Sauro Pierucci, Flavio Manenti, Giulia Bozzano, Davide Manca (Eds.)
Proceedings of the 30th European Symposium on Computer Aided Process Engineering
(ESCAPE30), May 24-27, 2020, Milano, Italy.
http://dx.doi.org/10.1016/B978-0-12-823377-1.50203-2

Advanced Process Control of an Industrial Depropanizer Column using Data-based Inferential Sensors

Martin Mojto[a], Karol L'ubušký[b], Miroslav Fikar[a] and Radoslav Paulen[a]

[a]*Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, Bratislava, Slovakia*
[b]*Slovnaft, a.s., Bratislava, Slovakia*
martin.mojto@stuba.sk

Abstract

Inferential sensors are used in industry to infer the values of the imprecisely and infrequently measured (or completely unmeasured) variables from measured variables (e.g., pressures, temperatures). This work deals with the design of inferential sensors suitable for an advanced process control of a depropanizer column of the Slovnaft refinery in Bratislava, Slovakia. We design linear inferential models of top and bottom product compositions. Model calibration is performed using historical production data. We study the effectiveness of several data-based methods (PCA, PLS, LASSO) for the design of inferential sensors. Our results show that the methods, which promote model sparsity are more suitable. Validation using a rigorous mathematical model shows that the designed inferential sensors are sufficient for the advanced process control of the column.

Keywords: Inferential Sensors, Data-based Models, Process Control.

1. Introduction

Majority of advanced controllers in industry is based on linear input-output models (Qin and Badgwell, 2003). This is usually sufficient because of the existence of corrective feedback actions, i.e., receding-horizon principle. In many cases, the input-output models do not include main process characteristics, e.g., product quality, since they are often expensive or impossible to measure. The process performance can thus be improved by designing soft-sensors (also called inferentials), which can infer the values of the unmeasured variables from other measured variables (e.g., pressures, temperatures) and improve the overall process management (Weber and Brosilow, 1972; Joseph and Brosilow, 1978; Joseph, 1999).

The inferential sensor stands for a model designed to predict hard-to-measure variable according to other easily measured variables. In chemical industry, hard-to-measure variables are typically concentrations of products, since online concentration sensors are expensive or too slow for effective monitoring and control. Several works were devoted to state estimation and control of (chemical) process using inferentials (Mejdell and Skogestad, 1991, 1993; Zhang, 2001; Kano et al., 2003; Parvizi Moghadam et al., 2019). The approaches are ranging from simple enhancement of monitoring and control—by heuristically combining several variables to compensate the sensor-noise or external disturbance (e.g., pressure-corrected temperature control)—to the use of

Acknowledgements: This research received funding from the Slovak Research and Development Agency under the project APVV 15-0007 and from the Scientific Grant Agency of the Slovak Republic under the grant 1/0004/17.

Table 1: Composition of feed stream of the depropanizer column.

C3 fraction	C4 fraction			C5 fraction
Propane	n-Butane	Trans-2-Butene	1-Butene	Isopentane
Propylene	Isobutane	Cis-2-Butene	Isobutene	

sophisticated data-mining techniques, which select the optimal structure of the inferential sensor using available measured outputs and their historical records (Morari and Stephanopoulos, 1980).

This work deals with the design of inferential sensors for advanced process control (APC) of a depropanizer column of the Slovnaft refinery situated in Bratislava, Slovakia. We design linear inferential models of top and bottom product composition in order to provide a simple and appropriate structure of the soft-sensors for the APC of the plant. Model calibration is carried out using historical production data. We use various data-based modeling techniques, including principal component analysis (PCA), partial least-squares (PLS), and Least Absolute Shrinkage and Selection Operator (LASSO). We compare the accuracy and the effectiveness of the designed inferential sensors using the validation dataset from historical plant data and also using a rigourous mathematical model of a depropanizer column via gPROMS ModelBuilder. We also assess the appropriateness of the use of a linear soft-sensor for the control of the product quality.

