Soft Sensors for Kerosene Properties Estimation and Control in Crude Distillation Unit

N. Bolf, G. Galinec, and M. Ivandić
Faculty of Chemical Engineering and Technology,
University of Zagreb, Department of Measurements and Process Control,
Savska c. 16/5a, 10 000 Zagreb, Croatia

Original scientific paper
Received: March 17, 2008
Accepted: June 4, 2009

Neural network-based soft sensors are developed for kerosene properties estimation, a refinery crude distillation unit side product. Based on temperature and flow measurements, two soft sensors serve as the estimators for the kerosene distillation end point (95 %) and freezing point.

Soft sensor models are developed using linear regression techniques and neural networks. After performing multiple linear regression analysis it is determined that it is not possible to realize linear models. Within MLP neural networks the number of neurons in the hidden layer are varied and different learning algorithms are used (back propagation with variations of learning rate and momentum, conjugate gradient descent, Levenberg-Marquardt) as well as pruning and Weigend regularization techniques. Bootstrap resampling with replacement and cross-validation resampling are used for improving generalization capabilities. Statistics and sensitivity analysis is provided for both models. Two developed soft sensors will be used in crude-oil unit as on-line estimators of kerosene properties, which so far were available only as infrequent and irregular laboratory analyzers.

Key words:
Crude distillation unit, kerosene, soft sensor, process monitoring and control, neural network

Introduction

Control systems and optimization procedures require regular and reliable measurements at the appropriate frequency. At the same time, legislation dictates strict product quality specifications and refinery emissions. As a result, greater number of process variables need to be measured and new expensive process analyzers need to be installed to achieve efficient process control. The quality measure may only be available as a laboratory analysis or very infrequent on-line measurement. This can lead to excessive off-specification products.[1]

Difficulties in measuring quality (primary) variables inevitably mean poor control or no control at all. Measurement difficulties may be caused by a variety of reasons, including:

– Lack of appropriate on-line instrumentation;

– Process operations depend on laboratory analyzers, which can be infrequent and irregular, in addition to long delays;

– Reliability of on-line instruments.

On-line sensors may be available but they may suffer from long measurement delays (e.g. gas chromatographs) or be subject to factors that affect the reliability of the sensor (e.g. drifts and fouling).[2,3]

In either case, on-line control or optimization schemes cannot be implemented. Measurement problems can limit the applicability of feedback control schemes, so common approach for resolving this problem is manual process control. Success of that strategy depends solely on the operator's training and experience.

In developing soft sensors, any modeling paradigm may be employed, including the development of first principles models. In many cases, only data based modeling methods are involved.[3,4] Using artificial neural network and genetic programming paradigms it is possible to capture non-linear process characteristics. If sufficiently accurate, the inferred primary output states may then be used as a feedback for automatic control and optimization.[5]

The application of soft sensors for estimating hard-to-measure process values is extremely interesting in the process industry, where there are usually a large number of values measured continuously and quickly, and they may serve as input signals for the soft sensor. In addition, processes in the chemical industry usually have relatively slow dynamic behavior with a major temporal delay, and,

* e-mails: bolf@fkit.hr, ggalin@fkit.hr, mivandic@fkit.hr
phone: + 385 1 4597 151, fax: + 385 1 4597 165

since the soft sensor is being realized in a computer, it may estimate values in advance, long before the real process output. In this way, it becomes possible to apply predictive control.[6]

They can work in parallel with real sensors, allowing fault detection schemes devoted to the sensor's status analysis to be implemented.[7,8] Also, they can take the place of sensors which have been taken off for maintenance, to keep control loops working properly and to guarantee product specification without undertaking conservative production policies, which are usually too expensive.[9]

Crude oil distillation process description

The crude distillation unit is one through which the entire crude entering a refinery must be processed. Because of the highly competitive market and stringent environmental laws, strict quality control of refinery products is essential. This requires that many properties should be measured online so that the unit can be effectively controlled through a feedback mechanism.[10,11]

Crude distillation unit consists of crude distillation column at atmospheric pressure, stripping unit, preheating section that uses a top and bottom pump-around and overhead condensing system. Products of the crude distillation unit are: unstabilized naphtha, heavy naphtha, kerosene, light gas oil (LGO), heavy gas oil (HGO), and atmospheric residue, as shown in Fig. 1. Kerosene is the second crude distillation column side-product. It is used for lighting and heating, and as fuel for jet and turbo-prop aircraft engines. Variables that directly or indirectly influence or indicate kerosene proper-

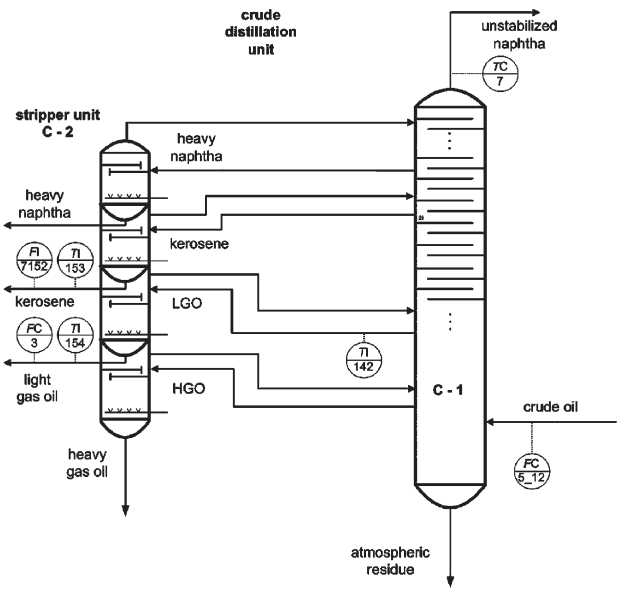


Fig. 1 – *Crude distillation unit*

ties are monitored continuously. Based on process expert knowledge, the following variables have been chosen as the input variables that could serve as estimators of distillation end point and kerosene freezing point:

- unstabilized naphtha temperature (*T*C7)
- kerosene temperature (*T*I153)
- kerosene flowrate (*F*I7152)
- LGO temperature (*T*I154)
- LGO flowrate (*F*C3)
- crude oil inlet flowrate (*F*C5_12).

