Predictive Modeling of Mercury Speciation in Combustion Flue Gases Using GMDH-Based Abductive Networks

R. E. Abdel-Aal

Department of Computer Engineering,
King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia

Abstract:

Modeling mercury speciation is an important requirement for estimating harmful emissions from coal-fired power plants and developing strategies to reduce them. First principle models based on chemical, kinetic, and thermodynamic aspects exist, but these are complex and difficult to develop. The use of modern data-based machine learning techniques has been recently introduced, including neural networks. Here we propose an alternative approach using abductive networks based on the group method of data handling (GMDH) algorithm, with the advantages of simplified and more automated model synthesis, automatic selection of significant inputs, and more transparent input-output model relationships. Models were developed for predicting three types of mercury speciation (elemental, oxidized, and particulate) using a small data set containing six inputs parameters on the composition of the coal used and boiler operating conditions. Prediction performance compares favourably with neural network models developed using the same dataset, with correlation coefficients as high as 0.97 for training data. Network committees (ensembles) are proposed as a means of improving prediction accuracy, and suggestions are made for future work to further improve performance.

Index Terms: Mercury speciation, Flue gases, Boiler emissions, Predictive modeling, inferential emission monitoring, Soft sensors, Abductive networks, GMDH algorithm, Neural networks, Network committees, Network ensembles.

Dr. R. E. Abdel-Aal,
P. O. Box 1759,
KFUPM, Dhahran 31261
Saudi Arabia
e-mail: radwan@kfupm.edu.sa
Phone: +966 3 860 4320, Fax: +966 3 860 3059

1. Introduction

Determining Mercury speciation is an important requirement for estimating mercury emission from combustion flue gases and the efficiency of control measures to reduce it. Major mercury sources from human activities are coal-fired electric utility boilers, where speciation depends on the operating conditions, including the type of coal used and flue gas temperature and composition. Mercury compounds from combustion sources consist mainly of gaseous elemental mercury (Hg^0), gaseous oxidized mercury (Hg^{2+}), and particle-bound mercury (Hg_p) [1]. Theoretical first-principle approaches have been used to study mercury speciation, including kinetic modeling [2] and thermodynamic equilibrium calculations [3]. Theoretical models rely on knowledge of the processes involved, which are often quite complex and highly nonlinear and therefore are difficult to describe accurately. Model development can also be expensive and time consuming. In addition, the above methods suffer from limited accuracy due to the lack of accurate rate constants of reaction mechanisms and to uncertainties caused by model assumptions and simplifications and incomplete understanding of mercury science [4]. Recently, data based modeling using machine learning techniques, such as neural networks, fuzzy logic, and genetic algorithms, has become a popular approach for solving complex non-linear problems without requiring exhaustive theoretical knowledge of the phenomenon being modeled. Such approaches depend primarily on experimental input-output data on the process, which are usually readily available in large quantities, rather than accurate theoretical knowledge. A model for the phenomenon considered is developed through training on input-output process data in the form of an adequate number of solved examples. Once synthesized, the model can be used to perform fast predictions of outputs corresponding to new cases previously unseen during training. The method offers a number of advantages over conventional approaches, including increased tolerance to noise and uncertainty, reduced need for knowledge on the modeled phenomenon, and the relative ease of developing and updating the model. In the last few years, neural networks have formed the basis of many soft (inferential) sensors for monitoring pollutant emissions [5-9]. Such sensors offer a cost-effective and reliable alternative

to expensive online analyzers in many areas of application, including mercury monitoring. Tsiros and Dimopoulos have used neural networks and other statistical and machine learning techniques to model atmospheric emission of gaseous soil mercury and identify critical factors for controlling such emissions [10]. Neural networks were used together with optical fiber chemical sensors for monitoring mercury and other heavy metals in aquatic samples [11]. Neural networks were used for modeling a spectrophotometric kinetic system and optimizing the experimental conditions for measuring traces of mercury in water [12]. Applications of neural networks to modeling mercury speciation in flue gases are relatively scarce in the literature, and only the work by Jensen et al. [13] could be cited at the time of writing.

In general, the neural network approach suffers from a number of limitations, including difficulty in determining optimum network topology and training parameters [14]. There are many choices to be made in determining numerous critical design parameters with little guidance available [15], and designers often resort to trial and error approaches which can be tedious and time consuming [16,17]. Such design parameters include the number and size of the hidden layers, the type of neuron transfer functions for the various layers, the learning rate and momentum coefficient, and training stopping criteria to avoid over-fitting and ensure adequate generalization with new data. Another limitation is the black box nature of neural network models that give little insight into the modeled relationship and the relative significance of various inputs, thus providing poor explanation facilities [18]. The acceptability of, and confidence in, automated prediction tools in areas such as electric load forecasting, pollution control and medical diagnosis are related to their transparency and their ability to justify results to human operators, experts and decision makers [19]. To overcome such limitations, we propose using self-organized abductive networks [20] based on the group method of data handling (GMDH) learning algorithm [21,22] as an alternative machine learning approach to modeling and estimating mercury speciation in the flue gasses of coal-fired power plants. We have previously used this approach in several weather prediction applications including

modeling and forecasting the minimum [23] and maximum [24] daily temperatures and the hourly temperature profile [25]. Compared to neural networks, abductive networks offer the advantages of faster model development requiring little or no user intervention, faster convergence during model synthesis without the problem of getting stuck in local minima, automatic selection of effective input variables, and automatic configuration of the model structure [14]. With the model represented as a hierarchy of polynomial expressions, resulting analytical model relationships can provide insight into the modeled phenomena, highlight contributions of various inputs, and allow comparison with previously used empirical or statistical models. The technique automatically avoids over-fitting by using a proven regularization criterion based on penalizing model complexity [22] without requiring a dedicated validation dataset during training, as is the case with many neural network paradigms.

