The Effects on Energy Markets Subjected to Regulatory Changes Using Neural Net Methodology[1]

RICHARD W. JANSON, CELAL BATUR, AND LALA KRISHNA, Department of Geography, Kent State University, Kent, OH 44240, Department of Mechanical Engineering, Department of Mathematical Sciences, University of Akron, Akron, OH 44325

ABSTRACT. Neural net methodology has been used to model alternative scenarios of fuel utilization. Regulation and legislation to address the problems of energy related pollution such as acid rain, nuclear waste, greenhouse gases, and tailpipe pollution, will alter fuel input ratios with consequential effects in the energy using sectors. Also, alternative input scenarios using clean coal technology, natural gas, and nuclear power have been modeled. Results indicate that large relative increases of coal or nuclear fuel inputs will cause similar substantial increases in electricity generation, and substitution effects will cause a shift of petroleum uses in final consumption from the commercial and residential sectors to the transport sector. Increasing the gas fuel input relative to other fuels causes little disturbance in using sectors. Incremental increases in fuel consumption maintaining constant relative fuel input shares causes little disturbance. On the other hand, massive increases in fuel consumption inputs maintaining constant input shares is likely to be disastrous public policy.

OHIO J. SCI. 94 (3): 60–69, 1994

INTRODUCTION

The interrelatedness of a commodity system, including supply and demand considerations embedded within a national economy, is understood in principle and even in detail through input-output analysis combined with inter-regional trade flows simulation and other methodologies. The difficulty of analysis may actually be increased by the plethora of relevant data. Models that offer intellectual closure are often technical, involved, and abstract with a result that limits the value to an elite group that understands the implications of the models, the precise significance of intermediate variables, and the structure of the data in the context of policy. Neural net analysis provides techniques with sufficient finesse to overcome many of these problems inherent in systems as complex as a national energy commodity market system.

A neural net is equivalent to a set of nonlinear equations that connect measured inputs representative of an entire focal system to corresponding measured outputs of the entire system. In many neural net models intermediate layers of variables are influenced in varying degree by all lower level (input side) variables and in turn influence all intermediate variables in all upper layers (output side) (see Lapedes and Farber 1988). In this research Energy Consumption by Source includes five categories: coal, natural gas, petroleum, nuclear, and hydropower (Energy Information Administration 1991). The five categories of energy sources are the five input variables to the neural net. Twelve output variables result from the three using sectors—1) Residential and Commercial Sector, 2) Industrial Sector, and 3) Transportation Sector—consuming energy in the following four forms: 1) coal, 2) natural gas, 3) petroleum, and 4) electricity. The interrelatedness of the many fuel and energy markets and the obvious functional relationships among the markets suggested the possibility that a neural network may adequately model the connections between inputs and outputs.

Once the neural network was established, scenarios of varying likelihood were modeled to anticipate effects on outputs that are expected to result from alternative legislative action or regulation of inputs. Much of this regulation may occur in order to address environmental considerations related to the type, location, and technology of energy consumption at electric generating plants, and the type, location, and sector of energy final usage, in an economic sense.

MATERIALS AND METHODS

A neural network model consists of an interconnected set of nonlinear processing elements that are called neurons. The output of each neuron is connected to every other neuron in the next layer and the strength of this connection is determined by an adjustable parameter called weight (Kosko 1988, 1992). A typical neural network model might contain three layers and six neurons (Fig. 1). In this simplified model, neurons on a given layer are not connected to each other. Furthermore, the output of each neuron is connected only to the neurons in the next forward layer. There are two inputs, x_1 and x_2, and one output $\hat{y}$ in this network. The basic idea in the neural network modeling is to adjust the weights w_i; $i = 1,2,...,9$ such that the network can estimate an unknown functional relationship between the inputs and the output. This unknown relationship is represented by

$$y = f(x_1,x_2) \quad (1)$$

Since the output of the network is a function of the weights, the neural network output is presented as

$$\hat{y} = g(x_1,x_2,w1,\ldots,w9) \quad (2)$$

The difference between the neural network output, $\hat{y}$,

[1]Manuscript received 22 February 1993 and in revised form 18 April 1994 (#93-08).

[2]This work was partially supported by the University of Akron (summer faculty research fellowship).

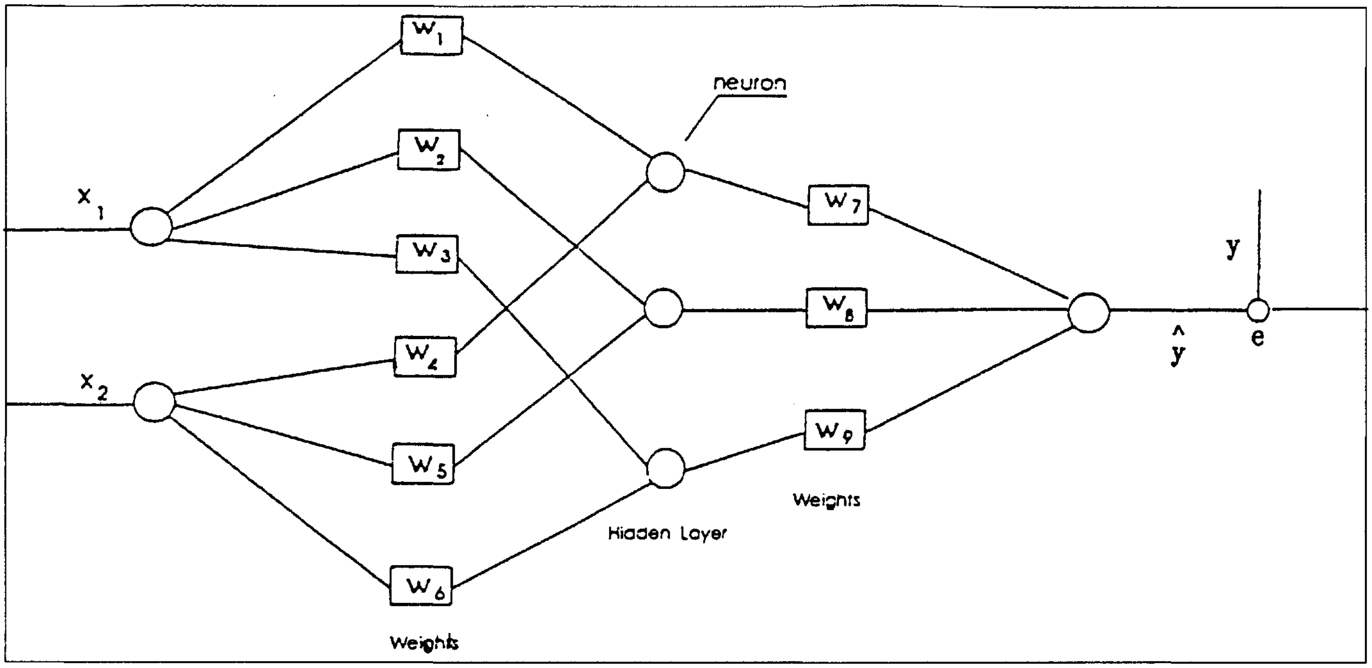

FIGURE 1. A neural network model with three layers and nine weights.

and the actual output of the unknown relationship, y, is defined as network error e given by

$$\begin{aligned} e &= y-\hat{y} \\ &= f(x_1,x_2) - g(x_1,x_2,w_1,\ldots,w9) \end{aligned} \tag{3}$$

The weights ($w_1,\ldots,w9$) must be determined in such a way that the error e is as small as possible. This task is known as the neural network training.

