

## Development of Soft Sensors for Debutanizer Product Quality Estimation and Control

Ivica Jerbić,<sup>a</sup> Nenad Bolf,<sup>b</sup> Hrvoje Pavelić<sup>a</sup>

*a INA, d.d – Sisak Oil Refinery, HR-44 000 Sisak, Croatia*

*b Faculty of Chemical Engineering and Technology, University of Zagreb, HR-10 000 Zagreb, Croatia*

### Abstract

This paper demonstrates soft-sensor design for product quality monitoring and process control of debutanizer column. The column is fed by unstabilized FCC gasoline, and products are Liquefied Petrol Gas (LPG) and stabilized FCC gasoline. Method of estimation of pentane fraction in liquefied petrol gas (LPG) and Reid vapor pressure of stabilized FCC gasoline using inferential model is elaborated. The aim is to control debutanizer thus pentane fraction in LPG is kept under 2 mass percent and RVP of FCC gasoline on desired value (50 kPa). Two neural soft sensor models are developed based on available process measurements and laboratory analysis – first for estimation of pentane fraction in LPG and second for estimation of RVP of stabilized FCC gasoline. For the building of the neural networks the cascade learning based on the cascade-correlation learning paradigm is developed. Developed soft sensors have been additionally validated by additional experimental data and achieved results have been analyzed and compared with laboratory analysis results.

**Keywords:** identification, process modeling, soft sensor, neural network, debutanizer column

## **1. Introduction**

One frequent problem in respect to industrial processes is the inability to measure key process variables in real time, especially the compositions of process streams and product properties. The increased productivity and improved quality of industrial chemical and biochemical processes in the most cases impose an increased demand for the development of advanced process sensor technology. Hence, the development of advanced sensors, which are based on new technologies of analytical chemistry and contemporary electronic using optical fibers and semiconductors, are a very important research field. An alternative solution of using secondary variables, which are easy measurable to infer the values of unmeasured process variables became an important field of interests. These kinds of measurements are called "inferential measurement", "virtual software sensors" or simply "soft sensors" (sensors implemented in software) [1].

Methods of chemometrics are directly related to soft sensors. They describe how data from process analyzers can be analyzed and modeled for use in process monitoring and control [2]–[3].

Control systems and optimization procedures require regular and reliable measurements at the appropriate frequency. Difficulties in measuring quality (primary) variables inevitably mean poor control or no control at all. Measurement difficulties can be caused by a variety of reasons, including:

- lack of appropriate on-line instrumentation

Process operation depend on laboratory analyzes, that results can be infrequent and irregular, in addition to long analysis 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 may be subject to factors that affect the reliability of the sensor (e.g. drifts and fouling).

In either case on-line control or optimization schemes cannot be implemented. Because measurement problems can limit the applicability of automatic feedback control schemes, a common approach to effecting control on the process is to control it manually. Such a strategy is usually adopted when the return of information for control purposes is slow and irregular. Its success depends on the operator's training and experience.

## **2. Soft Sensors**

An alternative approach to the installation of additional instrumentation is the more effective utilization of the measured information already available. Software sensors provide an elegant and effective way to improve the utilization of currently available information. They are primarily a means by which inferences can be made about the state of a process by «fusing» the available on-line, off-line, and historical process information. Inherent in software sensor technology is knowledge of the relationships and interdependencies of the system. The extraction of this knowledge and the form in which it is utilized is the key to the effectiveness of the software sensor technology [4].

The behavior of any process is indicated by the states of output variables, which are dependent on the operating conditions and the inputs to the process, Figure 1. However, productivity is quantified by a subset of these output variables; normally the specifications upon which the product is sold, e.g. purity, physical or chemical properties. These so called *primary* variables are often the ones that are difficult to measure on-line. Inferential measurement systems are designed to overcome such measurement problems. The other outputs (e.g. temperatures, flows and pressures) are called *secondary* variables and these are easily measured on-line. Due to the nature of chemical and process engineering systems, the states of many of the secondary variables reflect the states of primary variables. For example, pressures and temperatures define liquid composition while biomass growth is linked to carbon dioxide evolution and feed rate. Thus it should be possible to use the readily available secondary variables to infer the state of a quality or primary variable [5].



Fig. 1. Primary and secondary process outputs

In developing soft sensors the objective is to model the relationship between a primary output and secondary outputs and inputs. The model then can be used to estimate primary output at the frequency that easily measured inputs and secondary variables are measured. Thus any modeling paradigm may be employed, including the development of first principles models. In many cases it is concerned only data based modeling methods, since first principles modeling can be very difficult and time-consuming. Data based inferential measurement systems have been developed by means of time-series, artificial neural networks and genetic algorithms.