Kerosene properties were determined by carrying out laboratory assays based on the following standards:

- freezing point – ASTM D 2386 – 05
- distillation end point – EN ISO 3405

Soft sensor model development

Kerosene distillation end point and freezing point are the two properties which mainly depend on kerosene outlet temperature. The kerosene outlet temperature is not directly controlled, but it is influenced by the kerosene outlet flowrate. The kerosene properties partially depend on outlet temperatures and flowrates of its neighboring fractions. Therefore, the idea was to take the outlet temperatures and flowrates of light gas oil and heavy naphtha. Since heavy naphtha in this unit is not dragged from the column, unstabilized naphtha temperature from the top of column has been taken as an input. The last input variable for soft sensor models, crude oil inlet flowrate, has been taken because the flowrates of all outlet fractions depend on it along with the temperatures of those fractions and consequently their properties. Also, it should be kept in mind that periodically different mixes of crude oil are used in crude unit production which influences the work regime of the refinery.

Soft sensor models were developed using linear regression techniques and neural networks in Statsoft Statistica 7.1. Firstly, outliers were removed from raw data. Outliers were defined as values that are more than 1.5 times the interquartile range away from the 25th or 75th percentile. Data from plant database were scaled into a range appropriate for the network development. The minimax scaling (–1 to 1) method was used for input and output variables.

During preliminary testing, 20 data were resampled. To compare different structure of neural networks (radial basis function – RBF and multilayer perceptron – MLP) by implementation of random selection (Monte Carlo resampling) equalized division of data on training, selecting and testing subsets was obtained.[12,13] Within each resampling, 100 neural networks were tested from which 10 with best performance were kept. It was shown that the MLP network had the best properties. Therefore, they were used for additional testing to improve generalization capabilities.

During MLP neural networks testing, the number of neurons in the hidden layer was varied from 1 to 20 and different learning algorithms were used (back propagation with variations of learning rate and momentum, conjugate gradient descent, Levenberg-Marquardt) as well as *pruning* and *Weigend regularization* techniques. *Bootstrap resampling with replacement* and *cross-validation resampling* (4-fold, 5-fold and 10-fold)[14] were used with the aim to improve generalization. It was shown that the best results were achieved with cross-validation resampling. Also, the best results were achieved using a combination of back-propagation algorithm in the first, and conjugate gradient descent algorithm in the second stage of neural network training.

The full data set was divided randomly into five. One of the data parts was selected for testing, and the others were used for training and selecting in ratio 2:1. The rest went into ignore subset. Since cross-validation technique was used, the process was repeated five times. The same procedure was carried out with 10-fold technique. The train set is the set of points that are used to fit the parameters of the model. The select set is used as part of the model building process to prevent overfitting. The test set is used as an additional independent set for validation purpose.

Sensitivity analysis was carried out by treating each input variable in turn as if it were "unavailable".[14] Every model has defined a missing value substitution procedure, which is used to allow predictions to be made in the absence of values for one or more inputs. The basic measure of sensitivity is the ratio of the error with missing value substitution to the original error. The more sensitive the network is to a particular input, the greater the deterioration we can expect, and therefore the greater the ratio. It was shown that all six inputs must be taken into account for both soft sensors.

Results and discussion

Kerosene distillation end point and freezing point soft sensors were developed based on data sets consisting of 415 and 167 laboratory assays, respectively. The data were taken from distributed control system (DCS) in the period from July 2006 to September 2007. The reason for less data for estimation of kerosene freezing point lies in less frequent laboratory analysis. The sampling happens once daily, so models are necessarily steady state. The total number of samples after data preprocessing i.e. removing outliers was brought down to 357 and 142, respectively.

After performing multiple linear regression analysis, it was determined that due to small (around zero) correlation coefficients it was not possible to realize linear models for given inputs and outputs. So, neural network-based soft sensors were developed.

From a variety of different neural network structures – linear, RBFs and MLPs – the best results were achieved with MLPs for both of the soft sensor models.

Developed MLP neural network for kerosene distillation end point, T_{ed}, has 6-7-1 architecture, Fig. 2, and for freezing point, T_{fp}, 6-5-1 architecture, Fig. 3.

Using 5-fold cross validation resampling for kerosene distillation end point, T_{ed}, 357 data sets were distributed as follows: 189 data in training

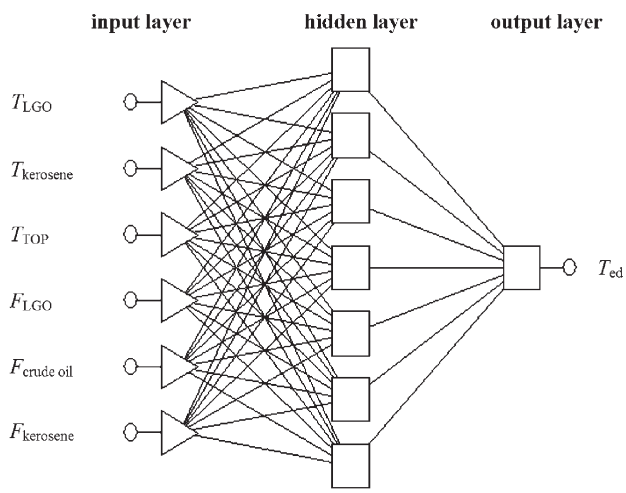


Fig. 2 – *Neural network architecture for kerosene distillation end point, T_{ed}*