Certain features of the network emerge from this simple example. The number of neurons and the number of layers determine how rich our neural network model is. For example, if we add an additional middle layer to our network, the number of weights increases from 9 to 18. The complexity of the nonlinear function g(.) can also be changed by the type of neuron in the model. A typical structure of the neuron is shone (Fig. 2). The neuron output, z, can be expressed as

$$z = h\left(\sum_{i=1}^{m} w_i u_i\right) \tag{4}$$

where u_i is the output of the neuron located on the backward layer, w_i is the weight associated with this connection and m is the total number of neurons connected. One common choice for the function h(.) is the sigmoid function.

$$h(v) = 1/(1+\exp(-\lambda v)) \tag{5}$$

where $\lambda > 0$ is the steepness factor.

This choice offers the advantage that the function is bounded and its derivatives can be evaluated in terms of h(.).

To train the network a training data set and a suitable minimization algorithm are required. Suppose that a set of N observations [$x_1(j),x_2(j),y(j)$; $j=1,2,\ldots N$] is available. As in Eqn. (3), the neural network modeling error can be determined as

$$e(j) = f(x_1(j),x_2(j))-g(x_1(j),x_2(j),w_1,\ldots,w_9) \tag{6}$$

To minimize this error, the following loss function is defined

$$V(j) = e^2(j) \tag{7}$$

A back propagation learning methodology minimizes the loss function with the gradient descent algorithm (Baum 1986). At each iteration j the weights are adjusted according to

$$w_i(j+1) = w_i(j) - \mu \left.\frac{\partial V(j)}{\partial w}\right|_{w=w_i} \qquad i=1,2,\ldots,9 \tag{8}$$

where μ is the user-specified learning rate and $\frac{\partial V}{\partial w}\big|\, w=w_i$ is the gradient of the loss function evaluated at w_i. The

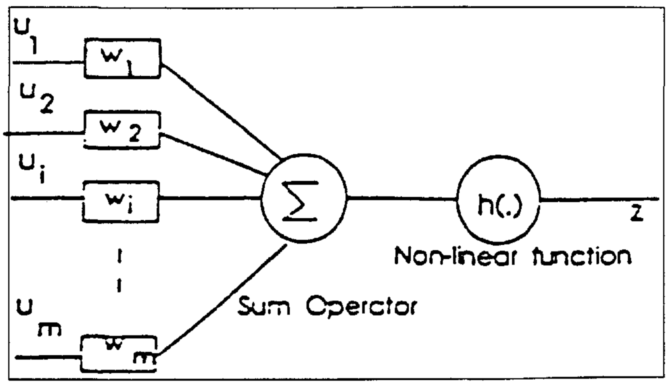

FIGURE 2. A typical structure of neuron.

gradient descent algorithm (Barto and Jordan 1987, Parker 1987) is essentially a simplified version of the celebrated Newton algorithm in a sense that the learning rate coefficient, μ, simply replaces a more complicated inverse Hessian matrix. The basic advantage of the algorithm is its simplicity. Furthermore, because of the network structure, the gradient term can be determined in terms of locally known quantities. However, it is also known that the convergence rate of the algorithm may be very slow close to a local minimum.

It is obvious from this simple example that a neural network is a nonlinear approximation function. Therefore, the performance of a neural network based model depends on how good the chosen model structure is as an approximator. In addition, the limitation of the gradient descent algorithm also determines the quality of approximation. It is assumed that any nonlinear relation can be approximated by a neural network to a given accuracy provided that the network has a sufficiently rich structure. An exact definition of rich structure is not specified. A three-layered network structure is the most commonly used neural network architecture.

The basic advantage of using a neural network based model is its flexibility with respect to model structure. Nonlinear models with increasing complexity can be generated by simply increasing the number of layers, the number of neurons, or the type of nonlinear function associated with the neuron (Lapedes and Farber 1988). This ability to adjust the model is the essential motivation behind modeling the energy data in terms of a neural network based model.

Bias Free Estimate

The power and potential of neural network modeling are suggested by noting that these models have been applied to problems as diverse as the following: stimulus-response experiments in psychology, machine health (using vibrations to monitor tolerance, for example) in automated factories, modeling brain function in physiology, scheduling work flow through job shops, and prediction of industrial bond ratings (Dutta and Shekham 1988, Amari 1967).

In all of these applications only inputs to the neural network model are required to estimate future outputs once learning has occurred. The data for inputs and outputs are considered to be bias free. The algorithm that models the pathways established by previous measured inputs to corresponding measured outputs provides a method to estimate future outputs from alternative inputs. An algorithm that computes the solutions (outputs) to a set of inputs is tantamount to a formula that connects inputs to outputs. Because prior sets of corresponding inputs and outputs represent bias-free samples, all of the power of statistical analysis is appropriate. When an algorithm works for known sets of inputs and outputs, it is presumed that running a given new set of inputs through the algorithm will yield a correct solution set of outputs.

At each node the output is computed from local information. The information required is the estimate of the node and the values of the inputs into the node. The salient point is that all of the information needed for a neuron to fire is local information; that is, available at the neuron (node). When a neuron fires, the very same signal is transmitted to all neurons in the next higher layer. For this reason an artificial neural net operation is also called parallel distributed processing.

Several aspects of this study require comment. The major purpose of the computation is to get a sense for the likely result of policy changes in energy markets. This implies that some inputs will have values outside the limits used in training the neural net. This is unavoidable, given the objective to clarify relationships among fuels, pollution, and policy. Scenarios that are likely to occur as a result of legislative action are simply unprecedented. Moreover, all models, not only neural networks, require a caveat of caution under these circumstances.

TRAINING OF THE NEURAL NET

The backpropagation network (shown in Fig. 3) is trained by the gradient descent algorithm as described by Eqn. (7) and is modified to

$$V(j) = \sum_{i=1}^{12} e_i^2 (j)$$

where $e_i(j)$ is the error associated with output (i) at iteration (j). The network has 5 inputs, 12 outputs, 2 hidden layers, and each hidden layer has 20 neurons. This network structure gives rise to 340 weights that need to be learned during this training process. The five inputs to the network are coal, gas, petroleum, nuclear, and hydro-electric energy sources. Following the notation of Eqn. 1 the input vector is

x = (Coal, Gas, Petroleum, Nuclear, Hydro-electric)

The outputs used during the training are the residential, industrial, and transportation usage of coal, gas, petroleum, and electricity energy sources. These are listed in vector y as:

y = (Residential, Commercial, Industrial, Transportation)

Note that each element in y vector has four different data points coming from coal, gas, petroleum, and electricity usage. A total of 18 past data points are available for inputs and outputs. Therefore, the training data can be expressed as

Training data = $(x\ y)_i$ = 1,2,...18

The training was accomplished using a commercial software package (NeuralWare 1993). During training, the data were continuously fed to the network and the weights were updated following the gradient descent algorithm. The learning parameter was initially 0.005. However, the NeuralWare software package gradually decreases this parameter for accuracy in convergence. The steepness parameter of the sigmoid function was 0.01. The learning algorithm was stopped when the loss function reached the value of 0.05. The final neural network structure presented in the paper was obtained by a trial and error procedure. Due to a limited amount of data,