Following a brief description of abductive networks and the underlying GMDH learning algorithm in Section 2, the mercury speciation dataset used in this study is described in Section 3. In Section 4, abductive network models for the three types of mercury speciation are described and their performance on both the training and evaluation sets is analyzed and compared with neural network results reported in the literature for the same dataset. Single (monolithic) abductive models of various levels of model complexity are presented. Modular network committees (ensembles) are also introduced as a means of improving prediction accuracy beyond that obtained with the monolithic models. Section 5 includes conclusions and suggestions for future work.

2. GMDH and AIM Abductive Networks

AIM (abductory inductive mechanism) [26] is a supervised inductive machine-learning tool for automatically synthesizing abductive network models from a database of inputs and outputs representing a training set of solved examples. As a GMDH algorithm, the tool can automatically synthesize adequate models that embody the inherent structure of complex and highly nonlinear systems. Automation of model synthesis not only lessens the burden on the analyst but also safeguards the model generated against influence by human biases and misjudgments. The GMDH approach is a formalized paradigm for iterated (multi-phase) polynomial regression capable of producing a high-degree polynomial model in effective predictors. The process is 'evolutionary' in nature, using initially simple (myopic) regression relationships to derive more accurate representations in the next iteration. To prevent exponential growth and limit model complexity, the algorithm selects only relationships having good predicting powers within each phase. Iteration is stopped when the new generation regression equations start to have poorer prediction performance than those of the previous generation, at which point the model starts to become overspecialized and therefore unlikely to perform well with new data. The algorithm has three main elements: representation, selection, and stopping. It applies abduction heuristics for making decisions concerning some or all of these three aspects.

To illustrate these steps for the classical GMDH approach, consider an estimation data base of n_e observations (rows) and $m+1$ columns for m independent variables (x_1, x_2, ..., x_m) and one dependent variable y. In the first iteration we assume that our predictors are the actual input variables. The initial rough prediction equations are derived by taking each pair of input variables (x_i, x_j ; $i,j = 1,2,...,m$) together with the output y and computing the quadratic regression polynomial [21]:

$$y = A + B\,x_i + C\,x_j + D\,x_i^2 + E\,x_j^2 + F\,x_i\,x_j \qquad (1)$$

Each of the resulting *m(m-1)/2* polynomials is evaluated using data for the pair of *x* variables used to generate it, thus producing new estimation variables (z_1, z_2, ..., $z_{m(m-1)/2}$) which would be expected to describe *y* better than the original variables. The resulting *z* variables are screened according to some selection criterion and only those having good predicting power are kept. The original GMDH algorithm employs an additional and independent selection set of n_s observations for this purpose and uses the regularity selection criterion based on the root mean squared error r_k over that dataset, where:

$$r_k^{\,2} = \sum_{\ell=1}^{n_s} (y_\ell - z_{k\ell})^2 \Big/ \sum_{\ell=1}^{n_s} y_\ell^{\,2}; \quad k = 1,2,...,m(m-1)/2 \tag{2}$$

Only those polynomials (and associated *z* variables) that have r_k below a prescribed limit are kept and the minimum value, r_{min}, obtained for r_k is also saved. The selected *z* variables represent a new database for repeating the estimation and selection steps in the next iteration to derive a set of higher-level variables. At each iteration, r_{min} is compared with its previous value and the process is continued as long as r_{min} decreases or until a given model complexity is reached. An increasing r_{min} is an indication of the model becoming overly complex, thus over-fitting the estimation data and performing poorly on the new selection data. Keeping model complexity checked is an important aspect of GMDH-based algorithms, which keep an eye on the final objective of constructing the model, i.e. using it with new data previously unseen during training. The best model for this purpose is that providing the shortest description for the data available [22]. Computationally, the resulting GMDH model can be seen as a layered network of partial quadratic descriptor polynomials, each layer representing the results of an iteration.

A number of GMDH methods have been proposed which operate on the whole training dataset thus eliminating the need for a dedicated selection set. The adaptive learning network (ALN) approach, AIM being an example, uses the predicted squared error (PSE) criterion [22] for selection and stopping to avoid model overfitting, thus solving the problem of determining when

to stop training in neural networks. The criterion minimizes the expected squared error that would be obtained when the network is used for predicting new data. AIM expresses the *PSE* as:

$$PSE = FSE + CPM\,(2K/N)\sigma_p^{\,2} \tag{3}$$

where *FSE* is the fitting squared error on the training data, *CPM* is a complexity penalty multiplier selected by the user, *K* is the number of model coefficients, *N* is the number of samples in the training set, and $\sigma_p^{\,2}$ is a prior estimate for the variance of the error obtained with the unknown model. This estimate does not depend on the model being evaluated and is usually taken as half the variance of the dependent variable *y* [22]. As the model becomes more complex relative to the size of the training set, the second term increases linearly while the first term decreases. *PSE* goes through a minimum at the optimum model size that strikes a balance between accuracy and simplicity (exactness and generality). The user may optionally control this trade-off using the *CPM* parameter. Larger values than the default value of 1 lead to simpler models that are less accurate but may generalize well with previously unseen data, while lower values produce more complex networks that may overfit the training data and degrade actual prediction performance.

AIM builds networks consisting of various types of polynomial functional elements. The network size, element types, connectivity, and coefficients for the optimum model are automatically determined using well-proven optimization criteria, thus reducing the need for user intervention compared to neural networks. This simplifies model development and considerably reduces the learning/development time and effort. The models take the form of layered feed-forward abductive networks of functional elements (nodes) [26], see Fig. 1. Elements in the first layer operate on various combinations of the independent input variables (*x*'s) and the element in the final layer produces the predicted output for the dependent variable *y*. In addition to the main layers of the network, an input layer of normalizers convert the input

variables into an internal representation as *Z* scores with zero mean and unity variance, and an output unitizer unit restores the results to the original problem space. AIM supports the following main functional elements:

(i) A white element which consists of a constant plus the linear weighted sum of all outputs of the previous layer, i.e.

$$\text{"White" Output} = w_0 + w_1x_1 + w_2x_2 + w_3x_3 + ... + w_nx_n \quad (4)$$

where $x_1, x_2,..., x_n$ are the inputs to the element and $w_0, w_1, ..., w_n$ are the element weights.