Building and testing the inferential measurement model present an iterative exercise and must be done rigorously prior to on-line implementation. Upon choosing potential secondary variables, we can either use all of them or a subset in building the model. If the modeling paradigms require them, as in the case of time-series, delays between primary output and secondary variable have to be specified. Data filter/smoothing constants have to be selected to attenuate the effects of noise. The parameters of the model are then determined using a suitable numerical optimization or search algorithm and the model validated against data [6].

In the adaptive framework shown at the Figure 2, the measured primary output is also fed to the parameter estimator, which updates the soft sensor model. The parameter estimator uses errors between estimated and measured primary outputs to lead the parameters of the inferential measurement model to the more representative values. Thus the adaptive inferential measurement scheme is usually a multirate system: estimates of the primary output are generated at the (faster) sampling rate of the secondary outputs and inputs, while adaptation occurs only at the (slower)

sampling rate of the primary variable, or whenever the primary variable becomes available [7]-[9].

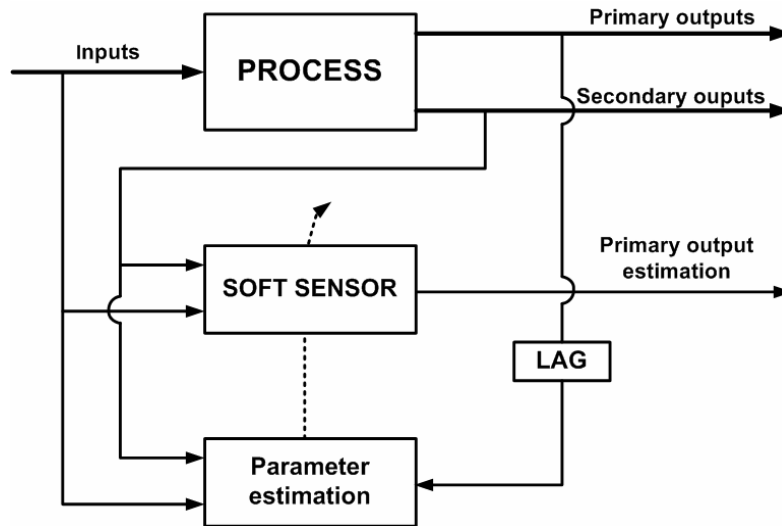


Fig. 2. The structure of an adaptive soft sensor

### 3. PROCESS

Debutanizer column is located at Gas Concentration Unit of FCC plant in INA Sisak Oil Refinery. It is used for FCC gasoline stabilization, precisely for separation of LPG from FCC gasoline. Process scheme is shown on Figure 3.

The column is fed by unstabilized FCC gasoline, and products are Liquefied Petrol Gas (LPG) and stabilized FCC gasoline. After treatment on DEA and MEROX plant LPG becomes commercial. Stabilized FCC gasoline is used as a component for gasoline blending.

The aim is to control debutanizer thus pentane fraction in LPG is kept under 2 mass percent (LPG quality specification) and RVP of FCC gasoline on desired value (50 kPa). Variables that affect debutanizer's product quality are temperatures on top and at bottom, temperatures on specific trays and reflux flow. Temperature of fifth tray is controlled by temperature control loop TRC-93 in cascade with reflux flow control loop FRC-101. Temperature at bottom and on 35<sup>th</sup> tray is controlled by hot flow through column reboiler. Disturbances are flow, temperature and composition variations of column feed.