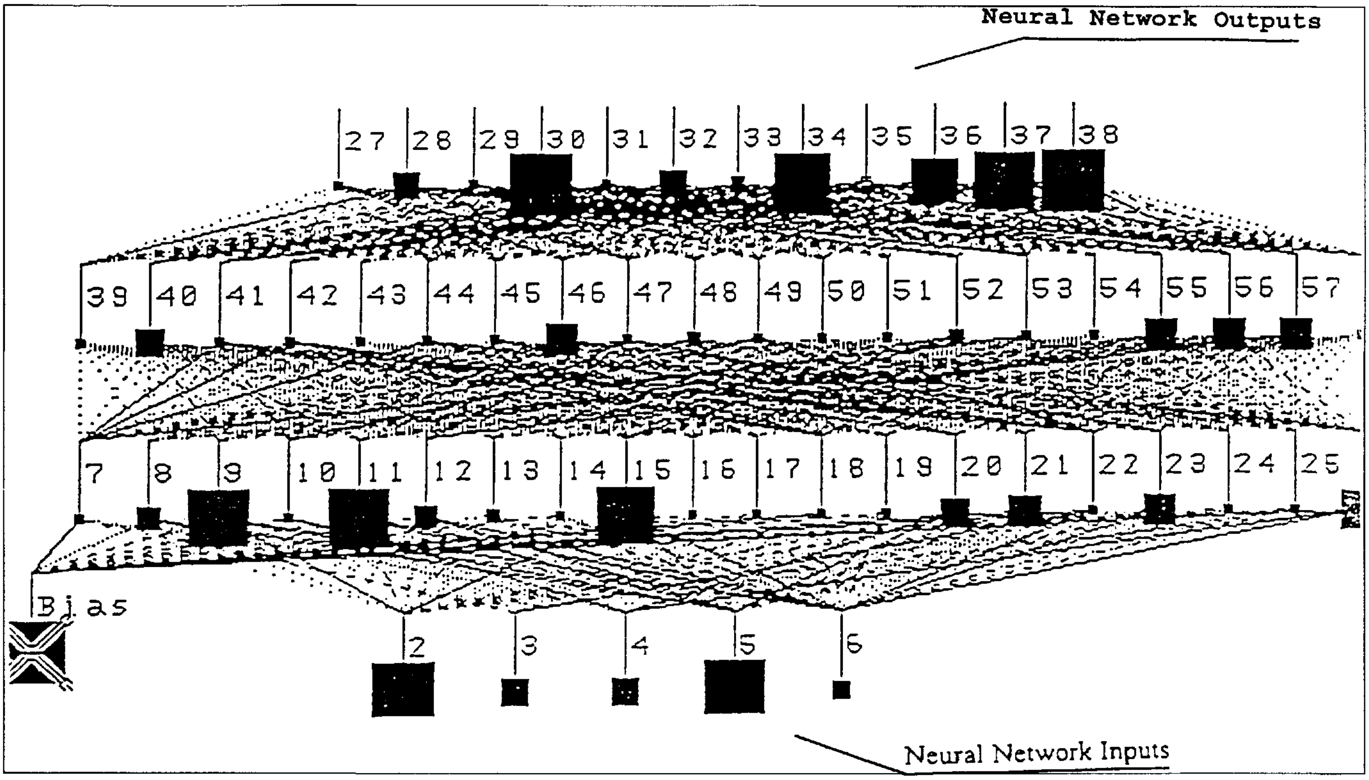

FIGURE 3. A neural network with two hidden layers is used to predict the outcomes corresponding to different scenarios. The size of the black boxes indicates the output magnitude of each neuron at a given time.

the training data set was not divided into two in order to use the first part for training and the second part for evaluation.

RESULTS AND DISCUSSION

Seven scenarios were specified in order to anticipate a range of possible energy strategies that could characterize the United States economy. Energy strategies are all dependent on legislative and administrative action by federal and state agencies. Legislation that defines the United States energy strategy is highly dependent on perceived environmental problems, international balance of payments considerations, farm policy issues such as requirements to add ethanol (manufactured from grain) to transport fuel for use in some of our largest urban centers, and resource utilization requirements or incentives. Tax credit for the use of Ohio coal in Ohio electric generating plants is an example of an incentive for a strategic energy policy to expand the market for high sulfur Ohio coal. Other strategies include the phase-out of nuclear power and the wider use of natural gas to reflect changing relative costs that more closely reflect total social cost. Each scenario is defined by corresponding inputs, which are compared with the measured inputs for the year 1992. This provides a standard for comparison (Fig. 4).

After the net has learned, that is, has been trained (in the sense that the algorithm had to compute a set of weights that accurately relates inputs to outputs), seven sets of input data for the various scenarios based on possible energy policy mandates were processed through the neural net in order to estimate the likely effects on outputs. The neural network used in the analysis of the scenarios is shown (Fig. 3).

SCENARIO 1; Acid Rain Reduction: Because of SO_2 pollution the coal input was reduced approximately 25% to 9.73 quadrillion Btu (Quads) and natural gas inputs were increased to 25.74 Quads. The total Btu requirement for the system was not materially changed (Fig. 5). As expected, industrial coal usage decreased (14%). Gas usage by industry was reduced (12%). Electricity production increased very slightly and petroleum usage very slightly declined. The disruptive effects of SCENARIO 1 policy are minimal and the conversion coefficient remains at the same level, approximately 81% (Fig. 5).

SCENARIO 2; Population Explosion: Each energy input was doubled to correspond roughly to the doubling of U.S. population by the year 2050 when world population will be 10+ billion and U.S. population will be around 500 million. (These numbers were based on rough trends for the last 50 years.) Relative share of inputs remained constant (Fig. 6). The neural net solution indicated that the result of this scenario would be disastrous. The energy conversion coefficient dropped from 81% to 48%. The output of the neural network may not be reliable for this 60 year time horizon. Nevertheless, it is clear that long term expansion of the economy must not be based on constant, non-changing production functions. The inefficiencies would be intolerable (Fig. 6).

SCENARIO 3; Tailpipe Pollution Reduction: The dependence of the United States economy on oil for automobile and truck transportation has already caused severe tailpipe pollution in our great urban centers.

