

Artificial Neural Networks Modeling to Reduce Industrial Air Pollution

Zvi Boger

OPTIMAL – Industrial Neural Systems Ltd.
54 Rambam St., Be'er Sheva 84243 Israel
and
Optimal Neural Informatics LLC
8203 Springbottom Way
Pikesville, MD, 21208, USA
Email: optimal@peeron.com

Abstract. Nitric acid production plants emit small amounts of nitrogen oxides (NO_x) to the environment. As the regulatory authorities demand the reduction of the resulting air pollution, existing plants are looking for economical ways to comply with this demand. Several Artificial Neural Networks (ANN) models were trained from several months of operating plant data to predict the NO_x concentration in the tail gas, and their total amount emitted the environment. The training of the ANN model was done by the Guterman-Boger algorithm set that generates a non-random initial connection weights, suggests a small number of hidden neurons, avoids, and escapes from, local minima encountered during the training. The ANN models gave small errors, 0.6% relative error on the NO_x concentration prediction and 0.006 kg/hour on daily emission in the 20-45 kg NO_x/hour range. Knowledge extraction from the trained ANN models revealed the underlying relationships between the plant operating variables and the NO_x emission rate, especially the beneficial effect of cooling the absorbed gas and reticulating liquids in the absorption towers. Clustering the data by the patterns of the hidden neurons outputs of auto-associative ANN models of the same data revealed interesting insights.

Keywords: Nitric acid plant modeling, NO_x reduction, Artificial neural network, Auto-associative clustering

Introduction

Nitric acid production plants emit small amounts of nitrogen oxides (NO_x) to the environment. As the regulatory authorities demand the reduction of the resulting air pollution, existing plants are looking for economical ways to comply with this demand. One way is to find out if there is a potential to optimize the current operating policies, by creating a model of the plant operation relationship with the NO_x emission. The suggested use of artificial neural networks (ANN) modeling techniques in industrial plants, in which the model is learned from data of the plant behavior, often arouses strong emotions. "No complicated equations! No man-years of development effort!" cheer the proponents. "No detailed equations? No reliability!" counter the opponents.

Even so, the use of ANN modeling in industrial plants is spreading, as other modeling methods are costly, both in resources and time, to fully meet the requirements of fault diagnosis or plant operation optimization.

"Soft sensor" is the accepted name for ANN model (or other model) that estimates the value of a plant variable based on other plant measurements. Such sensor, estimating the C5 impurity at the top of a distillation tower is described in [1]. Refinery NO_x emission modeled by an ANN is described in a recent paper [2].

An often-cited opposition to the use of ANN modeling in industrial diagnostics is the lack of "explanation" facility, the ability of the operator to understand the basis of the ANN recommendations. This paper shows that the "black-box" image of ANN model is misleading, and the trained ANN model can be analyzed to correctly explain the relationships between the plant operating variables and the NO_x emission rate.

The structure of the paper is as follows: A brief description of the nitric acid plant, a review of the Guterman-Boger algorithms for large-scale ANN modeling, the Causal Index (CI) method of analyzing trained ANN, and the use of the hidden neurons' output values as clustering tool. Because of non-disclosure agreements, the exact details of the plant are withheld.

1. A Brief Description of the Nitric Acid Plant

The structure of the nitric acid plant is formed from two major units – a reactor in which ammonia gas is reacted with compressed air, resulting in the formation of nitrogen dioxide NO₂. The resulting high-temperature gaseous stream, that contain also the nitrogen, is used to make steam, and then is cooled before the second major unit, the absorption of the NO₂ by water in two absorption towers. As the reaction of the NO₂ with water results in some formation of nitrogen oxide, NO, that is not absorbed by water; additional air is fed to the first absorption tower to re-oxidize the NO to the absorbable NO₂. The equations governing the gas reactions and absorption are

described in [3]. The gas exiting from the second absorption tower is sent to an expander, to utilize the high pressure, and then to the plant stack.

The allowed limit of the mixed nitrogen oxide species, NO_x, in the “tail gas” was 400 ppm, and the plant was required to meet a reduced limit of 200 ppm. An additional reactor was needed to achieve this NO_x concentration reduction, at great expense. Before deciding on this reactor, the plant management engaged the author to develop a model of the plant behavior, to find if changes in the plant operating variables would be able to achieve the required NO_x emission reduction.

