graham, W., and Brezonik, P. 2007. The WATERS Network: an integrated environmental observatory network for water research. Environmental Science and Technology, Volume 41, pp 6642–6647. Available at: http://pubs.acs.org/subscribe/journals/esthag/41/i19/pdf/100107feature_waters.pdf.
- Morris R.D., Naumova E.N. and Griffiths J.K. 1998. Did Milwaukee experience waterborne cryptosporidiosis before the large documented outbreak in 1993? Epidemiology, Volume 9, pp 264-270.
- Myers, M. D. 1997. Qualitative Research in Information Systems. MIS Quarterly, Volume 21, pp 241-242.
- Naiman, R.J., Bilby, R.E., and Bisson, P.A. 2000. Riparian ecology and management in the Pacific coastal rain forest, Bioscience, Volume 50, pp 996–1011.
- Naoum, S. G. 2001. Dissertation research and writing for construction students. Butterworth Heinmann, Oxford, ISBN 0-7506-2988-6.
- National Instruments Website, 2008: http://www.ni.com/.
- Niu, Z.G., Zhang, H.W., and Liu, H.B. 2006. Application of neural network to prediction of coastal water quality. J. Tianjin Polytechnic Univ., Volume 25, pp 89–92.
- Nolan, K.M., Bohman, L., Stamey, T., and Firda, G. 2005. Stage–Discharge Relations-Basic Concepts Training Class USGS Scientific Investigations Report 2005–5028, USGS Training Class SW1409. Available at: http://wwwrcamnl.wr.usgs.gov/sws/SWTraining/RatingsWeb/Index.html.
- Open University. 1975. Water Distribution, Drainage, Discharge and Disposal, PT272-7, Open University Press, Milton Keynes.

- Poff, N.L., Olden, J.D., Merritt, D.M., and Pepin, D.M. 2007. Homogenization of regional river dynamics by dams and global biodiversity implications, Proc. Natl. Acad. Sci., Volume 104, pp 5732–5737.
- Postel, S. and Richter, B. 2003. Rivers for life. Managing water for people and nature. Island Press, London.
- Prescott, L.M., Harley, J.P., and Klein, D.A. 1993, Microbiology, WCB Publishers, Iowa.
- Rahim, M.G., Goodyear, C.C., Kleijn, W.B., Schroeter, J., and Sondhi, M.M., 1993. On the use of neural networks in articulatory speech synthesis. Journal of the Acoustical Society of America, Volume 93, pp 1109-1121.
- Raman, H., and Chandramouli, V., 1996. Deriving a general operating policy for reservoirs using neural networks. J. Water Resource. Planning Management., Volume 122, pp 342–347.
- Reasoner, D. 1992. Pathogens in Drinking Water – Are there any new ones?. US Environmental Protection Agency, Washington, D.C.
- Ritchie, J.C. and Cooper, C.M. 1988. Comparison on measured suspended sediment concentrations with suspended sediment concentrations estimated from Landsat MSS data. International Journal of Remote Sensing, Volume 9, pp 379-387.
- Ritchie, J.C. and Cooper, C.M. 1991. An algorithm for using Landsat MSS for estimating surface suspended sediments. Water Resources Bulletin, Volume 27, pp 373-379.
- Ritchie, J.C. and Cooper, C.M. 2001. Remote sensing of water quality: Application to TMDL, pp 367-375. In: TMDL Science Issues Conference, Water Environment Federation, Alexandria, VA.

- Ritchie, J.C. and Schiebe, F.R. 2000. Water Quality, pp 287-303, 351-352. In: G.A. Schultz and E.T. Engman (eds.), Remote Sensing in Hydrology and Water Management, Springer-Verlag, Berlin, Germany.
- Ritchie, J.C., Cooper, C.M., and Schiebe, F.R. 1990. The relationship of MSS and TM digital data with suspended sediments, chlorophyll, and temperature in Moon lake, Mississippi. Remote Sensing Environment, Volume 33, pp 137-148.
- Ritchie, J.C., Schiebe, F.R., and McHenry, J.R. 1976. Remote sensing of suspended sediment in surface water. Photogrammetric Engineering and Remote Sensing, Volume 42, pp 1539-1545.
- Ritchie, J.C., Schiebe, F.R., Cooper, C.M., and Harrington Jr., J.A. 1994. Chlorophyll measurements in the presence of suspended sediment using broad band spectral sensors aboard satellites. Journal of Freshwater Ecology, Volume 9, pp 197-206.
- Rogers, L.L., and Dowla, F.U. 1994. Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resource Research., Volume 30, pp 457–481.
- Rouse R., 2001. New Drinking Water Regulations in the UK. London , Drinking Water Inspectorate.
- Sarle, W.S. 1994. Neural Networks and Statistical Models. Proceedings of the Nineteenth Annual SAS Users Group International Conference, Cary, NC: SAS Institute, pp 1538-1550
- Schalkoff, R., 1992. Pattern Recognition: Statistical, Structural and Neural Approaches. Wiley, NY.
- Schalles, J.F., Schiebe, F.R., Starks, P.J., and Troeger, W.W. 1997. Estimation of algal and suspended sediment loads (singly and combined) using hyperspectral sensors and

integrated mesocosm experiments. Proceedings of the 4th International Conference on Remote Sensing for Marine and Coastal Environments, Volume 1, pp 247-248.

- Schiebe, F.R., Ritchie, J.C., and McHenry, J.R. 1976. Influence of suspended sediments on the temperature of surface waters in reservoirs. Verh. International Ver. Limnology, Volume 19, pp 133-136.
- Scholefield, D., Le Goff, T., Braven, J., Ebdon, L., Long, T., and Butler, M. 2005. Concerted diurnal patterns in riverine nutrient concentrations and physical conditions. Science of the Total Environment, Volume 344, pp 201–210.
- Schuster C.J., Ellis A.G., Robertson W.J., Charron D.F., Aramini J.J., Marshall B.J. and Medeiros D.T. 2005. Infectious disease outbreaks related to drinking water in Canada, 1974-2001. Can J Public Health, Volume 96, pp 254-258.
- Schwartz J., Levin R. and Goldstein R., 2000. Drinking water turbidity and gastrointestinal illness in the elderly of Philadelphia. J Epidemiol Community Health, Volume 54, pp 45-51.
- Schwartz J., Levin R. and Hodge K. 1997. Drinking water turbidity and paediatric hospital use for gastrointestinal illness in Philadelphia. Epidemiology, Volume 8, pp 615-620.
- Shu, J., 2006. Using neural network model to predict water quality. North Environ. Volume 31, pp 44–46.
- Sinclair M.I. and Fairley C.K. 2000. Drinking water and endemic gastrointestinal illness. J Epidemiol Community Health, Volume 54, pp 728.
- Skalski, J.R., and McKenzie, D.H. 1982. A design for aquatic monitoring programs. J. Environ. Manag. Volume 14, pp 237–251.

- Smith, D.G., and McBride, G.B. 1990. New-Zealand National Water-Quality Monitoring Network – design and 1st year's operation. Water Resources Bull. Volume 26, pp 767–775.
- Smith, H.V. 1992. Cryptosporidium and Water: a review, Journal of the Institute of Water Engineers and Scientist, Volume 6, pp 443-51.
- Smits, J.R.M., Breedveld, L.W., Derksen, M.W.J., Katerman, G., Balfoort, H.W., Snoek, J. and Hofstraat, J.W. 1992. Pattern classification with artificial neural networks: classification of algae, based upon flow cytometer data. Analatical Chemistry. Acta, Volume 258, pp 11–25.
- Solt, G.S., and Shirely, C.B. 1991. An Engineer's Guide to Water Treatment, Avebury Technical, Aldershot.
- Spellman, F.R. 1999. Spellman's Standard Handbook for Wastewater Operators, Vol. 1, Technomic Publ., Lancaster, PA.
- Starodubstev, V.M., Fedorenko, O.L. and Burlibaev, M.Z. 2004. Assessment of the influence of river runoff regulation on ecological situation. Risk Assessment as a tool for water resources decision-making in central Asia, Volume 34, pp 279-304.
- Stevenson, D.G. 1998. Water Treatment Unit Processes, Imperial College Press, London.
- Stone, M. and Brooks, R. 1990. Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. Journal of the Royal Statistical Society, Series B, Volume 52, pp 237-269.
- Strobl, R.O. and Robillard, P.D. 2008. Network design for water quality monitoring of surface freshwaters: a review. J. Environ. Management., Volume 87, pp 639–648.
- Tebbutt, T.H.Y. 1979. Principles of Water Quality Control, Pergamon Press, London.

- Tomlinson, M.S. and De Carlo, E.H. 2003. The need for high resolution time series data to characterize Hawaiian streams. JAWRA Journal of the American Water Resources Association, Volume 39, pp 113–123,
- Twort, A.C., Crowley, F.C., Lawd, F.M, and Ratnayka, P.P. 1994. Water supply, 4th Edn, Arnold, London.
- U. S. Environmental Protection Agency. 2000. Nutrient Criteria Technical Guidance Manual. Lakes and Reservoirs. EPA- 822-F-00-002.
- U.S. Environmental Protection Agency Office of Water. 1998. The Quality of Our Nation s Waters: 1996, Executive Summary, April 1998 - EPA842-F-99-0003D.
- Ullidtz, P. 2002. Analytical tools for design of flexible pavements, Keynote Address, 9th International Conference on Asphalt Pavements, Copenhagen, Denmark.
- van den Wollenberg, A.L. 1977. Redundancy Analysis--An Alternative to Canonical Correlation Analysis. Psychometrika, Volume 42, pp 207-219.
- van der Voet, H. 1994. Comparing the predictive accuracy of models using a simple randomization test. Chemometrics and Intelligent Laboratory Systems, Volume 25, pp 313-323.
- Vigneswaran, S., and Visvanathan, C. 1995. Water Treatment processes: simple options, CRC Press, Boca Raton, Florida.
- Vincenti, W. G. 1990. What Engineers Know and How They Know It. Baltimore, MD: John Hopkins University Press.
- Vörösmarty, C., Lettenmaier, D., Leveque, C., Meybeck, M., Pahl-Wostl, C., Alcamo, J., Cosgrove, W., Grassl, H., Hoff, H., Kabat, P., Lansigan, F., Lawford, R. and Naiman, R. 2004. Humans transforming the global water system, EOS, Volume 85, pp 509–520.

- Voss, C., Tsikriktsis, N. & Frohlich, M. 2002. Case Research in operations management, International Journal of Operations & Production Management, Volume 22, pp 195-219.
- Wen, C.W., Lee, C.S. 1998. A neural network approach to multi-objective optimization for water quality management in a river basin. Water Resource Research., Volume 34, pp 427–436.
- Wendlandt, W. W. and Hencht, H. G. 1996. Reflectance Spectroscopy , John Wiley and Sons, New York.
- Wenning, R.J. and Erickson, G.A. 1994. Interpretation and analysis of complex environmental data using chemometric methods, Trends in Analytical Chemistry, Volume 13, pp 446–457.
- Whitfield, P.H., 1988. Goals and data-collection designs for water-quality monitoring. Water Resour. Bull., Volume 24, pp 775–780.
- Whitlock, C.H., Kuo, C.Y., and LeCroy, S.R. 1982. Criteria for the use of regression analysis for remote sensing of sediment and pollutants. Remote Sensing Environment, Volume 12, pp 151-168.
- WHO, 1993. Guidelines for Drinking Water Quality, Vol. 1: Recommendations, 2nd edn., World Health Organisation, Geneva.
- Wilkinson, S.N., Prosser, I.P., Rustomji, P., and Read, A.M. 2009. Modelling and testing spatially distributed sediment budgets to relate erosion processes to sediment yields. Environmental Modelling & Software, Volume 24, pp 489–501.
- Wohlin C. 2000. Experimentation in Software Engineering: An introduction, Springer.
- WRc, 1991. Recovery of Cryptosporidium from Water, FR 0189, Foundation for Water Research, Marlow.

- WRc, 1994. Removal of Cryptosporidium oocysts by Water Treatment Processes, FR 0457, Foundation for Water Research, Marlow.
- Wu, H.J., Lin, Z.Y., and Guo, S.L., 2000. The application of artificial neural networks in the resources and environment. Resource Environment Yangtze Basin, Volume 9, pp 237–241.
- Xiang, S.L., Liu, Z.M., and Ma, L.P., 2006. Study of multivariate linear regression analysis model for ground water quality prediction. Guizhou Sci., Volume 24, pp 60–62.
- Yang, W., Nan, J. and Sun, D., 2008. An online water quality monitoring and management system developed for the Liming River basin in Daqing, China. J. Environ. Management., Volume 88, pp 318–325.
- Yeomans, S.G. 2005. ICT-Enabled collaborative working methodologies in construction. CICE, Loughborough University Thesis.
- Yin, R. K. 1984. Case Study Research: Design and Methods, 2nd Edition, Sage Publications.

APPENDIX A (PAPER 1)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2008. The impact of bad sensors on the water industry and possible alternatives, ITcon, 13, Special Issue, Sensors in Construction and Infrastructure Management, pp. 166-178.

INVESTIGATION: THE IMPACT OF BAD SENSORS ON THE WATER INDUSTRY AND THE POSSIBLE ALTERNATIVES

Revised: 01 June 2007

A. Moustafa, A. El-Hamalawi and A. Wheatley

Civil & Building Engineering Department, Loughborough University, Loughborough LE11 3TU, England (UK)
Email: (A.Moustafa@lboro.ac.uk; A.El-Hamalawi@lboro.ac.uk; a.d.wheatley@lboro.ac.uk)

Keywords: Acoustic-Doppler, Online, Drinking Water, Sensors.

The advanced monitoring of water quality for the performing of a real-time hazard analysis prior to Water Treatment Works (WTW) is more of a necessity nowadays both to give warning of any contamination and also to avoid downtime of the WTW. Downtimes could be a major contributor to risk. Any serious accident will cause a significant loss in customer and investor confidence. In this paper, two case studies treatment plants were examined. One was a groundwater WTW and the other a river WTW. The results showed that a good correlation existed between the controlling parameters measured at the river WTW, but there was a lack of good warning parameters correlation for the groundwater WTW. Both of the case studies highlighted the need for a new non-invasive/remote sensor measurement and some new investment in information technology infrastructure. Results emphasised the value of backup monitoring and self-adjusting automation processes that are needed also to counteract the rise in power costs. The study showed that a relationship between the different types of sensors and/or measured parameters can be deduced in order to cross-check the sensors performance and be used as a guide to when maintenance is really needed. Operating hierarchal procedures within the WTWs could also be used to cut costs, by affecting conditioning monitoring.

1. INTRODUCTION

The water quality objectives for drinking water treatment works are regulated by the EU Drinking Water Directive (1998). A water treatment works must be able to produce a consistently high quality regardless of the quality of the intake or how great the demand might be. Water treatment consists of a range of unit processes, used in a multi barrier series and this provides some design and operational flexibility to maintain water quality. The treatment required will depend upon its initial quality, which is normally related to its source. In other words, the cleaner the raw water, the fewer treatment steps that are required, and hence the overall cost of water is less (Hughes, 2004).

In 2004, 375 raw water sites were monitored for compliance with the Surface Water Abstraction Directive (75/440/EEC) in England and Wales (UK Environment Agency, 2006). Of these, 155 sites failed to comply with the Directive. However, over 90% of these “failures” were due to insufficient sampling. These sampling shortfalls occur for a number of reasons, such as abstractions not being operated at the time of sampling, analysis problems at the laboratory, and sampling error. The quality of abstracted water was reported to have generally improved since 1993.

In order to comply with these regulations on a 95% basis, and because of spatial and time dependent variability of water characteristics, on-line monitoring will be an advantage. Current techniques for measuring water quality involve in-situ measurements and/or the collection of water samples for subsequent laboratory analyses. While these technologies provide accurate measurements for a point in time and space, they are expensive, and do not provide either the spatial or temporal view of water quality needed for monitoring, assessing, or managing water quality for an individual water body or for multiple water bodies across the landscape (Ritchie & Cooper, 2001).

Knowledge of mass concentration of suspended solids (SSC) and Turbidity levels are necessary to understand and know prior to entry to any WTW. The primary traditional measurement technique has been to take periodic water samples for later analysis. This method may be adequate for some applications but has many limitations because of the changeable character of suspended materials. Even collecting frequent water samples cannot accurately define a time series of suspended material that is often highly (spatially and temporally) variable and is modified by tidal currents, water depth, and wind effects (Gartner, 2004).

The automation of WTW systems is not as developed as other process industries. This is thought to be due to the harsh environment in which sensors have to be located. The lack of sensors suitable for on-line real-time monitoring and/or control is often reported to be due to sensors inconsistency and decay with time, while more sensitive and standardised laboratory-based methods/techniques are time consuming, and require sample collection and retrospective analysis (Bourgeois et al., 2001).

Monitoring systems also helps the operator in decision making when performing supervisory control tasks, rather than being used to fully automate processes. The diagnostics system, built within the monitoring system, uses sensor readings to assess the state of the system to detect abnormalities is raw water and should help to identify a route cause and remedy to the abnormalities detected. Where the sensors are not significantly more reliable than the systems being monitored, the indication of an abnormal state may be the result of a sensor failure rather than a system failure. Failure to identify the source of the indication of an "abnormal state" and take appropriate corrective action could result in expensive and unnecessary system shutdowns or maintenance (Alag et. al., 2001).

This paper will look at two case studies and discuss the sensors measurements taken and ways of validating these measurements and other new techniques that could help the operators improve on the WTW's processes. The decision to look at case studies was made in order to find out how the water industry is coping with the flux of data that is being generated by sensors and to see what sort of improvements and/or recommendations can be made based on these studies (Dolgonosov & Korchagin, 2005).