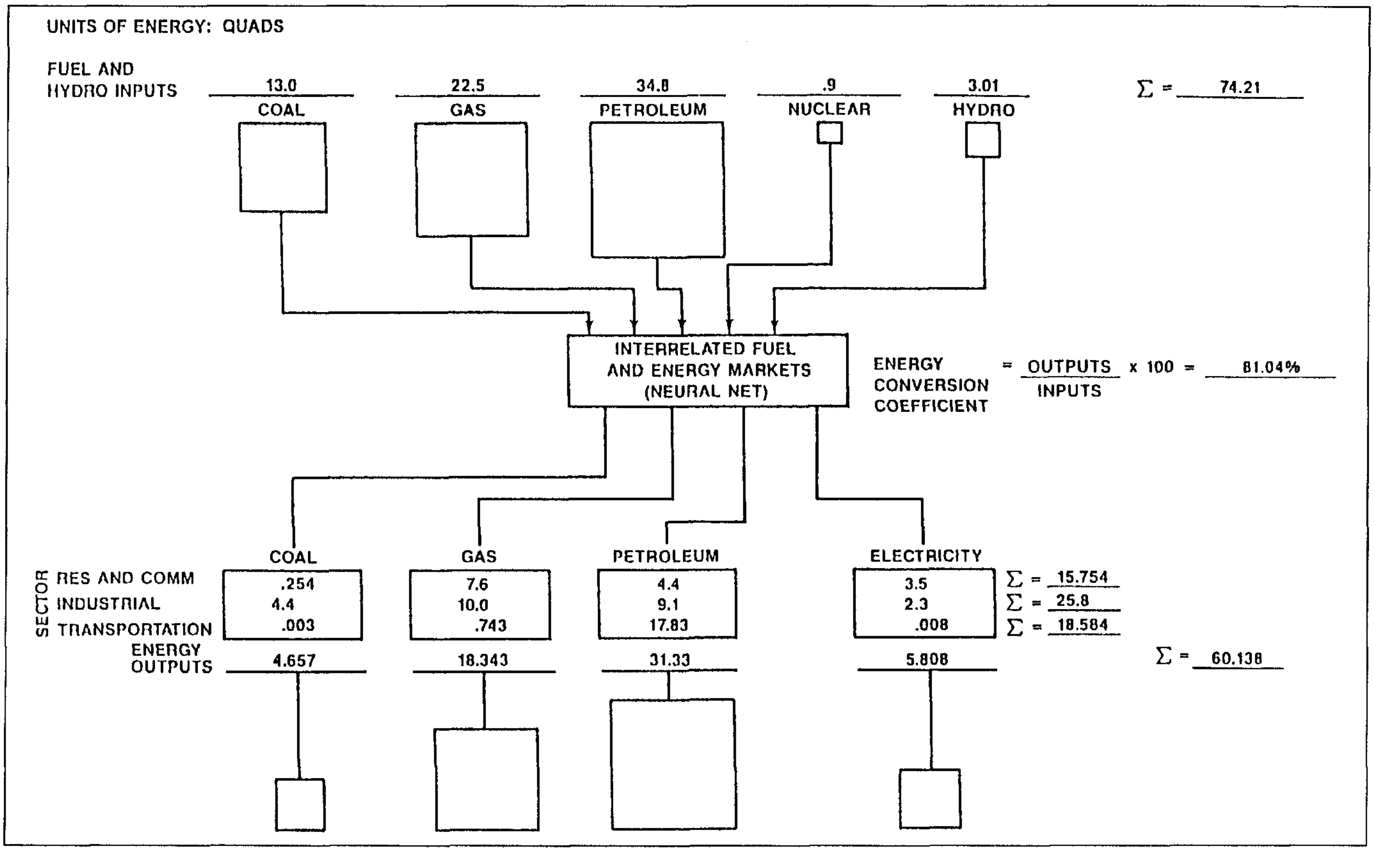

Figure 4. Scenario 1992 Actual. The Standard.

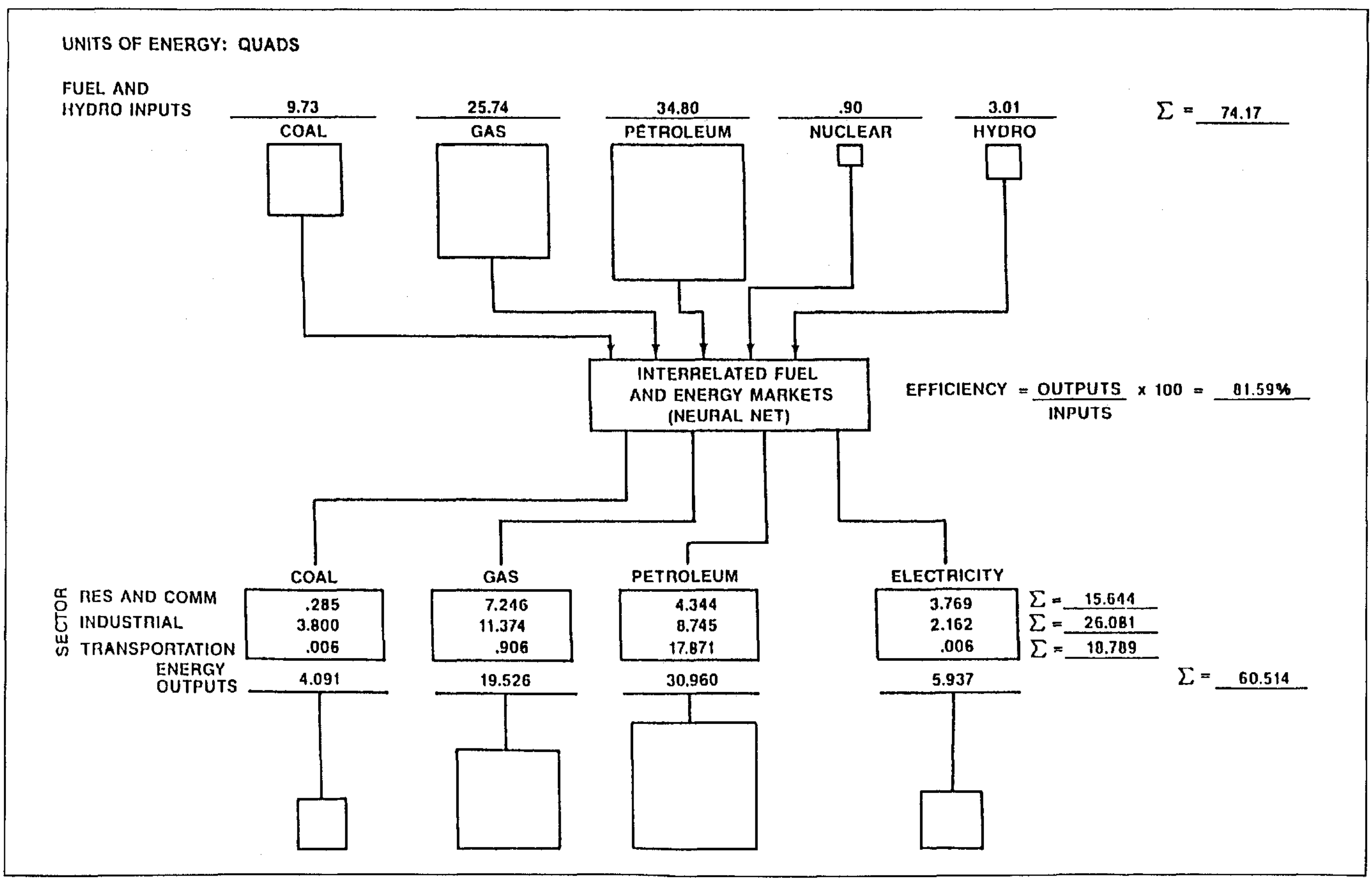

Figure 5. Scenario 1.

Petroleum inputs were reduced by 5.0 Quads and natural gas inputs were increased by the same amount (Fig. 7). Petroleum usage by the industrial sector was increased (13%), and petroleum usage by the transportation sector was reduced (7%). The conversion coefficient actually increased by one percentage point. This scenario suggests a reasonable policy to increase petroleum for industrial energy supply while reducing energy supplied (liquid fuel) to the transportation sector (Fig. 7).

Scenario 4; Greenhouse Gas Stabilization: Energy consumption as an input was increased by 15%, all of which was produced by nuclear technology. This scenario models a long run approach to environmental problems caused by unwanted greenhouse gases that are by-products of electric generating plants that use fossil fuel inputs (Fig. 8). (Nuclear plants do not produce products of combustion such as carbon dioxide and sulfur dioxide.) Electricity energy outputs increased 59% and all other energy outputs except coal increased as well. Electricity was substituted for other forms of energy used by the residential and commercial sector, releasing both gas and petroleum to the industrial sector. Also, substantially more electricity was consumed by the industrial sector (53%). The conversion coefficient dropped from 81% to 74%. This probably reflects the low thermal efficiency of nuclear plants and the disruptions entailed in the increase of nuclear inputs (13 times the 1992 level) (Fig. 8).

Scenario 5; Clean Coal Technologies: Coal inputs were doubled to model breakthroughs in clean coal use technology. The growth in energy usage is all accounted for by coal (Fig. 9). The effect was a dramatic increase in electricity output (61%). In this respect dramatic increase in coal input had the same effect as dramatic increase in nuclear input. The effect was different in the using sectors. The increase in coal fuel input caused a substantial increase in residential and commercial usage of electricity (68%). Also, gas consumption in the residential and commercial sector increased substantially (41%). Petroleum usage by the transport sector increases significantly (24%). This scenario suggests a strategy for a mature society where energy usage by the industrial sector is actually decreasing, because of change in product mix as a region's real income increases (less heavy industry), and the implementation of energy saving processes in production is taking place. The energy conversion coefficient dropped from 81% to 74% (Fig. 9).