The chief plant operating engineer provided a database of 40 plant instrument measurements, saved every 5 minutes by the process computer, during the preceding six months. Included in this database was the measured NO_x concentration in the tail gas that was to be the output of the ANN model.

3. A Short Overview of Artificial Neural Networks Modeling

An ANN model is trained by learning from known examples. A network of two layers of simple mathematical “neurons” is connected by weights. Data inputs are connected to the neurons in the first layer (called “hidden” neurons), which are connected in turn to the second layer of “output” neurons. Adjusting the values of the weights between the “neurons” during the training of the ANN is done by “back-propagation” of the errors between the output neurons and the known data outputs. Once the ANN is trained, it is verified by presenting examples not used in the training. The ANN may then be used to generate model outputs from the new examples presented to it. More information can be found in many books, such as [4] and journal articles, and in a review of the ANN literature, which is published in the comp.ai.neural-nets discussion group [5].

There are several obstacles in applying ANN to systems containing a large number of inputs and outputs. Most ANN training algorithms need thousands of repeated presentations (“epochs”) of the training examples to finally achieve small modeling errors. Large ANN models tend to get stuck in local minima during the training. As most ANN training starts from random initial connection weight sets, and the number of neurons in the hidden layer are usually determined by heuristic rules, many re-training trials are needed to achieve good models. The Guterman-Boger (GB) training algorithm set [6] can easily train large scale ANN models, as it starts from non-random initial connection weights, obtained by the assumption that the inputs and outputs of the training data set are linearly related. The number of major PCA dimensions in the data recommends the number of hidden neurons (typically five), and the ANN is trained by the conjugate gradient method [7] with algorithms that avoid, and escape, local minima. It was found that even ANN with thousands of features could be trained in a matter of few hours on modern PC computers, even when the GB algorithm set is operating in the interactive MATLAB environment [8]. One of the algorithm set allows the identification of the more relevant features, and previous experience showed that a reduced dimensional ANN model is giving better results [9]. In this case, the number of features was small, 39, and the trained ANN gave good results, and thus no feature reduction was made.

Once a trained ANN is available, it can be analyzed for knowledge extraction. A causal index (CI) algorithm was proposed in [10] and found to be very useful in relating each input change influence on the relative magnitude and sign changes of each output [11]. The causal index method is an easily, somewhat qualitatively, method for rule extraction. The CI is calculated as the sum of the product of all “pathways” between each input to each output,

$$CI = \sum_{j=1}^h W_{kj} * W_{ji}$$

where there are h hidden neurons, W_{kj} are the connection weights from hidden neuron j to output k , W_{ji} are the connection weights between input i to hidden neuron j .

Examining the CI for each output as a function of the inputs' number reveals the direction (positive or negative) and the relative magnitude of the relationship of the inputs on the particular output. Although somewhat heuristic, it is more reliable than local sensitivity checks. Their advantage is that they do not depend on a particular input vector, but on the connection weight set that represents all the training input vectors. This is also one of their limitations, as a local situation may be lost in the global representation.

Another useful analysis is to identify clusters in the data, is by training an unsupervised auto-associative ANN (AA-ANN), in which the features as presented both as inputs and outputs to the ANN. As there is no direct connections between the inputs and the outputs of the AA-ANN, and if the deviation between the real input feature vectors and the “predicted” input features is small, it means that the “binary” hidden neurons outputs are

representing the essential information in the dataset in order to generate the correct outputs of the ANN model. It was found [11] that in a well-trained ANN, the hidden neurons' output values tend to be close to either one or zero. Thus they can be rounded into binary patterns, giving a maximum of 2^h possible classes, if h is the number of hidden neurons. These "binary" values generate the minimum entropy (or the maximum information content) [12]. Thus, all data examples that generate the same hidden neurons output pattern are likely to belong to the same cluster. The values of features of each cluster are averaged and then divided by the average of the feature values of the full dataset. Feature ratios that are significantly different from unity are those that make each cluster distinct from other clusters. More information on the use of these techniques in industrial settings can be found in references [13,14];

4. ANN Model Training and Analyzing

The saved database, at 5 minutes interval collected in the January-July 2005 period, was cleaned by eliminating periods in which the plant operated at less than 100% capacity, or when alternative experimental operating policies were tried. In some cases when a process variable is measured by duplicate sensors, their readings were combined by averaging. The data were preprocessed by zero centering (subtracting the mean of each feature) and unit scaling (dividing by the standard deviation of each feature). The outputs of the ANN (and AA-ANN) models were further re-scaled into the [0.1 – 0.9] range. The two numbers of hidden neurons were selected, five and six, and the five hidden neurons model was found to give smaller modeling error.