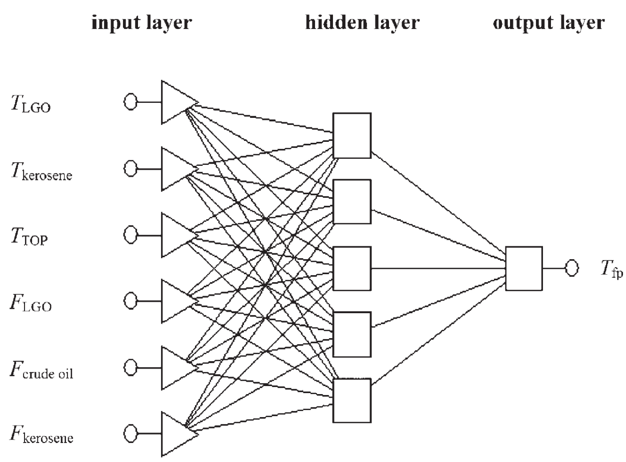


Fig. 3 – *Neural network architecture for freezing point, T_{fp}*

Table 1 – *Model summary report for kerosene distillation end point, T_{ed}*

Profile, T_{ed}	Train perf.	Select perf.	Test perf.	Train error	Select error	Test error	Training/members
MLP 6-7-1	0.8083	0.8445	0.8704	0.1469	0.1558	0.1420	BP100,CG61b

Table 2 – *Model summary report for kerosene freezing point, T_{fp}*

Profile, T_{fp}	Train perf.	Select perf.	Test perf.	Train error	Select error	Test error	Training/members
MLP 6-5-1	0.7851	0.8407	0.8457	0.1329	0.1546	0.1585	BP100,CG61b

Perf. – equal to σ ratio – Error σ and Data σ ratio

Error – the sum of the squared differences between the target and actual output values on each output unit. This is the standard error function used in regression problems.

BP100CG161b signifies "one hundred iterations of back propagation, followed by sixty-one iterations of conjugate gradient descent, at which point training was terminated due to over-learning and the best network in the training run retrieved."

dataset, 95 data in selection and 70 in test datasets. For kerosene freezing point, T_{fp}, 10-fold cross validation resampling was used and 142 data sets were distributed accordingly: 84 data in training dataset, 42 in selection and 13 in test dataset. Selection of data in subsets was done randomly and with respect to normal distribution.

Model summary report tables are given for both models which show performances and errors of each individual subset of data, train, select and test data sets, Tables 1 and 2.

The best neural networks, for both, kerosene distillation end point, T_{ed} and kerosene freezing point, $T_{fp,}$ were chosen according to smallest select error, but it was also important that Train Perf., Select Perf. and Test Perf. have approximately equal values for all three data subsets, which indicates that behaviour of both neural network models in all three data subsets is equable and both neural networks have better generalization abilities.

Neural network weights for T_{ed} and T_{fp} are given in Tables 3 and 4, respectively.

In order to investigate the influence of each input variable on the soft sensor outputs sensitivity tests were carried out.

The ratios for both, T_{ed} and T_{fp}, Tables 5 and 6, are somewhat higher than one, which indicates that all input variables have to be taken into account when developing both models, but also that their quantitative influence on the output is scarce.

Table 3 – *Neural network weights for kerosene distillation end point, T_{ed}*

T_{ed}	2.1	2.2	2.3	2.4	2.5	2.6	2.7	3.1
Thresh	0.73344	0.24877	0.83489	–1.14004	–2.23430	–1.20312	–1.22159	–1.30029
1.1	–1.99096	0.20779	0.58647	–1.57480	–1.41092	0.11584	–1.90024	
1.2	0.36353	–0.41322	–1.26291	–1.18202	2.17593	1.17536	1.23081	
1.3	0.67389	1.03383	2.41328	–2.86689	0.06285	–0.65123	1.25566	
1.4	–1.33421	–0.90676	–2.89738	0.33265	–0.02141	–0.14694	1.90959	
1.5	0.21360	0.57702	1.65236	–1.04197	0.83528	–0.02365	–0.45406	
1.6	–0.57435	–1.86089	2.50138	1.49651	0.17057	0.68137	–0.71128	
2.1								1.07914
2.2								–0.41635
2.3								–0.12893
2.4								–0.22975
2.5								–0.60760
2.6								1.49473
2.7								–1.05711

Table 4 – *Neural network weights for kerosene freezing point, T_{fp}*

T_{fp}	2.1	2.2	2.3	2.4	2.5	3.1
Thresh	–0.16071	–0.03047	0.79049	–0.33202	–0.43763	–1.14966
1.1	–3.79436	–3.49958	2.42142	–1.17872	1.52527	
1.2	–1.29148	–0.99620	2.34736	–3.25617	6.95270	
1.3	–3.21757	–0.74459	3.62094	0.83042	–4.08097	
1.4	0.17657	0.47599	0.46709	1.04728	–1.45442	
1.5	–0.64563	–1.18868	0.85322	–0.56919	1.36945	
1.6	8.10975	–3.21323	–8.80043	–2.13861	–0.09024	
2.1						–0.36557
2.2						0.45066
2.3						–0.39957
2.4						0.54328
2.5						0.31215

Table 5 – *Sensitivity analysis for kerosene distillation end point, T_{ed}*

T_{ed}	T_{LGO}	$T_{kerosene}$	T_{TOP}	F_{LGO}	$F_{crude\ oil}$	$F_{kerosene}$
Ratio	1.209777	1.089305	1.159017	1.092740	1.056956	1.058474
Rank	1.	4.	2.	3.	6.	5.