(ii) Single, doublet, and triplet elements which implement a third-degree polynomial expression with all possible cross-terms for one, two, and three inputs respectively; for example,

$$\text{"Doublet" Output} = w_0 + w_1x_1 + w_2x_2 + w_3x_1^2 + w_4x_2^2 + w_5x_1x_2 + w_6x_1^3 + w_7x_2^3 \quad (5)$$

3. The dataset

We used a dataset of experimental measurements compiled by Jensen et al. [13]. The dataset consists of 82 samples of measurements, each representing a coal-fired power plant in the USA as part of the Information Collection Request (ICR) scheme run by the Environmental Protection Agency (EPA) in 1998 [13]. Each record contains six explanatory input variables representing the coal properties and the boiler operating conditions that are known to affect mercury speciation emission, as well as three outputs representing elemental, oxidized, and particulate mercury in the flue gas of the boiler. Table 1 describes all inputs and outputs in the dataset.

4. Abductive Network Models for Mercury Speciation

We have developed abductive network models for the relationship between each of the three types of mercury speciation and the six input parameters in the dataset. The 82 data samples were randomly split into 72 samples for model training and 10 samples for model evaluation. The effect of varying the complexity for single (monolithic) models on the prediction accuracy

with both the training and evaluation sets was investigated and models giving optimum performance were identified. We study the significance of input selections made automatically during model synthesis and relate such findings with those reported in the literature. The use of network committees (ensembles) consisting of different models was explored as a means of improving prediction performance beyond that obtained by single models.

4.1. Single Abductive Network Models for Mercury Speciation

Abductive network models were developed for each of the three types of mercury speciation (elemental (EL), oxidized (OX), and particulate (PA)). All models were trained on the 72 training samples and their performance evaluated on both the training set and the remaining 10-sample evaluation set. It is noted that this evaluation subset of the data was not previously seen during training and did not contribute to model development in any way. Therefore, obtained evaluation results should give a good indication of model performance with new cases under similar prevailing conditions. It would be interesting to see how the abductive network models perform on data for power plants in the ICR scheme other than the 82 plants for which data was given in [13]. As a performance metric, we used the mean absolute error (MAE) defined as the average of the absolute values of the error between the actual and predicted values taken over all the samples considered. Table 2 shows model structures synthesized at various levels of model complexity (CPM values) for elemental mercury (EL) together with their performance on both the training and evaluation sets. Increased model complexity is manifested in more complex functional elements (e.g. Doublet instead of Linear), larger number of selected inputs, and larger number of layers in the model. The model structure indicates the input variables automatically selected by the abductive learning algorithm from the 6 inputs provided. The variable number is the same as the input variable number indicated in Table 1. As model complexity increases (lower CPM values), the MAE on both the training and evaluation set initially decrease as the model better fits the training data. Further increase in model complexity beyond CPM = 0.2 degrades performance on the evaluation set due to overfitting. Since

performance on previously unseen data is the prime concern, the model generated with CPM = 0.2 is considered the optimum model in this case. Variables 1 (Coal heat), 2 (Coal Hg), and 6 (Outlet temperature) are used by all models synthesized, which indicates their importance as explanatory inputs for elemental mercury. These three input variables exhibit the largest correlation with the elemental mercury output, with the Pearson correlation coefficients using the full dataset being -0.31, 0.49, and 0.55, respectively. In addition to these variables, the optimum model at CPM = 0.2 also uses variable 5 (Ash %). It is noted that variable 4 (Coal S) does not feature in any of the elemental mercury models, which indicates that the coal sulfur content does not have a strong influence on elemental mercury emission. Moreover, variable 3 (Coal Cl) is not selected by 50% the models shown in Table 2, including the optimum model with CPM = 0.2. The corresponding correlation coefficients between the EL output and each of variables 4 and 3 are 0.15 and -0.19, respectively.

Table 3 shows model structures synthesized at various levels of model complexity for oxidized mercury (OX) together with their performance on both the training and evaluation sets. All models with CPM ≤ 2 use variables 3 (Coal Cl) and 4 (Coal S). Of all input variables, variable 4 has the highest correlation coefficient with the OX output, value = 0.62. Sulfur is known to be a major factor in the oxidation of elemental mercury, and experimental results suggest that both HCl and SO_2 may contribute directly to the mercury oxidation mechanism [27]. The optimum OX model at CPM = 0.5 uses only four input variables, including variable 1 (Coal Heat) and 6 (Outlet temperature) in addition. Table 4 shows the results for particulate mercury (PA). All models with CPM ≤ 2 require the contribution of all six input variables. The optimum model at CPM = 2 has a 6-input 3-layer structure which is more complex than the 4-input 2-layer structure obtained for the optimum models for EL and OX models. This indicates that modeling particulate mercury using the given six inputs appears to be a more complex process compared to modeling elemental and oxidized mercury.

The cross plots in Fig. 2 indicate the correlation between the actual and predicted values for the three types of mercury speciation when the optimum abductive models derived from Tables 2-4 are evaluated on the full training set of 72 samples in each case. The correlation coefficients range from 0.932 to 0.972 which compare favorably with the value of 0.987 quoted by [13] for a much smaller (13-sample) randomly selected subset of training data. Performance of the three optimum models on the external evaluation set of 10 samples is shown in Fig. 3 where the MAE error for the elemental (EL), oxidized (OX), and particulate (PA) mercury are given as 0.62, 0.29, and 0.55 lb/10^{12} Btu of coal, respectively. Inspite of the complexity of the PA model, the plot in Fig. 3(c) shows that 50% of the evaluation samples have practically zero prediction error. For evaluation sample number 6, the absolute errors for the three speciation types have the values 0.23, 0.01, and 0.07, respectively. In the work by Jensen et al. [13], only one sample was used for external evaluation of the neural network models developed, and the corresponding absolute errors reported have the values 0.006, 0.14, and 0.30, respectively. Our results indicate comparable prediction performance for the abductive and neural network models, bearing in mind that the abductive network models were trained on 72 samples while the neural network models were developed using the larger number of 81 samples. Overall, results show that prediction accuracy is poorest for the oxidized mercury species. Bench-scale investigations reported in [28] indicate that NO, NO_2, hematite (α-Fe_2O_3), maghemite (γ-Fe_2O_3), and HCl promote the conversion of elemental mercury to oxidized and/or particulate mercury. Absence of such information from the dataset used is expected to degrade the predictive performance of the resulting models for all three species.