Initially, the ANN model was trained with the 5 minute data to predict the NO_x concentration. When it was found by the subsequent trained model analysis that the gas absorption temperature is one of the more important operating parameter affecting the NO_x emission, an ANN model based on the daily averages was trained, thus eliminating the diurnal temperature change effect.

The 5 minute NO_x concentration at the stack modeling results are shown in Figure 1. It can be seen that the mean average error between the actual measurements and the ANN model is 0.6%, with a standard deviation of 6.7%.

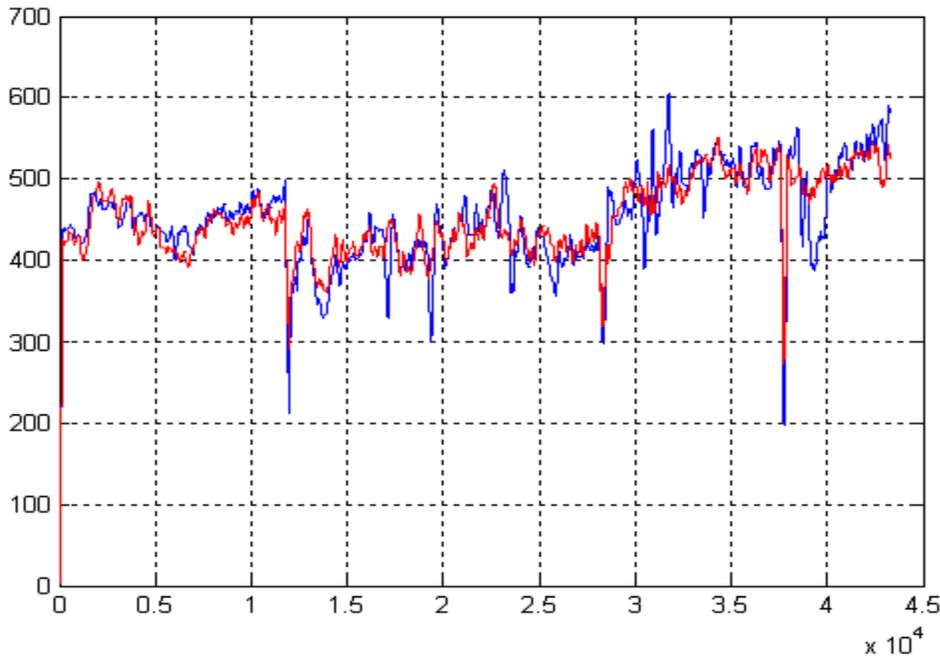


Figure 1: ANN modeling of the 5 minute data. Mean relative error 0.6%, standard deviation 6.7% . Y scale - NO_x concentration (ppm), blue trace - measurements, red trace ANN model output. X scale – sample number.

The daily average ANN model was trained to give the total NO_x emission, and the results are shown in Figure 2. The mean model error is 0.006 Kg/Hr NO_x, with a standard deviation of 0.61 Kg/Hr.