Table 6 – *Sensitivity analysis for kerosene freezing point, T_{fp}*

T_{fp}	T_{LGO}	$T_{kerosene}$	T_{TOP}	F_{LGO}	$F_{crude\ oil}$	$F_{kerosene}$
Ratio	1.083358	1.073875	1.155833	1.036225	1.011410	1.329413
Rank	3.	4.	2.	5.	6.	1.

After sensitivities had been calculated for all variables, they were ranked in order and, as shown in Table 5 T_{LGO} and T_{TOP} had the greatest influence on T_{ed} followed by F_{LGO}, $T_{kerosene}$, $F_{kerosene}$ and $F_{crude\ oil}$. Also, Table 6 shows that T_{fp} is most sensitive to $F_{kerosene}$ and T_{TOP}, followed by T_{LGO}, $T_{kerosene}$, F_{LGO} and $F_{crude\ oil}$.

Regression tables for both soft sensor models show data mean, data standard deviation, error mean, error standard deviation, absolute error mean, standard deviation ratio and correlation values. Regression tables are given for each individual data subset (train, select and test data set), Tables 7 and 9, and overall neural networks, Tables 8 and 10. Data

Table 7 – *Regression table for each individual data subset for kerosene distillation end point, T_{ed}*

T_{ed}	Train	Select	Test
Data mean	229.0370	228.6000	229.1714
Data σ	3.0884	3.1065	2.7722
Error mean	–0.0329	0.3645	0.0552
Error σ	2.4963	2.6235	2.4128
Abs. e. mean	1.8671	2.1808	1.8657
σ ratio	0.8083	0.8445	0.8704
correlation, R	0.5891	0.5410	0.5465

Table 8 – *Regression table for overall neural network for kerosene distillation end point, T_{ed}*

T_{ed}	Overall
Data mean	228.9328
Data σ	3.0616
Error mean	0.1012
Error σ	2.5600
Abs. e. mean	1.9731
σ ratio	0.8361
correlation, R	0.5545

Table 9 – *Regression table for each individual data subset for kerosene freezing point, T_{fp}*

T_{fp}	Train	Select	Test
Data mean	–50.8929	–50.5714	–50.6154
Data σ	1.3542	1.4662	1.4956
Error mean	0.0005	–0.0991	–0.0954
Error σ	1.0632	1.2326	1.2647
Abs. e. mean	0.8524	1.0112	1.0680
σ ratio	0.7851	0.8407	0.8457
correlation, R	0.6197	0.5415	0.5337

Table 10 – *Regression table for overall neural network for kerosene freezing point, T_{fp}*

T_{fp}	Overall
Data mean	–50.7465
Data σ	1.4164
Error mean	–0.0702
Error σ	1.1477
Abs. e. mean	0.9336
σ ratio	0.8103
correlation, R	0.5860

mean is defined as arithmetic mean of experimental values:

$$\bar{y}_{\mathrm{exp}} = \frac{\sum_{i=1}^{n} y_{\mathrm{exp},i}}{n} \tag{1}$$

Data σ represents standard deviation of experimental values:

$$\text{Data } \sigma = \sqrt{\frac{\sum_{i=1}^{n} (y_{\mathrm{exp},i} - \bar{y}_{\mathrm{exp}})^2}{n}} \tag{2}$$

Errors are defined as differences between the corresponding values that are predicted by the model and the experimental values:

$$e_i = \hat{y}_i - y_{\mathrm{exp},i} \quad i = 1, 2, \ldots, n \tag{3}$$

Regression tables give error mean and error standard deviations. Error mean is arithmetic mean of neural network error:

$$\bar{e} = \frac{\sum_{i=1}^{n} e_i}{n} \tag{4}$$

Error σ is standard deviation of neural network error:

$$\text{Error } \sigma = \sqrt{\frac{\sum_{i=1}^{n} (e_i - \bar{e})^2}{n}} \tag{5}$$

Abs e. mean is absolute error mean of neural network calculated for all data:

$$\text{Abs. error mean} = \frac{1}{n} \cdot \sum_{i=1}^{n} \left| \hat{y}_i - y_{\mathrm{exp},i} \right| \tag{6}$$

σ ratio represents Error σ and Data σ ratio and correlation is Pearson product-moment correlation coefficient (Pearson R – correlation coefficient).

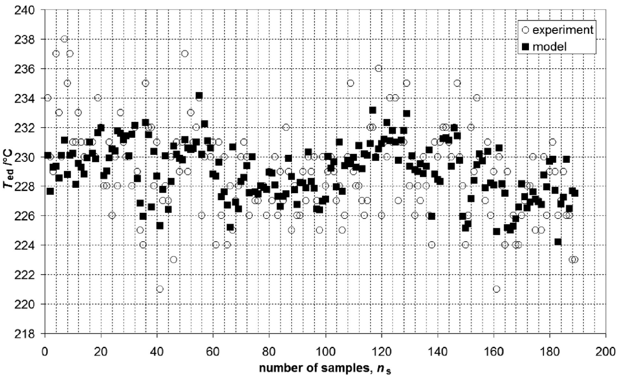
σ ratio in Tables 7 and 9 for individual data subsets corresponds to values of Train Perf., Select Perf. and Test Perf. in model summary report tables, Tables 1 and 2.

Absolute error mean for developed neural network for kerosene distillation end point, T_{ed}, given for all data subsets and for overall neural network model, indicates that differences between the corresponding values predicted by the model and the observed experimental values are around 2 °C. Also, standard deviation ratio for that neural network indicates that standard deviation error of developed neural network is smaller than standard deviation of experimental values which satisfies the required accuracy of the model.