4.2. Abductive Network Ensembles for Improved Prediction Performance

In quest for higher prediction accuracies than those provided by single models, the concept of network committees (ensembles) has been adopted in many disciplines, e.g. [29-31]. With this approach, a number of networks are used simultaneously and their outputs combined to produce the final predicted committee output, see Fig. 4. The output combination module often performs

simple functions on the outputs of individual members, such as weighted averaging [32]. When member networks are independent, the resulting diversity in the prediction process is expected to boost the generalization performance, thus improving the accuracy, robustness, and reliability of the committee predictions. Neural networks allow great diversity in the available architectures (multi-layer perceptron (MLP), radial basis function (RBF), etc.), learning algorithms (back propagation, simulated annealing, etc.), and in the parameters that can be varied during training (e.g. network topology, neuron transfer functions, initial random weights, learning rate, momentum, stopping criteria, etc.). This allows many possibilities for constructing individual committee members that are reasonably independent using the same training data. Although neural network committees have been reported for many applications, there appears to be little mention of GMDH-based abductive network committees in the literature. Due to the self-organizing and self-stopping nature of such networks, the absence of initial random weights, and the little room for user intervention during training, there is less diversity in the models that can be synthesized using the same training data. Abductive network committees reported for improved classification of medical data have used member networks developed using different model complexities [33], different subsets of training data [33], and different subsets of input variables [34].

We have investigated the use of abductive network committees for improving the accuracy of predicting mercury speciation beyond that obtained using the single models described in Section 4.1 above. Due to the small size of the dataset available, it was not practical to train models on different training subsets, and therefore we utilized variations in model complexity and input variables to ensure a reasonable degree of independence among resulting committee members. Table 5 lists details of the abductive network committees developed for the EL, OX, and PA together with their performance on the evaluation set and the percentage reduction in prediction error compared to optimum models reported in Section 4.1. All network members are trained on the full training set. Each of the committees for EL and OX consists of two member networks

trained on different subsets of input variables as indicated in Table 5. The committee for PA consists of three member networks trained on all the input variables at different CPM values. In all cases, the committee output is obtained by simple averaging of the outputs of individual members. Improvements as high as 21% in the prediction accuracy compared to the optimum single model are achieved through the use of network committees. However, statistical mean comparison tests indicate that the decrease in the mean absolute error obtained using the committees described is not statistically significant at the 95% confidence level.

5. Conclusions

Abductive network machine learning has been demonstrated as an alternative approach for modeling and predicting mercury speciation in the flue gasses of coal-fired boilers. Compared to the neural networks approach, the proposed method simplifies model development, automatically selects effective inputs and gives greater insight into the modeled phenomena which should prove useful in developing strategies to reduce emissions. Models were synthesized at various levels of model complexity for each of the mercury speciation types in terms of the properties of the coal used and operating temperature. Model prediction performance was tested on the training set as well as an external evaluation subset previously unseen during training. Performance compares favourably with neural network models developed using the same data. Up to 21% reduction in prediction error was achieved through the use of simple network committees (ensembles) trained using different subsets of input variables and utilizing variations in model complexity. In general, the limited prediction accuracy obtained is partly attributed to the absence of important input features from the data set used, such as trace metal concentrations in coal and ash, excess air, and HCl and NO_x content. Future work would seek further improvement in prediction accuracy by including such important inputs and using a larger dataset. Classifiers predicting exceedances of permitted pollution thresholds for the three speciation types can also be developed using the same technique.

Acknowledgements

The author wishes to acknowledge the support of the Department of Computer Engineering of King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia.

References

[1] S.J. Lee, Y.-C. Seo, H.-N. Jang, K.-S. Park, J.-I. Baek, H.-S. An, K.-C. Song, Speciation and mass distribution of mercury in a bituminous coal-fired power plant, Atmospheric Environment 40 (2006) 2215-2224.

[2] M. Xu, Y. Qiao, C. Zheng, L. Li, J. Liu, Modeling of homogeneous mercury speciation using detailed chemical kinetics, Combustion and Flame 132 (2003) 208-218.

[3] T. Nishitani, I. Fukunaga, H. Itoh, T. Nomura, The relationship between HCL and mercury speciation in flue gas from municipal solid waste incinerators, Fuel Processing Technology 63 (2000) 197-213.

[4] C.-J. Lin, P. Pongprueksa, S.E. Lindberg, S.O. Pehkonen, D. Byun, C. Jang, Scientific uncertainties in atmospheric mercury models I: Model science evaluation, Atmospheric Environment 40 (2006) 2911-2928.

[5] H. Zhou**,** K. F. Cen, J. R. Fan**,** Multi-objective optimization of the coal combustion performance with artificial neural networks and genetic algorithms, International Journal of Energy Research 29 (2005) 499-510.

[6] L. Fortuna, A. Rizzo, M. Sinatra, M. G. Xibilia, Soft analyzers for a sulfur recovery unit, Control Engineering Practice 11 (2003) 1491-1500.

[7] H.C. Krijnsen, J.C.M. van Leeuwen, R. Bakker, C.M. van den Bleek, H.P.A. Calis, Optimum NO_x abatement in diesel exhaust using inferential feedforward reductant control, Fuel 80 (2001) 1001-1008.

[8] M. Canakci, A. Erdil, E. Arcaklioglu, Performance and exhaust emissions of a biodiesel engine, Applied Energy 83 (2006) 594-605.

[9] Y. Chi, J.M. Wen, D.P. Zhang, J.H. Yan, M.J. Ni, K.F. Cen, HCl emission characteristics and BP neural networks prediction in MSW/coal co-fired fluidized beds, Journal of Environmental Sciences-China 17 (2005) 699-704.