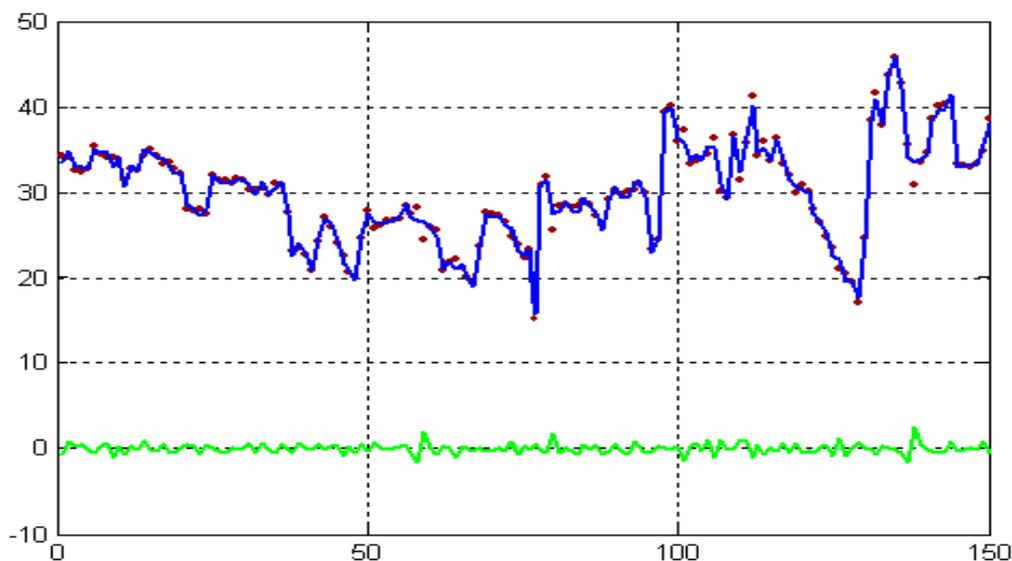


Figure 2: ANN modeling of the daily average data of total NO_x emission. X scale –day number
Y scale – Total NO_x emission (Kg/Hr), blue trace - measurements,
red dot ANN model output, Green trace model-plant deviation (Kg/Hr)..

The daily ANN model was analyzed by the Causal Index method. It was found that the NO_x amount sent to the stack was positively dependant both on the reactor reactant flows, and the absorption tower temperatures. Both relationships are consistent with chemical engineering considerations. Some unexpected findings were found by the Causal Index values, but the reasons for these findings are explained in the Discussion section.

AA-ANN was then trained from the 5 minute data, presenting the pre-processed plant features (without the NO_x measurements) both as inputs and outputs, again with five hidden neurons. After the training, the hidden neurons' outputs were rounded to one or zero, and the all data that had the same "binary" pattern were grouped into clusters.

When the full dataset was used to train the AA-ANN, 28 such clusters were identified, and the feature ratios results of the 22 clusters with non-trivial number of examples are shown in Table 1.