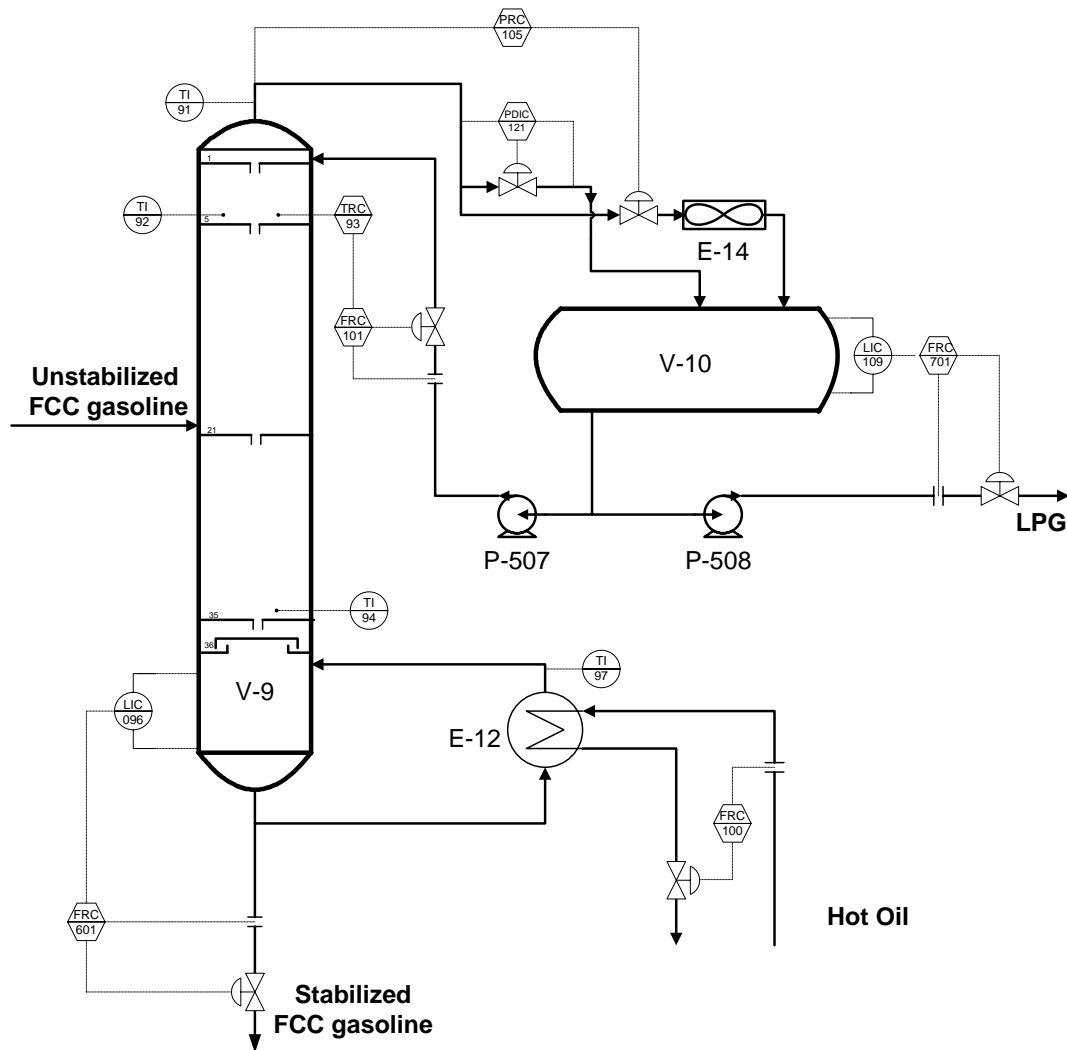


Fig. 3. Diagram of the debutanizer column

## 4. Results and Discussion

### 4.1 Neural Network-Based model

Neural soft sensor models are developed based on available process measurements and laboratory analysis and using software simulation model. Process variables are measured continuously and saved in DCS memory so there is available temperature and flow database. Two neural soft sensor models are developed for continuous product quality monitoring. First one for estimation of pentane fraction in LPG and second for estimation of RVP of stabilized FCC gasoline. Neural network structure used for RVP estimation is shown on Figure 4. It is constructed by input and output layer and hidden layer with five neurons. Network input vector consist of the temperatures on the column's top,  $T_t$ , on the 5<sup>th</sup> tray,  $T_5$ , on the 35<sup>th</sup> tray,  $T_{35}$ , and

reflux flow rate. Output is the RVP value of stabilized FCC gasoline. Similarly, neural network for estimation of pentane fraction in LPG is constructed. Pentane fraction in LPG and RVP value of FCC gasoline database is limited by dynamics of laboratory analysis (once per day).