2. Problem Description

The studied depropanizer column (see scheme in Fig. 1) contains 40 trays, operates in above-the-atmospheric pressure, and its feed composition is shown in Tab. 1. The column serves to separate feed mixture to C3-fraction-rich distillate product and to C4/C5-fraction-rich bottom product. The available operational degrees of freedom are feed flowrate F, bottom product flowrate B, distillate flowrate D, reflux flowrate R, heat duty in the reboiler Q_B, and heat duty in the condenser Q_D. Several of these variables are available as historical data. These are marked correspondingly in Fig. 1. The plant measurements, also available from historical data, are pressure at the top of the column p_D, pressure at the bottom of the column p_B, and temperatures of distillate T_D, at the 10$^{\text{th}}$ tray T_{10}, at the 37$^{\text{th}}$ tray T_{37} and at the bottom T_B.

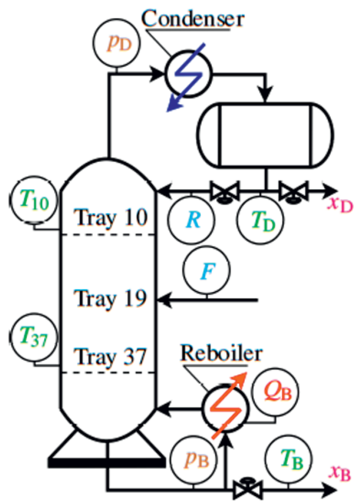


Figure 1: A schematic diagram of the depropanizer column.

It is evident that the use of thermodynamic properties model to monitor top/bottom stream compositions is prohibitive in this case, even under any appropriate ideality assumptions. This is because there are too many degrees of freedom for a seven-component mixture that cannot be inferred from plant data. According to previous paragraph, there are eleven possible input variables to select for the design of an inferential sensor. Current inferentials (denoted as ref), used in the refinery for monitoring/control of concentrations of C4/C5 fraction in the distillate (x_D) and of C3 fraction in the bottom stream (x_B), are designed according to King (2011) in the following form:

$$x_D = a_1 p_D + a_2 T_{10} + a_3 \frac{R}{F}, \qquad x_B = a_1 p_B + a_2 T_{37} + a_3 \frac{Q_B}{F}. \tag{1}$$

Our goal will be to identify models of inferential sensors in the following form

$$x = (a_1, a_2, \ldots, a_{n_p}) \left(F, R, Q_{\mathrm{B}}, p_{\mathrm{D}}, p_{\mathrm{B}}, T_{\mathrm{D}}, T_{\mathrm{B}}, T_{10}, T_{37}, \frac{R}{F}, \frac{Q_B}{F}\right)^T = a^T m, \tag{2}$$

where m is the vector of available input variables for the inferential sensor and $a \in \mathbb{R}^{n_p}$ is the vector of parameters of the inferential sensor.

3. Identification of Inferential Models

Given n data points, we use various statistical methods in order to design the inferential sensors.

- The basic method is Ordinary Least-Squares (OLS) regression. This method estimates the parameters of an inferential sensor according to

  $$\min_a \sum_{i=1}^{n} (x_i - a^T m_i)^2, \tag{3}$$

  which minimizes sum of squared errors between measured compositions and compositions estimated from inferential sensors.
- Principal Component Analysis (PCA) (Pearson, 1901) is a method of identifying a $\tilde{n}_p$-dimensional subspace ($\tilde{n}_p \leq n_p$) of orthogonal coordinates that exhibit a maximum variation in a given dataset. The principal components are identified by SVD decomposition of the covariance matrix of the input dataset MM^T, $M := (m_1, m_2, \ldots, m_n)$, taking the eigenvectors (for subset definition) and the associated eigenvalues (for measure of variance). The (PCA) regression is then done over the selected subspace using (3).
- Partial Least Squares (PLS) regression (Wold et al., 1984) is similar to PCA (Dunn et al., 1989) with the difference that it takes a cross-covariance matrix between inputs and outputs MX^T, $X := (x_1, x_2, \ldots, x_n)$.
- Least Absolute Shrinkage and Selection Operator (LASSO) (Santosa and Symes, 1986) is a method that simultaneously identifies the structure of the model and model parameters by