2. LITERATURE REVIEW

One of the pioneering control systems for WTW was installed by East Worcestershire Water Board in the late 1960s, now absorbed into Severn Trent Water, and was described in the IEE evening paper at Savoy Place in 1973. As well as describing the telemetry system, the paper presented work being done with the Department of Mathematics at Cambridge University to calculate optimal routing and pump control strategies using an on-line model to achieve efficiency savings. The principal objectives were to optimise the decision making process and to remove the human operator from the loop.

The early control systems were designed around telemetry systems, which have been progressively extended and amalgamated to form very large, highly complex distributed monitoring systems with several thousand remote telemetry units (RTU) and a wide array of different protocols and interfaces. Top-end master stations have generally been replaced on a five-to-eight year cycle due to the dependency on software, database and computer technology.

The drive for greater operational efficiency came with privatisation in 1990. An unprecedented level of capital investment was required in new water and wastewater treatment processes to meet higher water quality and river quality standards set by the European Union. The investment in new treatment plants was accompanied by high level of investment in ICA (Instrumentation Control and Automation), predominantly based on PLC technology.

A recent survey (National Instruments, 2003) of test and measurement engineers revealed that nearly 20% of the total cost of most data acquisition applications is spent on hardware/sensor set up and configuration (fig 1). The estimated worldwide sensor market is expected to exceed $40bn per year by 2008 (Market Report, 1999).

Consequently, this would indicate that sensing technology (hardware) and application complexity dictate the cost of an overall measurement - the more advanced the technology and intricate the application, the more expensive the sensor and resulting measurement. However, with the introduction of the Sensors "Plug & Play" (based on the IEEE 1451.4 standard), as means of standardising the sensors design industry, this has worn out this idea by providing an inexpensive method for improving analogue sensor accuracy and usability, while reducing overall measurement costs.

This can be done with smart sensors contain a Transducer Electronic Datasheet (TEDS), which can be queried by a data acquisition system to obtain the sensor's unique configuration parameters without manual interference from the system user. Once the smart sensor is connected, configuration information is electronically transferred to the data acquisition system and the sensor is automatically set up. In addition to reducing set-up and configuration time, sensor measurement accuracy also improves. But the question remains, to what extent the water industry is catching up with their infrastructure and asset developments using smart sensors

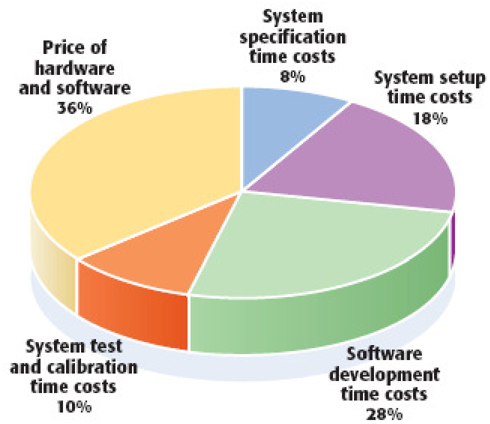


Figure 11 The total cost of data acquisition

Even with the introduction of "Plug & Play" technology, there are still a number of factors that make the process of sensor data validation and sensor failure detection difficult. From the process control point of view the plant and process supervision has to fulfil three major tasks (Enste & Uecker, 2000):

1. Validation of process information, which indicates the information quality and usability. Process information is usually used without any validation procedure and process control functions are very sensitive to faulty or doubtful process information. Therefore, one wrong measurement or a set of inconsistent measurements can disturb the process and result in a loss of productivity.
2. Supervision of process progress, which ensures the desired behaviour of process values that are relevant to product quality.
3. Functional reserve, which detects in advance unexpected process disturbances in the future.

In order for the plant and process supervision to be able to achieve the three tasks, the sensor network must be trustworthy and dependable. But factors that can be considered as major obstacles to realising this are, firstly, sensor failures that are masked by normal system manoeuvres or deviations. Subtle sensor failures such as drift are particularly difficult to detect. Secondly, the imperfect nature of the sensors adds noise to the sensors readings (Alag et al., 2001).

Key features such as the cost of ownership, ease of use, placement of the sensor, as well as the sensor response time are major factors influencing the water industry acceptance of new and improved sensors technology. Other technical aspects such as the principle of measurement, reliability, accuracy and detection limits will also dictate whether or not the technology will be accepted and promoted as a standard. Therefore, it is clear that both the performance characteristics (range, linearity, accuracy, response time, limit of detection, etc.) as well as the fundamental properties of the sensors (single or multi-parameter, need for external sampling and filtration, intrusive/non-invasive) are of major importance when looking at suitable and new methodologies.

3. KEY PARAMETERS TO MONITOR

There are numerous physical, optical & electrochemical water parameters that must be monitored prior to entry to any WTW or after being processed by the WTW. Table 1 lists some of these key parameters.

Parameter	Description
Salinity	A measure of salt concentration, and is calculated from conductivity and chloride readings. Salinity measurement assesses the purity of drinking water, monitor salt water intrusion into fresh water marshes and groundwater aquifers
Chloride, Cl	A highly soluble and ubiquitous form of chlorine. Chloride is one of the most common ions found in natural waters. Chloride is abundant in all living cells to maintain osmotic pressure and control communication, e.g. neural transmission.
Conductivity	A measure of electrical current flow through a solution. In addition, because conductivity and ion concentration are highly correlated, conductivity measurements

	are used to calculate ion concentration in solutions. Conductivity readings determine the purity of water, to watch for sudden changes in natural water, and to determine how the water sample will react in other chemical analyses.
Colour	Coloured water can be typically caused by dissolved organic matter, which absorbs visible light. Apparent colour is due to both light absorption and light scattering. Dissolved matter exclusively causes true colour. Organic matter, which absorbs light within the 300 to 400 nm wavelengths, and fluoresces in the range of 200 to 400 nm, creates the appearance of colour in water. Typically, these materials are organic in nature and contain aromatic rings.
Flow rate	A rate at which a volume of water moves or flows across a certain cross-sectional area during a specified time, and is typically measured in cubic meter per second. Water flow is measured in order to estimate pollutant spread, to monitor groundwater flow, to measure river discharge, to manage water resources, and to evaluate the effects of flooding.
Dissolved Oxygen	A certain level of dissolved oxygen is required to maintain respiration in the animal life of natural waters.
Potential of Hydrogen (pH)	Expressed as the logarithm of the reciprocal of the hydrogen ion concentration, and is measured on a scale of 0 to 14: 7 being neutral, less than 7 acidic, and more than 7 basic. Acid & basic conditions are highly corrosive and intervene with transport of water.
Temperature	A measure of the kinetic energy of water. Water temperature varies according to season, depth, and in some cases, time of day. Temperature affects the water's ability to dissolve gases, including oxygen. The lower the temperature, the higher the solubility. Thermal pollution, the artificial warming of a body of water because of industrial discharge of cooling water. This artificially heated water decreases the amount of dissolved oxygen. Typically in Northern Europe water temperature is between 8 -15°C.
Turbidity	A measure of the content of suspended solids in water, and is also referred to as the "cloudiness" of the water. Turbidity is measured by shining a beam of light into the solution. The light scattered off of the particles suspended in the solution is then measured, and the turbidity reading is most often given in Nephelometric Turbidity Units (NTU), which is calibrated against clay standard of Formazin. Turbidity readings help to monitor dredging and construction projects, examine microscopic aquatic plant life, and to monitor surface, storm, and wastewater.
ORP (Oxidation Reduction Potential)	A measure of the difference in electrical potential between a relatively chemically inert electrode and an oxidizable electrode placed in a solution. Researchers use ORP to measure the activity and strength of oxidizers (those chemicals that accept electrons) and reducers (those that lose electrons) in order to monitor the reactivity of drinking water and groundwater. ORP potential is temperature-dependent.

Table 1: Key monitored parameters of water prior to entering Water Treatment Works

Water pollutants are both suspended sediments (turbidity) as particles and dissolved. Remote sensing applications to determine water quality are to date limited to measuring those substances or conditions that influence and change optical and/or thermal characteristics of the surface water properties (Ritchie & Cooper, 2001). Remote sensor design is made more difficult by the highly variable sizes of these potential pollutants. Suspended sediments, chlorophylls, DOM, temperature, and oil are water quality indicators that can change the spectral and thermal properties of surface waters and are most readily measured by remote non-invasive sensing techniques. Substances (i.e., nutrients, metals) that do not change the optical and/or thermal characteristics of surface waters can only be inferred by measuring surrogate properties (i.e., chlorophylls) which may have responded to an input of marker chemicals. These remote sensing techniques should improve our abilities to monitor changes in the water topography and contents (Martinez et al, 2004; Jane & Martinez, 2005).

4. CASE STUDIES

Two case studies were investigated. The sites chosen were based on two completely different types of water sources. One site, which can be referred to as the groundwater site, has its main water source from boreholes. The second site investigated was a lowland river.

This combination of ground and surface water WTW enabled us to compare the demands on the sensors used in the two most common water resources. The hypothesis was to identify whether a possible combination of sensors would make it possible to operate the processes of the WTW more efficiently or whether new alternative sensors would be necessary. Alag et al (2001) reported that by combining information from many different sources, it would then be possible to decrease the uncertainty and ambiguity inherent in processing the information from a single sensor source. A large number of sensors measuring many different variables can collectively achieve a high level of accuracy and reliability, depending on their accuracy and reliability.

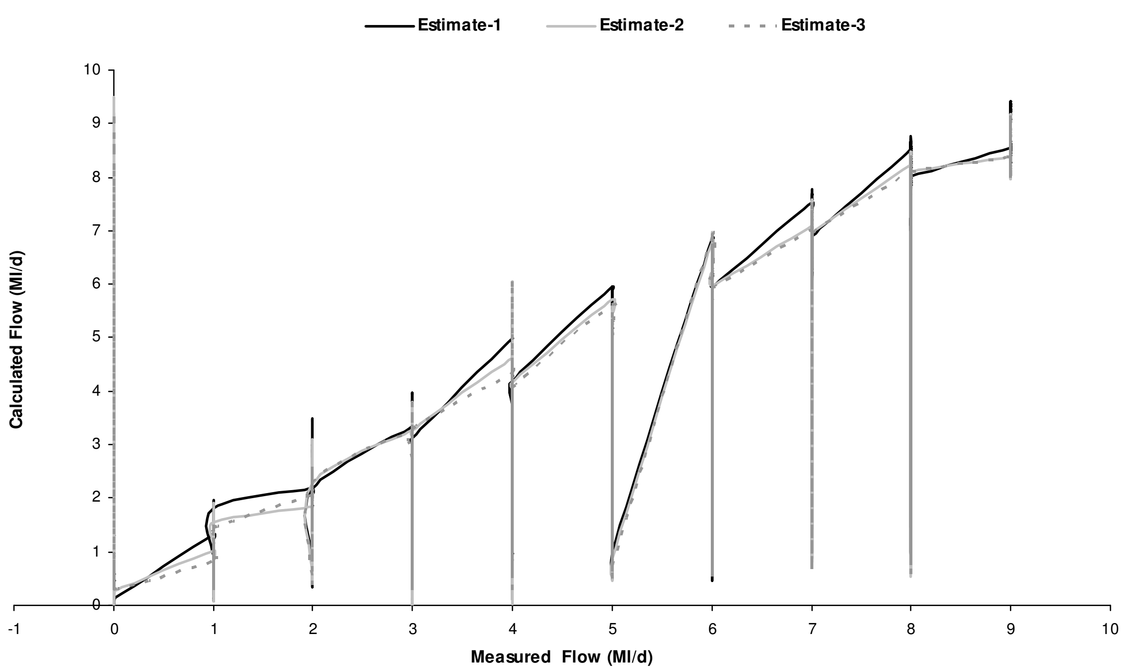
4.1 Groundwater WTW

The project was initially launched aimed at resolving problems that were hampering the continuous operation at the groundwater WTW. The problems were causing shutdown of the WTW from reporting false alarms. This needed continuous intervention by the operating manager to resolve these problems. The operators covered a large area and a number of remote sites and so the frequency of false alarms was unmanageable and the sensors switched off.

After careful examination of the data retrieved from the site in text format, it was noticed that there was no consistency in the qualitative data monitored from the different boreholes. The only two parameters that were archived were Fluoride and Phosphate dosage levels to comply with most recent requirements set by the regulators. This raised the question as to what parameters were required to comply with regulation and what was necessary for control.

Thus it was decided to try and reprocess the data measured, using the Partial-Least Squares method to determine if these two sensors could provide more general information about reliability of output. Three modelled equations were extracted for the Borehole1 Flow. Graph 1 below illustrates an example of results of correlations of flow and qualitative data for the three models, with the measured flow versus the calculated flow (for the surrogate sensors) in mega litres per day.

Partial least squares (PLS) is a method for constructing predictive models when the factors are many and highly collinear (developed in the 1960's by Herman Wold as an econometric technique). It is common in chemical engineering processes. PLS has been applied to monitoring and controlling industrial processes; a large process can easily have hundreds of controllable variables and dozens of outputs (Dijkstra, 1983; Geladi & Kowalski, 1986; Stone & Brooks, 1990).



Graph 1: Illustration of the three models for Groundwater Borehole Flow

The graphs drawn out are highly similar in response. Estimate-1 is a calculated response for Borehole1 Flow using all of the other measured readings as dependent factors, e.g Reservoir Level, Borehole Flow, etc. In other words, calculation made included the maximum amount of data independent of it origin. Estimate 2 & 3 used lesser independent data inputs and the resulting correlations seem weaker.

The three graphs showed a strong coherence and indicated the possible use of a very simple equation to predict the total outflow from the WTW and therefore cross check sensor output. It could help cut the maintenance cost and improve reliability by discarding spurious data under normal operational conditions. The sudden drops in the three graphs were the downtime of the WTW during the experimental period.

The analysis indicates that it must also be possible to improve the quality of sensor information by using this type of expert system where a strong correlation can be established.

4.2 River Water WTW

Part of the objectives of the project was to improve the operational running of the SCADA systems. The WTW is a river abstraction of around 100ML/day from a low land river (of 30-100 m^3/s)

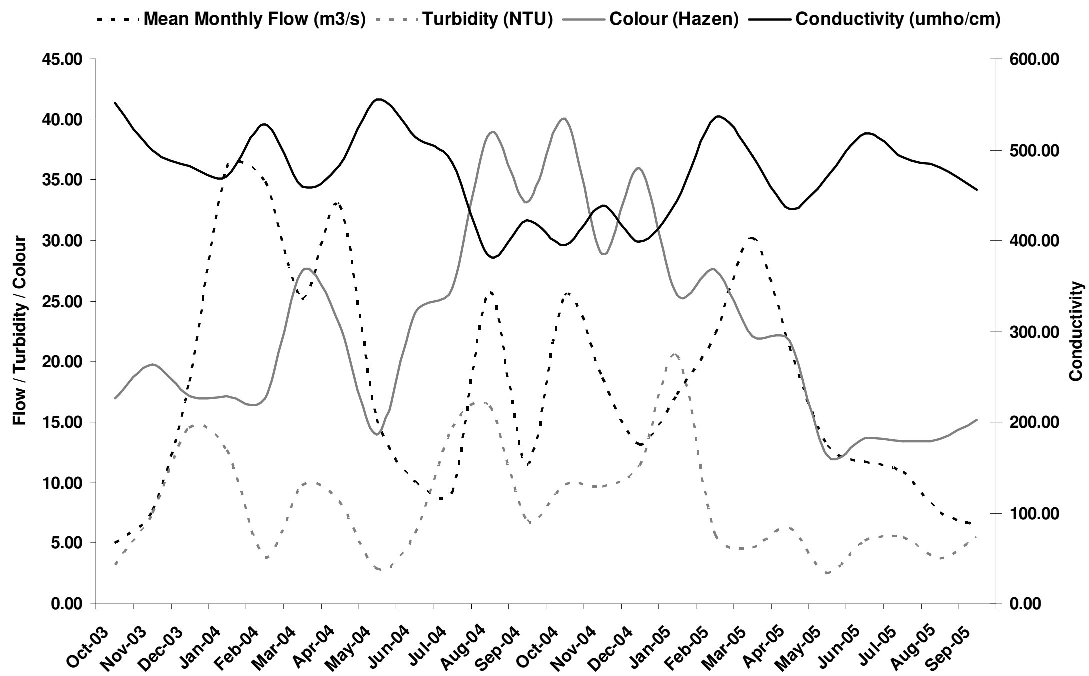
The site monitors Turbidity, Colour, pH, Conductivity, Ammonia & Temperature. These parameters were chosen on the basis of their availability and also their relevance to water quality management. Using this data, a partial least square analysis was performed to try and correlate these data together. The very positive link found was between turbidity & colour. This is understandable; when turbidity is high then other parameters will also increase since sediment load will contribute to the colouring & turbidity in the water. It was also found that the temperature was linked to conductivity readings. This is as it would be anticipated that temperature would affect solubility of contaminants and therefore conductivity.

From the analysis, it is also noticeable that there are relations between all the parameters including flow rate. This should not be surprising since an increase in flow is likely to increase the transport and erosion of materials.