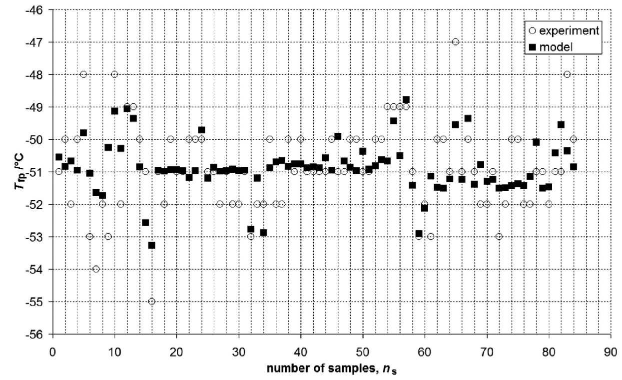
Even better results were achieved with developed neural network for kerosene freezing point, T_{fp}. With absolute error mean of around 1 °C (for all data subsets and for overall neural network model), and smaller standard deviation ratio developed neural network for kerosene freezing point, T_{fp} is even more accurate than neural network for kerosene distillation end point, T_{ed}.

Correlation coefficient for both neural networks is relatively small, around 0.5, but when taken into account that the value of ratio for all six inputs in both neural networks is around 1, in sensitivity analysis table, Tables 5 and 6, it is to be expected that correlation coefficients can not get any higher.

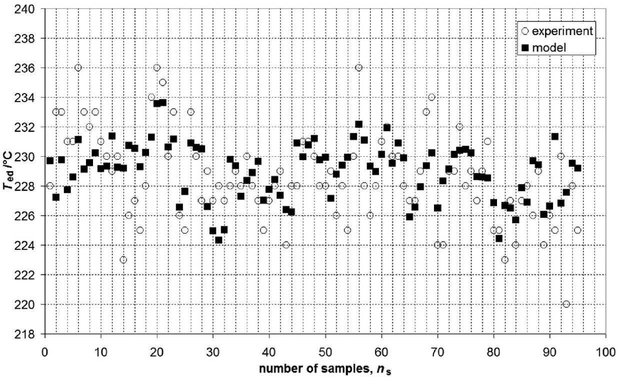
The comparison of results achieved by laboratory analysis and model prediction for both models and each data subset, train, select and test are shown in Figs. 4 through 9. Number of samples is the number of data in each subset with frequency of laboratory analysis once a day. T_{ed} and T_{fp} represent distillation end point and freezing point temperatures determined experimentally (○) and evaluate by soft sensor (■), respectively. These data were omitted during model building procedure for model



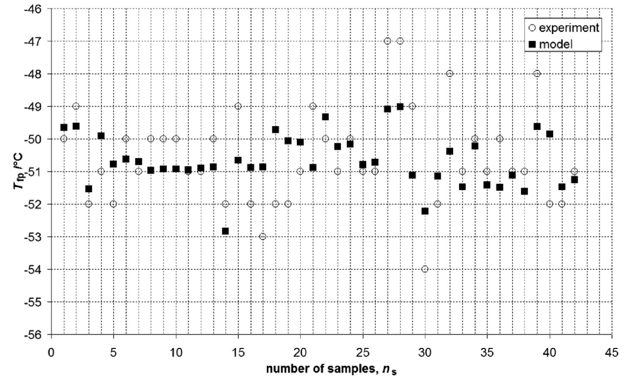
F i g . 4 – *Comparison of laboratory assays and model results for kerosene distillation end point for training data set*



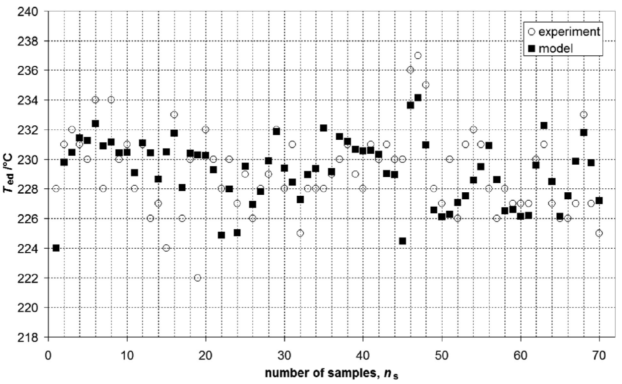
F i g . 7 – *Comparison of laboratory assays and model results for kerosene freezing point for training data set*



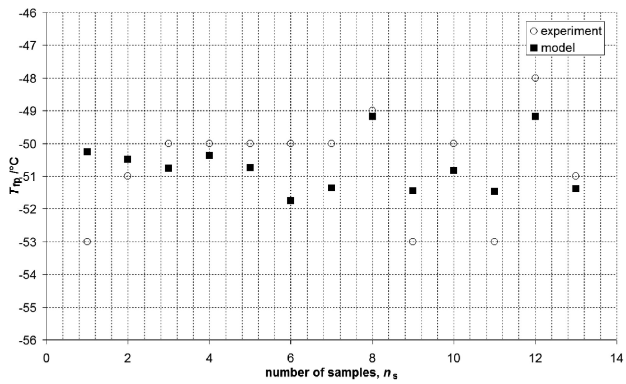
F i g . 5 – *Comparison of laboratory assays and model results for kerosene distillation end point for selection data set*



F i g . 8 – *Comparison of laboratory assays and model results for kerosene freezing point for selection data set*



F i g . 6 – *Comparison of laboratory assays and model results for kerosene distillation end point for testing data set*