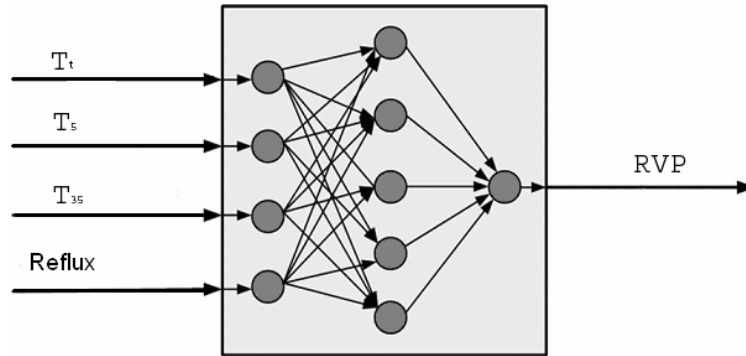


Fig. 4. Applied neural network structure for RVP prediction

#### 4.2 Building of the Neural Network by Cascade Learning

For the building of the neural networks the cascade learning based on the cascade-correlation learning paradigm is developed. Cascade learning starts off with no hidden nodes. The only connections are direct connection from the input layer (and bias) to the output layer. Hidden nodes are added one at a time, and the purpose of each new hidden node is to predict the current remaining output error in the network. Hidden nodes receive input from all previous hidden nodes as well as from the input buffer; in other words, the hidden layer has cascaded connections. For the training adaptive gradient method described in the literature [10] is applied.

In this application the cascade-correlation algorithm works as follows:

1. Train the direct connections from the input layer and bias to the output layer. Train until the RMS (root mean square) output error stabilizes.
  2. Iterate on the following steps:
    - Train a new hidden node so as to maximize a measure of the correlation between its output and residual error at the output for the current training vector. The untrained hidden node is referred to as a “candidate”. When training has stabilized, or after a given number of training iterations, learning is permanently disabled for the incoming connections to that node. At this point, the hidden node is said to be “tenured”;
    - connect the newly tenured hidden node to all nodes in the output layer, and randomly initialize the weights on those connections;
    - train all the weights on all connection from the input layer, bias, and tenured hidden nodes to the output layer. Train until the RMS output error stabilizes.
- The iterative steps are repeated until performance of the network (RMS error measured on a test set) no longer shows any improvement.

The software package [10] provides mechanisms to automatically transform data into formats suitable for neural network training. A variety of analyses are performed

to determine how it might be transformed to optimize the performance of a network. In our application continuous transforms of all input variables are scaled within the range - 1 to +1. This technique is called the zero-mean normalization method. For the most appropriately transformed inputs, the data is concentrated around the average or mean. Mapping the mean to zero, and the balance of the range into the - 1 to +1 interval provides maximum gradient adaptation as the input deviates from its average. Also, during the training of the model, the real world target outputs are transformed to internal target outputs for training the neural network. Neural network generated outputs are transformed to model outputs by putting them through the inverse of the transform that was used to map real world targets to neural network targets.

Figure 5 shows results of testing of the neural model for prediction RVP value in stabilized FCC gasoline. These data are used for the training, i.e. the neural network-based model. As expected, accordance is acceptably.

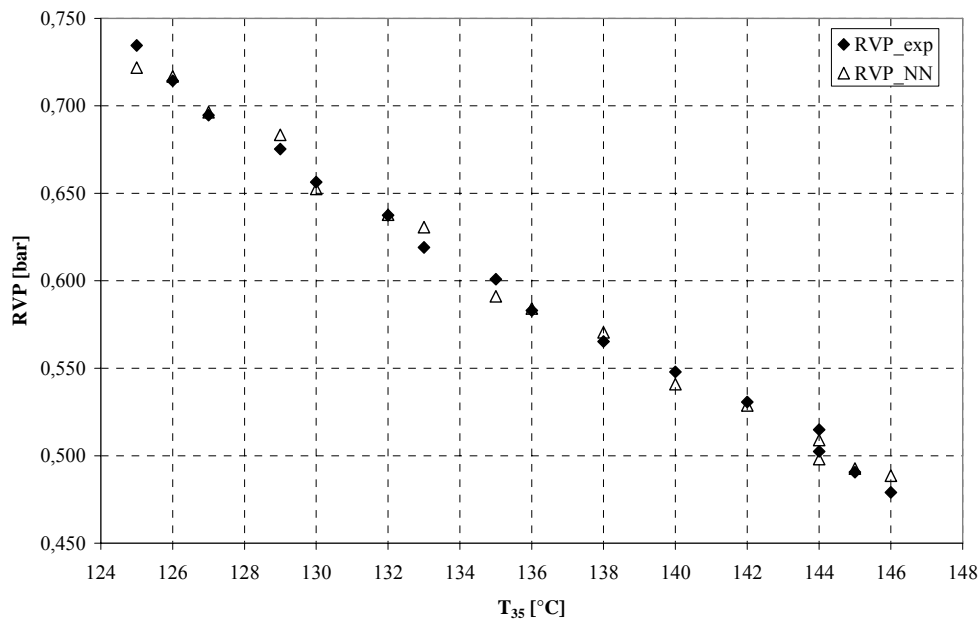


Fig. 5. Results of the soft sensor model testing for predicting of RVP values in stabilized FCC gasoline

Tables 1 and 2 show the comparison of average absolute errors, RMS errors, maximal absolute errors and confidence interval for neural network models. In this context, the training set is the set of points that are used to fit the parameters of the model. The test set is used as part of the model building process to prevent overfitting. The validation set is used as an additional independent test set for validation issues.