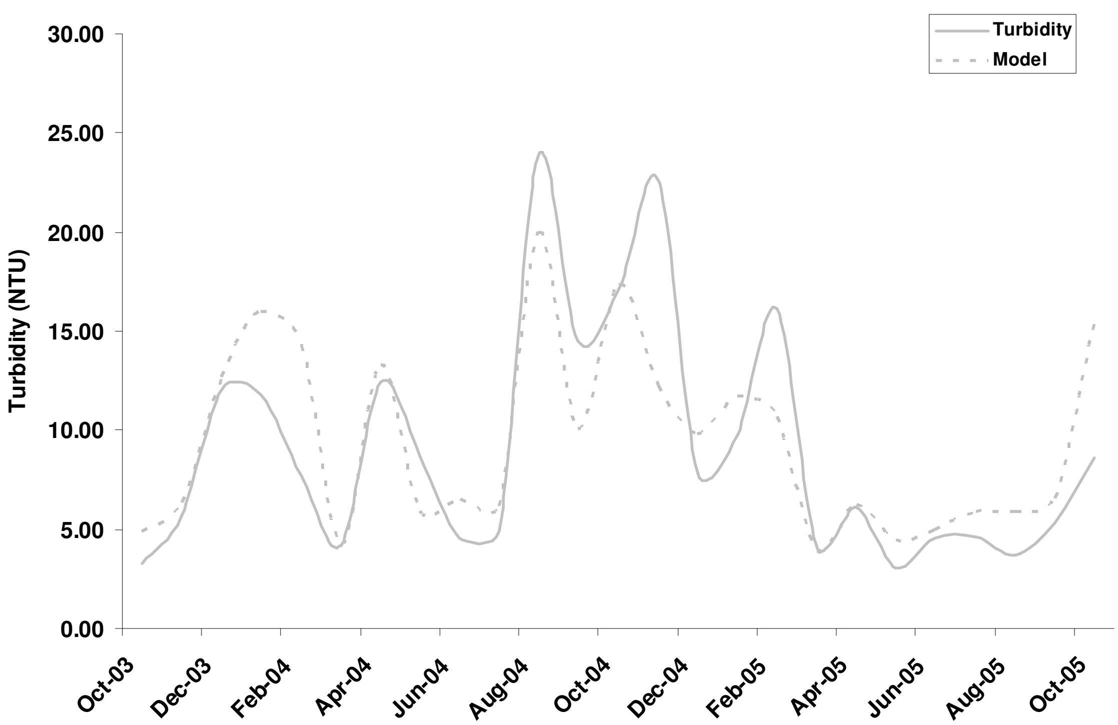
Previous literature correlating different types of sensors in drinking water were not found and are clearly not really used fully. The interrelationship between sensors should be established to crosscheck the other sensor performance. Data gathered also does not show how it benefits SCADA or how it links with it. Finally, data from sensors are often disregarded; since operators intuitively find data from grab wet samples more reliable.

During the research of the flow data available, it emerged that there were two river Derwents whose flow rates were continuously monitored, St Mary's Bridge in Derby, which is considered to be the closest monitoring station to the site being investigated, and a Yorkshire Derwent flow rate at Buttercrambe in York, which is approximately 95 miles to the North-East of Derby. The statistical correlation between the two flow rates was very strong, 0.84. This has implied that there might be less local variations in rainfall than might have been thought. There are also a lot of river obstructions to both rivers which should dampen this correlation.

Using the Least-Squares technique and using the Flow rate of the River Derwent at St Mary's Bridge it was possible to predict the Turbidity trend. Graph 2 shows the Flow, Turbidity, Conductivity & Colour levels. Graph 3 illustrates the finding in deducing the trend shape of the Turbidity level from just the flow rate at St Mary's Bridge.

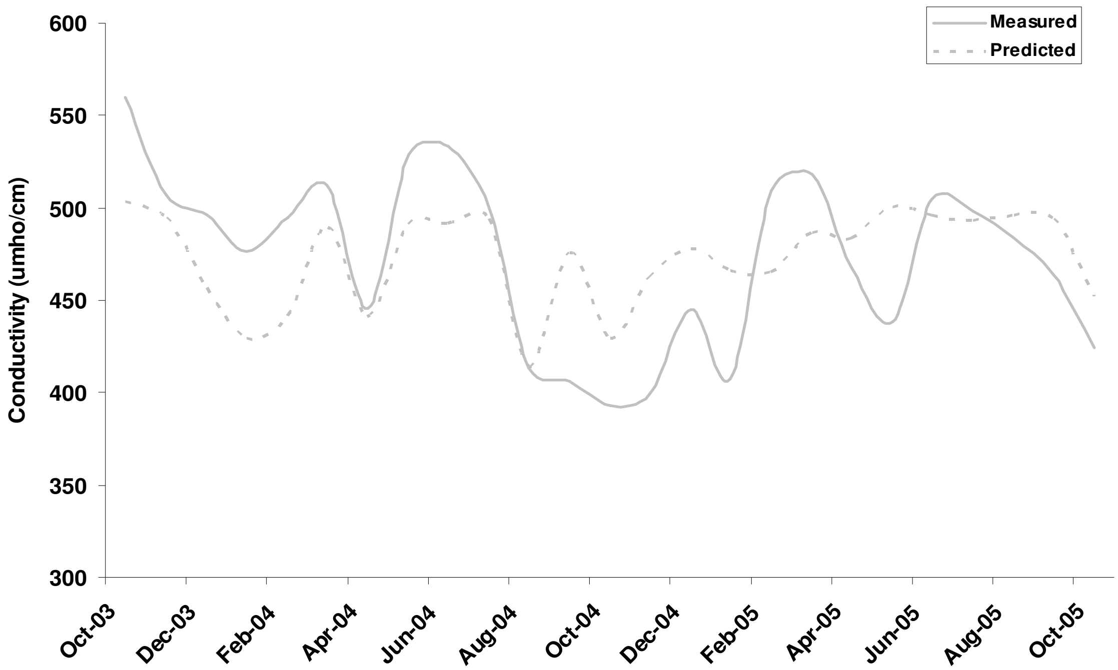


Graph 2: Comparison between Flow, Turbidity, Conductivity & Colour levels measured.



Graph 3: Measured Turbidity levels versus predicted levels extracted from the flow rate at St Mary's Bridge.

Following the derivation of the Turbidity level trend, the data is fed back into the Least-Squares technique again together with the flow data and the predicted Conductivity trend was calculated, graph 4. In short, Flow rate with Turbidity level will give Conductivity value.



Graph 4: graph of the predicted Conductivity value against the measured value.

The same process was repeated again by feeding the Conductivity value back into the Least-Squares method to derive the Colour level. In other words, Flow Rate, together with Turbidity Level & Conductivity Value can give Colour Level.

It would be assumed that the calculated interpretations of the data for Turbidity, Conductivity & Colour could be more realistic of the real values in the River Derwent. But it should be stressed that these relationships should be developed as moving correlations over several years and they will not be appropriate coefficients for other WTW. These relationships were derived using the data supplied from one site, which means that it's more related to this site rather than being generalised. There is at least the prospect of using data from a much wider field of similar rainfalls but with different weightings in the model.

One possible explanation on why the predicted values differ from the measured ones is because the measured values were lab measurements, rather than online measurements. Many operators have more confidence in the representative value of grab samples. It is also likely that the correlation coefficient would vary according to flow, season (e.g. leaf fall), antecedent dry period and impoundment and other river engineering.

5. LABORATORY ANALYSIS OF SENSORS: ULTRASONIC DOPPLER

Further investigative study has led the research to look at the possibility of having at least one type of non-intrusive sensor where the previous modelling technique can rely on as an independent-reliable source of data. As the previous sections have shown, it would be logical to look at flow measuring sensors as the possible independent, reliable & non-intrusive sensor.

New innovative remote sensors are being developed and a review of the different types of non-invasive techniques was made. These included Acoustic, Fluorescence, Laser & X-Ray techniques were all examined. A comparison was made and concluded that Acoustic-Doppler techniques would suit the objective most. This was based on the fundamental basis of operation, since the technique uses the particles suspended in water to determine the flow velocity of the water.

Use of in situ optical instruments such as optical back scatter (OBS) sensors (Downing et al., 1981; Downing, 1983) and transmissometers with the capability of producing time series of high-frequency measurements of suspended material help address the variable nature of SSC & Turbidity (Gartner, 2004). However, calibration of these instruments is complicated because the response function of the OBS sensor depends on grain size and is nonlinear with concentration (Downing, 1996).

In addition, optical sensors are extremely sensitive to biological fouling problems (Hamilton et al., 1998; Bourgeois et al., 2001). Often, only a few days of data are usable from records collected in highly active reservoirs/lakes unless optical sensors are frequently cleaned.

Alternatively, acoustic sensors that are far less susceptible to effects of biological fouling (Downing, 1996; Gartner, 2004) have shown promise for determining reliable estimates of suspended solids (e.g., Thorne et al., 1991; Hay and Sheng, 1992; Osborne et al., 1994).

Thevenot and Kraus (1993) and Hamilton et al. (1998) provide extensive comparisons of the strengths and weaknesses of optical and acoustic methods for monitoring suspended materials. While many early studies primarily dealt with suspensions of sand-size material, some later studies (e.g., Hamilton et al., 1998; Jay et al., 1999) examine the potential for determining suspended cohesive sediment concentration.

As use of acoustic Doppler current profilers (ADCPs) has become more widespread, so have attempts to characterize suspended material from acoustic backscatter intensity measurements made by those acoustic instruments used to measure water velocity (e.g., Thevenot et al., 1992; Reichel and Nachtnebel, 1994). In addition to being less susceptible to biological fouling, commercially available ADCPs may provide non-intrusive estimates of SSC & Turbidity profiles concurrent with measurements of velocity profiles using the same instrument. However, the process of converting backscatter intensity to mass concentration is not straightforward. Among other things, complex acoustic transmission losses from beam spreading and attenuation must be accounted for correctly. They depend on multiple factors including environmental characteristics such as suspended material and the salinity, temperature, and pressure (of the water), and instrument characteristics such as power, transducer size, and frequency. While most early studies utilizing acoustic backscatter to estimate suspended solids typically include beam spreading and (water) absorption in the calculation of acoustic transmission losses, they often omit corrections for attenuation from suspended particles and non-spherical spreading in the transducer near field. More recent studies have begun to include these factors into consideration.

Jay et al. (1999) apply a correction function for improved calculation of beam spreading losses in the ADCP transducer near field to account for the complex acoustic beam pattern, and Holdaway et al. (1999) account for sediment attenuation in their evaluation of ADCPs to estimate suspended sediment concentration.

5.1 Experiment & Methods

The first experiment, were simple as a controlled amount of Bentonite clay powder were added to a water container (approximately 20l of tap water). The added amounts were controlled by adding the right dosage equivalent to 5 mg/l at each stage. The turbidity was increased from approximately 0 NTU (clear tap water) to approximately 30 NTU, which is double the average river turbidity reading measured, fig 3. The ultrasonic Doppler sensor and an electric stirrer were used to keep the clay in suspension.

The Ultrasonic Doppler used was not a non-invasive model (Nortek Vectrino Velocimeter). Unfortunately project resources were not available to purchase one especially.

Velocimeters use acoustic sensing techniques to measure flow in a remote sampling volume. The measured flow is practically undisturbed by the presence of the probe. Data are available at an output rate of 25 Hz. The 3-D velocity range is 2.5 m/s, and the velocity output has no zero-offset.

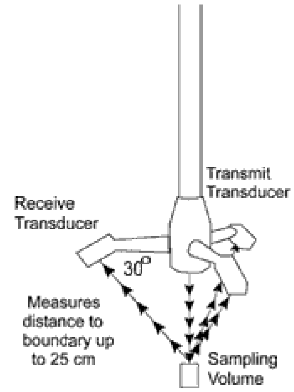
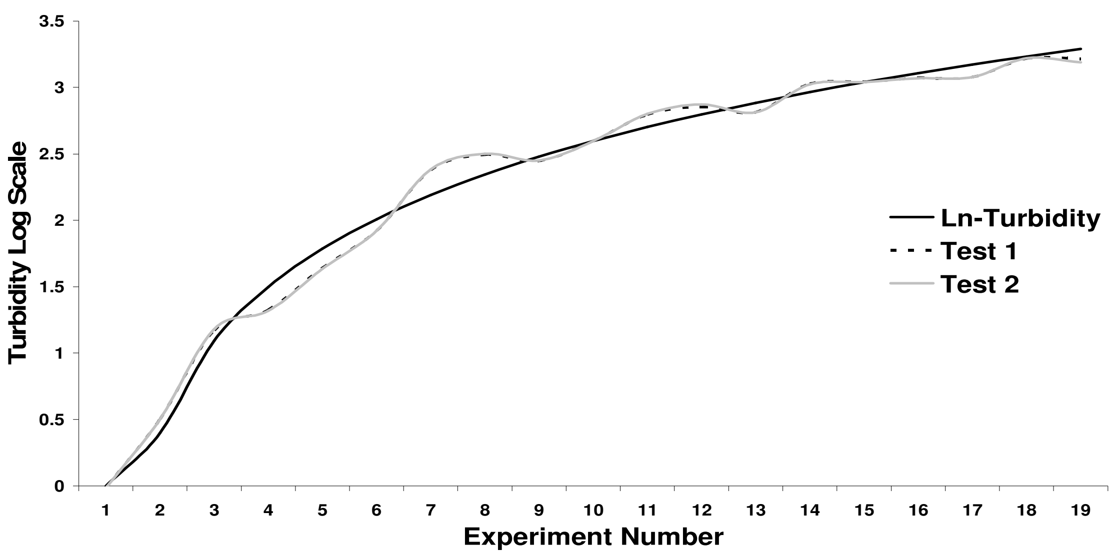


Figure 2: The acoustic sensor used has one transmit transducer and three receive transducers. The sampling volume is located away from the sensor to provide undisturbed measurements. Doppler velocity is derived from signals scattered by small particles. In natural bodies of water (streams, lakes, rivers, oceans, etc.) the correlation of particles is sufficient for proper operation. In model tanks with running water (flumes, open channels, closed pipes, etc.) microscopic bubbles in the water column tend to act as natural seeding. In very clean, quiescent water (ship models, tow tanks, and some wave flumes), seeding materials must be added to concentrations of approximately 10 mg/l.

The stirrer position was fixed during the entire experimental setup close to the bottom of the tank as much as possible; so as to ensure that there is no precipitation of the particles at the bottom. The stirrer speed was also fixed in order to ensure that data differences can be attributed to a change in turbidity levels only. Another critical factor during the setup of the probe was to ensure that the side benchmarked as the X-direction of the probe should always be fixed in the direction of the flow. This ensures similar bias from any differences in the sensors.

5.2 Results & Discussion



Graph 5: comparison between the measured turbidity levels versus readings taken by the Ultrasonic probe after applying PLS to the readings

Graph 5 demonstrates that the probe was capable of detecting the amount of turbidity particulates flowing by it. But the major setback was even after repeated trials, the test graphs could not be improved. This still leaves the possibility that the sensitivity was too low to identify minuet changes in turbidity. In other words, the probe was not designed to investigate this type of experiment; it was rather designed to measure the flow velocity of the particles.

Following the success of the basic trials, focus the next experiments are to initiate test of sensors placed by the WTW at the intakes. The probe was placed in a laboratory flume in which it was possible to vary velocity, colour and turbidity.

6. CONCLUSION

In conclusion, it is sufficient to say that there is a lack of confidence in sensors and data collected. In one of the case studies, the sensors were actually a major burden on resources (false alarms) rather than being the dependent analyser. Correlations were established between the different types of sensors, i.e. chemical and physical (flow) which does offer the possibility of better models with inbuilt deletion of spurious data. This would eventually lead to enhanced control of the water treatment works. Correlations were also found between river flows in two different rivers 150km apart indicating the possibility of value in a National network of sensors. A modern non invasive probe technology, in this case Acoustic Doppler, was investigated to provide data which could be interrogated differently to achieve qualitative as well quantitative data. Recommended further work should include the possible use of Object Oriented Modelling or Artificial Neural Networks in sensors correlations.

7. ACKNOWLEDGEMENT

The authors would like to thank the Engineering & Physical Sciences Research Council (EPSRC) in the UK and Myriad Vision Ltd for their support and funding of this project. A special appreciation is also in place for Severn Trent Plc for the support given in accessing and gathering of the data for the two sites.

8. REFERENCES

Alag S., Agogino A.M., Morjaria M. (2001). A methodology for intelligent sensor measurement, validation, fusion and fault detection for equipment monitoring and diagnostics. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 15, 307–320.

Bourgeois, W., Burgess, J. E., Stuetz, R.M. (2001). Review: On-line monitoring of wastewater quality: a review. Journal of Chemical Technology & Biotechnology 76, pp. 337 – 348.

Council Directive 98/83/EC on the quality of water intended for human consumption. (1998). Official Journal L 330, 003 -0054. http://eur-ex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31998L0083:EN:HTML

Dijkstra, T. (1983) Some comments on maximum likelihood and partial least squares methods. Journal of Econometrics, 22, 67-90.

Dolgonosov, B. M., Korchagin, K. A. (2005). Statistical assessment of relationships between water flow in a river and water turbidity in water intakes. Water Resources, vol. 32, no. 2, pp. 175-182.

Downing, J.P. (1983). An optical instrument for monitoring suspended particles in ocean and laboratory. OCEANS 1983, pp. 199 – 202.

Downing, J.P. (1996). Suspended sediment and turbidity measurements in streams: What they do and do not mean. Automatic Water Quality Monitoring Workshop, Richmond, B.C.

Downing, J.P., Sternberg, R.W., Lister, C.R.B. (1981). New instrumentation for the Investigation of sediment suspension processes in the shallow marine environment. Marine Geology 42, pp. 19 – 34.

Enste, U., Uecker, F. (2000). Use of supervisory information in process control. Special Feature Intelligent Sensors, IEE Computing & Control Engineering Journal, pp. 234 – 241.

Environment Agency Website. (2004). http:// www.environment-agency.gov.uk/.

Gartner J.W. (2004). Estimating suspended solids concentration from backscatter intensity measured by acoustic Doppler current profiler in San Francisco Bay, California. Marine Geology 211, pp. 169-187.

Geladi, P, Kowalski, B. (1986) Partial least squares regression: A tutorial. Analytical Chimica Acta, 185, pp. 1-17.

