Journal of Engineering Science and Technology
Vol. 3, No. 1 (2008) 40 - 47


# DRIFT ANALYSIS ON NEURAL NETWORK MODEL OF HEAT EXCHANGER FOULING

M. RAMASAMY*, A. SHAHID, H. ZABIRI

Chemical Engineering Department, Universiti Teknologi PETRONAS
Bandar Seri Iskandar, 31750 Tronoh, Perak, MALAYSIA
*Corresponding Author: marappagounder@petronas.com.my

**Abstract**

Neural Networks (NN) provide a good platform for modeling complex and poorly understood systems in many different fields. Due to the empirical nature of NN, it is typically valid only for small operating windows. As the process drifts, the prediction accuracy of such models deteriorates very much rendering the models unfit. An on-line mechanism to follow the drift in the process is necessary in order to retrain the NN models. Information Criteria have been reported to be used for the selection of relevant input variables and determination of optimal NN model structures. This paper proposes the use of information criteria for tracking the model prediction accuracy and provides an algorithm for retraining the model. A heat exchanger in a refinery Crude Preheat Train (CPT) has been used as a case study. The operational problems of heat exchangers in CPT are compounded by the varying nature of crude blends and the complex fouling phenomenon. Fouling develops slowly and therefore the drift in the process occurs on a slower scale. The performance of a NN fouling model, developed using industrial data is investigated for drift. Model performance at different operating conditions is evaluated and it has been shown that drifts do occur in the process. An algorithm for retraining NN model has been proposed.

Keywords: Neural Networks, Information Criteria, Heat Exchanger, Drift Analysis.

## 1. Introduction

Recently, there has been growing interest in modelling of nonlinear and complex systems in order to optimise and efficiently control the plant operation. Varying

**Nomenclatures**

| | |
|---|---|
| $D$ | Direction of change |
| $K$ | Number of free model parameters |
| $N$ | Number of observations |
| $t$ | Time instant |
| $y$ | Actual measured value |
| $\hat{y}$ | Predicted value |
| *Greek Symbols* | |
| $\varepsilon$ | Error of the regression |
| $\sigma$ | Standard deviation |
| *Subscripts* | |
| $i$ | Observation number |
| $t$ | Time instant |

nature of feed-stocks and frequent changes in the economic scenario forces the plant operators to operate the plant at widely varying optimal operating conditions. Mathematical models are generally used for determining optimal operating conditions and prediction for planning and scheduling. It is well known that theoretical models are difficult to develop for highly nonlinear, complex and poorly-understood systems. Neural Networks (NN) provide a good platform for modelling such systems. However, NN models suffer due to their empirical nature that they are valid only for the range of operation that is used for training the model. Moreover, process characteristics change slowly over a period of time due to factors such as deactivation of catalysts in reactors, fouling in heat exchangers, etc. As the process characteristics drift, the accuracy of NN models deteriorates and need to be retrained.

Information Criteria have been used for the selection of appropriate network model for a given problem [4]. They provided a methodology to choose important input variables and the number of hidden neurons and their connections thereof. This paper proposes the use of information criteria to follow the process drift and to establish when to retrain the model for improving its accuracy.

Modelling of heat exchangers in refinery Crude Preheat Train (CPT) using NN have been reported in the literature [10]. The NN model reported in their paper basically captures the complex fouling phenomenon. But the challenges in a CPT include the frequent changes in crude and crude blends with varying fouling characteristics. This results in poor prediction accuracies and makes it necessary to retrain the model. However, there is no systematic approach to determine when exactly the retraining process should start.

This paper investigates the occurrence of process drifts in a heat exchanger in CPT using the feedforward NN model based on data collected from a real plant for different periods of operation. Based on the information criteria, an algorithm for retraining the NN model has been proposed in order to maintain the prediction accuracy of the model in spite of process drifts. Section 2 explains the two important information criteria, namely Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC). A brief explanation on the heat exchanger

NN model is provided in Section 3. Results from the simulation studies on the heat exchanger and discussions are presented in Section 4 while the retraining algorithm is presented in Section 5.