[10] I.X. Tsiros, I.F. Dimopoulos, A Preliminary Study of the Application of Some Predictive Modeling Techniques to Assess Atmospheric Mercury Emissions from Terrestrial Surfaces, Journal of Environmental Science and Health, Part A: Toxic/Hazardous Substances & Environmental Engineering 38 (2003) 2495 – 2508.

[11] I.M. Raimundo, Jr., R. Narayanaswamy, Simultaneous determination of Zn(II), Cd(II) and Hg(II) in water, Sensors and Actuators B: Chemical 90 (2003) 189-197.

[12] M. Kompany-Zareh, H. Tavallali, M. Sajjadi, Application of generalized artificial neural networks coupled with an orthogonal design to optimization of a system for the kinetic spectrophotometric determination of Hg(II), Analytica Chimica Acta 469 (2002) 303-310.

[13] R.R. Jensen, S. Karki and H. Salehfar, Artificial neural network-based estimation of mercury speciation in combustion flue gases, Fuel Processing Technology 85 (2004) 451-462.

[14] A.P. Alves da Silva, U.P. Rodrigues, A.J. Rocha Reis, L.S. Moulin, NeuroDem - a neural network based short term demand forecaster**.** Presented at the IEEE Power Technical Conference, Portugal, 2001.

[15] H.S. Hippert, C.E. Pedreira, R.C. Souza, Combining neural networks and ARIMA models for hourly temperature forecasts, Proceedings of the IEEE International Joint Conference on Neural Networks, 2000, pp. 414-419.

[16] W. Charytoniuk, M.S. Chen, Neural network design for short-term load forecasting**.** Proceedings of the International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, 2000, pp. 554 –561.

[17] I, Tassadduq, S. Rehman, K. Bubshait, Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia, Renewable Energy 25 (2002) 545-554.

[18] T. Matsui, T. Iizaka, Y. Fukuyama, Peak load forecasting using analyzable structured

neural network**,** Proceedings of the IEEE Power Engineering Society Winter Meeting, 2001, pp. 405–410.

[19] H.W. Lewis III, Intelligent hybrid load forecasting system for an electric power company, Proceedings of the IEEE Mountain Workshop on Soft Computing in Industrial Applications, 2001, pp. 23-27.

[20] G.J. Montgomery, K.C. Drake, Abductive networks, Proceedings of the SPIE Applications of Artificial Neural Networks Conference, Orlando, Florida, 1990, pp. 56-64.

[21] S.J. Farlow, The GMDH algorithm, In: Farlow SJ, ed. Self-Organizing Methods in Modeling: GMDH Type Algorithms, Marcel-Dekker, New York, 1984, pp. 1-24.

[22] A.R. Barron, Predicted squared error- a criterion for automatic model selection. In: S.J. Farlow, ed. Self-Organizing Methods in Modeling: GMDH Type Algorithms, Marcel-Dekker, New York, 1984, pp. 87-103.

[23] R.E. Abdel-Aal, M.A. Elhadidy, A machine-learning approach to modelling and forecasting the minimum temperature at Dhahran, Saudi Arabia, Energy - The International Journal 19 (1994) 739-749.

[24] R.E Abdel-Aal, M.A. Elhadidy, Modeling and forecasting the maximum temperature using abductive machine learning, Weather and Forecasting 10 (1995) 310-325.

[25] R.E. Abdel-Aal, Hourly temperature forecasting using abductive networks, Engineering Applications of Artificial Intelligence 17 (2004) 543-556.

[26] AbTech Corporation, Charlottesville, VA, AIM User's Manual, 1990.

[27] S. Kellie, Y. Cao, Y.F. Duan, L.C. Li, P. Chu, A. Mehta, R. Carty, J.T. Riley, W.P. Pan, Factors affecting mercury speciation in a 100-MW coal-fired boiler with low-NOx burners, Energy & Fuels 19 (2005) 800-806.

[28] K.C. Galbreath, C.J. Zygarlicke, J.E. Tibbetts, R.L. Schulz, G.E. Dunham, Effects of NO_x, α-Fe_2O_3, γ-Fe_2O_3, and HCl on mercury transformations in a 7-kW coal combustion system, Fuel Processing Technology 86 (2005) 429-448.

[29] Z.-H. Zhou, Y. Jiang, Y.-B. Yang, S.-F. Chen, Lung cancer cell identification based on artificial neural network ensembles, Artificial Intelligence in Medicine 24 (2002) 25-36.

[30] R.E. Abdel-Aal, Improving electric load forecasts using network committees, Electric Power Systems Research 74 (2005) 83-94.

[31] A.M. Fernandes, A.B. Utkin, A.V. Lavrov, R.M. Vilar, Development of neural network committee machines for automatic forest fire detection using lidar, Pattern Recognition 37 (2004) 2039-2047.

[32] D. Jimenez, Dynamically weighted ensemble neural networks for classification. IEEE World Congress on Computational Intelligence, 1998, pp. 753-756.

[33] R.E. Abdel-Aal, Abductive network committees for improved classification of medical data, Methods of Information in Medicine 43 (2004) 192-201.

[34] R.E. Abdel-Aal, Improved classification of medical data using abductive network committees trained on different feature subsets, Computer Methods and Programs in Boimedicine 80 (2005) 141-153.

Table 1. List of input and output variables for the mercury speciation dataset used.

Input Variable Number	Description	Output	Description
1	Coal heat, Btu/lb, dry	EL	Outlet elemental Hg, lb/10^{12} Btu
2	Coal Hg, lb/10^{12} Btu	OX	Outlet oxidized Hg, lb/10^{12} Btu
3	Coal Cl, dry ppm, mass	PA	Outlet particulate Hg, lb/10^{12} Btu
4	Coal S, mass % dry		
5	Ash %, dry		
6	Outlet temperature, °F		

Table 2. Model structures synthesized for elemental mercury (EL) at various levels of model complexity, together with their performance on both the training and evaluation sets. Optimum model is marked with *. Input variable numbers are the same as those in Table 1.