Hamilton, L.J., Shi, Z., Zhang, S.Y. (1998). Acoustic backscatter measurements of estuarine suspended cohesive sediment concentration profiles. Journal of Coastal Research 14, pp. 1213 – 1224.

Hay, A.E., Sheng, Jinyu. (1992). Vertical profiles of suspended sand concentration and size from multi-frequency. Journal of Geophysical Research 97, 15, pp. 661 – 15,677.

Holdaway, G.P., Thorne, P.D., Flatt, David, Jones, S.E., Prandle, David. (1999). Comparison between ADCP and transmissometer measurements of suspended sediment concentration. Continental Shelf Research 19, pp. 421 – 441.

Hughes, G. (2004). Standardising Control for the Water Industry. Computing and Control Engineering Journal, vol 15, issue 5.

Jane, K.H., and Martinez, K (2005). Environmental sensor networks, Eos, Volume 86(16) page 162.

Jay, D.A., Orton, Philip, Kay, D.J., Fain, Annika, Baptisa, A.M. (1999). Acoustic determination of sediment concentrations, settling velocities, horizontal transports and vertical fluxes in estuaries. Proceedings: 6th Working Conference on Current Measurement, pp. 258– 263.

Market Report. (1999). Sensor Markets 2008. By Intechno Consulting.

Martinez, K, Hart, J. K., Ong, R. (2004). Environmental Sensor Networks. Computer, vol. 37, no. 8, pp. 50-56.

National Instruments survey of data acquisition engineers and technicians. (2003).

Osborne, P.D., Vincent, C.E., Greenwood, B. (1994). Measurement of suspended sand concentrations in the nearshore: field intercomparison of optical and acoustic backscatter sensors. Continental Shelf Research 14, pp. 159 – 174.

Reichel, G., Nachtnebel, H.P. (1994). Suspended sediment monitoring in a fluvial environment: advantages and limitations applying an acoustic Doppler current profiler. Water Research 28, pp. 751–761.

Ritchie, J.C. and Cooper, C.M. (2001). Remote sensing for determining water quality: Application to TMDL, pp. 367-375. In: TMDL Science Issues Conference, Water Environment Federation, Alexandria, VA, USA.

Stone, M. and Brooks, R. (1990). Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. Journal of the Royal Statistical Society, Series B, 52(2), 237-269.

Thevenot, M.M., Kraus, N.C. (1993). Comparison of acoustical and optical measurements of suspended material in the Chesapeake Estuary. Journal of Marine Environmental Engineering 1, 65– 79. Gordon and Breach Science Publishers.

Thevenot, M.M., Prickett, T.L., Kraus, N.C. (1992). Tylers Beach, Virginia, dredged material plume monitoring project 27 September to 4 October 1991. Dredging Research Program Technical Report DRP-92-7, US Army Corps of Engineers, pp. 204.

Thorne, P.D., Vincent, C.E., Harcastle, P.J., Rehman, S., Pearson, N. (1991). Measuring suspended sediment concentrations using acoustic backscatter devices. Marine Geology 98, pp. 7 – 16.

APPENDIX B (PAPER 2)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2007. Case studies on the need for monitoring of water quality in the UK. Eleventh International Water Technology Conference, IWTC11 2007 Sharm El-Sheikh, Egypt, pp. 389-994.

CASE STUDIES ON THE NEED FOR MONITORING OF WATER QUALITY IN THE UK

Ahmed Moustafa, Ashraf El-Hamalawi[1] and Andrew Wheatley.

Civil & Building Engineering Department, Loughborough University, Loughborough, LE11 3TU, England (UK)

ABSTRACT

The advanced monitoring of water quality and performing a real-time hazard analysis prior to entering Water Treatment Works (WTW) is very much a necessity nowadays in order to give warning of any contamination and avoid downtime of the WTW. Two case studies from the UK were examined, one being a groundwater WTW and the other a river WTW. Measured data from both sites were analysed. The results showed that no good correlation existed between the controlling parameters measured at the river WTW, but showed a good correlation for the groundwater WTW. The case studies highlighted the need for a new non-invasive type of measurement and for water companies to invest in their information technology infrastructure.

Keywords

Groundwater, Water Quality Analysis, Water Monitoring, Sensors, Turbidity.

1. INTRODUCTION

The objective of a water treatment works is to produce an adequate and continuous supply of water that is chemically, bacteriologically and aesthetically pleasing. More specifically, water treatment works must produce water that is (American Waterworks Association [1], EU Drink Water Directive [2]):

a) Palatable with no unpleasant taste,
b) Safe; does not contain pathogens or chemicals harmful to the consumer,
c) Clear; free from suspended solids and turbidity,
d) Colourless and odourless; aesthetic to drink,
e) Reasonably soft; allows consumers to wash clothes, dishes, themselves, without use of excessive quantities of detergents or soap,
f) Non-corrosive; to protect pipework and prevent leaching of metals from tanks or pipes,
g) Low organic content (high organic content results in unwanted biological growth in pipes and storage tanks that often affects quality).

A water treatment works must be able to produce a finished product of consistently high quality regardless of how great the demand might be. Like waste-water treatment, water treatment consists of a range of unit processes, usually used in series and this provides some design and operational flexibility to achieve this. The treatment required by water prior to being delivered to consumers will depend upon its initial quality, which is normally related to its source. In other words, the cleaner the raw water, the fewer treatment steps that are required, and hence the overall cost of water is less (Hughes, [3]).

The most expensive operations in conventional treatment are sedimentation and filtration, while water softening can also be very expensive. Groundwater is generally much cleaner than surface water and thus does not require the same degree of treatment, apart from aeration and disinfection before supply. Naturally occurring substances that may need to be reduced or removed in groundwater include iron, hardness (if > 300 mgl^{-1} as $CaCO_3$) and carbon dioxide. Compounds originating from urban activity are becoming increasingly common in groundwater and those requiring treatment include nitrates, pathogens and trace organics such as pesticides. Surface water requires more complex treatment due to its complex nature, although the quality of surface water can be very high, for example upland reservoir (Geldreich, [4]; Reasoner, [5]).

Monitoring and assessing the quality of waters in streams, reservoirs, lakes, and estuaries is critical to improve water quality. Current techniques for measuring water quality involve *insitu* measurements and/or the collection of water samples for subsequent laboratory analyses. While these technologies provide accurate measurements for a point in time and space, they are expensive, and do not provide either the spatial or temporal view of water quality needed for monitoring, assessing, or managing water quality for an individual water body or for multiple water bodies across the landscape (Ritchie & Cooper, [6]).

In 2004, 375 sites were monitored for compliance with the Surface Water Abstraction Directive (75/440/EEC) in England and Wales. Of these, 155 sites failed to comply with the Directive. However, over 90% of these 'failures' were due to insufficient sampling. These sampling shortfalls occur for a number of reasons, such as abstractions not being operated at the time of sampling, problems at the laboratory, and sampling error. The quality of abstracted water generally improved since 1993. It was found that levels of coloration, nitrate and polycylic aromatic hydrocarbons (PAHs) most commonly exceeded the Directive's standards in 2004 (Environment Agency, [7]).

Major water pollutants are suspended sediments (turbidity), pathogens, nutrients, metals, dissolved organic matter (DOM), pesticides, chlorophylls (algae, plants), temperature, and oils. Remote sensing applications to determine water quality are limited to measuring those substances or conditions that influence and change optical and/or thermal characteristics of the surface water properties (Ritchie & Cooper, [6]). Suspended sediments, chlorophylls, DOM, temperature, and oil are water quality indicators that can change the spectral and thermal properties of surface waters and are most readily measured by remote non-invasive sensing techniques. Substances (i.e., nutrients, metals) that do not change the optical and/or thermal characteristics of surface waters can only be inferred by measuring surrogate properties (i.e., chlorophylls) which may have responded to an input of chemicals. These remote sensing techniques should improve our abilities to monitor changes in the water topography and contents (Martinez et al, [8], Jane et al [9]).

The decision to look at case studies was made in order to find out how the water industry is coping with the flux of data that is being generated by sensors and to see what sort of improvements and/or recommendations can be made based on these studies (Dolgonosov & Korchagin, [10]).

2. CASE STUDIES

Two case studies were investigated. The sites chosen were based on two completely different types of water sources. One site, which can be referred to as the groundwater site, has its main water source from boreholes. The second site investigated was a lowland river of 30-100m^3/s flow rate.

This combination of ground and surface water WTW would enable us to set the benchmark on the reliability of sensors used in both cases and also examine hypothesis concluded from each site if applicable to the other. The hypotheses are based on the possible combination of sensors that would make it possible to operate the processes of the WTW or find alternative types of sensors.

2.1 Groundwater WTW

The project was initially aimed at resolving problems that were hampering the continuous operation at the groundwater WTW. The problems were caused by the shutdown of the WTW and the reporting of false alarms. This needed continuous intervention by the operating company to resolve these problems.

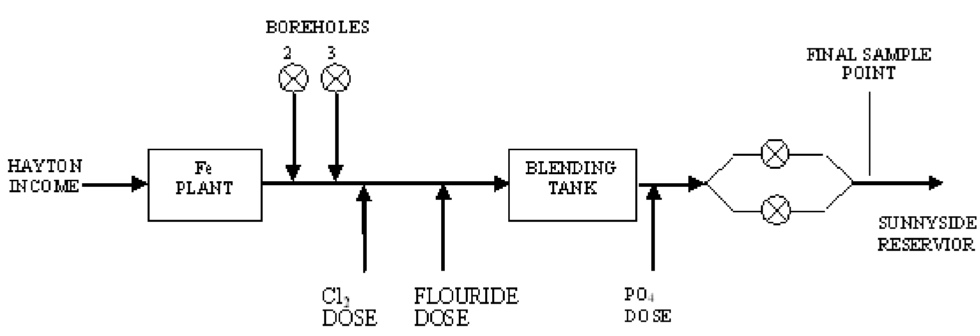


Figure 1: Groundwater WTW process diagram

2.1.1 Data Stored

The WTW operating system is configured to store several types of data that go through it and all causes of alarms raised. Some of the data stored were Fluoride & Phosphate levels, Borehole Flow 1, Borehole Flow 2, Incoming Flow, No.2 Borehole Flow, Total Borehole Flow, Booster Pump No.1 Delivery Pressure, and Booster Pumps Total Outflow. Other readings were also stored by the system indicating the cause of alarms or failures, e.g. power failure, maintenance worker on site, intruder alert, etc

After careful examination of the data retrieved from the site in text format, it was noticed that there was not much qualitative monitored data from the boreholes. The only two parameters that were stored were Fluoride and Phosphate dosage levels. This raises the question of what can be considered as minimum in terms of water monitoring WTW processes. It also highlights the issues involving legal requirements by the operating company to ensure operational standards are being met.

Table 1 illustrates the correlation coefficients found between the parameters. The correlation coefficient is a measure of the extent to which two measurement variables "vary together". The correlation coefficient is scaled so that its value is independent of the units in which the two measurement variables are expressed. The value of any correlation coefficient must be between -1 and +1 inclusive. If positive correlation is found, it implies that large values of one variable tend to be associated with large values of the other; if negative correlation, then small values of one variable tends to be associated with large values of the other; and if correlation is near zero, then value of both variables tend to be unrelated.

	Borehole Flow 1	*Borehole Flow 2*	*Incoming Flow*	*No.2 Borehole Flow*	*Total Borehole Flow*	*Booster Pump No.1 Delivery Pressure*	*Fluoride Level*	*Phosphate Level*	*Booster Pumps Total Outflow*
Borehole Flow 1	1.00								
Borehole Flow 2	0.47	1.00							
Incoming Flow	0.46	**0.98**	1.00						
No.2 Borehole Flow	**0.85**	0.39	0.39	1.00					
Total Borehole Flow	**0.93**	0.50	0.50	**0.86**	1.00				
Booster Pump No.1 Delivery Pressure	0.38	0.51	0.51	0.33	0.40	1.00			
Fluoride Level	0.12	0.14	0.14	0.05	0.12	0.13	1.00		
Phosphate Level	-0.18	0.00	-0.01	-0.13	-0.03	-0.03	-0.10	1.00	
Booster Pumps Total Outflow	**0.74**	**0.83**	**0.84**	0.65	**0.77**	0.56	0.15	-0.04	1.00

Table 1: Correlation coefficients between measured parameters at the surface water WTW

From Table 1, a number of operational remarks can be made:

1. The Borehole Flow 2 with Incoming Flow are very much correlated, which suggests that they measure the same flow.
2. No. 2 Borehole Flow with Borehole Flow 1 and the Total Borehole Flow indicate that the major contributor to the groundwater WTW is actually Borehole 2. This could help the operator in determining the optimum location of water quality monitoring sensors that would yield the best possible operational gain.
3. Booster Pumps Total Outflow correlations indicate that its operational mode is very much dependent on the Incoming Flow more than on its outgoing one.
4. The flow rate is not connected to the Fluoride or the Phosphate levels. This could simply be caused by failure of the dosage sensor or an outside controller that adjusts the dosage levels.
5. Borehole Flow 1 is very much dependent on three main contributors to its readings, No 2 Borehole Flow, Total Borehole Flow and Booster Pumps Total Outflow.

Following on from the last operational remark made, it was decided to try and reprocess the data measured, using the Partial-Least Squares method. Three modelled equations were extracted for the Borehole Flow 1. Figure 2 below illustrates the results of the three models, with the measured flow versus the calculated flow in mega litres per day.

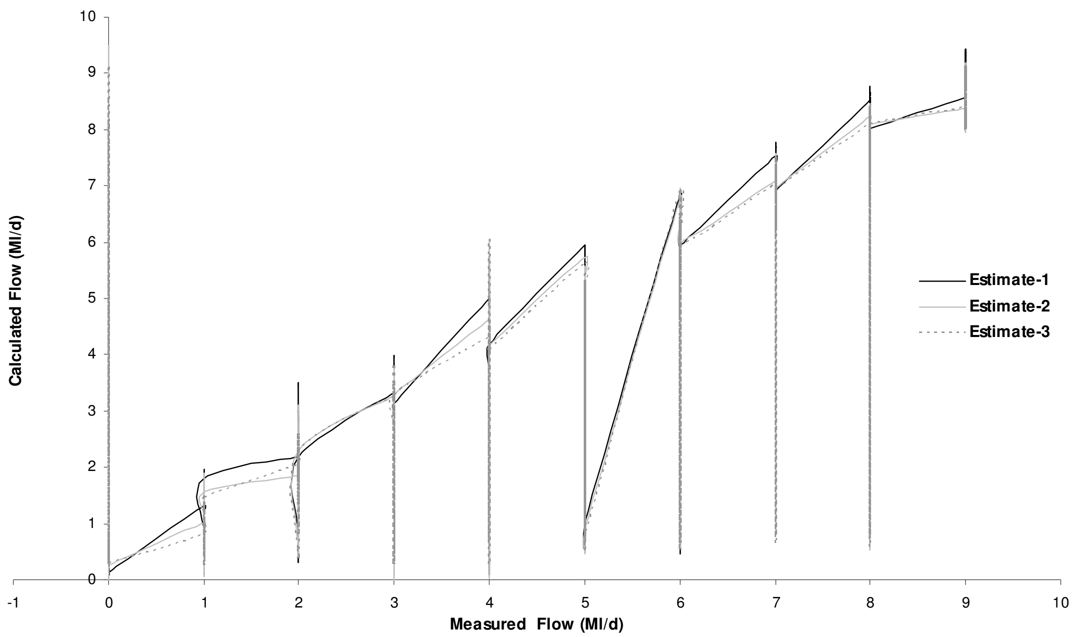


Figure 2: Illustration of the three models for Groundwater Borehole Flow

Partial least squares (PLS) is a method for constructing predictive models when the factors are many and highly collinear. PLS was developed in the 1960's by Herman Wold as an econometric technique, but some of its most avid proponents are chemical engineers and chemometricians. PLS has been applied to monitoring and controlling industrial processes; a large process can easily have hundreds of controllable variables and dozens of outputs (Dijkstra, [11]; Geladi & Kowalski, [12]; Stone & Brooks, [13]).

The graphs drawn out are highly similar in response. Estimate-1 is a calculated response for Borehole Flow 1 using all of the measured readings in table 1 as dependent factors. Estimate-2 is calculated using only the highly positive correlations associated with Borehole Flow 1, which are Borehole Flow 2, No. 2 Borehole Flow, Total Borehole Flow and Booster Pumps Total Outflow. Estimate-3 was drawn out from only No. 2 Borehole Flow, Total Borehole Flow and Booster Pumps Total Outflow which have a correlation level of 0.75 or over.

The three graphs showed a strong coherence and indicated the possible use of computations to predict the total outflow from the WTW. It can also help cut the maintenance cost of the plant by computing the output and discarding the other flow meters under normal operational conditions. The sudden drops in the three graphs were accredited to the downtime the WTW faced during the period when the flow data was measured.

Figure 2 highlights the relevance reliability of the measured flow at the groundwater WTW and the need to further enhance the capabilities of interpreting this data to run the site and be able to, as a key parameter, audit its operations, whether raised by alarms or detection in faulty sensors. The other recorded data were the Fluoride and Phosphate dosage levels. Figure 3 shows the two measurements for comparisons.