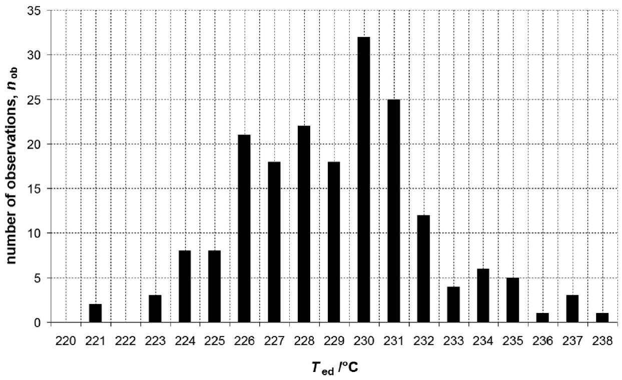


F i g . 9 – *Comparison of laboratory assays and model results for kerosene freezing point for testing data set*

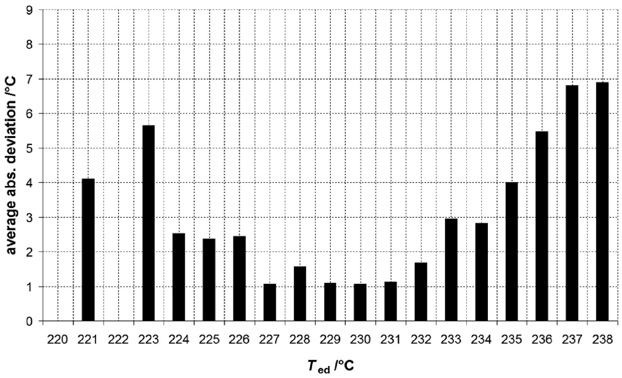
training, prevention of overfitting and validation purpose.

Distillation end point model results deviate from experimental data with average absolute deviations of around 2 °C. Also, average absolute deviations from laboratory determined freezing point temperatures is around 1 °C.

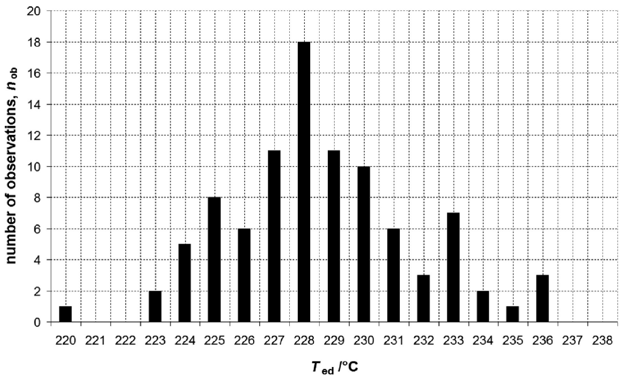
Figs. 10, 12 and 14 represent frequency of occurrence for kerosene distillation end point temperature, T_{ed}, obtained by laboratory assays for each data subset. One can conclude that results have normal distribution (Gaussian distribution) of around 228 °C in the range from 220 to 238 °C. The average absolute temperature deviations for kerosene distillation end point obtained by the soft sensor model in dependence of distillation end point temperatures, T_{ed}, are shown in Figs. 11, 13 and 15, for each neural network data subset. Average absolute



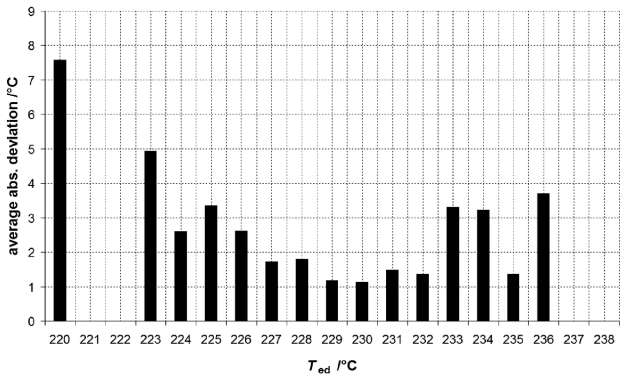
F i g . 1 0 – *Distribution of laboratory assay results for kerosene distillation end point temperature for training data set*



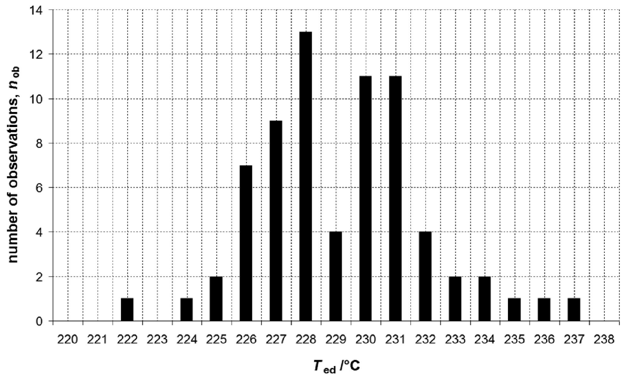
F i g . 1 1 – *Average absolute deviations of kerosene distillation end point temperature obtained by soft sensor model and laboratory assays (training data set)*



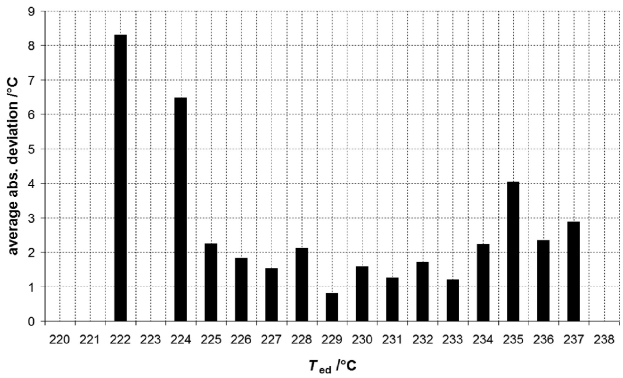
F i g . 1 2 – *Distribution of laboratory assay results for kerosene distillation end point temperature for selection data set*