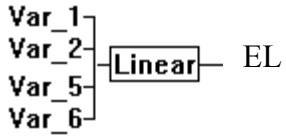
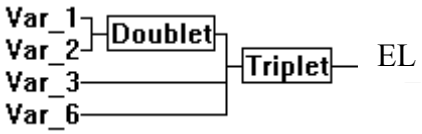
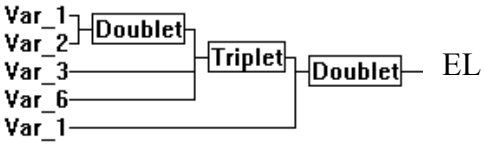
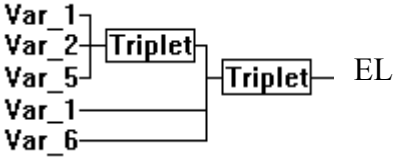
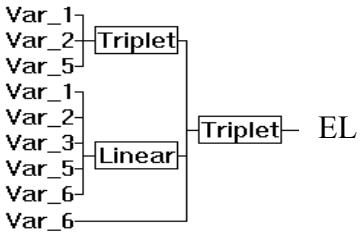
CPM	Resulting Model Structure	MAE on Training set (72 samples)	MAE on external Evaluation set (10 Samples)
5		1.22	1.60
2		0.74	0.90
1		0.62	0.73
0.5		0.56	0.67
0.2*		0.50	0.62
0.1		0.50	0.65

Table 3. Model structures synthesized for oxidized mercury (OX) at various levels of model complexity, together with their performance on both the training and evaluation sets. Optimum model is marked with *. Input variable numbers are the same as those in Table 1.

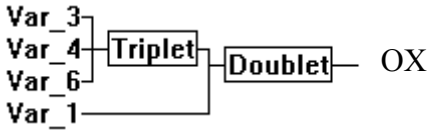
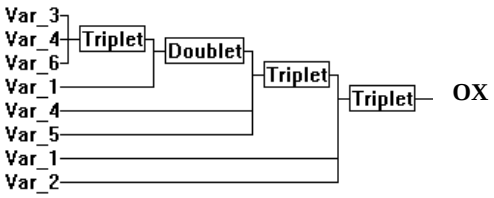
CPM	Resulting Model Structure	MAE on Training set (72 samples)	MAE on external Evaluation set (10 Samples)
5		1.00	0.79
2		0.63	0.49
1		0.63	0.61
0.5*		0.46	0.29
0.2		0.32	0.49

Table 4. Model structures synthesized for particulate mercury (PA) at various levels of model complexity, together with their performance on both the training and evaluation sets. Optimum model is marked with *. Input variable numbers are the same as those in Table 1.

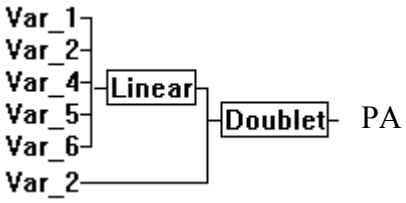
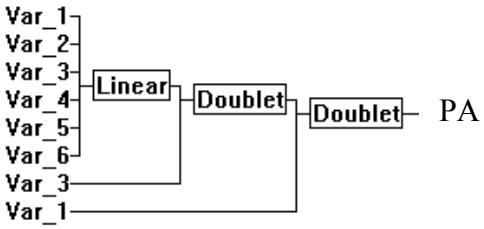
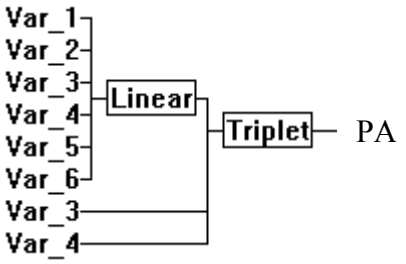
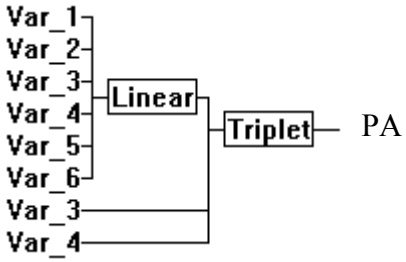
CPM	Resulting Model Structure	MAE on Training set (72 samples)	MAE on external Evaluation set (10 Samples)
5		0.73	0.94
2*		0.65	0.55
1		0.57	1.00
0.5		0.57	1.00

Table 5. Composition and performance of network committee models developed for the three types of mercury speciation. All member modules are trained on the full training set of 72 samples. Input variable numbers are the same as those in Table 1.

Speciation Type	Committee Structure				MAE for Committee on External Evaluation Set (10 samples)	% Reduction in Prediction Error by Committee Compared to Optimum Single Model
		Member # 1	Member # 2	Member # 3		
Elemental	Input Variables	1,3,5	2,4,6	_	0.49	21%
	CPM	0.2	0.1	_		
Oxidized	Input Variables	1,2,5	3,4,6	_	0.25	14%
	CPM	0.5	0.2	_		
Particulate	Input Variables	All six inputs	All six inputs	All six inputs	0.52	5.5%
	CPM	0.1	0.2	2.0		