## 2. Information Criteria

The underlying idea of information criteria is to find an optimal trade-off between an unbiased approximation of the underlying model and the loss of accuracy caused by estimating an increasing number of parameters [4] in the context of model selection criteria. Information criteria combine some measure of fit with a penalty term to account for model complexity. If it is assumed that an appropriate model (fixed complexity) of a given system is available, then the same criteria can be used for detecting the process drift.

### 2.1 Akaike information criteria (AIC)

The most prominent and widely used criterion is the AIC [1], which in principle applies to any model estimated by maximum likelihood. The AIC is defined as

$$AIC = ln\left(\frac{\varepsilon\varepsilon'}{N}\right) + \frac{2K}{N} \tag{1}$$

where $\varepsilon$ is the error of the regression, $K$ is the number of free model parameters and $N$ is the number of observations. However, this criterion is not valid for mis-specified models. If the network be able to map the true function exactly, then the AIC formula was given by Amemiya [2]:

$$AIC = MSE + \sigma^2 \frac{2K}{N} \tag{2}$$

where MSE is the mean square error and $\sigma^2$ is the variance. The AIC gives an indication as to whether or not the NN contains irrelevant hidden units; if it does, the AIC takes enormously high values.

### 2.2 Bayesian information criteria (BIC)

Bayesian Information Criterion [7, 8] is similar to AIC except the modification in the likelihood factor. The equation for BIC is given by

$$BIC = MSE + log(N)\sigma^2 \frac{2K}{N} \tag{3}$$

Generally, both AIC and BIC are used for optimising the structure of the models. It may be noted that for the purposes of drift analysis, the AIC and BIC not only depend on the MSE but also on the variance of the error. This is clearly an advantage in using AIC or BIC for the drift analysis over the RMSE values alone.

## 2.3 Other measures

Apart from information criteria, there exist many alternative criteria which are commonly used during training and validation of NN models such as (i) Root Mean Square Error (RMSE), (ii) Correct Directional Change (CDC) and (iii) coefficient of determination $R^2$.

### *2.3.1 Root Mean Square Error (RMSE)*

RMSE is generally used to determine the goodness of fit of the correlation and is defined by

$$RMSE = \frac{\sqrt{\varepsilon\varepsilon'}}{N} \tag{4}$$

### *2.3.2 Correct Directional Change (CDC)*

CDC is a measure of the capability of the model to predict the correct direction of change in a variable and is defined in Eq. (5).

$$CDC = \frac{100\%}{N-1}\sum_{t=2}^{N} D_t \tag{5}$$

with $D_t = 1$, if $\left[y_i(t) - y_i(t-1)\right]\times\left[\hat{y}_i(t) - \hat{y}_i(t-1)\right] > 0$

$D_t = 0$, if $\left[y_i(t) - y_i(t-1)\right]\times\left[\hat{y}_i(t) - \hat{y}_i(t-1)\right] < 0$

where $y_i$ is the $i^{\text{th}}$ actual measured value and $\hat{y}_i$ is the $i^{\text{th}}$ predicted value.

### *2.3.3 Coefficient of Determination, $R^2$*

The coefficient of determination, $R^2$, is one of the statistical methods to check model performance. The coefficient of determination is defined by

$$R^2 = 1 - \frac{\sum_{t=1}^{T}\left(y_t - \hat{y}_t\right)^2}{\sum_{t=1}^{T}\left(y_t - \bar{y}_t\right)^2} \tag{6}$$

It measures the goodness of fit of the regression: closer the value of the coefficient of determination to one, better the fit; and closer the value to zero, worse the fit.