F i g . 1 3 – *Average absolute deviations of kerosene distillation end point temperature obtained by soft sensor model and laboratory assays (selection data set)*



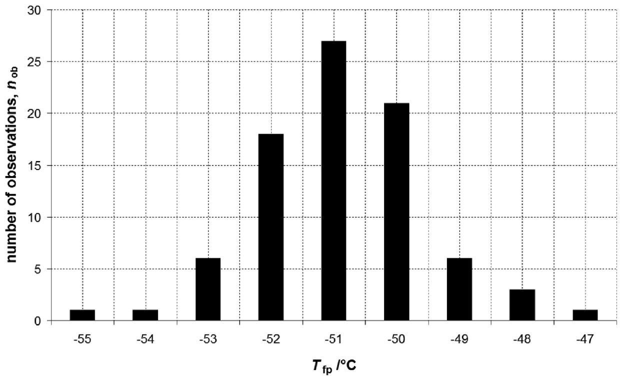
F i g . 1 4 – *Distribution of laboratory assay results for kerosene distillation end point temperature for testing data set*



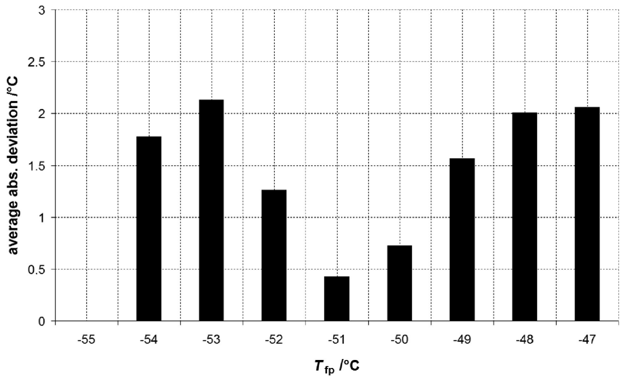
F i g . 1 5 – *Average absolute deviations of kerosene distillation end point temperature obtained by soft sensor model and laboratory assays (testing data set)*

deviations are calculated the same way as the absolute error mean, Tables 7 through 10, but for each individual temperature interval. From the diagrams, it is apparent that the lowest soft sensor model deviations are obtained in the temperature range between 225 and 232 °C. This was expected since large amount of experimental data is found in that range. When approaching margins, the networks' capability to predict significantly decreases.

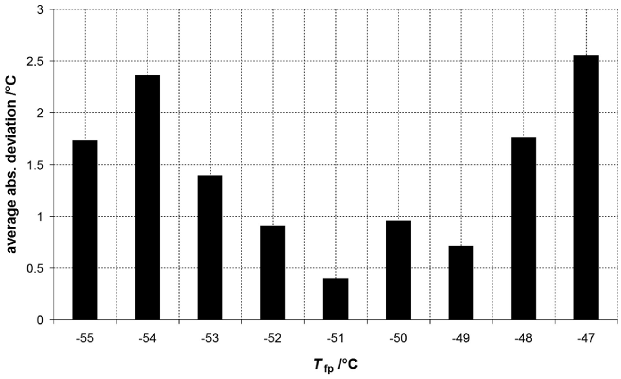
Figs. 16, 18 and 20 represent frequency of occurrence for freezing point, T_{fp}, obtained by laboratory assays for each data subset. The results are normally distributed around −51 °C in the temperature range from −55 to −47 °C. The only deviation from this is in the testing subset. Since 10-fold cross validation resampling was used for freezing point, T_{fp}, in 142 data it leaves only 13 data in the testing subset. This is a small data subset and since



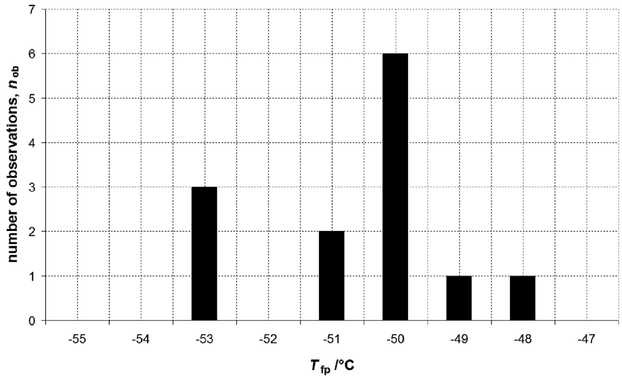
F i g . 1 6 – *Distribution of laboratory assay results for kerosene freezing point temperature for training data set*



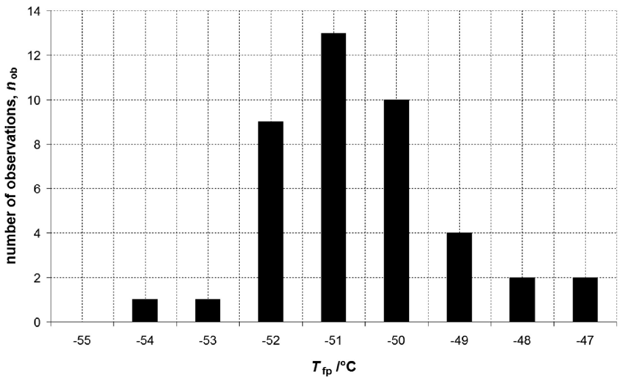
F i g . 1 9 – *Average absolute deviations of kerosene freezing point temperature obtained by soft sensor model and laboratory assays (selection data set)*