Scenario 6; Short Run Expansion: All energy inputs increased 10%. This scenario models a growing economy without any relative change in input structure (Fig. 10). The outputs increased in a somewhat similar way. All using sectors accounted for more of the energy outputs (residential and commercial use increased 12%; industrial usage increased 16%; and transportation use increased 1%). More gas was used in the residential and commercial sector (24%); and more petroleum was used in the industrial sector (35%). The petroleum usage by the transportation sector remained almost constant. The energy conversion coefficient also remained almost constant (Fig. 10).

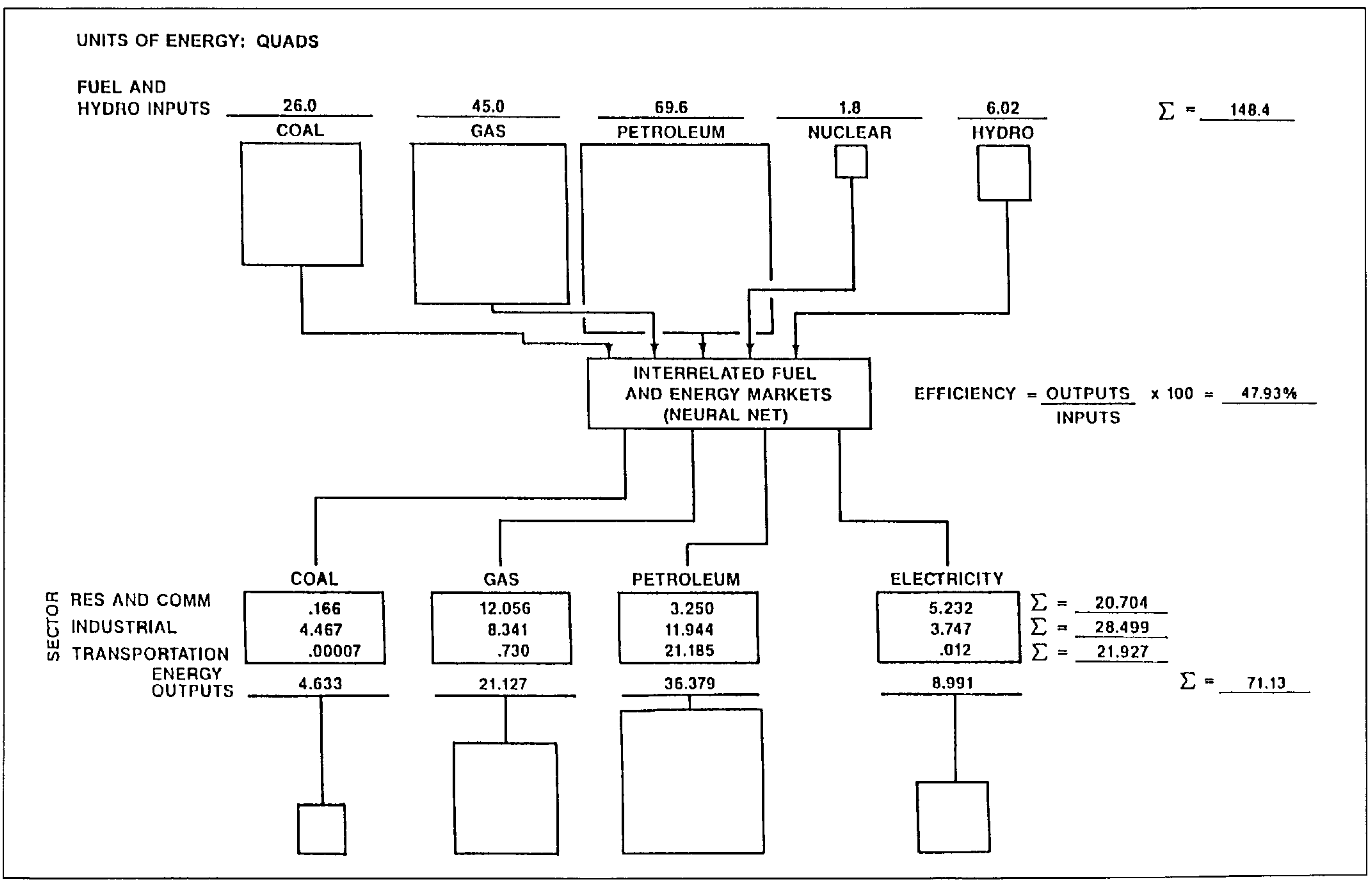

Figure 6. Scenario 2.