Table 1: Feature ratios of the major clusters

cluster #	1	2	3	4	6	8	9	10	11	12	13	14
# in cluster	4901	6907	1192	3389	1429	1374	2108	1215	1358	1103	751	1147
KgNOx/hr	33.07	29.74	30.92	28.56	21.23	23.71	28.96	24.63	26.01	17.89	21.18	28.49
NFT1105	1.04	1.05	0.97	0.98	0.93	0.95	1.02	0.92	0.92	0.93	0.99	0.92
NFT1104	1.04	1.05	0.97	0.99	0.93	0.95	1.02	0.92	0.92	0.93	0.99	0.92
NFT1103	1.04	1.04	0.98	0.98	0.94	0.96	1.02	0.92	0.93	0.94	1.00	0.94
NFT1102	1.04	1.04	0.98	0.98	0.94	0.96	1.02	0.92	0.93	0.94	0.99	0.94
NTT1113	1.00	1.00	1.00	1.00	1.00	1.01	1.00	1.00	1.00	1.00	1.00	1.01
NTT1112	1.01	1.01	1.01	1.01	1.00	1.01	1.00	1.00	1.00	1.00	1.00	1.00
NT110050	1.01	1.01	1.00	1.01	0.99	1.00	1.01	0.99	0.99	0.99	1.00	0.99
NTT11007	1.01	1.01	1.00	1.01	0.99	1.00	1.01	0.99	0.99	0.99	1.00	0.99
NT110018	1.00	0.99	0.99	0.97	0.97	0.98	1.01	0.97	0.99	0.98	0.99	0.99
N2RATIO2	0.96	0.95	0.97	0.96	0.97	0.97	0.96	0.97	0.97	0.98	0.97	0.98
N2RATIO3	1.01	1.00	1.01	1.00	1.01	1.01	1.01	1.01	1.02	1.02	1.01	1.02
NTT11009	0.99	0.98	0.97	0.96	0.95	0.97	0.99	0.95	0.97	0.95	0.98	0.98
NFT1205	1.02	1.04	0.99	0.98	0.93	0.94	1.02	0.93	0.93	0.92	0.99	0.94
NTT1206	0.97	0.98	0.99	0.95	0.93	0.98	1.00	0.98	1.02	0.97	0.96	1.03
NT110010	0.99	0.98	0.97	0.95	0.93	0.98	0.98	0.94	0.99	0.96	0.98	1.01
NT110012	0.97	0.96	0.93	0.92	0.90	0.93	0.96	0.92	0.97	0.92	0.96	0.98
NT110011	0.97	0.96	0.93	0.92	0.90	0.93	0.96	0.92	0.97	0.92	0.96	0.99
NT110020	0.96	0.92	0.92	0.90	0.90	0.92	0.96	0.94	0.97	0.91	0.95	0.96
NFT1255	1.11	1.05	1.04	1.00	0.83	0.94	1.01	0.87	0.93	0.71	0.88	1.03
NPT1252	1.00	1.01	1.00	1.02	0.98	1.03	0.99	0.95	0.98	0.97	0.99	1.02
NTT1253	0.94	0.89	0.91	0.85	0.88	0.90	0.97	0.92	0.98	0.92	0.96	0.97
NPDT1250	1.05	1.03	0.97	0.94	1.00	0.87	1.08	0.98	0.91	0.94	1.09	0.85
NDPX1107	1.04	1.05	0.97	0.96	0.92	0.94	1.02	0.90	0.91	0.92	1.00	0.93
NAT1133	1.01	0.94	1.00	0.93	0.89	0.80	0.98	1.03	0.98	0.87	0.84	0.92
NAT1130	0.97	1.03	0.96	1.04	0.89	0.96	0.92	1.00	0.87	0.89	1.00	0.86
NTT1262	0.91	0.85	0.91	0.79	0.87	0.89	1.01	0.94	1.03	0.93	0.98	0.98
NLT1258	0.96	0.99	1.02	1.00	1.04	1.04	1.02	1.00	1.04	1.03	1.00	1.06
NLT1208	0.96	1.01	1.02	1.00	1.03	1.03	1.02	1.03	1.03	1.03	1.03	1.03
NTT11004	0.91	0.85	0.91	0.79	0.87	0.89	1.01	0.95	1.03	0.94	0.98	0.98
NT110017	0.97	0.95	0.96	0.92	0.93	0.95	1.01	0.95	1.00	0.97	0.99	0.99
NTT11005	1.00	1.00	1.00	1.00	0.99	0.99	1.01	0.99	0.99	0.99	1.00	1.00
NTT701	1.00	1.00	1.00	0.99	0.99	0.99	1.01	0.99	0.99	0.99	1.00	0.99
NTT702	0.99	0.95	1.03	0.99	1.05	1.01	1.05	1.10	1.08	1.05	1.03	1.05
NT110046	1.00	0.97	1.00	0.97	1.03	0.97	1.04	1.06	1.03	1.03	1.04	0.97
NPT603	1.00	1.01	1.00	1.01	0.97	1.02	0.99	0.94	0.96	0.97	0.99	1.00
NFT1215A	1.03	1.04	0.97	0.98	0.93	0.95	1.01	0.92	0.92	0.93	0.99	0.92
NPT1101	1.02	1.00	0.98	0.97	0.98	0.96	1.03	0.98	0.99	0.95	1.03	0.98
NFT1206	1.07	1.09	0.99	1.00	0.95	0.94	1.03	0.92	0.93	0.95	1.01	0.93
N2NOX	1.12	1.01	1.05	0.97	0.72	0.80	0.98	0.83	0.88	0.61	0.72	0.96

Table 1: Feature ratios of the major clusters (cont.)