## 3. Heat Exchanger Model

The refinery CPT consists of 11 heat exchangers, each of them with 2 or 4 shells operating in series/parallel modes. The heat exchanger considered in this study is the fourth in the CPT. In this particular heat exchanger, the crude oil flows through the shell side and the heating medium, low sulphur waxy residue (LSWR), flows through the tube side. Data around this heat exchanger was collected for a period of three years on a daily average basis starting from June 2002 following a turn-around operation in the plant during which the heat exchangers were mechanically cleaned and the heat transfer efficiency was the maximum. The data collected include the flow rates, inlet and outlet temperatures, the composition of crude blend and the assay of crude. The data was analyzed for outliers through principal component analysis (PCA) [6] and all the outliers were removed. The important relevant input variables that contribute to the fouling in the heat exchanger were identified using projection to latent structures (PLS) [6]. A NN model was developed using feedforward back-propagation algorithm correlating the selected input variables to the outlet temperatures in the shell-side and the tube-side. The NN model had an input layer with 25 neurons, a hidden layer with 15 neurons and an output layer with 2 neurons. For more details on model development and the list of input variables selected, please refer to Radhakrishnan, et al. [10].

## 4. Results and Discussion

The NN model for the heat exchanger has been trained using the data starting from June, 2002 to March, 2004. The data was randomised and divided into three sets, namely, training data set (50%), validation data set (40%) and test data set (10%). The performance of the model is tested mainly based on the prediction error for the test data set. Figure 1 shows the comparison between the predicted and actual shell-side outlet temperatures. The performance of the model in terms of RMSE and CDC for the test data set are $1^{o}$C and 89%, respectively, which indicate that the model is reasonably accurate.

The performance of any NN model developed using data from a certain window of operation will deteriorate when the operating window changes. This means that an empirical model such as NN model cannot be reliably used for long periods of time by training only once. Especially, systems like heat exchangers in CPT have high tendency to foul and reduce its efficiency in relatively shorter periods of time. In this work, AIC and BIC, in addition to RMSE and CDC, were used as indicators of the performance of the model. Data from two subsequent periods of operation, (i) a period of 50 days immediately following the training data, and (ii) a period of 3 months subsequent to the first period were collected and used for analysing the drift in the performance of the developed model.

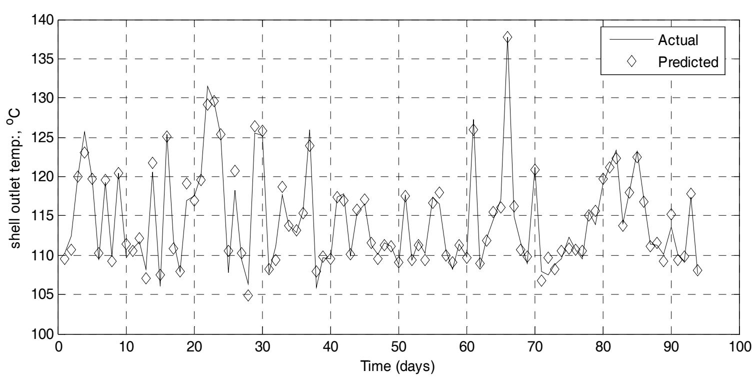


**Fig.1. Comparison between the Actual and Predicted Shell Outlet Temperatures during the Model Development Phase.**

## 4.1 Drift analysis I

The data for the period for the first drift analysis starts from 1 April 2004 immediately following the data set used in the model development phase. It may be noted that the original NN model was developed using the data until the end of March 2004. Figure 2 shows the comparison between the actual and the predicted shell outlet temperatures. It is observed that the performance of the NN model has deteriorated which indicates that there is a drift in the process. The corresponding RMSE and CDC values are 10.9$^{o}$C and 85%, respectively, and prove the poor performance of the model.