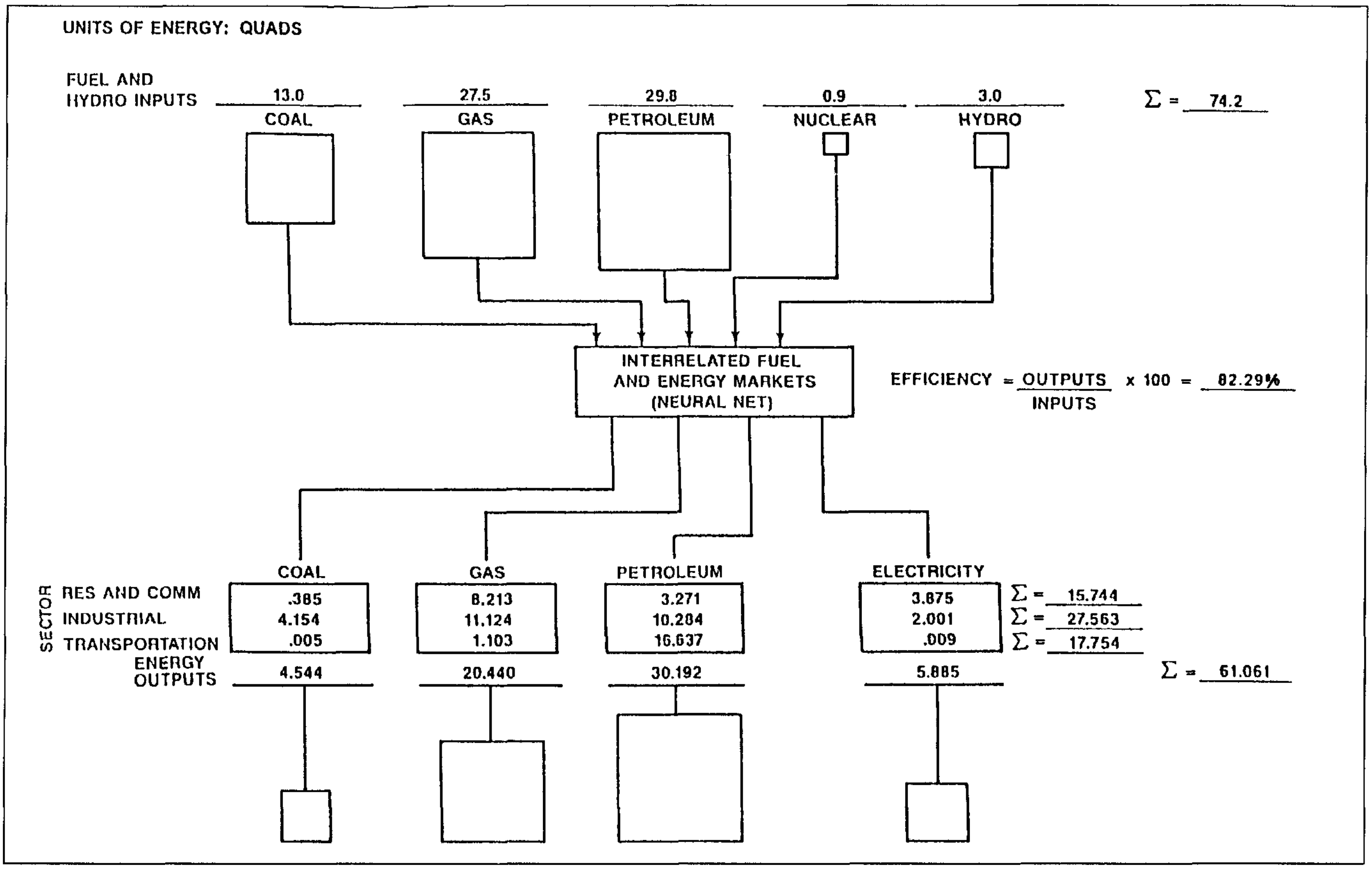

FIGURE 7. SCENARIO 3.

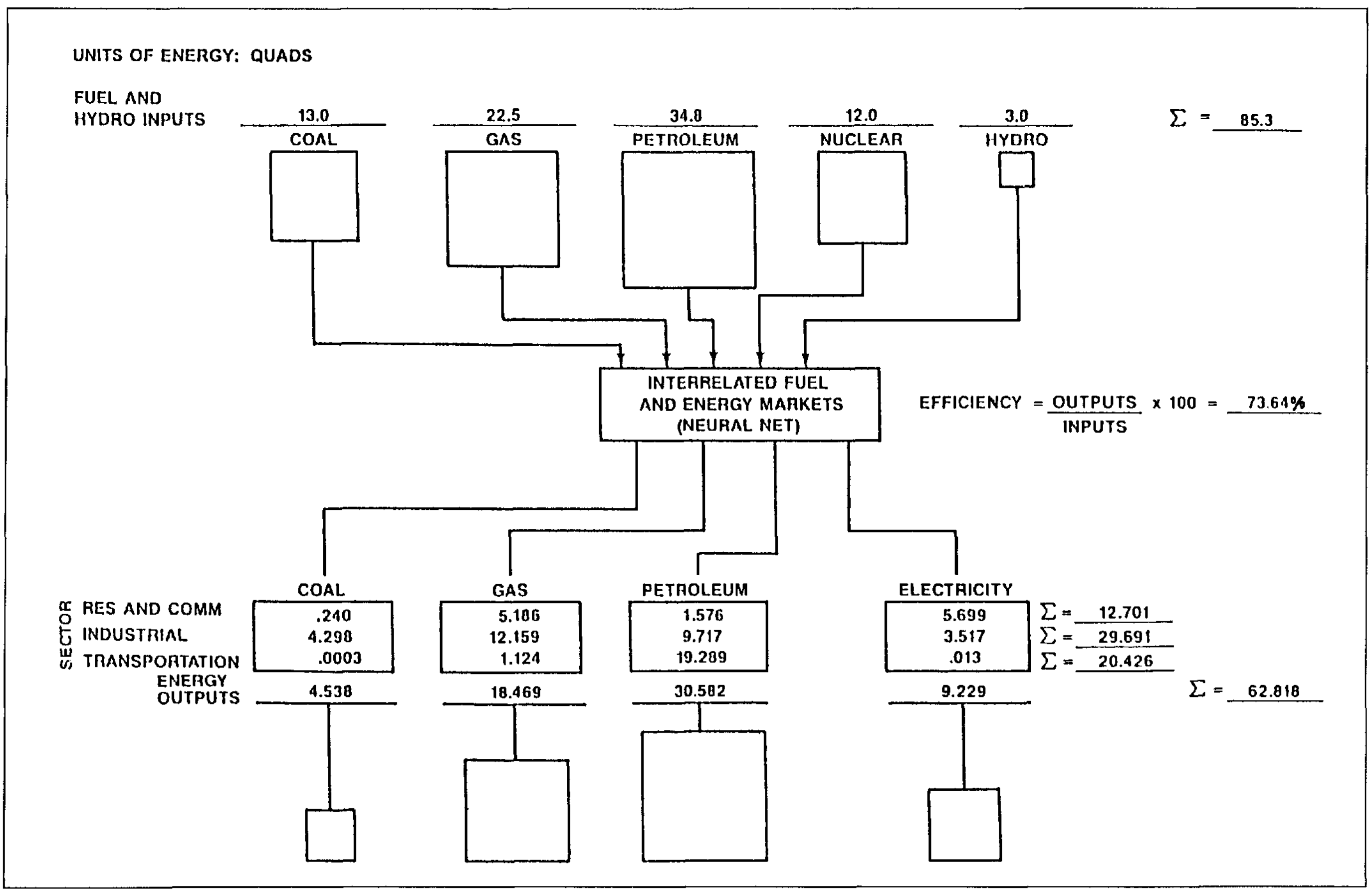

FIGURE 8. SCENARIO 4.