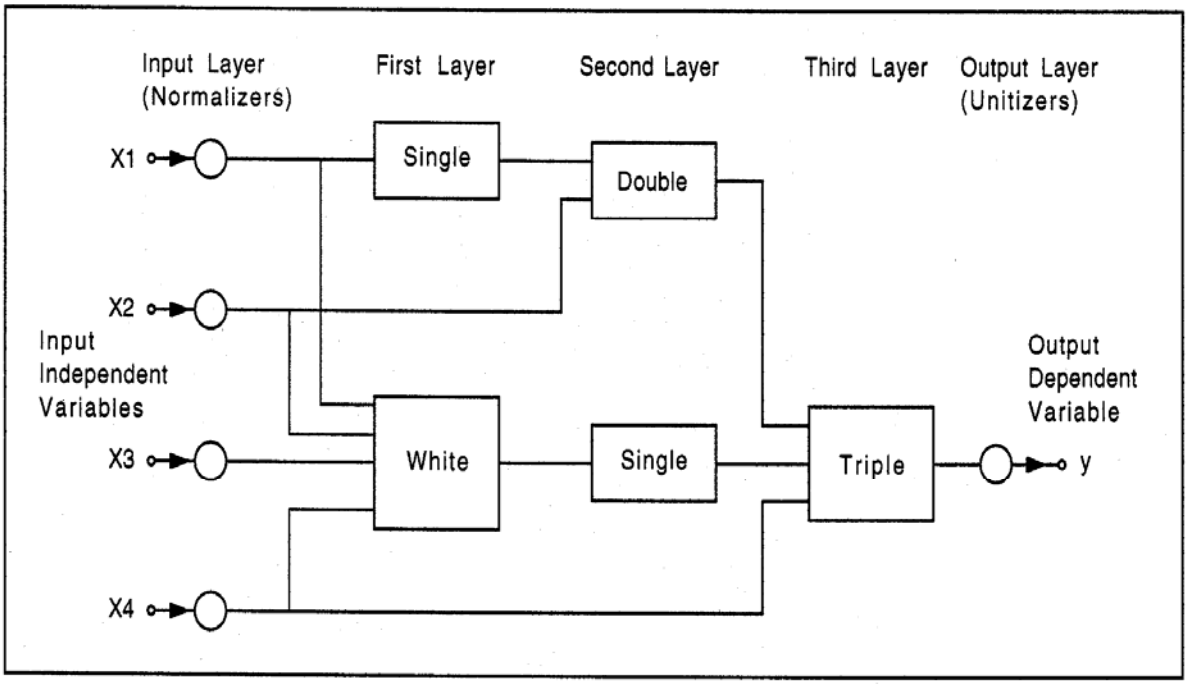


Fig. 1. A typical AIM abductive network model showing various types of functional elements.

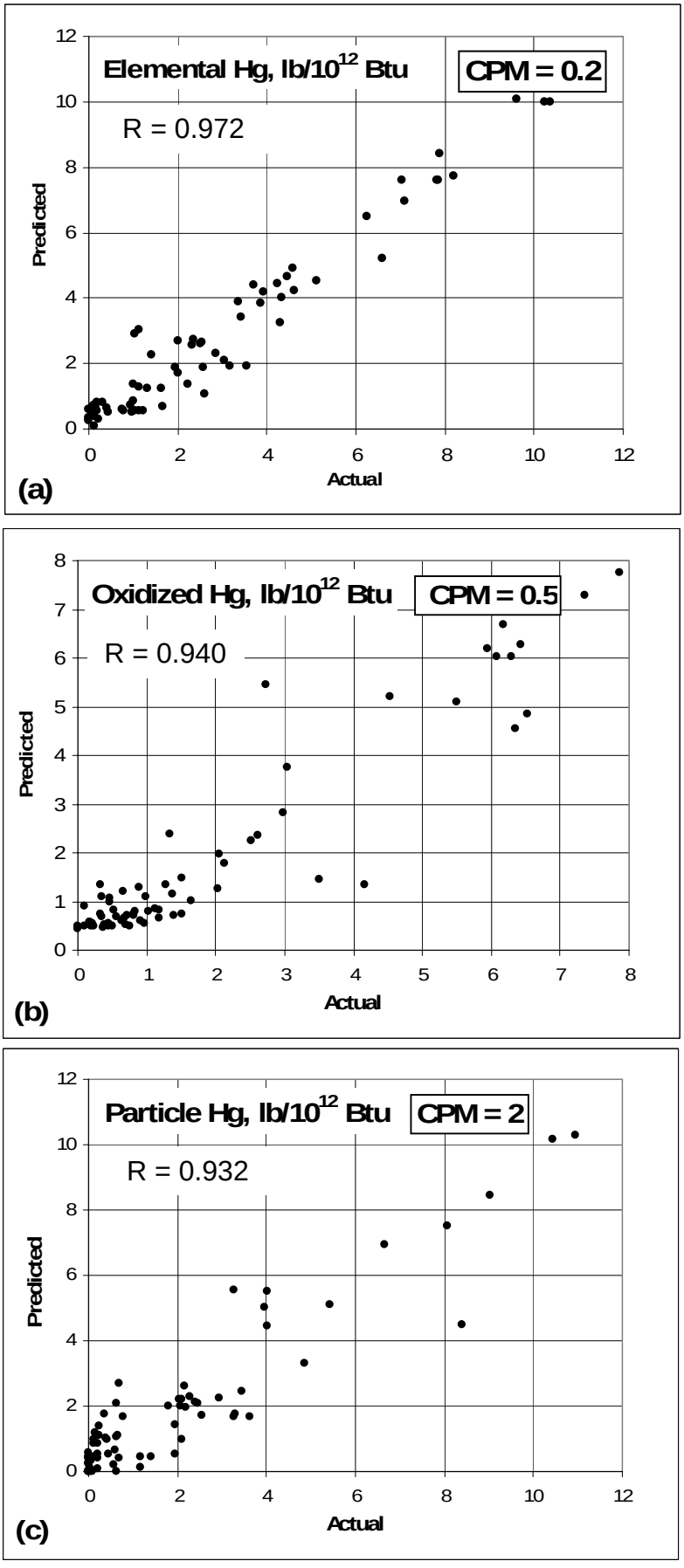


Fig. 2. Scatter plots showing the performance of the optimum single models for the three types of mercury speciation on the training set.