  $$\min_a \sum_{i=1}^{n} (x_i - a^T m_i)^2 + \lambda \|a\|_1, \tag{4}$$

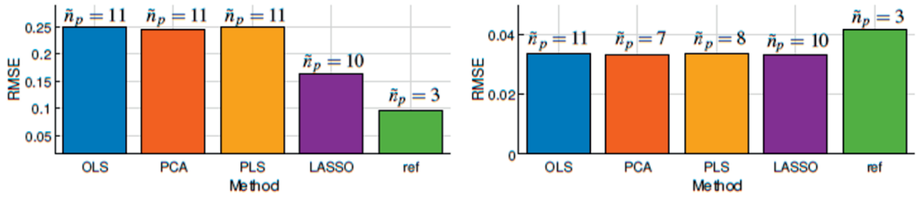
  where λ is a weight between model accuracy and model over-parameterization.

Note that among the presented methods, the PCA method comprises a certain distinct feature. As it does not require any output data, it can potentially be applied on much larger datasets in our general problem setup. This comes since the industrial data are usually available with much finer time granularity for the online sensors rather than for infrequent and expensive measurements.

4. Design of Inferential Sensors using Industrial Data

The design of inferential sensors is performed according to the data from more than two years of production (13.12.2016–21.2.2019) of the depropanizer column. The distillate composition is measured once per month and therefore only 28 measurements is available for the given period. Bottom inferential sensor is design according to 176 measurements as the concentration of this product is measured approximately once per week. The input data provided by online sensors are measured every 30 minutes, which gives an input dataset of 38,360 measurements.

A gross error detection and its subsequent reduction was performed on the given dataset by performing an SVD decomposition of the covariance matrix MM^T and by retaining only those data



(a) Distillate composition. (b) Bottom composition.

Figure 2: Accuracy of the designed inferential sensors according to the industrial validation data. The number of model input variables ($\tilde{n}_p$) is shown on top of each bar.

points that are inside the $\pm 3\sigma$ confidence ellipsoid given by the covariance matrix. The excluded points consist mostly of plant start-up, shutdown and upsets data.

When applying the methods mentioned in Section 3, we use 50 % of the data for training and 50 % of the data for validation. We use the implementations of the design methods available within MATLAB. We are interested in simplest possible structures, which is interesting for APC applications. For this reason, when using PCA and PLS to identify the model structure, we select only those input variables whose principal components explain 90 % of variance in the data. We use root mean square error (RMSE) to assess the accuracy of the designed inferential sensors.

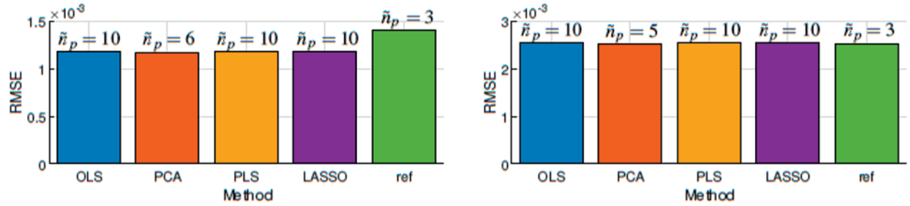
We designed inferential sensors according to the OLS, PCA, PLS and LASSO. We repeated the design 1,000 times, each time randomly distributing the available data among the training and validation datasets. Presented results stand for mean values obtained. The results are shown in Fig. 2 in terms of accuracy of the sensors on the validation dataset for distillate and for bottom composition. Accuracy of the model is influenced by the ability of the method to escape from model over-parameterization. We show the number of variables that the model uses on the top of the bar that represents model accuracy.

Overall, we notice that the nature of the problem of designing the inferential sensors differs between distillate and bottom sensor. While the best distillate sensor uses only few (three) input variables, the most accurate bottom soft-sensors retain 7–11 out of eleven available input variables (all the temperatures and pressures). This is attributed to the small dataset available for the tuning of the distillate sensor and is subject of further analysis.