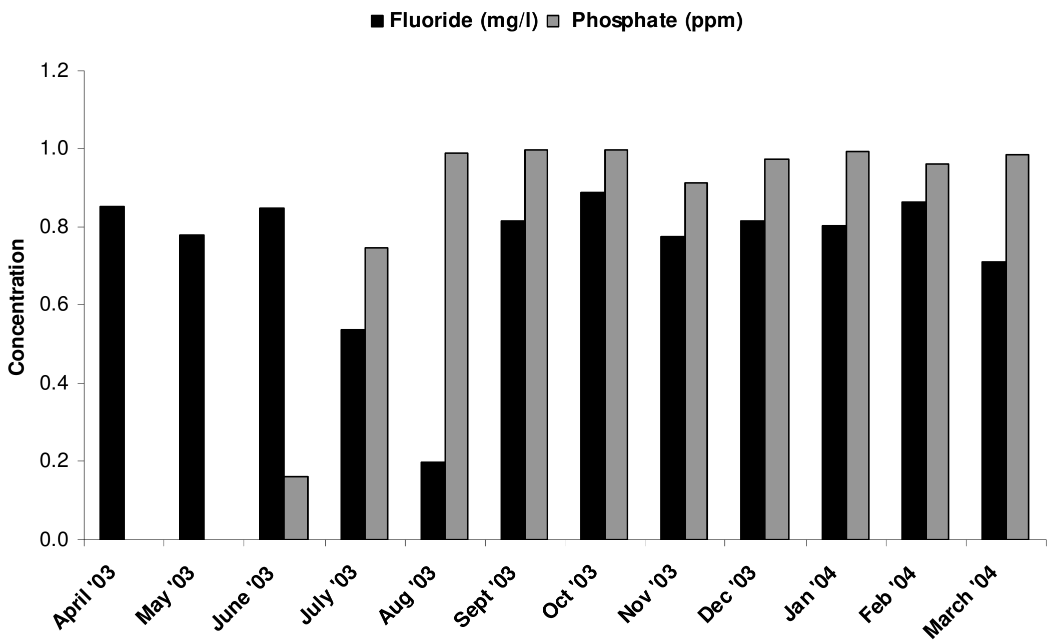


Figure 3: Fluoride and Phosphate dosage levels measured.

The measured dosage levels lead to a number of conclusions. Firstly at the start of April ’03 and May ’03, there was no evidence of Phosphate and this would in turn raise the question of when the legal maximum limit of around 1ppm was being met by the WTW and of when the equipment might have been installed. Secondly, a drop in Fluoride dosage was noticeable during the two month period of July ’03 and Aug ’03. There was no clear explanation that could have caused this drop in dosage levels. Despite comparing the data with the flow rate of the plant, still no explanation was given by the operator. It is clearly visible that the dosage method is almost in continuous sympathy with the flow.

Based on these conclusions drawn out from the graphs, it can be said that the data is still insufficient to run the WTW. The data is also insufficient and/or missing to confirm compliance; water quality measurements are not available to determine the efficiency of the WTW processes. The inaccuracy and discrepancies in the data caused by down time in the WTW is hampering the processes work and increases cost and time. No link with the SCADA (surveillance control and data acquisition) system can be drawn, except an indication that the chemical dosage levels are controlled from outside the plant works system.

2.2 River Water

Part of the objectives of the project was to improve the operational running of the SCADA system of the WTW through the use of an advanced sensors network and automated processes. A second case of surface water WTW was therefore chosen. This was to ensure that findings and obstacles faced with the first case study were very much real and unnoticed. The WTW is a river abstraction of around 100ML/day.

2.2.1 Data Stored

The site monitors Turbidity, Colour, pH, Conductivity, Ammonia & Temperature. These parameters were chosen on the basis of their availability and also their relevance to each other. Using this data, a statistical analysis was performed to try and correlate these data together.

A deterministic value of how much correlation existed between the different parameters, the covariance value between each parameter was calculated (Table 2).

The covariance tool is used to examine each pair of measurement parameters to determine whether the two measurement parameters tend to move collectively— that is, whether large values of one parameter tend to be associated with large values of the other (positive covariance), whether small values of one variable tend to be associated with large values of the other (negative covariance), or whether values of both variables tend to be unrelated (covariance near zero).

	Turbidity	*Colour*	*pH*	*Conductivity*	*Ammonia*	*Temperature*
Turbidity						
Colour	**81.223**					
pH	-0.532	-0.483				
Conductivity	**-362.84**	**-388.41**	4.068			
Ammonia	0.334	0.105	-0.002	-0.004		
Temperature	-5.748	-10.514	0.007	17.486	-0.002	

Table 2: Covariance coefficients between the parameters monitored

It is clear from the table that the most obvious positive link exists between turbidity & colour. This is acceptable as turbidity contributes largely to the colouring in the water as a result of sediments floating in the water. Figure 4 shows the Turbidity & Colour readings changing with time. The higher the covariance coefficients between the parameters monitored, the stronger the correlation between these parameters. It is also clear that the temperature has some effect on the conductivity readings. This sounds logical, as it would be anticipated that temperature would affect solubility of contaminants and therefore conductivity.

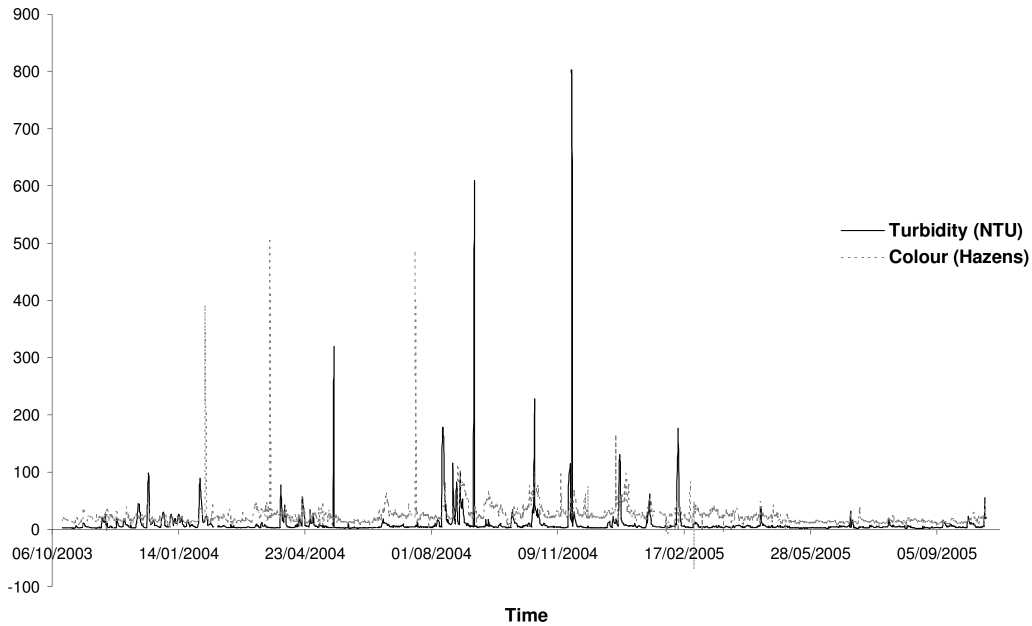


Figure 4: Changes of Turbidity & Colour with Time.

It is also noticeable from Table 2 that the covariance between turbidity & colour at one end, and conductivity at the other end. As stated before, negative covariance's imply small values of one variable tend to be associated with large values of the other. In other words, a decrease in turbidity & colour increases conductivity. Figure 5 illustrates these findings.

Figure 5 reveals the distinct characteristics of the river. The river conductivity reading mainly lies between 250 and 650 (umho/cm). These readings are circumstantial depending on other factors, like river flow, rain fall, etc. Salts, minerals, and even dissolved gases contribute to the conductivity of a given solution. This means that the conductivity can be used as an indicator of the amount of dissolved materials in a solution, thus how reliable an indicator of other parameters is.

Since the covariance analysis cannot be taken as a conclusive result, because it does not account for the different dimensionality of the parameters, another statistical analysis was used to double check the findings. This time the correlation coefficient was used (Table 3).

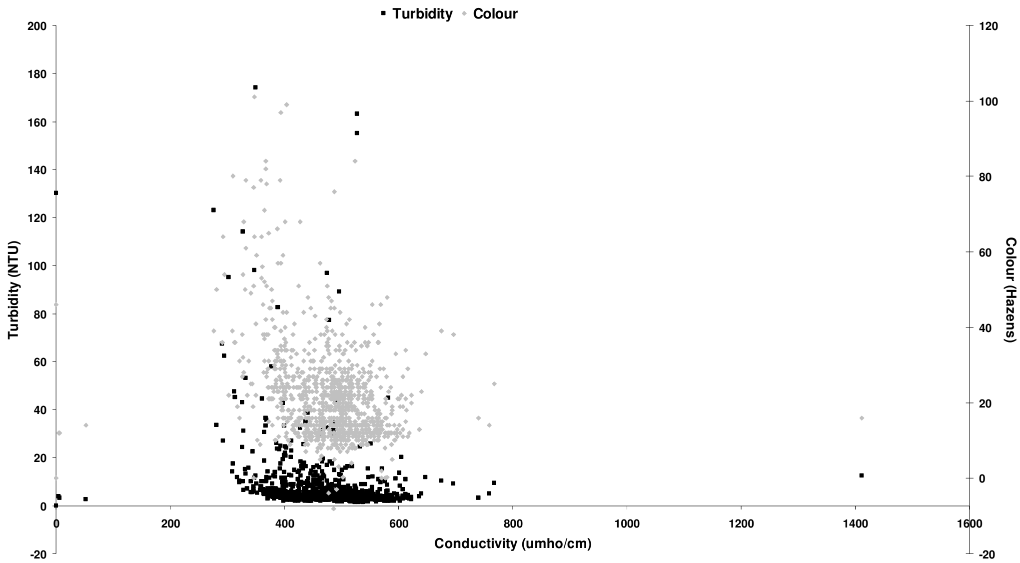


Figure 5: Conductivity readings with changes in Turbidity & Colour

	Turbidity	*Colour*	*pH*	*Conductivity*	*Ammonia*	*Temperature*
Turbidity	1.000					
Colour	**0.117**	1.000				
pH	-0.067	-0.084	1.000			
Conductivity	-0.146	-0.218	**0.199**	1.000		
Ammonia	**0.229**	0.100	-0.127	0.000	1.000	
Temperature	-0.041	-0.191	0.006	0.055	-0.010	1.000

Table 3: Correlation coefficient between the parameters monitored

The correlation coefficients calculated for the measured parameters at the river WTW have very small significance. This is an indication that all the parameters measured are not related to each other and cannot be interpolated from other types of sensors. The only explanation that can be noted from the correlation is that the relationship between Turbidity and Ammonia can be attributed to the rainfall effect on farmlands adjacent to the river. The rain is expected to wash some of the fertilizers into the river stream and causing the turbidity to rise at the same time.

From the above, it can be concluded that the data is not reliable enough to run the plant; each parameter measured is very much independent on its own of other types of parameters and their sensors. Sensors are clearly not really used fully; the interrelationship between sensors should be established to crosscheck the other sensor performance. Data also does not show how it benefits SCADA or how it links with it. Finally, data from sensors are disregarded as valuable information, since data used in the analysis were measured, not monitored.

3. Conclusion

Two case studies were looked at to examine the reliability and confidence of the sensors readings on the running of Water Treatment Works. Groundwater and surface water sites were both taken as a first step in analysing their measurements taken. The data represented by the two sites showed the lack of use, which is currently widespread in the water industry, for the full potential of an alternative measuring and monitoring techniques that can be implemented within the industry. It also emphasised the issue of backup monitoring and self adjusting automation processes that are needed within the industry to face the huge rise in power consumption. The study also showed that a relationship is needed to be found between the different types of sensors and/or measured parameters in order to cross check the sensors performance and be used as a guide of when maintenance procedures are needed. Operating procedures within the WTWs are also required to be improved to cut costs; for example, the use of artificial neural network would enhance the work rate of the site simply by detecting when it needs to reset itself without the support of an external operator.

Acknowledgement

The authors would like to thank the Engineering & Physical Sciences Research Council (EPSRC) in the UK and Myriad Vision Ltd for their support and funding of this project. A special appreciation is also in place for Severn Trent Plc for the support given in accessing and gathering of the data for the two sites.

References

[1] American Water Works Association, Water Quality and Treatment: A Handbook of Community Water Supplies, McGraw-Hill, New York, 1990.

[2] http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31998L0083:EN:HTML

[3] Hughes, G. Standardising Control for the Water Industry. The IEE Computing and Control Journal, 2004.

[4] Geldreich, E.E., Microbial Quality of Water Supply in Distribution Systems, Lewis Publishers, Bocta Raton, Fl, USA, 1996.

[5] Reasoner, D., Pathogens in Drinking Water – Are there any new ones? US Environmental Protection Agency, Washington, D.C, USA, 1992.

[6] Ritchie, J.C. and Cooper, C.M. Remote sensing for determining water quality: Application to TMDL, pp. 367-375. In: TMDL Science Issues Conference, Water Environment Federation, Alexandria, VA, USA, 2001.

[7] Environment Agency, 2004. http:// www.environment-agency.gov.uk/.

[8] Martinez, K, Hart, J. K. and Ong, R. Environmental Sensor Networks. Computer, vol. 37, no. 8, pp. 50-56. 2004

[9] Jane, K.H., and Martinez, K, *Environmental sensor networks*, Eos, Volume 86(16) page 162, 2005.

[10] Dolgonosov, B. M., and Korchagin, K. A. Statistical assessment of relationships between water flow in a river and water turbidity in water intakes. Water Resources, vol. 32, no. 2, pp. 175-182.

[11] Dijkstra, T. Some comments on maximum likelihood and partial least squares methods. Journal of Econometrics, 22, 67-90, 1983.

[12] Geladi, P, and Kowalski, B. Partial least squares regression: A tutorial. Analytical Chimica Acta, 185, 1-17, 1986.

[13] Stone, M. and Brooks, R. Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. Journal of the Royal Statistical Society, Series B, 52(2), 237-269, 1990.

APPENDIX C (PAPER 3)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2007. Online Laboratory Investigation on the use of Acoustic Doppler in Turbidity Measurement. 8th UK's National Young Water Professionals Conference - Guildford, UK.

Online Laboratory Investigation on the use of Acoustic Doppler in Turbidity Measurement

A. Moustafa*, A. El-Hamalawi* and A. Wheatley*

* Civil & Building Engineering Department, Loughborough University, Loughborough, LE11 3TU, England (UK)
(E-mail: a.moustafa@lboro.ac.uk; a.el-hamalawi@lboro.ac.uk; a.d.wheatley@lboro.ac.uk)

Abstract The advanced monitoring of water quality and performing a real time analysis prior to entering Water Treatment Works (WTW) is a real need nowadays in order to give warning of any contamination and provide better control to avoid downtime of the WTW. Two case studies were examined, one being a groundwater WTW and the other a river WTW. Measured data from both sites were analysed. The results showed that the quality of data from the groundwater case study was poor and no good correlation existed between the controlling parameters and standard water quality indicators. The data from the river water WTW was better. The case studies highlighted the need for a new non-invasive/remote sensor type of measurement and for water companies to improve their information technology infrastructure. Results also emphasised the issue of backup monitoring, indirect parameters and self adjusting automation processes that are needed within the industry to improve performance especially to face the huge rise in power costs. The study also showed that a model relationship was needed and would be found between the different types of sensors and/or measured parameters in order to cross check the sensors performance and be used as a guide of when maintenance procedures are needed.
Keywords Acoustic-Doppler, Groundwater, Online, Remote-Sensing

INTRODUCTION

The water quality objectives for drinking water treatment works are regulated by the EU Drinking Water Directive (1998). A water treatment works must be able to produce a consistently high quality regardless of the quality of the intake or how great the demand might be. Water treatment consists of a range of unit processes, usually used in series and this provides some design and operational flexibility to achieve this. The treatment required will depend upon its initial quality, which is normally related to its source. In other words, the cleaner the raw water, the fewer treatment steps that are required, and hence the overall cost of water is less (Hughes, 2004).

The most expensive operations in conventional treatment are sedimentation (assisted by coagulation) and filtration, while water softening can also be very expensive. These are good targets for better automation and there has been much research on developing real time control of coagulation (IWA handbook). Groundwater is generally much cleaner than surface water and thus does not require the same degree of solid removal, aeration and disinfection we often suffice before supply. Naturally occurring substances that may need to be reduced or removed in groundwater include iron, hardness (if > 300 mgl-1 as $CaCO_3$) and carbon dioxide. Compounds originating from urban activity however are becoming increasingly a problem in groundwater and those requiring treatment include nitrates, pathogens and trace organics such as pesticides. Surface water has always required more complex treatment due to its complex nature, although the quality of surface water can also be very high, for example upland reservoirs, although expansion of the resources is difficult (Geldreich, 1996; Reasoner, 1992).

Monitoring and assessing the quality of waters in streams, reservoirs, lakes, and estuaries is critical to improve water quality. Current techniques for measuring water quality include insitu measurements but are also reliant on water samples for subsequent laboratory analyses. While these technologies provide accurate measurements for a point in time and space, they are expensive, and do not provide either the spatial or temporal view of water quality needed for continuous monitoring, assessing, or better managing of water quality (Ritchie & Cooper, 2001).

In 2004, 375 sites were monitored for compliance with the Surface Water Abstraction Directive (75/440/EEC) in England and Wales. Of these, 155 sites failed to comply with the Directive. However, over 90% of these 'failures' were due to insufficient sampling. These sampling shortfalls occur for a number of reasons, such as abstractions not being operated at the time of sampling, problems at the laboratory, and sampling error. The quality of abstracted water has improved since 1993. It was found that levels of colour, nitrate and polycylic aromatic hydrocarbons (PAHs) were parameter that most commonly exceeded the Directive's standards in 2004 but allother had improved (Environment Agency, 2004).