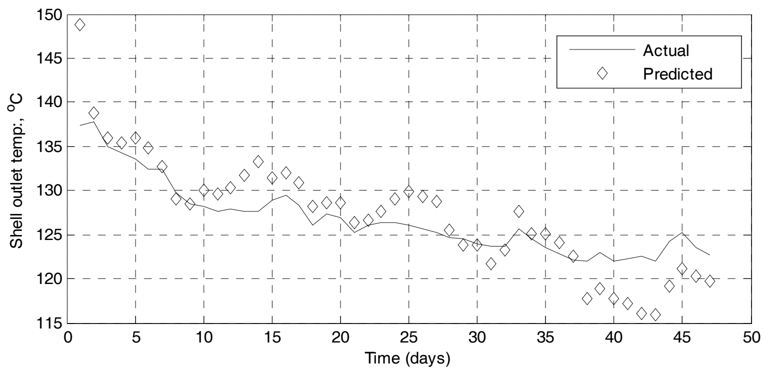


**Fig.2. Comparison between the Actual and Predicted Shell Outlet Temperatures during the First Drift Analysis Period.**

## 4.2 Drift analysis II

The period for the second drift analysis starts from 21 May 2004. For this period, the outlet temperatures in the shell and tube sides are predicted by simulating the model developed originally. From the simulation results shown in Figure 3, it is observed that the predictive ability of the model has deteriorated considerably and the model needs to be retrained using the new data available. The reasons for the deterioration in the model performance may be attributed to variations in the crude blend processed during this period resulting in different rates of fouling.

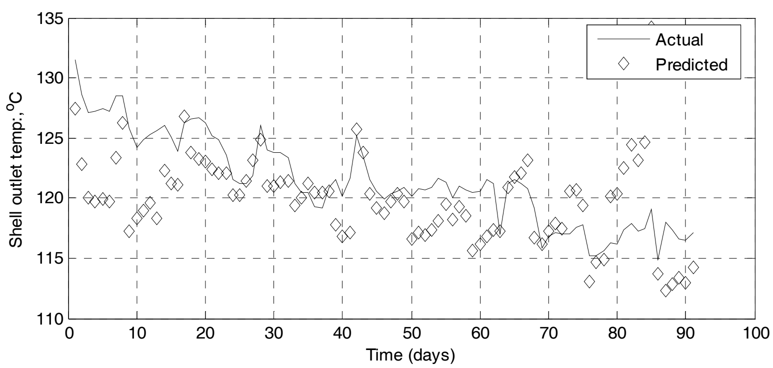


**Fig.3. Comparison between the actual and predicted shell outlet temperatures during the second drift analysis period.**

The different measures of model inaccuracy due to the process drift are compiled for the model development phase, and the two subsequent drift analysis periods as shown in Table 1. It has been found that the drift in the process can be well detected from the increasing values of RMSE, AIC, and BIC and the decreasing values of CDC and $R^2$.

**Table 1. Comparison of Model Performance during Training, Drift Analysis I and Drift Analysis II**

| | Training | | Drift Analysis I | | Drift Analysis II | |
|---|---|---|---|---|---|---|
| **Statistical Test** | **Shell Side** | **Tube side** | **Shell Side** | **Tube side** | **Shell Side** | **Tube side** |
| **RMSE, $^o$C** | 1.008 | 3.176 | 10.88 | 71.8 | 15.89 | 111.17 |
| **CDC, %** | 89.3473 | 89.247 | 84.78 | 15 | 63.6 | 2 |
| **$R^2$** | 0.987 | 0.913 | 0.913 | 0.853 | 0.52 | 0.15 |
| **AIC** | 24.6223 | 7.2E+3 | 4.2E+5 | 4.7E+8 | 1.4E+6 | 2.5E+9 |
| **BIC** | 54.64 | 1.6E+4 | 8.2E+6 | 9.5E+8 | 2.7E+6 | 4.9E+9 |

## 5. Model Retraining Algorithm

Drift in the model was observed over period of different operating conditions. A retraining algorithm has been proposed to detect the drift and update the model parameters automatically to improve the overall model performance. Figure 4 shows the flow chart of the proposed algorithm.