When comparing the different methods, we can find that OLS is not able to reduce the number of parameters/inputs of the inferential model in both cases. PCA and PLS perform similarly, although the promotion of model sparsity improves with the number of available data (bottom sensor). LASSO achieves slightly better results than OLS in both cases by being able to sparsify the model slightly, even if the dataset is small. The performance of the current on-site inferential sensor (ref) is very good in case of distillate composition inferential. However, there seems to be a room for improvement (around 20 %) of the inferential sensor for bottom composition.

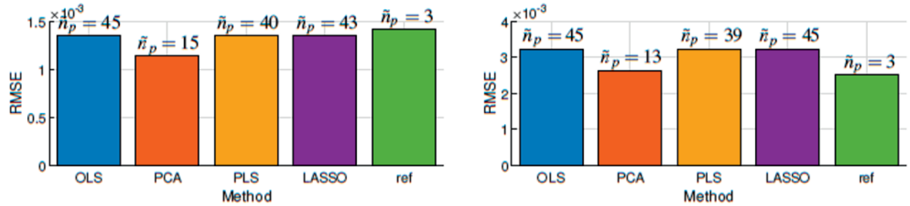
5. Analysis of Inferential Sensors using Synthetic Data

In order to confirm our findings from the previous chapter and to analyse the further possible enhancements of inferential sensors, we designed the mathematical model of depropanizer column in gPROMS ModelBuilder. The model is built according to the parameters from the technical documentation of the depropanizer in the Slovnaft refinery and validated in the simulations with several step changes. We use 200 different (steady state) measurements for the design of inferential sensors. Compared to the historical data from the refinery, this is much larger dataset with many more different operating points (steady-states). The simulated measurements are corrupted with normally distributed random noise of similar magnitudes as present in the industrial data.



(a) Distillate composition. (b) Bottom composition.

Figure 3: Accuracy of the designed inferential sensors according to synthetic validation data. The number of model input variables ($\tilde{n}_p$) is shown on top of each bar.



(a) Distillate composition. (b) Bottom composition.

Figure 4: Accuracy of the designed inferential sensors according to synthetic validation data for 49 possible input variables. The number of model input variables ($\tilde{n}_p$) is shown on top of each bar.

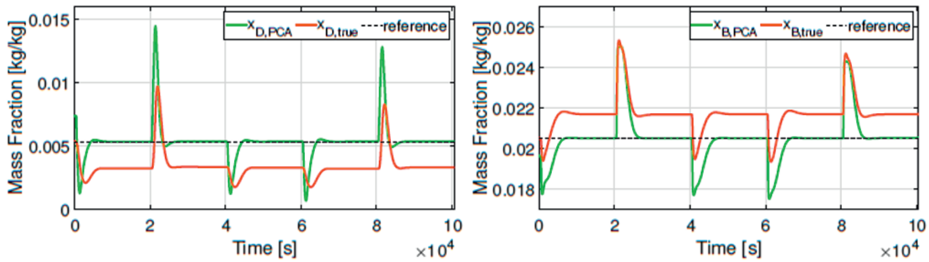
First, we use the information from the same sensors as in the previous section to mimic the case of studied column instrumentation. The results are shown in Fig. 3. We can mostly see that they confirm our previous conclusions. As the noise present in the data corresponds to the assumptions of the least-squares regression and variance analysis, all the methods exhibit an improved performance. Even the OLS method is able to sparsify the model to the extent of PLS and LASSO method. PCA can go even further without compromising the model quality. The slight improvement of the PCA-based distillate sensor w.r.t. ref is caused by adding variables R and F into the sensor structure.

Next, we consider data from more online sensors, 49 instead of 11. The input dataset is enriched by all the tray temperatures ($T_{1\text{–}40}$). In this case, we can conclude that bottom sensor cannot be much improved using new instrumentation. The optimal distillate inferential sensor seems to be 21 % more accurate that the current ref model but it is also much more complicated, one would need at least nine new sensors for the model by PCA. Lastly, we also see that the LASSO method does not cope well with situations when much dimension-reduction is needed.