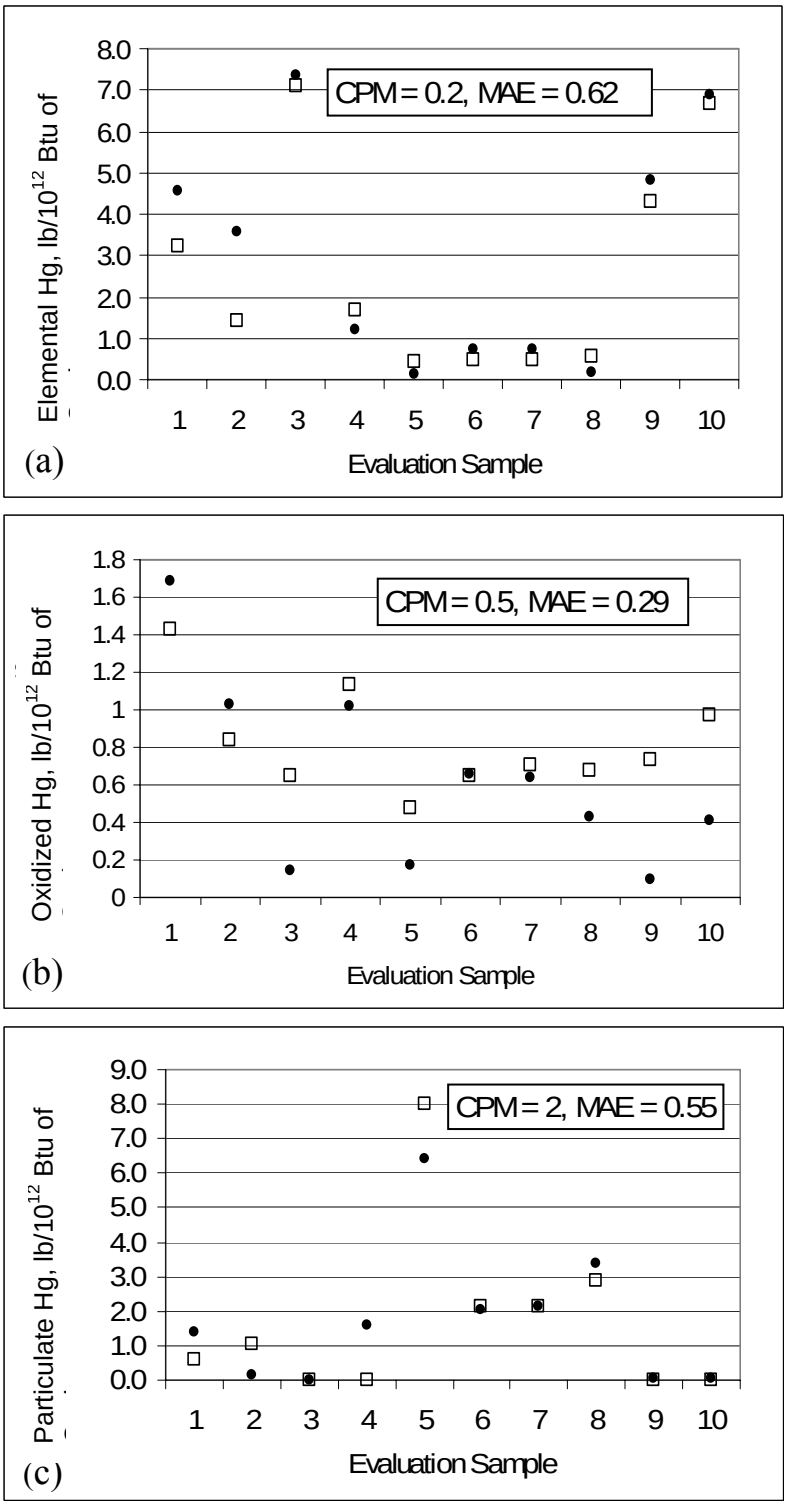


Fig. 3. Plots showing the performance of the optimum single models for the three types of mercury speciation on the evaluation set. •: Actual, □: Predicted

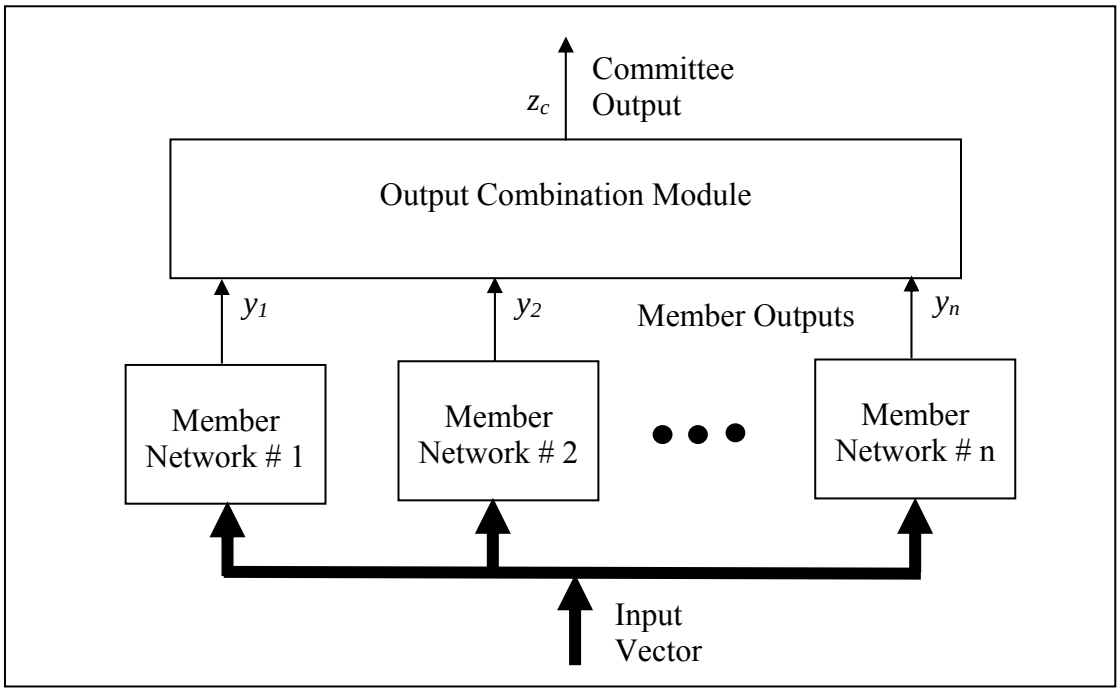


Fig. 4. Schematic of a network committee.