cluster #	15	16	17	18	19	20	21	22	24	25	27	28
# in cluster	1332	1062	1627	1597	846	2446	1241	5373	345	206	105	12
KgNOx/hr	25.47	25.15	20.04	34.99	32.84	31.47	33.55	38.13	30.38	40.72	20.16	34.59
NFT1105	1.03	1.05	1.02	1.05	0.91	1.03	0.92	1.03	1.04	1.04	1.02	0.99
NFT1104	1.03	1.05	1.02	1.05	0.91	1.03	0.91	1.02	1.04	1.04	1.02	0.99
NFT1103	1.03	1.05	1.03	1.05	0.91	1.03	0.92	1.02	1.03	1.03	1.02	0.98
NFT1102	1.04	1.05	1.03	1.05	0.91	1.03	0.92	1.02	1.03	1.03	1.02	0.99
NTT1113	1.01	1.00	1.01	1.01	1.00	1.01	1.01	1.01	1.01	1.01	1.01	1.01
NTT1112	1.01	1.00	1.01	1.01	1.00	1.01	1.00	1.00	1.01	1.00	1.00	1.00
NT110050	0.99	1.00	0.99	1.01	0.96	0.99	0.96	0.99	0.99	1.01	0.98	0.99
NTT11007	0.99	1.01	0.99	1.01	0.96	0.99	0.96	0.99	0.99	1.01	0.99	0.98
NT110018	1.03	1.01	1.02	1.02	1.00	1.03	0.99	1.03	1.03	1.03	1.02	1.01
N2RATIO2	0.96	0.96	0.97	0.96	0.97	0.96	0.96	0.96	0.95	0.95	0.96	0.95
N2RATIO3	1.01	1.01	1.01	1.01	1.00	1.00	1.00	1.00	0.99	0.99	1.00	0.99
NTT11009	1.05	0.99	1.02	1.04	1.03	1.06	1.03	1.06	1.05	1.04	1.05	1.03
NFT1205	1.03	1.06	1.03	1.04	0.94	1.02	0.95	1.05	1.06	1.03	1.04	0.94
NTT1206	1.06	0.95	1.03	1.04	1.05	1.06	1.03	1.08	1.07	1.05	1.08	0.89
NT110010	1.05	0.99	1.02	1.04	1.05	1.06	1.04	1.07	1.05	1.03	1.05	1.04
NT110012	1.09	0.96	1.03	1.09	1.08	1.11	1.07	1.12	1.09	1.08	1.09	1.10
NT110011	1.09	0.96	1.03	1.09	1.08	1.11	1.08	1.12	1.09	1.08	1.09	1.10
NT110020	1.11	0.91	1.04	1.07	1.16	1.14	1.16	1.16	1.12	1.09	1.13	1.12
NFT1255	0.78	0.90	0.65	1.09	1.01	0.96	1.03	1.16	1.19	1.08	0.67	1.00
NPT1252	1.00	1.01	1.00	1.03	0.98	1.00	1.00	1.02	1.03	1.00	0.99	0.98
NTT1253	1.14	0.92	1.08	1.10	1.19	1.18	1.17	1.19	1.14	1.11	1.17	1.13
NPDT1250	1.01	1.02	0.96	1.05	0.95	1.09	0.95	1.00	0.97	1.02	0.97	0.93
NDPX1107	1.14	1.06	1.09	1.06	0.86	1.04	0.86	1.00	1.00	1.05	1.14	0.97
NAT1133	1.14	0.93	1.06	1.07	1.17	1.14	1.14	1.12	0.85	1.29	1.08	1.22
NAT1130	0.96	1.00	1.02	0.92	1.09	1.05	1.11	1.06	1.14	1.15	1.21	1.20
NTT1262	1.17	0.96	1.11	1.08	1.26	1.21	1.19	1.22	-0.15	10.93	-0.52	48.54
NLT1258	1.05	1.02	1.01	1.02	0.97	1.01	0.96	0.99	0.90	1.04	1.03	1.02
NLT1208	1.11	1.03	1.06	1.06	0.91	1.08	0.87	0.93	1.14	1.15	1.21	1.20
NTT11004	1.16	0.96	1.11	1.08	1.26	1.21	1.19	1.22	1.06	1.07	1.08	1.05
NT110017	1.07	1.00	1.05	1.04	1.07	1.09	1.05	1.09	1.00	1.02	1.00	1.00
NTT11005	1.01	1.00	1.00	1.01	0.98	1.01	0.98	1.01	1.00	1.02	1.00	1.00
NTT701	1.01	1.00	1.00	1.01	0.98	1.01	0.98	1.01	0.88	0.96	1.03	1.07
NTT702	1.03	0.92	1.00	0.98	1.07	1.04	1.03	0.96	0.95	1.02	1.03	1.06
NT110046	1.03	0.97	1.01	1.01	1.02	1.04	0.99	0.99	1.03	1.01	1.00	0.97
NPT603	1.00	1.02	1.01	1.03	0.96	1.01	0.98	1.02	1.04	1.04	1.02	0.97
NFT1215A	1.03	1.05	1.02	1.04	0.91	1.02	0.92	1.02	1.00	1.05	1.02	1.05
NPT1101	1.01	0.98	1.01	1.00	1.00	1.03	0.98	1.01	0.98	1.08	0.98	0.95
NFT1206	1.00	1.07	1.00	1.01	0.85	0.98	0.83	0.97	1.03	1.38	0.68	1.17
N2NOX	0.86	0.85	0.68	1.18	1.11	1.06	1.13	1.29	1.03	1.38	0.68	1.17

It can be seen that some clusters (# 6, 8, 12, 13, 17, 27) have lower NOx emission amounts, and the identification of the feature ratios that are much smaller (or higher) than unity may explain these results. What was more surprising was the fact that some of the examples in these clusters are contiguous in time, very close to the duration of a complete shift (that is morning, afternoon or night shift). This raises the possibility that some shift managers are more efficient in running the plant and thus reducing the NOx emission.