Table 1. Neural network-based model statistics for predicting of  $C_5$  in LPG

$C_5$ in LPG	Avg. Abs.	Max. Abs.	RMS	Conf. Int. (95%)
Training	0.10106	1.152207	0.2917485	0.576
Testing	0.101553	1.152207	0.2957075	0.586
Validation	0.099062	1.126737	0.2751214	0.579

*Avg. Abs.* - the average absolute error between predicted output values and the corresponding target values;

*Max. Abs.* - the maximum absolute error between predicted output values and the corresponding target values;

*RMS* - the root mean square error between predicted output values and the corresponding target values;

*Conf. Interval* - the range [target value  $\pm$  confidence interval] within which the corresponding predicted output occurs 95% of the time.

Table 2. Neural network-based model statistics for predicting RVP of FCC gasoline

$C_5$ in LPG	Avg. Abs.	Max. Abs.	RMS	Conf. Int. (95%)
Training	0.005574	0.012725	0.006647	0.013121
Testing	0.005368	0.012725	0.006405	0.012691
Validation	0.006408	0.012203	0.00755	0.015875

Developed soft sensor neural models have been additionally validated by experimental data saved in DCS memory and achieved results have been compared with laboratory analysis results. Figure 6 shows comparison of laboratory analysis and results achieved by soft sensor for predicting RVP value of stabilized FCC gasoline. Soft sensor follows satisfactory the trend of changing RVP values, but certain deviations still persist. These deviations can be explained with the model nonperfection and also with the fact that the values of majority experimental data are situated around 50 kPa. Therefore, the neural network has better approximation characteristics within these values.

Analogously, Figure 7 shows comparison for soft sensor model for predicting of pentane fraction in LPG. Majority of available experimental data was around 0 mass. percent hence the approximation capabilities of the neural network are better at lower values of pentane fraction. Statistics for both model validation sets are shown in Table 3.



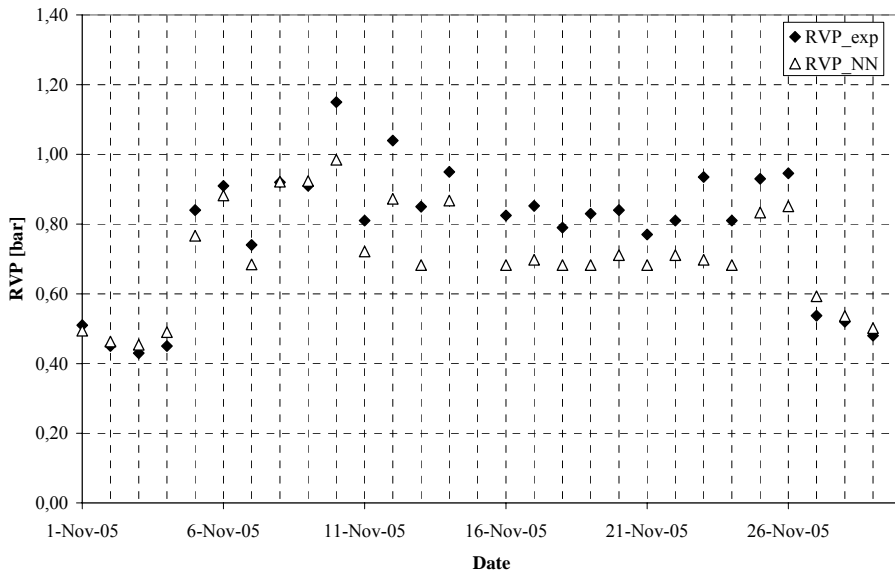


Fig. 6. Comparison of laboratory analysis and results achieved by soft sensor model for prediction of RVP value in stabilized FCC gasoline