Water pollutants are both suspended sediments (turbidity) as particles and dissolved. Remote sensing applications to determine water quality are to date limited to measuring those substances or conditions that influence and change optical and/or thermal characteristics of the surface water properties (Ritchie & Cooper, 2001). Remote sensors design is made more difficult by the hughly variable sizes of these potential pollutants. Suspended sediments, chlorophylls, DOM, temperature, and oil are water quality indicators that can change the spectral and thermal properties of surface waters and are most readily measured by remote non-invasive sensing techniques. Substances (i.e., nutrients, metals) that do not change the optical and/or thermal characteristics of surface waters can only be inferred by measuring surrogate properties (i.e., chlorophylls) which may have responded to an input of marker chemicals. These remote sensing techniques should improve our abilities to monitor changes in the water topography and contents (Martinez et al, 2004; Jane et al 2005).

This research has analysed two case studies in order to find out how well the current generation of sensors is coping with the flux of data that is being generated and to see what sort of improvements and/or recommendations can be made based on these studies (Dolgonosov & Korchagin, 2005).

CASE STUDIES

Two case studies were investigated. The sites chosen were based on two completely different types of water sources. One site, which can be referred to as the groundwater site, has its main water source from boreholes. The second site investigated was a lowland river.

This combination of ground and surface water WTW would enable us to compare the demands on the sensors used in the different water quality cases. The hypothesis was to identify whether a possible combination of sensors would make it possible to operate the processes of the WTW more efficiently or whether new alternative sensors would be necessary.

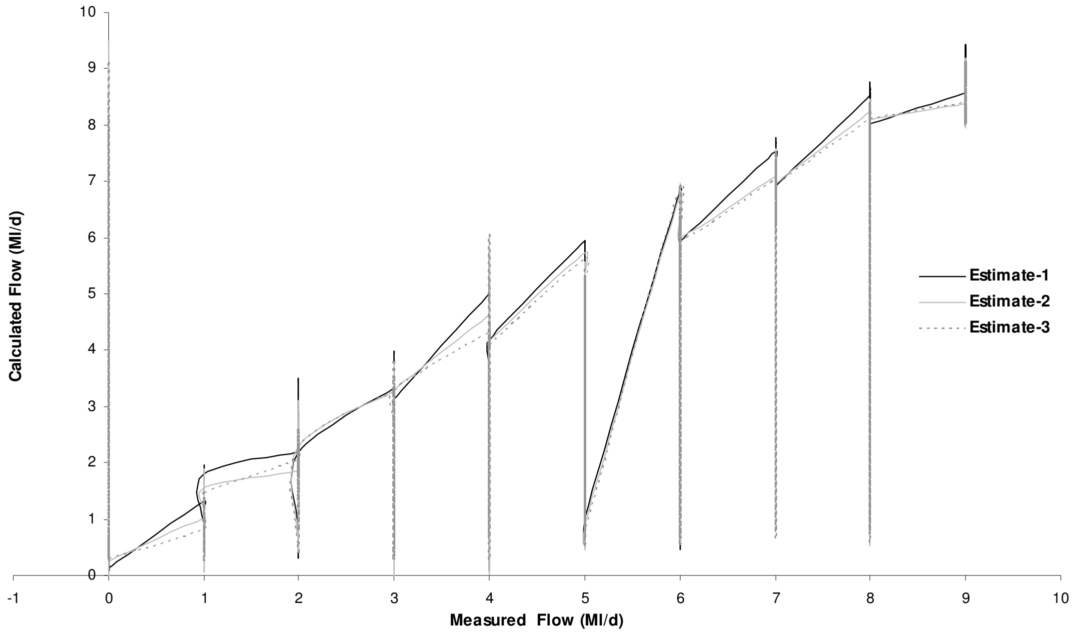
Groundwater WTW

The project was initially aimed at resolving problems that were hampering the continuous operation at the groundwater WTW. The problems were causing shutdown of the WTW and reporting false alarms. This needed continuous intervention by the operating manager to resolve these problems. The operators covered a large area and a number of obstructions and so the frequency of false alarms was unmanageable and the sensors switched off.

After careful examination of the data retrieved from the site in text format, it was noticed that there was not much qualitative monitored data from the boreholes. The only two parameters that were archived were Fluoride and Phosphate dosage levels to comply with requirements set by the regulators. This raised the question as to what parameters were required for regulation and what was necessary for control.

Thus it was decided to try and reprocess the data measured, using the Partial-Least Squares method to determine if these two sensors could provide more general information about reliability of output. Three modelled equations were extracted for the Borehole Flow 1. Graph 1 below illustrates an example of results of correlations of flow and qualitative data for the three models, with the measured flow versus the calculated flow (for the surrogate sensors) in mega litres per day.

Partial least squares (PLS) is a method for constructing predictive models when the factors are many and highly collinear (developed in the 1960’s by Herman Wold as an econometric technique). It is common in chemical engineering processes. PLS has been applied to monitoring and controlling industrial processes; a large process can easily have hundreds of controllable variables and dozens of outputs (Dijkstra, 1983; Geladi & Kowalski, 1986; Stone & Brooks, 1990).



Graph 1: Illustration of the three models for Groundwater Borehole Flow

The graphs drawn out are highly similar in response. Estimate-1 is a calculated response for Borehole Flow 1 using all of the measured readings as dependent factors. In other words, includes the maximum amount of data independent of it origin. Estimate 2 & 3 use less data and the correlations are weaker.

The three graphs showed a strong coherence and indicated the possible use of a very simple equation to predict the total outflow from the WTW. It could help cut the maintenance cost and improve reliability by discarding the other flow meters and sensors under normal operational conditions. The sudden drops in the three graphs were the downtime of the WTW during the experimental period.

The analysis indicates that it would be possible to reduce the number of sensors where a strong correlation can be established but further sensors analysis is necessary to draw links to water quality.

River Water WTW

Part of the objectives of the project was to improve the operational running of the SCADA systems. The WTW is a river abstraction of around 100ML/day from a low land river (of 30-100 m^3/s)

The site monitors Turbidity, Colour, pH, Conductivity, Ammonia & Temperature. These parameters were chosen on the basis of their availability and also their relevance to water quality management. Using this data, a partial least square analysis was performed to try and correlate these data together. The very positive link found was between turbidity & colour. This is understandable; when turbidity is high then other parameters will also increase since sediment load will contribute to the colouring & turbidity in the water. It was also found that the temperature was linked to conductivity readings. This is as it would be anticipated that temperature would affect solubility of contaminants and therefore conductivity.

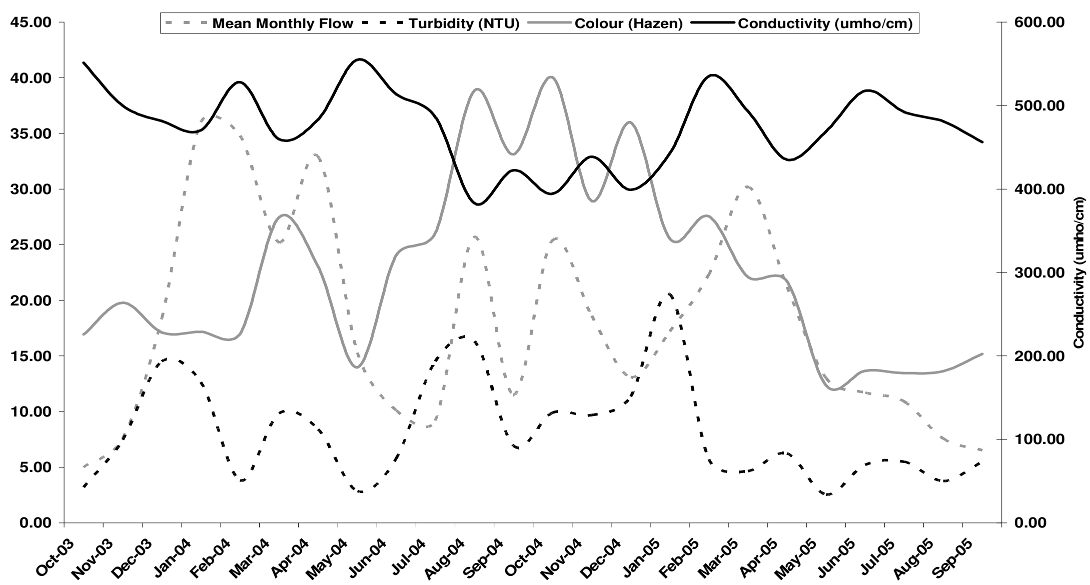
From the detail analysis, it is noticeable that there are relations between all the parameters including flow rate. This should not be surprising since an increase in flow is likely to increase the transport and erosion of materials.

Previous literature correlating different types of sensors were not found and are clearly not really used fully. The interrelationship between sensors should be established to crosscheck the other sensor performance. Data gathered also does not show how it benefits SCADA or how it links with it. Finally, data from sensors are often disregarded as valuable, since data from grab wet often though more reliable.

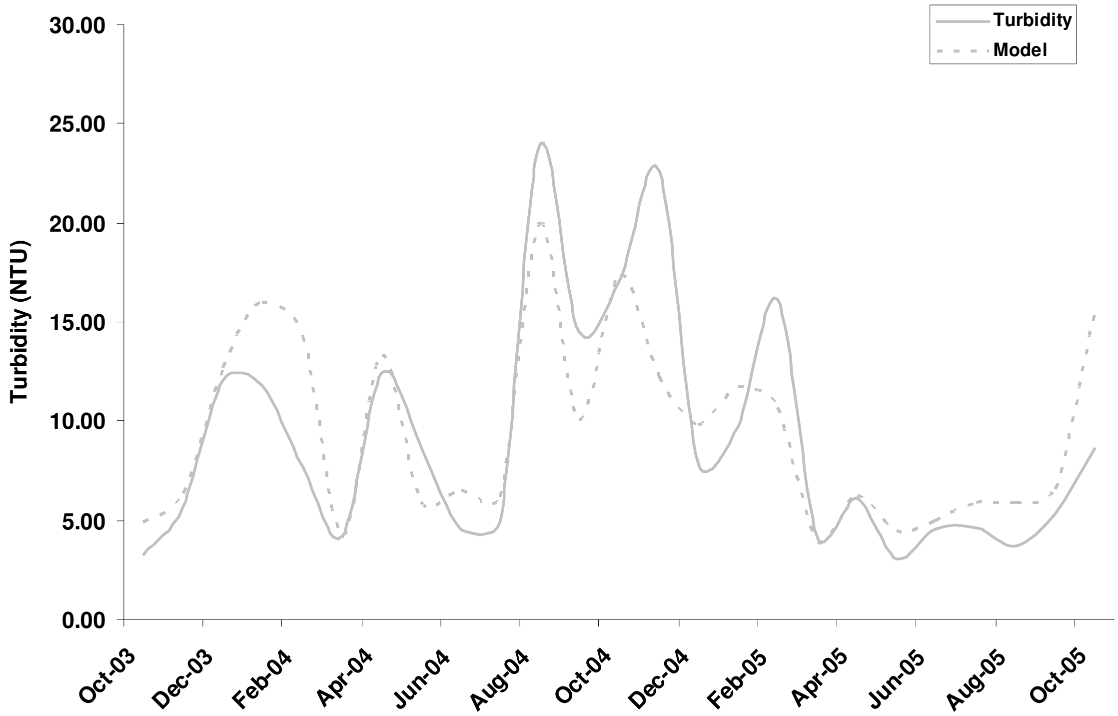
RESULTS & DISCUSSION

During the research of the flow data available, it emerged that there were two river Derwents whose flow rates were continuously monitored, St Mary's Bridge in Derby, which is considered to be the closest monitoring station to the site being investigated, and a Yorkshire Derwent flow rate at Buttercrambe in York, which is approximately 95 miles to the North-East of Derby. The statistical correlation between the two flow rates was very strong, 0.84. This has implied that there might be less local variations in rainfall than might have been thought. There are also a lot of obstructions to the Derby River which should cancel this correlation.

Using the Least-Squares technique and using the Flow rate of the River Derwent at St Mary's Bridge it was possible to predict the Turbidity trend. Graph 2 shows the Flow, Turbidity, Conductivity & Colour levels. Graph 3 illustrates the finding in deducing the trend shape of the Turbidity level from just the flow rate at St Mary's Bridge.

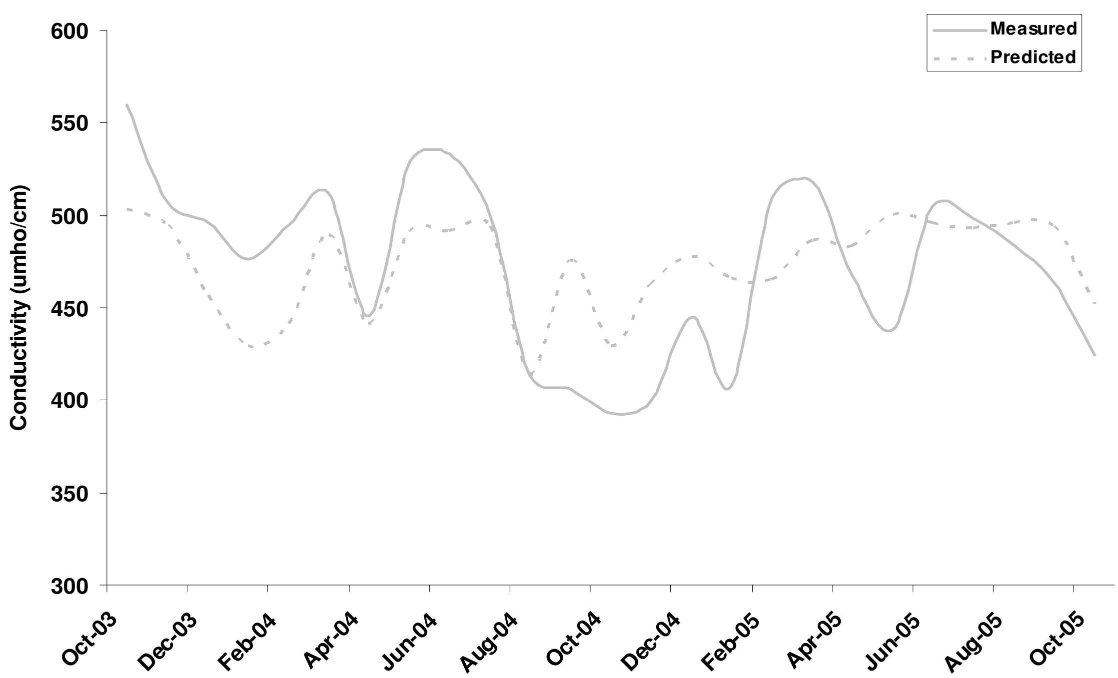


Graph 2: Comparison between Flow, Turbidity, Conductivity & Colour levels measured.



Graph 3: Measured Turbidity levels versus predicted levels extracted from the flow rate at St Mary's Bridge.

Following the derivation of the Turbidity level trend, the data is fed back into the Least-Squares technique again together with the flow data and the predicted Conductivity trend was calculated, graph 4.
In short, Flow rate + Turbidity level => Conductivity value.



Graph 4: graph of the predicted Conductivity value against the measured value.

The same process was repeated again by feeding the Conductivity value back into the Least-Squares method to derive the Colour level. In other words, Flow Rate + Turbidity Level + Conductivity Value => Colour Level.

It would be assumed that the calculated interpretations of the data for Turbidity, Conductivity & Colour could be more realistic of the real values in the River Derwent. But it should be stressed that these relationships should be developed as moving correlations over several years and may not be appropriate for other WTW. These relationships were derived using the data supplied from one site, which means that it's more related to this site rather than being generalised. There is at least the prospect of using data from a much wider field of similar rainfalls but with different weightings in the model.

One possible explanation on why the predicted values differ from the measured ones, is because the measured values were lab measurements, rather than online measurements. Many operators have more confidence in the representative value of grab samples. It is also likely that the correlation coefficient would vary according to flow, season (e.g. leaf fall), antecedent dry period and impoundment.

Laboratory analysis of sensors: Ultrasonic Doppler
New innovative remote sensors are being developed and a review of the different types of non-invasive techniques was made. These included Acoustic, Fluorescence, Laser & X-Ray techniques were all examined. A comparison was made and concluded that Acoustic-Doppler techniques would suit the objective most. This was based on the fundamental basis of operation, since the technique uses the particles suspended in water to determine the flow velocity of the water.

The first experiment, were simple as a controlled amount of Bentonite clay powder were added to a water container (approximately 20*l* of tap water). The added amounts were controlled by adding the right dosage equivalent to 5 mg/*l* at each stage. The turbidity was increased from approximately 0 NTU (clear tap water) to approximately 30 NTU, which is double the average river turbidity reading measured. The ultrasonic Doppler sensor and an electric stirrer were used to keep the clay in suspension.

The Ultrasonic Doppler used was not a non-invasive model. Unfortunately project resources were not available to purchase one especially.

Velocimeters use acoustic sensing techniques to measure flow in a remote sampling volume. The measured flow is practically undisturbed by the presence of the probe. Data are available at an output rate of 25 Hz. The 3-D velocity range is 2.5 m/s, and the velocity output has no zero-offset.