5. Discussion

Reviewing the ANN modeling analysis revealed a major lack of information in the plant data collection scheme – no information on the control loop set points values. These are changed by the shift managers, responding to plant upsets or transients. As these changes are not recorded, some cause and effects may be mis-understood. For instance, if the NO_x absorption seems insufficient, the set point of the absorbing water is increased. Subsequent analysis will relate high NO_x emission with increased absorption water flow. Thus the daily feature averages is much more reliable than 5 minute data.

If the control loop set points changes were available, the insight and experience of the better shift managers may be learned and incorporated in the computer control scheme, or at least known to the less experienced shift managers.

The major finding of the ANN modeling, that reducing the absorption tower operating temperature, was not helpful to solve the NO_x emission issue, because the cooling water supply was outside the control of the plant management. Eventually, another NO_x reducing technique was successfully adopted.

6. Conclusion

The ANN modeling of the nitric acid production plant predicted the NO_x emission amounts and concentration in the tail gas with small errors. The analysis of the ANN and AA-ANN models revealed some known, and some previously unknown, relationships in the plant operation.

The ANN models trained from daily plant feature averages proved more informative than the 5 minute data, although it may be the results of the importance of the diurnal temperature changes in this plant.

The inclusion of the control loop set points in the plant database may provide more information for future analysis that will improve the operational knowledge for better efficiency.

Acknowledgements

Thanks are due to the chief operation engineer, Mr. A.R., for providing the plant data and helpful discussions, and to the plant management for allowing the publication of this paper.

References

1. Boger, Z., Guterman, H., Segal, T.: Application of large-scale artificial neural networks for modeling the response of a naphtha stabilizer distillation train. *Proceedings of the AIChE annual meeting*, Chicago (1996), <http://www.che.wisc.edu/aiche/1996/10b-abstracts/19.html>
2. Fortuna, L., Graziani, S., Xibilia, M.G: Virtual instruments in refineries. *IEEE Instrumentation & Measurement Magazine*, 8, (4), 26-34, (2005)
3. Sweeney, J.A., Liu, J.A.: Use of Simulation to Optimize NO_x Abatement by Absorption and Selective Catalytic Reduction, *Ind. Eng. Chem. Res.*, 40 (12), 2618 -2627, (2001)
4. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, (1997)
5. Sarle, W.S.: Frequently Asked Questions. comp.ai.neural-nets Users Group, <ftp://ftp.sas.com/pub/neural>. (2005).
6. Guterman, H.: Application of Principal Component Analysis to the Design of Neural Networks. *Neural, Parallel and Scientific Computing*, 2, 43-54 (1994)
7. Leonard, J., Kramer, M.A.: Improvement of the Back-Propagation Algorithm for Training Neural Networks. *Comput. Chem. Engng.* 14, 337-341 (1990)
8. Boger, Z.: Who is Afraid of the Big Bad ANN? *Proc. Intl. Joint Conf. Neural Networks*, 2000-2005 (2002)
9. Boger, Z.: Selection of the quasi-optimal inputs in chemometric modeling by artificial neural network analysis. *Analytica Chimica Acta* 490 (1-2), 31-40, (2003)

10. Baba, K., Enbutu, I., Yoda, M.: Explicit Representation of Knowledge Acquired from Plant Historical Data using Neural Network. Proc. of the Intl. Joint Conference on Neural Networks, vol. 3, 155-160, (1990)
11. Boger, Z., Guterman, H.: Knowledge Extraction from Artificial Neural Networks Models. Proc. IEEE Intl. Conf. Systems Man and Cybernetics, 3030-3035 (1997)
12. Kamimura, R. and Nakanishi S. Hidden information maximization for feature detection and rule discovery, Network Computation in Neural Systems 6, 577-602 (1995)
13. Boger, Z. Artificial Neural Networks Modeling as a Diagnostic and Decision Making Tool. D. Ruan and P.F. Fantoni (eds.): Power Plant Surveillance and Diagnostics, Modern Approaches and Advanced Applications, Physica-Verlag, Chapter 16, 243-252(2002)
14. <http://optimalneural.com>