At last, we tested the inferential control in the scenario, where the plant model is controlled by PI controllers using the following CV-MV pairs: p_{D}-Q_{C}, h_B-B, h_C-D, x_{D}-R, and x_{D}-Q_{B}. The results from the simulation of control using PCA-based inferential sensors are shown in Fig. 5. The prediction and steady-state error is indicated by evaluating the true plant response. We observe 90–95 % accuracy of the control, which is correspondence with the results from prediction accuracy analysis. The designed sensors are thus appropriate for control (APC) purposes.

6. Conclusions

We studied the design of linear inferential sensors for the top and bottom product composition of an industrial depropanizer column. We studied the accuracy and the effectiveness of several advanced data-based methods (PCA, PLS, LASSO). We used historical plant data for model calibration.



(a) The distillate product. (b) The bottom product.

Figure 5: The comparison of product concentrations from the simulation of depropanizer column with implemented PCA inferential sensors.

Further we used a rigorous process model to confirm our findings and to analyze appropriateness of inferential control of the depropanizer column and its potential further improvements. The results show that the designed inferential sensors are sufficient for the advanced process control of the column. The methods that separate model-structure identification and the model regression are found more suitable.

References

W. Dunn, D. Scott, W. Glen, 1989. Principal components analysis and partial least squares regression. Tetrahedron Computer Methodology 2 (6), 349 – 376.

B. Joseph, June 1999. A tutorial on inferential control and its applications. In: Proceedings of the 1999 American Control Conference. Vol. 5. pp. 3106–3118.

B. Joseph, C. B. Brosilow, 1978. Inferential control of processes: Part i. steady state analysis and design. AIChE Journal 24 (3), 485–492.

M. Kano, N. Showchaiya, S. Hasebe, I. Hashimoto, 2003. Inferential control of distillation compositions: selection of model and control configuration. Control Engineering Practice 11 (8), 927 – 933.

M. King, 2011. Process Control: A Practical Approach. John Wiley & Sons Ltd., iSBN: 978-0-470-97587-9.

T. Mejdell, S. Skogestad, 1991. Composition estimator in a pilot-plant distillation column using multiple temperatures. Industrial & Engineering Chemistry Research 30 (12), 2555–2564.

T. Mejdell, S. Skogestad, 1993. Output estimation using multiple secondary measurements: High-purity distillation. AIChE Journal 39 (10), 1641–1653.

M. Morari, G. Stephanopoulos, 1980. Studies in the synthesis of control structures for chemical processes: Part iii: Optimal selection of secondary measurements within the framework of state estimation in the presence of persistent unknown disturbances. AIChE Journal 26 (2), 247–260.

R. Parvizi Moghadam, J. Sadeghi, F. Shahraki, 2019. Optimization of time-variable-parameter model for data-based soft sensor of industrial debutanizer. Optimal Control Applications and MethodsIn Press.

K. Pearson, 1901. Liii. on lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 2 (11), 559–572.

S. J. Qin, T. A. Badgwell, 2003. A survey of industrial model predictive control technology. Control Engineering Practice 11 (7), 733 – 764.

F. Santosa, W. W. Symes, 1986. Linear inversion of band-limited reflection seismograms. SIAM Journal on Scientific and Statistical Computing 7 (4), 1307–1330.

R. Weber, C. Brosilow, 1972. The use of secondary measurements to improve control. AIChE Journal 18 (3), 614–623.

S. Wold, A. Ruhe, H. Wold, W. J. Dunn, III, 1984. The collinearity problem in linear regression. the partial least squares (PLS) approach to generalized inverses. SIAM Journal on Scientific and Statistical Computing 5 (3), 735–743.

J. Zhang, 2001. Inferential feedback control of distillation composition based on PCR and PLS models. In: Proceedings of the 2001 American Control Conference. Vol. 2. pp. 1196–1201.