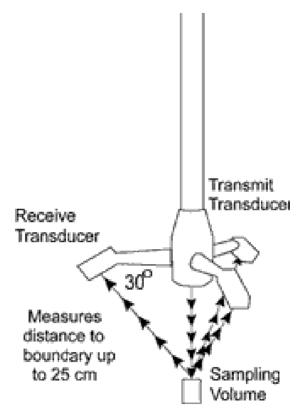
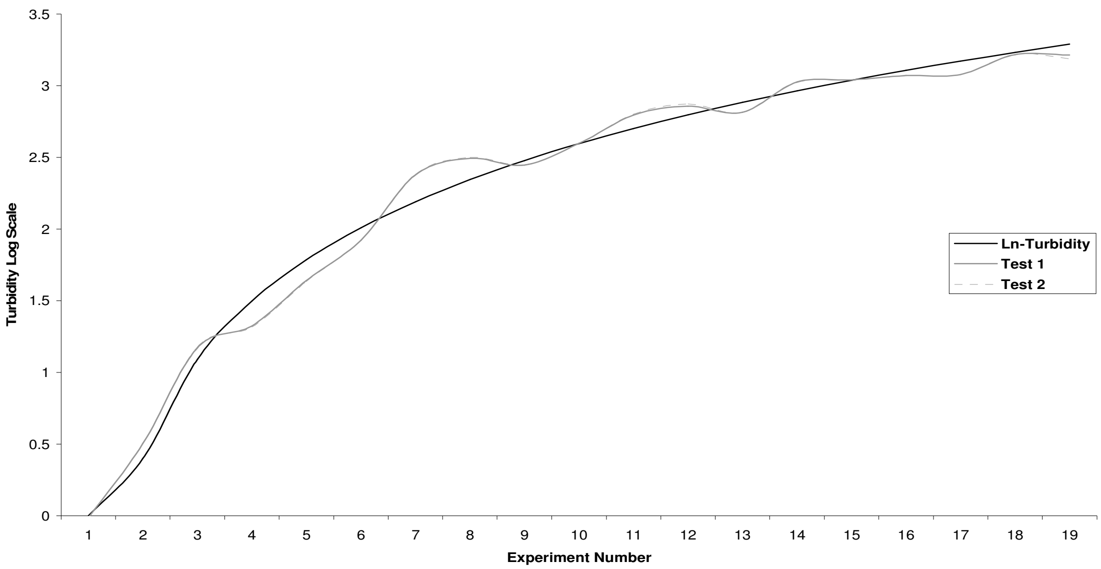


Figure 2: The acoustic sensor used has one transmit transducer and three receive transducers. The sampling volume is located away from the sensor to provide undisturbed measurements. Doppler velocity is derived from signals scattered by small particles. In natural bodies of water (streams, lakes, rivers, oceans, etc.) the correlation of particles is sufficient for proper operation. In model tanks with running water (flumes, open channels, closed pipes, etc.) microscopic bubbles in the water column tend to act as natural seeding. In very clean, quiescent water (ship models, tow tanks, and some wave flumes), seeding materials must be added to concentrations of approximately 10 mg/l.

The stirrer position was fixed during the entire experimental setup close to the bottom of the tank as much as possible; so as to ensure that there is no precipitation of the particles at the bottom. The stirrer speed was also fixed in order to ensure that data differences can be attributed to a change in turbidity levels only. Another critical factor during the setup of the probe was to ensure that the side benchmarked as the X-direction of the probe should always be fixed in the direction of the flow. This ensures similar bias from any differences in the sensors.



Graph 5: comparison between the measured turbidity levels versus readings taken by the Ultrasonic probe after applying PLS to the readings

Graph 5 demonstrates that the probe was capable of detecting the amount of turbidity particulates flowing by it. But the major setback was even after repeated trials, the test graphs could not be improved. This still leaves the possibility that the sensitivity was too low to identify impartial changes in turbidity. In other words, the probe was not designed to investigate this type of experiment, it was rather designed to merely measure the sounds coming off and back.

Following the success of the basic trials, focus the next experiments are to initiate test of sensors placed by the WTW at the intakes. The probe was placed in a laboratory flume in which it was possible to vary velocity, colour and turbidity.

CONCLUSIONS

- There is a lack of confidence in sensor data. In one of the case studies, the sensors were actually a major burden on resources (false alarms).
- Correlations were established between the different type of sensors, i.e. chemical and physical (flow) which does offer the possibility of better models with inbuilt deletion of spurious data.
- Correlations were found between river flows in two different rivers 150km apart indicating the possibility of value in a National network of sensors.
- Modern non invasive probe technology, in this case Acoustic Doppler, does provide data which could be interrogated differently to provide qualitative as well quantitative data.
- Further work should include the possible use of Object Oriented Modelling or Artificial Neural Networks in sensors correlations.

ACKNOWLEDGEMENT

The authors would like to thank the Engineering & Physical Sciences Research Council (EPSRC) in the UK and Myriad Vision Ltd for their support and funding of this project. A special appreciation is also in place for Severn Trent Plc for the support given in accessing and gathering of the data for the two sites.

REFERENCES

American Water Works Association, Water Quality and Treatment (1990) A Handbook of Community Water Supplies, McGraw-Hill, New York.

Council Directive 98/83/EC on the quality of water intended for human consumption (1998); Official Journal L 330, 003 -0054. http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31998L0083:EN:HTML

Dijkstra, T. (1983) Some comments on maximum likelihood and partial least squares methods. Journal of Econometrics, 22, 67-90.

Dolgonosov, B. M., and Korchagin, K. A. (2005) Statistical assessment of relationships between water flow in a river and water turbidity in water intakes. Water Resources, vol. 32, no. 2, pp. 175-182.

Environment Agency Website (2004). http:// www.environment-agency.gov.uk/.

Geladi, P, and Kowalski, B. (1986) Partial least squares regression: A tutorial. Analytical Chimica Acta, 185, 1-17.

Geldreich, E.E. (1996) Microbial Quality of Water Supply in Distribution Systems, Lewis Publishers, Bocta Raton, Fl, USA.

Hughes, G. (2004) Standardising Control for the Water Industry. Computing and Control Engineering Journal, vol 15, issue 5.

Jane, K.H., and Martinez, K (2005) Environmental sensor networks, Eos, Volume 86(16) page 162.

Martinez, K, Hart, J. K. and Ong, R. (2004) Environmental Sensor Networks. Computer, vol. 37, no. 8, pp. 50-56.

Reasoner, D. (1992) Pathogens in Drinking Water – Are there any new ones? US Environmental Protection Agency, Washington, D.C, USA.

Ritchie, J.C. and Cooper, C.M. (2001) Remote sensing for determining water quality: Application to TMDL, pp. 367-375. In: TMDL Science Issues Conference, Water Environment Federation, Alexandria, VA, USA.

Stone, M. and Brooks, R. (1990) Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and principal components regression. Journal of the Royal Statistical Society, Series B, 52(2), 237-269.

APPENDIX D (PAPER 4)

Moustafa, A, El-Hamalawi, A, Wheatley, A. 2008. Deciphering river flow data to determine river properties. Twelfth International Water Technology Conference, IWTC12 2008, Alexandria, Egypt, pp. 305-315.

DECIPHERING RIVER FLOW DATA TO DETERMINE RIVER PROPERTIES

Ahmed Moustafa, Ashraf El-Hamalawi[1] and Andrew Wheatley

Civil & Building Engineering Department, Loughborough University, Loughborough, LE11 3TU, England (UK)

ABSTRACT
The advanced monitoring of water quality and performing a real time hazard analysis prior to entering Water Treatment Works (WTW) is very much a necessity nowadays in order to give warning of any contamination and avoid downtime of the WTW. Ultrasonic flow sensors are considered to be the most stable & reliable sensors to be used within the water & wastewater industry, as they are not easily affected by the weather conditions of the rivers' varying temperatures. Two case studies were looked at from two UK Rivers; one being the River Derwent in the East Midlands, Derbyshire, and the other being the River Severn in the West Midlands, Gloucestershire. Correlations between the flow rates of each river and the turbidity rates at a Water Treatment Works (WTW) near the gauging stations on the rivers were established. The river flow rates were then used to estimate the turbidity levels in the rivers prior to entering the WTW. The results showed that correlations existed between the rivers flow rates and its properties such as turbidity, colour & conductivity. The case studies highlighted the need for a new, intelligent and smart non-invasive type of measurement device to be introduced to enable the WTW to operate with optimum performance.

Keywords

Water Monitoring, Sensors, Turbidity, Flow Rate

1. INTRODUCTION
Water treatment plants must be able to produce a finished product of consistently high quality regardless of how great the demand might be. Like wastewater treatment, water treatment consists of a range of unit processes, usually used in series, which provides some design and operational flexibility. The treatment required by water prior to being delivered to consumers will depend upon its initial quality, which is normally related to its source. In other words, the cleaner the raw water, the fewer treatment steps that are required, and hence the overall cost of water is less. The most expensive operations in conventional treatment are sedimentation and filtration, while water softening can also be very expensive. Groundwater is generally much cleaner than surface water and so does not require the same degree of treatment, apart from aeration and disinfection before supply. Naturally occurring substances that may need to be reduced or removed in groundwater include iron, hardness (if > 300 mg/l as $CaCO_3$) and Carbon dioxide. Substances originating from humans are becoming increasingly common in

groundwater and those requiring treatment include nitrates, pathogens and trace organics such as pesticides. Surface water requires more complex treatment due to its complex nature, although the quality of surface water can be very high (for example upland reservoir) (Geldreich, [1]; Reasoner, [2]).

In 2004, 375 sites were monitored for compliance with the Surface Water Abstraction Directive (75/440/EEC) in England and Wales. Of these, 155 sites failed to comply with the Directive. However, over 90% of these 'failures' were due to insufficient sampling. These sampling shortfalls occur for a number of reasons, such as abstractions not being operated at the time of sampling, problems at the laboratory, and sampling error. The quality of abstracted water generally improved since 1993. It was found that levels of colouration, nitrate and polycyclic aromatic hydrocarbons (PAHs) most commonly exceeded the Directive's standards in 2004 (Environment Agency, [3]).

Monitoring and assessing the quality of waters in streams, reservoirs, lakes, and estuaries are critical to improve water quality. Current techniques for measuring water quality involve insitu measurements and/or the collection of water samples for subsequent laboratory analyses. While these technologies provide accurate measurements for a point in time and space, they are expensive, and do not provide either the spatial or temporal view of water quality needed for monitoring, assessing, or managing water quality for an individual or multiple water bodies across the landscape. Remote sensing of indicators of water quality offers the potential of relatively inexpensive, frequent, and synoptic measurements using non-invasive sensors.

Major pollutants are suspended sediments (turbidity), pathogens, nutrients, metals, dissolved organic matter (DOM), pesticides, chlorophylls (algae, plants), temperature, and oils. Remote sensing applications to determine water quality are limited to measuring those substances or conditions that influence and change optical and/or thermal characteristics of the surface water properties (Ritchie & Cooper, [4]). Suspended sediments, chlorophylls, DOM, temperature, and oil are water quality indicators that can change the spectral and thermal properties of surface waters and are most readily measured by remote sensing techniques. Substances (i.e., nutrients, metals) that do not change the optical and/or thermal characteristics of surface waters can only be inferred by measuring surrogate properties (i.e., chlorophylls) which may have responded to an input of chemicals. These remote sensing techniques should improve our ability to monitor changes in the water topography and contents.

In this paper, correlations between the flow rates and various other properties such as turbidity, colour and conductivity for two UK Rivers were computed, and then used to predict these properties' levels at various other locations and points in time for the rivers and associated WTW/gauge stations. The case studies highlighted the need for a new, intelligent and smart non-invasive type of measurement device to be introduced to enable the WTW to operate with optimum performance. The decision to look at case studies was made in order to find out how the water industry is coping with the flux of data that is being generated by sensors and to see what sort of improvements and/or recommendations can be made based on these studies (Dolgonosov & Korchagin, [5]).

2. CASE STUDIES

The two sites examined were based on two completely different river sources. One site, which can be referred to as the St Mary's Bridge site, has its main water source from the River Derwent in East Midlands. The second site investigated, can be referred to as Saxons Lode, has it main water source from the River Severn in West Midlands. This combination of two different river sources and locations enabled us to compare the demands on the flow sensors used in most river water resources. The hypothesis was to identify whether a possible combination, with a minimal number of sensors, would make it possible to operate the processes of the WTW close to the two case studies more efficiently or whether new alternative sensors would be necessary. Alag et al [6] reported that by combining information from many different sources, it would then be possible to decrease the uncertainty and ambiguity inherent in processing the information from a single sensor source. A large number of sensors measuring many different variables can collectively achieve a high level of accuracy and reliability, depending on their accuracy and reliability.

2.1 St Mary's Bridge

Using one of the River Derwent gauging stations, operated by the Environmental Agency, the daily mean flow rate was attained at St Mary's Bridge in Derbyshire County. The WTW being investigated has a river abstraction of around 100 Ml/day output from a lowland river with a flow of 30 to 100 m^3/s and is located south of the gauging station on the River Derwent.

The project was initially launched in order to improve the monitoring of the river water prior to entering the WTW to ensure the continuous operation at the WTW. Several problems had to be resolved, the most important of which were the shutdown of the WTW because of false alarms, and the time taken to analyse the water quality during which the water characteristics would have changed. This needed continuous intervention by the operating manager to resolve these problems. The operators covered a large area and a number of remote sites and so the frequency of false alarms was unmanageable and the sensors switched off.

The main finding of the analyses was a correlation between the daily mean flow rate and the turbidity measurements. Consequently, using a partial-least squares statistical method; it was found that turbidity levels in the rivers could be estimated using the current daily mean flow. The Partial least squares (PLS) method, developed in the 1960's by Herman Wold as an econometric technique, is a method for constructing predictive models when the factors are many and highly collinear. It is more commonly used in chemical engineering processes, but less so in civil engineering. PLS has been applied to monitoring and controlling industrial processes, where hundreds of controllable variables and dozens of outputs need to be processed (Dijkstra,[7]; Geladi & Kowalski, [8]; Stone & Brooks, [9]). Figure 1 illustrates the resulting correlation between the flow and the turbidity levels on the River Derwent, both actual and estimated, using the correlated river flow. Figure 2 depicts measured conductivity values against approximated values using the same statistical concept.

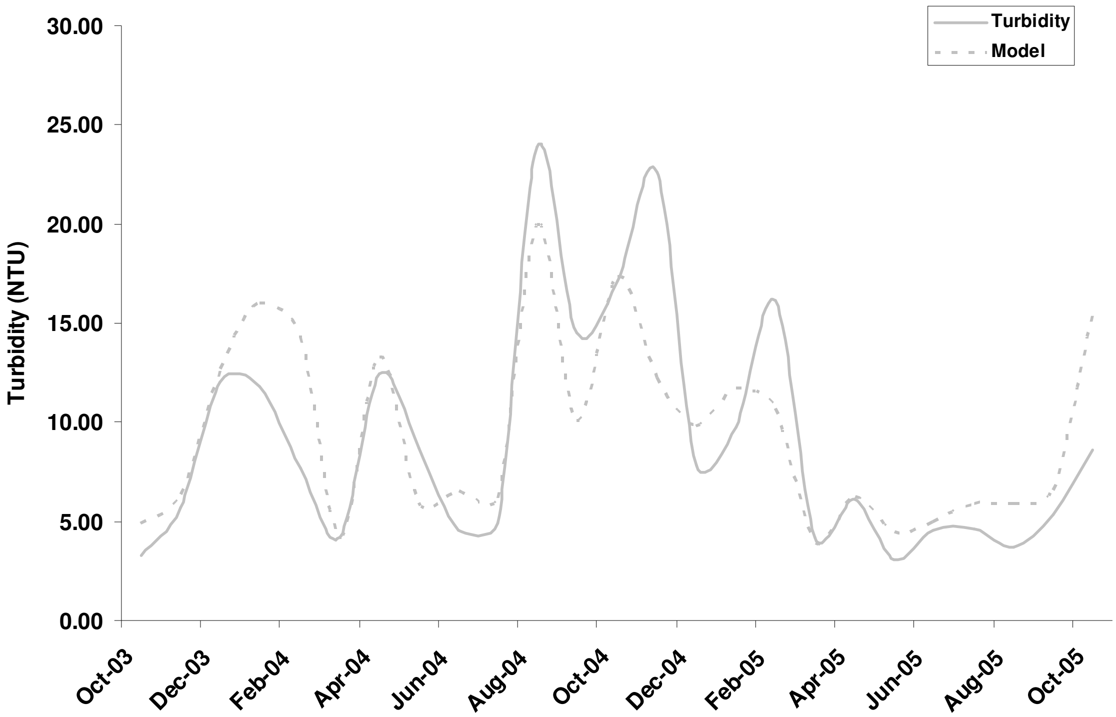


Figure 1: Measured Turbidity levels versus predicted levels extracted from the flow rate at St Mary's Bridge on the River Derwent.

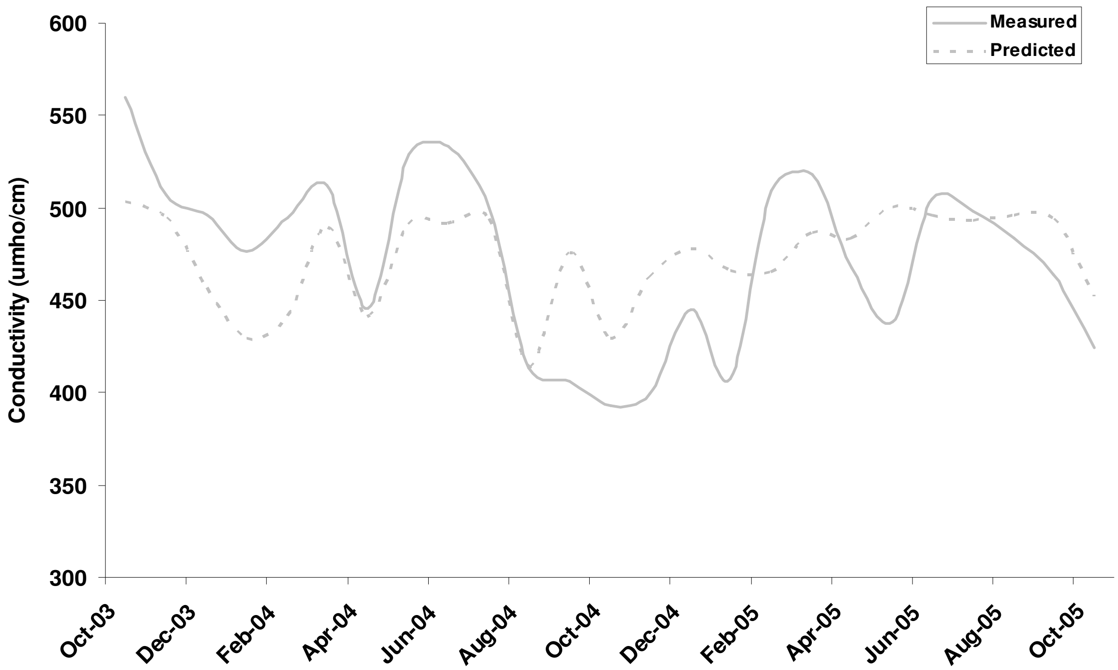


Figure 2: Measured Conductivity value against the predicted value extracted from the flow rate at St Mary's Bridge on the River Derwent.