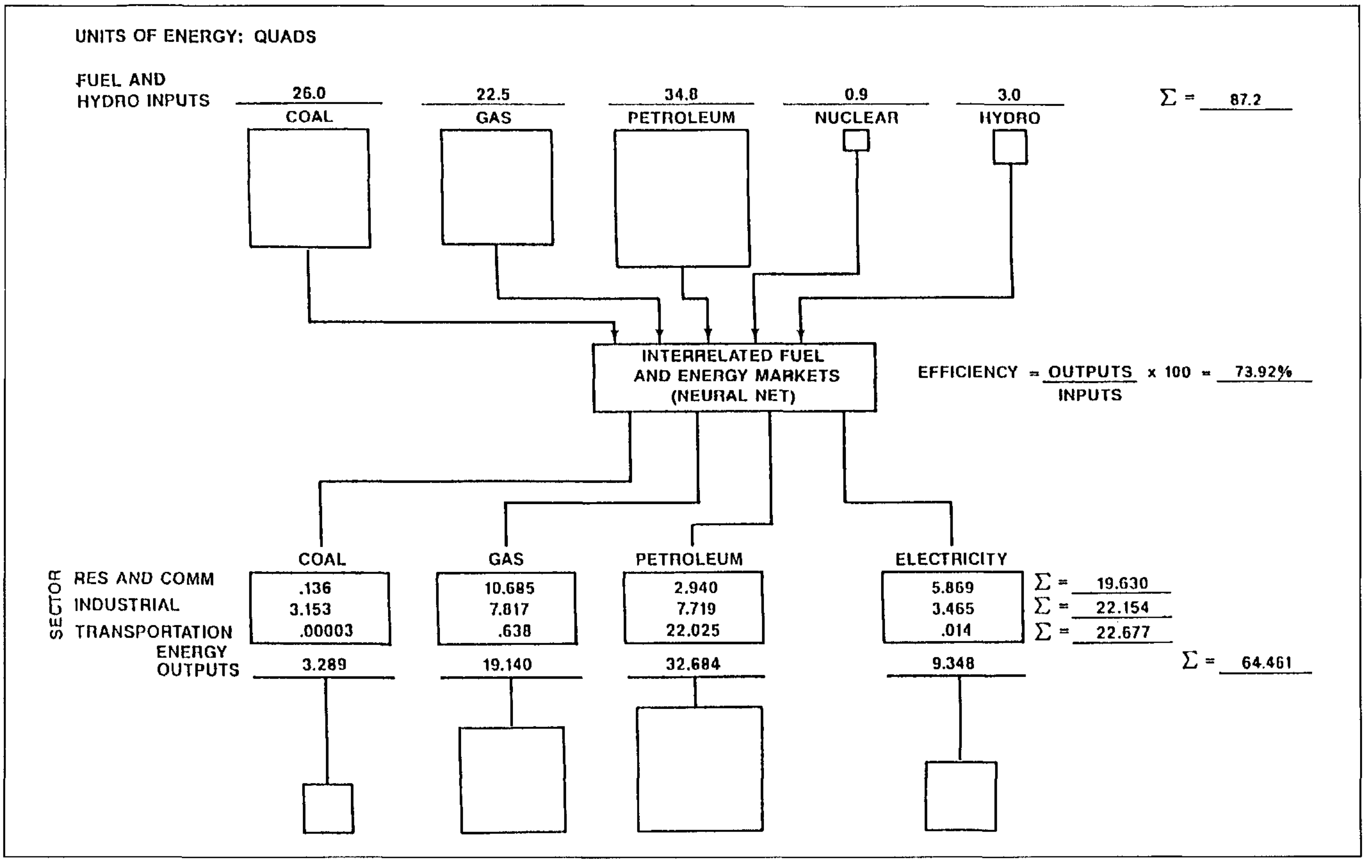

FIGURE 9. SCENARIO 5.

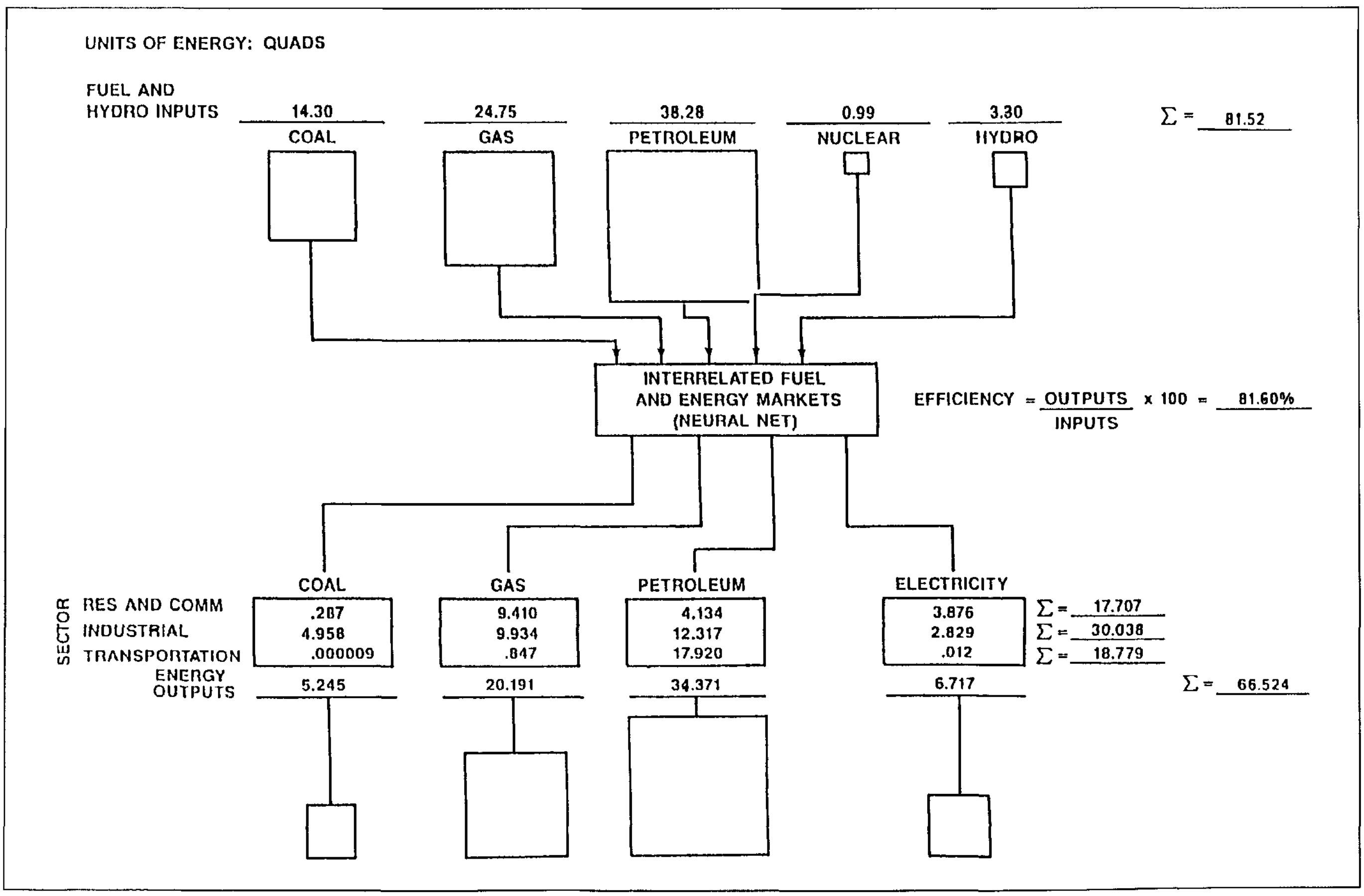

FIGURE 10. SCENARIO 6.

SCENARIO 7; Nuclear Phase-out and Coal Expansion: In this scenario environmental concerns, political gridlock, and increasing relative costs phase-out nuclear electric generating plants. Coal fuel inputs were increased from 13.0 Quads to 33.44 Quads. Petroleum inputs were reduced 12%, natural gas usage was increased by 10%, and hydro inputs were increased 10% also. This is a possible long term scenario that is not dependent on nuclear power and is heavily dependent on coal inputs (Fig. 11). The result was that electricity production increased 78%. Energy usage increased in the residential and commercial sector (21%) and also increased in the transportation sector (17%). Energy usage in the industrial sector actually declined (12%). More gas and substantially more electricity were used in the combined residential and commercial sector. Also more petroleum was used in the transportation sector (18%). This scenario suggests a high consumption economy with successful energy conservation by industrial firms. The energy conversion coefficient dropped from 81% to 69%. This is a growth scenario that depends on coal for the energy growth component along with a conscious effort to conserve petroleum (Fig. 11).

Overall Implications

Very large proportional increases in coal use and very large proportional increases in nuclear power generation have some similar economic effects. A substantial increase in the generation of electricity is the most obvious adaptation expressed by the neural net. This causes a substitution effect in the energy markets and more electricity is used, especially in the residential and commercial markets. This substitution effect shifts some petroleum out of the residential and commercial sector into the transportation sector allowing actual increases in energy use for transport. All current (state-of-practice) transportation engine technologies require liquid fuel to facilitate distribution of the fuel within the engine.

Substitution of natural gas for other fuels, indicated by a proportional increase in gas-fuel inputs, causes very little disturbance to the fuel-energy system. This is important because the trend to internalize environmental costs favors the expansion of natural gas inputs. Also, across-the-board increases in fuel inputs on the order of 10% with proportions held constant caused only modest adaptations in the using sectors market. The coefficient of energy conversion remained high under both of these conditions. On the other hand, doubling of all fuel (and hydro) inputs appeared to have unacceptable wasteful effects on the energy conversion coefficient. If fuel inputs are doubled in the long term (say 50 years), production functions must be altered to adapt to the changing conditions. The neural net analysis implies that any effort to maintain linear homogeneous production functions in the long run will be disastrous. The analysis lends weight to the notion that population growth and economic growth require continuous adaptations in all production activities to avoid unacceptable consequences in fuel and energy use.