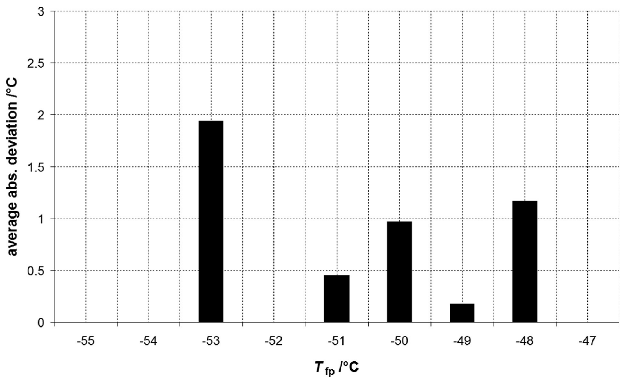
F i g . 1 7 – *Average absolute deviations of kerosene freezing point temperature obtained by soft sensor model and laboratory assays (training data set)*



F i g . 2 0 – *Distribution of laboratory assay results for kerosene freezing point temperature for testing data set*



F i g . 1 8 – *Distribution of laboratory assay results for kerosene freezing point temperature for selection data set*



F i g . 2 1 – *Average absolute deviations of kerosene freezing point temperature obtained by soft sensor model and laboratory assays (testing data set)*

the data is chosen randomly it is, at the same time, hard to achieve normal distribution. It should be kept in mind that 10-fold cross-validation technique randomly distributes data ten times and only data subsets of developed neural network with the best statistical parameters are shown.

Figs. 17, 19 and 21 depict average absolute temperature deviations for kerosene freezing point obtained by the soft sensor model in dependence of freezing point temperatures, T_{fp}, for all data subsets. The lowest soft sensor deviations are obtained in the temperature range between −52 and −50 °C. This range provides the majority of experimental data.

To improve the model performance ten suboptimal models were generated, sorted and selected on the basis of their Pearson R – correlation coefficient and aggregated in a network ensemble using simple average and non linear combination via neural networks. It was observed that Pearson R – correlation coefficient increased negligibly.

Conclusion

Kerosene distillation end point and freezing point soft sensors are developed based on refinery data from DCS and laboratory assays.

Since the results achieved by laboratory analysis for distillation end point and freezing point temperature are reported with the measurement uncertainty ± 1 °C, average absolute deviations of both model results are acceptable for implementation. The reason for this small value of correlation coefficient could be in systematic error, which is the subject of future research. Also, increased predicting capability on the operating window edges will be possible when new process data and analysis will be available.

The design of soft sensors in the presence of small data sets, that is common in industrial practice, posses a challenge as it regards the verification and validation of model performance, when the performance are evaluated on a very small test set.

The implementation of soft sensors in refinery plants is a challenging task. It involves synergy between plant experts, system analysts and process operators. Neural networks play an important role in the development of soft sensors. Well-trained neural networks can be employed as soft sensors for on-line estimation and prediction of key process parameters.

ACKNOWLEDGMENT

The authors wish to thank Mr. Darko Klarić and Mr. Krešimir Jednačak with INA-Refinery Rijeka who contributed to the work through advice and counsel.

This paper is a result of the scientific project ZP-125-1963-1964 Soft sensors and analyzers for process monitoring and control, supported by the Croatian Ministry of Science, Education and Sports.

List of symbols

F – volumetric flowrate, $m^3\ h^{-1}$
n_{ob} – number of observations
n_s – number of samples
T – temperature, °C
R – correlation coefficient
σ – standard deviation

Abbreviations

DCS – distributed control system
ed – distillation end point
fp – freezing point
HGO – heavy gas oil
LGO – light gas oil
MLP – multi-layer perceptron
RBF – radial basis function

References

1. *Tham, M. T., Morris, A. J., Montague, G. A.*, Chem. Eng. Res. Des. **67** (1989) 547.
2. *Martin, G. D.*, Chem. Eng. Prog. **93** (1997) 66.
3. *Quek, C. J., Balasubramanian, R., Rangaiah, G. P.*, Hydrocarbon Process., Int. Ed. **79** (2000) 101.
4. *Bolf, N., Kopčić, N., Briški, F., Gomzi, Z.*, Chem. Pap. **61** (2007) 98.
5. *Jerbić, I., Pavelić, H., Bolf, N.*, Fuels and Lubricants **46** (2007) 149.
6. *Willis, M. J., Montague, G. A., Di Massimo, C., Tham, M. T., Morris, A. J.*, Automatica **28** (1992) 1181.
7. *Rizzo, A., Xibilia, M. G.*, IEEE Transactions on Control System Technology **10** (2002) 421.
8. *Buceti, G., Fortuna, L., Rizzo, A., Xibilia, M. G.*, Fusion Eng. Des. **60** (2002) 381.
9. *Fortuna, L., Rizzo, A., Sinatra, M., Xibilia, M. G.*, Control Engineering Practice **11** (2003) 1491.
10. *Cerić, E.*, Petroleum – Processes and products, INA and Kigen, Zagreb, 2006 (in Croatian)
11. *Chatterjee, T., Saraf, D. N.*, J. Process Control **14** (2004) 61.
12. *Licitra, S., Xibilia, M. G., Graziani, S.*, A neural based sensor for a debutanizer distillation column, in Proc. of the 3rd European IFS Workshop, Intelligent Forecasting Systems for Refineries and Power Systems, June 2000, Santorini (Greece).
13. *Willis, M. J., Di Massimo, C., Montague, G. A., Tham, M. T., Morris, A. J.*, Inferential measurement via artificial neural networks, in Proc. IFAC Symp. ITAC'91, Jan. 1991, Singapore.
14. Statistica Electronic Manual, version 7.1, Statsoft Inc., Tulsa, 2006.