Graphs 1 & 2 show a strong coherence and indicated the possible use of a very simple technique to predict the river water constituents and therefore cross check sensors output. It could help cut the maintenance cost and improve reliability by discarding spurious data under normal operational conditions. The analyses also indicate that it must also be possible to improve the quality of sensor information by using this type of expert system where a strong correlation can be established. PLS regression is an extension of the multiple linear regression models (e.g., Multiple Regression or General Stepwise Regression). In its simplest form, a linear model specifies the (linear) relationship between a dependent (response) variable Y, and a set of predictor variables, the X's, so that

$$[Y] = b_0 + b_1[X_1] + b_2[X_2] + ... + bp[Xp] \quad (1)$$

In this equation b_0 is the regression coefficient for the intercept and the bi values are the regression coefficients (for variables 1 through p) computed from the data. To put it in perspective, the dependent variable would be turbidity level, and the predictor variable is the daily mean flow rate.

PLS regression extends multiple linear regression without imposing the restrictions employed by discriminate analysis, principal components regression, and canonical correlation. In partial least squares regression, prediction functions are represented by factors extracted from the Y'XX'Y matrix. The number of such prediction functions that can be extracted typically will exceed the maximum of the number of Y and X variables. In short, PLS is the least restrictive of the various multivariate extensions of the multiple linear regression models. This flexibility allows it to be used in situations where the use of traditional multivariate methods is severely limited, such as when there are fewer observations than predictor variables. Furthermore, partial least squares regression can be used as an exploratory analysis tool to select suitable predictor variables and to identify outliers before classical linear regression.

2.2 Saxons Lode

A second case study was initiated to cross-validate the findings of the first case study. Attention was turned to the largest UK River, the River Severn. Again, the WTW concerned in this case was also downstream from the Saxons Lode gauging station. The fluctuations in the turbidity levels experienced in both case studies can be attributed to the fluctuations in the daily mean flow rate. This is a natural phenomenon and has more than one cause for its rise and fall. Some of the factors attributing to this would be rainfall, wind, flood seasons, seasonal temperature changes. Figure 3 illustrates the resulting correlation between the flow and the turbidity levels on the River Severn, both actual and estimated, using the correlated river flow.

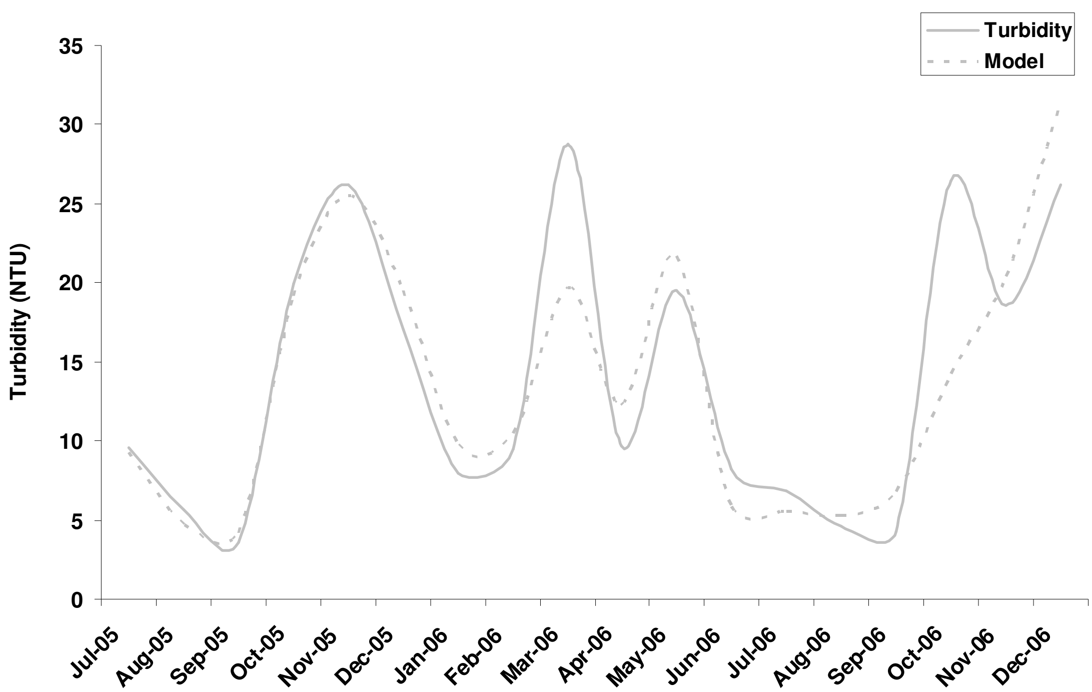


Figure 3: Measured Turbidity levels versus predicted levels extracted from the flow rate at Saxons Lode on the River Severn.

3. RESULTS & DISCUSSION

For each of the WTW that were investigated; each site monitors Turbidity, Colour, pH, Conductivity, Ammonia & Temperature. Using this data, a partial least squares analysis was performed to try and correlate these data together. A very positive link was found between turbidity & colour, fig 4. This relation between turbidity & colour is logical; when turbidity is high then other parameters, e.g. colour, will also increase, since sediment load will contribute to the colouring & turbidity in the water. It was also found that the temperature was linked to conductivity readings, fig 5. This is as expected, since temperature would affect the solubility of contaminants, and therefore conductivity. The shift in time variation between the two graphs is attributed largely due to operator and sampling errors. For example, it is impractical to measure temperature of running water by sampling the water in the lab and not at the source.

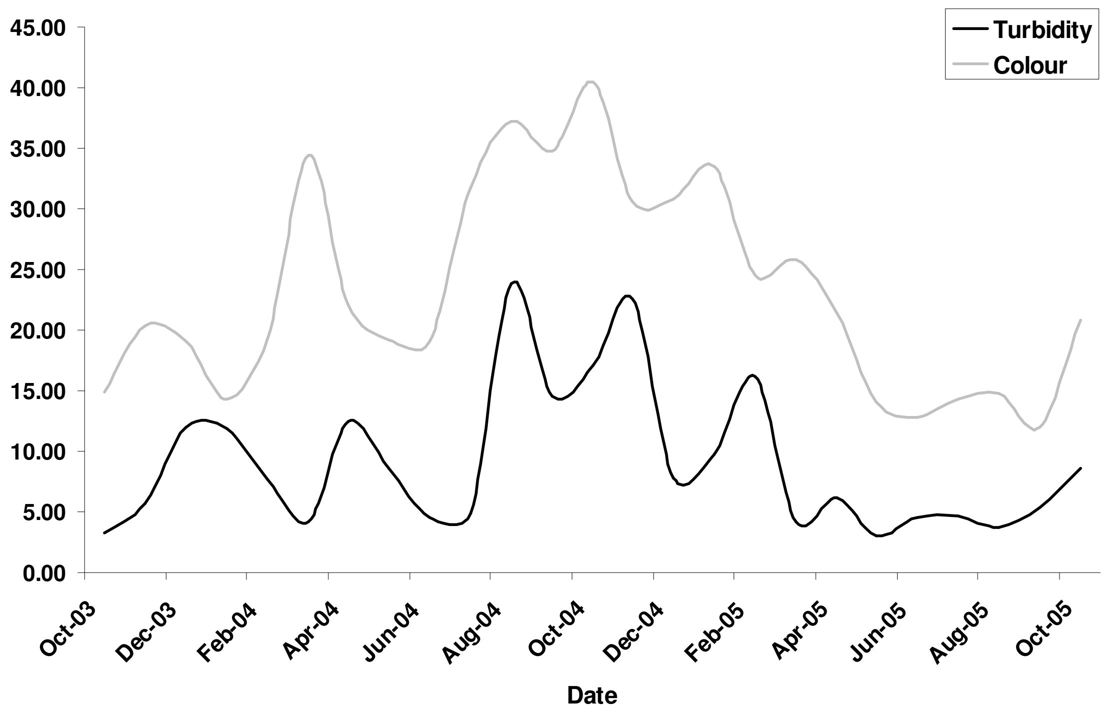


Figure 4: Graphical comparison between the Turbidity & Colour measurements at St Mary's Bridge on the River Derwent

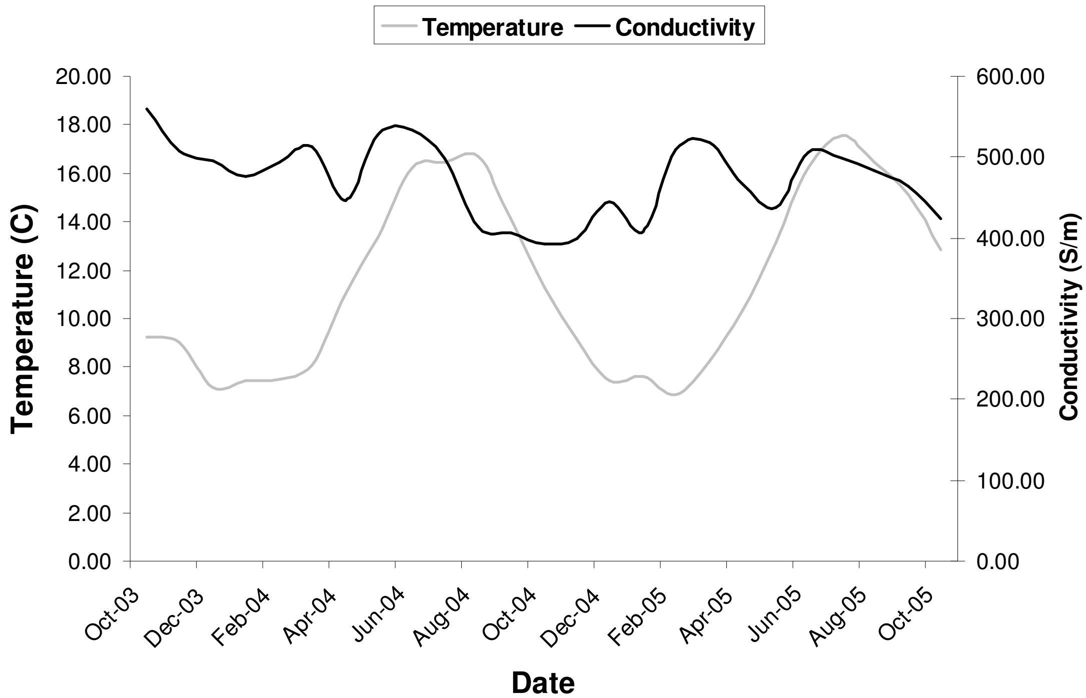


Figure 5: Graphical comparison between the Conductivity & Temperature measurements at St Mary's Bridge on the River Derwent.

From the analysis, it is also noticeable that there are relations between all the parameters including flow rate. This should not be surprising since an increase in flow is likely to increase the transport and erosion of materials. Figure 6 shows the complete data collected from the WTW with the daily mean flow rate. Following the derivation of the Turbidity level trend from the flow rate at St Mary's bridge, the data was fed back into the PLS regression, together with the flow data, and the predicted conductivity trend was calculated. The same process was repeated again by feeding the conductivity values back into the PLS method to derive the colour level. In simpler terminology, once the turbidity level is estimated, and using PLS, data fed back into PLS used conductivity as the dependent variable, while turbidity and flow account for the predictor variables. This process is repeated with colour as the dependent variable and conductivity is added to the predictor variables list.

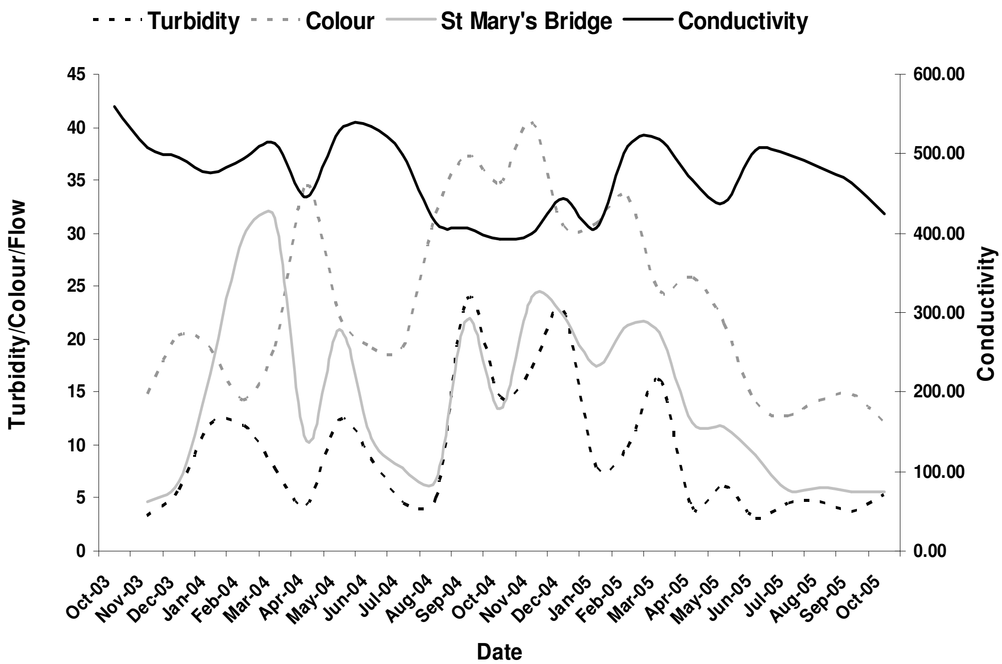


Figure 6: Graphical comparison between the three measured variables (turbidity, conductivity & colour) together with the daily mean flow rate at St Mary's Bridge, on the River Derwent, for two consecutive years.

It would be assumed that the calculated interpretations of the data for Turbidity, Conductivity & Colour could be more realistic of the real values in the River Derwent. It should however be stressed that these relationships should be developed as moving correlations over several years and they will not be appropriate coefficients for other WTW. These relationships were derived using the data supplied from one site. Different sites will have different data and correlations, but the same process could be used to save time and money, in order to increase the efficiency of predicting various parameters based on the flow going through WTWs. These correlation coefficients would vary according to flow, season (e.g. leaf fall), antecedent dry period and impoundment and other river engineering.

The interrelationship between sensors should be established to cross-check the other sensor performance. Data gathered also does not show how it benefits SCADA (Supervisory Control And Data Acquisition) or how it links with it. Finally, data from sensors are often disregarded; since operators intuitively find data from grab-wet samples more reliable.

4. CONCLUSION

Two case studies were looked at to try and examine the relationship between the rivers' flow rates and the rivers' constituents. Two river abstraction Water Treatment Works were probed, one WTW on the River Derwent and the other on the River Severn. The findings presented for the two sites showed the lack of understanding of the rivers flow rates by the operators of the two WTW's and the potential use &

benefits if it had been investigated much earlier. It also highlights the full potential of an alternative measuring and monitoring technique that can be implemented within the industry. Results emphasised the issue of backup monitoring and self adjusting automation processes that are needed within the industry. The study revealed that a relationship is needed to be found between the different types of sensors and/or measured parameters in order to cross-check the sensors' performance and be used as a guide of when maintenance procedures are needed.

ACKNOWLEDGEMENT

The authors would like to thank the Engineering & Physical Sciences Research Council (EPSRC) in the UK and Myriad Vision Ltd for their support and funding of this project. A special appreciation is also in place for Severn Trent Plc for the support given in accessing and gathering of the data for the two sites.

References

[1] Geldreich, E.E., Microbial Quality of Water Supply in Distribution Systems, Lewis Publishers, Bocta Raton, Fl, USA, 1996.

[2] Reasoner, D., Pathogens in Drinking Water – Are there any new ones? US Environmental Protection Agency, Washington, D.C, USA, 1992.

[4] Ritchie, J.C. and Cooper, C.M. Remote sensing for determining water quality: Application to TMDL, pp. 367-375. In: TMDL Science Issues Conference, Water Environment Federation, Alexandria, VA, USA, 2001.

[3] Environment Agency, 2004. http:// www.environment-agency.gov.uk/.

[5] Dolgonosov, B. M., and Korchagin, K. A. Statistical assessment of relationships between water flow in a river and water turbidity in water intakes. Water Resources, vol. 32, no. 2, pp. 175-182.

[6] Alag S., Agogino A.M., Morjaria M. A methodology for intelligent sensor measurement, validation, fusion and fault detection for equipment monitoring and diagnostics. Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 15, 307–320, 2001.

[7] Dijkstra, T. Some comments on maximum likelihood and partial least squares methods. Journal of Econometrics, 22, 67-90, 1983.

[8] Geladi, P, and Kowalski, B. Partial least squares regression: A tutorial. Analytical Chimica Acta, 185, 1-17, 1986.

[9] Stone, M. and Brooks, R. Continuum regression: Cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares, and

principal components regression. Journal of the Royal Statistical Society, Series B, 52(2), 237-269, 1990.

Myriad Vision Ltd
58 Regent Road
Leicester
LE1 6YJ
UK

Centre for Innovative and Collaborative Engineering
Department of Civil & Building Engineering
Loughborough University
Loughborough
Leicestershire, LE11 3TU