## 6. Conclusions

Neural network model performances were analyzed during the model development phase with a test data set and for two subsequent periods. It was observed that the model performance drifts over time for varying operating conditions, due to the empirical nature of neural networks. For improving the prediction accuracy of the model, it is proposed to retrain the model whenever the performance indicators such as RMSE, AIC or BIC crosses certain threshold values. A retraining algorithm has been proposed.

## Acknowledgements

Authors of this paper thankfully acknowledge the support provided by Universiti Teknologi PETRONAS for carrying out this research work.

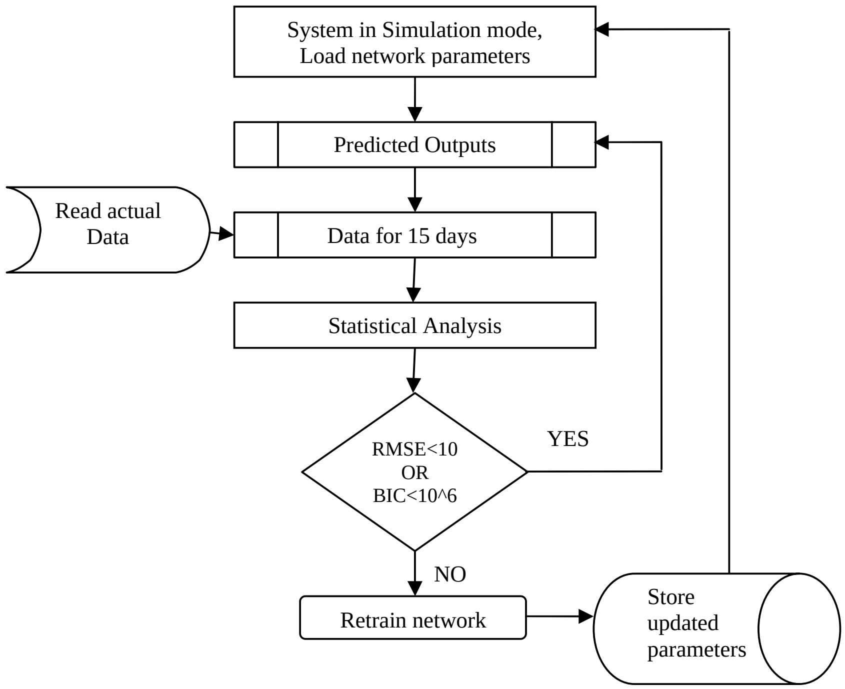


**Fig. 4. Flow Chart for Model Retraining Algorithm.**

## References

1. Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on automatic control*, AC19, 716-723.
2. Amemiya, T. (1980). Selection of regressors, *International Economic Review*, 21, 331-354.
3. Anders, U. (1996). Statistical Model Building for Neural Networks. <www.actuaries.org/AFIR/ colloquia/ Nuernberg/Anders.pdf>.
4. Anders, U. & Korn, O. (1999). Model selection in neural networks. *Neural Networks*, 12, 309-323.
5. Stergiou, C. & Siganos, D., Neural Networks. <http://www.doc.ic.ac.uk>
6. Eriksson, L., Johansson, E., Kettaneh-Wold, N. & Wold, S. (2001). *Multi- and Megavariate Data Analysis: Principle and Applications*. Umetrics Inc.
7. Hastie, T., Tibshirani, R. & Friedman, J. (2001). *The elements of statistical learning: Data mining, Inference, and Prediction*. New York, Springer.
8. Helie, S. (2006). An introduction to Model Selection: Tools and Algorithms. *Tutorials in Quantitative Methods for Psychology*, 2, (1), 1-10.
9. Miller, I., Freund, J.E. & Johnson, R.A. (1990). *Probability and Statistics for Engineers*. 4th Ed., Prentice Hall Inc.
10. Radhakrishnan, V.R., Ramasamy, M., Zabiri, H., Do Thanh, V., Tahir, N.M., Mukhtar, H., Hamdi, M.R. & Ramli, N. (2007). Heat exchanger fouling model and preventive maintenance scheduling tool. *Applied Thermal Engineering*, 27, 2791-2802.