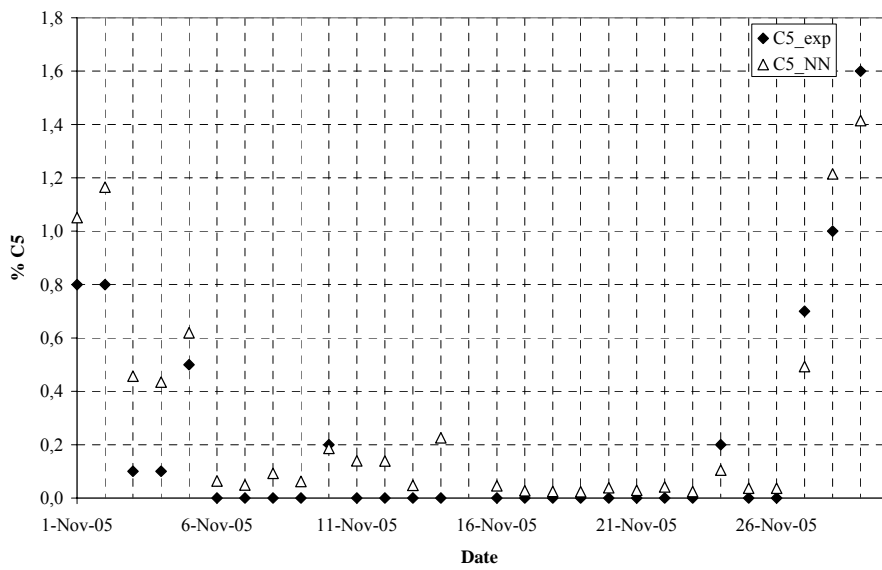


Fig. 7. Comparison of laboratory analysis and results achieved by soft sensor model for prediction of pentane fraction in LPG

Table 3. Neural network-based model statistics for validation sets

C <sub>5</sub> in LPG	Avg. Abs.	Max. Abs.	RMS	Conf. Int. (95%)
RVP	0.08	0.24	0.5660	0.1050
C <sub>5</sub>	0.1	0.4	0.8396	0.1559

## 5. Conclusion

In this paper the neural network-based software sensors are developed with the aim of estimation debutanizer product properties that are not continuously measured. The software sensors estimate pentane fraction in LPG and Reid vapor pressure of stabilized FCC gasoline.

Developed models have proved neural network applications as intelligent sensors for reliable and adequately accurate estimations of primary process outputs, based on secondary process outputs.

Soft sensors are shown to be a good alternative to hardware analyzers for debutanizer products and can be built by using data from existing plant. They are relatively simple for development and coupled with their accuracy and reliability makes them a substantive mechanism for complex process monitoring.

## References

- [1] Martin, G. D., Consider Soft Sensors, *Chem. Eng. Prog.*, July 1997., pp. 66–70, 1997.
- [2] Brown, S. D. (1998). Information and Data Handling in Chemistry and Chemical Engineering: The State of the Field from the Perspective of Chemometrics, *Computers and Chem. Eng.*, 23, 203
- [3] Seborg, D. E., T. F. Edgar and D. A. Mellichamp (2004). Process Dynamics and Control, John Wiley & Sons, New York
- [4] Montague, G.A., Morris, A.J. and Tham, M.T. (1992). Enhancing bioprocess operability with generic software sensors, *J. Biotechnology*, 25, pp. 183-201
- [5] Bolf, N., Glasner M., Jerbić I., Application of Soft Sensors and Inferential Control in Chemical & Biochemical Process Control, *Proceeding of the International Conference "Intelligent Control Systems"*, Brno, Czech Republic, pp. 194-198, 2005.
- [6] Petrović I., Domijan P., Jelavić M., Estimation of Lube Oil Viscosities on a Vacuum Distillation Column, *ICIT 2003*, Maribor, pp. 265-270, 2003.
- [7] Quek C. J., R. Balasubramanian, G. P. Rangaiah (2000). Consider using «soft analysers» to improve SRU control, *Hydrocarbon processing*, pp. 101-106
- [8] Brunet-Manquat, F., Willis, M.J. and Tham, M.T. (1995). Integrated inferential measurement and control, *Proc. ACC'95*, June, Seattle.
- [9] Dufour P., S. Bhartiya, P. S. Dhurjati, F. J. Doyle III (2005). Neural network-based software sensor: training set design and application to a continuous pulp digester, *Control Eng. Practice*, 13, pp.135-143
- [10] Neuralware, *Neural Computing - A Technology Handbook for Professional II/Plus and NeuralWorks Explorer*, NeuralWare Inc., Pittsburgh, USA, 2001.