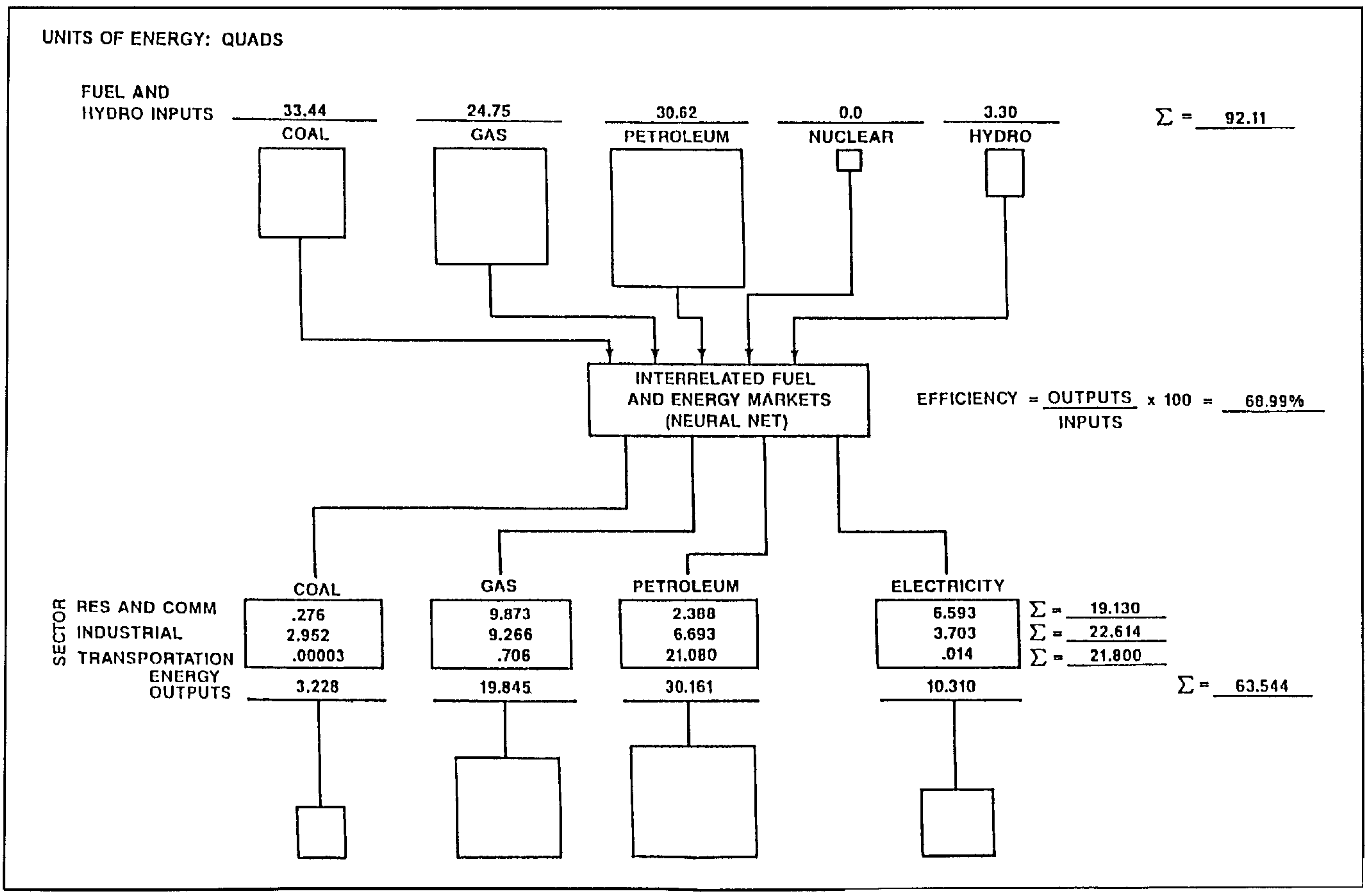

FIGURE 11. SCENARIO 7.

Neural network models are presumed to correspond in a rough way with the patterns of information flow through the neural system of the brain (Rumelhart and McClelland 1986). Various assumptions regarding the signal flow through the neural net correspond to various models of network behavior. All of the models have in common the idea of a black box with an internal mechanism to convert input signals to output signals. Data records of inputs and outputs of a process are used as unbiased samples that reflect the internal mechanism of the process.

The iterations during the training phase correspond to optimization by trial and error episodes (Ramanujam and Sadayappan 1988). Almost any arrangement of neurons and pathways that is sufficiently complex could be used to estimate the weights, provided certain mathematical conditions are met (Sietsma and Dow 1991). If there is a strong functional relationship, many very complex mazes will work. The same set of inputs will traverse alternative pathways to and through the layers of internal neurons, finally producing the set of outputs that correctly corresponds to the set of inputs by using the gradient descent algorithm to minimize the error. If there is not a strong functional relationship, no maze will work.

In this research the five energy inputs must be consistent with the 12 energy outputs. Our scenarios are equivalent to solving a set of 17 nonlinear equations. The neural network methodology, the black box approach, is valid because there is a correspondence with reality, assured by the neural net training. We presume that true answers are obtained via simulation using the trained neural network.

LITERATURE CITED

Amari, S-I. 1967 A theory of adaptive pattern classifiers. IEEE Transactions on Electronic Computers. EC-16: 299-307.

Barto, A. and M. Jordan 1987 Gradient following without backpropagation in layered networks. Proceedings of the IEEE First International Conference on Neural Networks, Vol. II. pp. 629-636.

Baum, E. 1986 Generalizing backpropagation to computation. *In:* J. Denker (ed.), AIP Conference Proceedings 151: Neural Networks for Computing. American Institute of Physics, New York, NY. pp. 47-52.

Dutta, S. and S. Shekham 1988 Bond rating: A non-conservative application of neural networks. Proceedings of the IEEE International Conference on Neural Networks, Vol. II. pp. 443-458.

Energy Information Administration 1991 Section 2: Energy Consumption. Monthly Energy Review. December. pp. 21-37.

Kosko, B. 1988 Feedback stability and unsupervised learning. Proceedings of the 1987 IEEE Conferences on Neural Networks, Vol. I. pp. 141-152.

___ 1992 Neural Networks and Fuzzy-Systems. Prentice-Hall, Englewood Cliffs, NJ. pp. 181-188.

Lapedes, A. and R. Farber 1988 How neural nets work. *In:* D. Anderson (ed.), Proceedings of the 1987 IEEE Conference on Neural Processing Systems—Natural and Synthetic. American Institute of Physics, New York, NY. pp. 442-456.

Moody, J. and J. Denker 1988 Speedy alternatives to backpropagation. Neural Networks Supplement: INNS Abstracts. 1: 202.

NeuralWare 1993 Pittsburgh, PA.

Parker, D. 1987 Optimal algorithms for adaptive networks: Second order backpropagation, second order direct propagation, and second order Hebbian learning. Proceedings of the IEEE First International Conference on Neural Networks, Vol. II. pp. 593-600.

Rumelhart, D. and J. McClelland 1986 Parallel Distributed Processing Explorations in the Micro-structure of Cognition. Vols. 1 and 2. Bradford Books/MIT Press, Cambridge, MA. pp. 318-362.

Ramanujam, J. and P. Sadayappan 1988 Optimization by neural networks. Proceedings of the IEEE International Conference on Neural Networks, Vol. II. pp. 325-332.

Sietsma, J. and R. J. F. Dow 1991 Neural net pruning—why and how. Proceedings of the IEEE International Conference on Neural Networks, Vol. I. pp. 325-333.