Identification of a Pilot Scale Distillation Column: A Kernel Based Approach

B. Huyck [*,**,***] **K. De Brabanter** [***] **F. Logist** [**]
J. De Brabanter [*,***] **J. Van Impe** [**] **B. De Moor** [***]

[*] *Department of Industrial Engineering, KaHo Sint Lieven, Gent, Belgium (e-mail: Bart.huyck@esat.kuleuven.be)*
[**] *Department of Chemical Engineering (CIT - BioTeC), K.U.Leuven, Leuven, Belgium*
[***] *Department of Electrical Engineering (ESAT - SCD), K.U.Leuven, Leuven, Belgium*

Abstract: This paper describes the identification of a binary distillation column with *Least-Squares Support Vector Machines (LS-SVM)*. It is our intention to investigate whether a kernel based model, particularly an LS-SVM, can be used for the simulation of the top and bottom temperature of a binary distillation column. Furthermore, we compare the latter model with standard linear models by means of mean-squared error (MSE). It will be demonstrated that this nonlinear model class achieves higher performances in MSE than linear models in the presence of nonlinear distortions. When the system is close to linear, the performance of the LS-SVM is only slightly better than the linear models.

Keywords: Chemical Industry; Distillation columns; kernel based system identification

1. INTRODUCTION

In a world where economic and environmental issues become more and more important, efficient knowledge of the behavior of a process has become indispensable. Mathematical models are heavily exploited for the predicting of process behaviour, e.g., in view process monitoring and control. In the case of control, prediction and simulation is mostly done by linear models (Qin and Badgwell, 2003). In the academic world, however, an evolution towards nonlinear models can be observed. Both linear and nonlinear models can be built on mechanistic knowledge (white-box models) or available input-output data (black-box models). As methods of the latter class can generally be employed flexibly and without a large effort, these (linear) black-box are often preferred in industrial practice. In contrast to linear systems where black-box system identification techniques are well understood and described (Ljung, 1999), for nonlinear systems a variety of possible model structures and techniques exists, e.g., neural networks, wavelets, fuzzy models and Least Squares Support Vector Machines. In this paper we focus on the applicability of LS-SVMs for black-box system identification of a pilot scale binary distillation column. As most of industrial process control applications use linear models, the LS-SVM models are compared to standard linear techniques as transfer function models, subspace state-space models and the Box-Jenkins type models. This paper is structured as follows. The next section focuses on the model structure of an LS-SVM. Section 3 introduces the binary distillation column. Section 4 presents the identification procedure. In Section 5 the results are presented. Finally, Section 6 summarises the main conclusions.

2. LEAST SQUARES SUPPORT VECTOR MACHINES

The standard framework for LS-SVM is based on a primal-dual formulation. Given a training data set $\mathcal{D}_n = \{(u_k, y_k) : u_k \in \mathbb{R}^d, y_k \in \mathbb{R}; k = 1, \ldots n\}$ of size n:

$$y_k = m(u_k) + \epsilon_k, \qquad k = 1, \ldots, n, \tag{1}$$

where $\epsilon_k \in \mathbb{R}$ are assumed to be independent and identically distributed zero mean random errors with finite variance. The optimization problem of finding the vector w and $b \in \mathbb{R}$ for regression can be formulated as follows (Suykens et al., 2002):

$$\begin{aligned} \min_{w,b,e} \mathcal{J}(w,e) &= \tfrac{1}{2} w^T w + \tfrac{\gamma}{2} \sum_{k=1}^{n} e_k^2 \\ \text{s.t.} \quad y_k &= w^T \varphi(u_k) + b + e_k, \quad k = 1, \ldots, n, \end{aligned} \tag{2}$$

where $\varphi : \mathbb{R}^d \to \mathbb{R}^{n_h}$ is the feature map to the high dimensional feature space (Vapnik, 1999), unknowns $w, b \in \mathbb{R}$ and residual e. However, we do not need to evaluate w and $\varphi(\cdot)$ explicitly. By using Lagrange multipliers for the optimization problem (2):

$$\begin{aligned} \mathcal{L}(w,b,e;\alpha) = \frac{1}{2} w^T w + \frac{\gamma}{2} \sum_{k=1}^{n} e_k^2 \\ - \sum_{k=1}^{n} \alpha_k \{ w^T \varphi(u_k) + b + e_k - y_k \}, \end{aligned}$$

where α_k are the Lagrange multipliers, the Karush-Kuhn-Tucker (KKT) conditions for optimality are given by $\frac{\partial \mathcal{L}}{\partial w} = \frac{\partial \mathcal{L}}{\partial b} = \frac{\partial \mathcal{L}}{\partial e_k} = \frac{\partial \mathcal{L}}{\partial \alpha_k} = 0$. After elimination of the variables w and e, the solution is given by the linear system (3):

$$\left(\begin{array}{c|c} 0 & 1_n^T \\ \hline 1_n & \Omega + \frac{1}{\gamma} I_n \end{array}\right) \left(\begin{array}{c} b \\ \hline \alpha \end{array}\right) = \left(\begin{array}{c} 0 \\ \hline Y \end{array}\right), \tag{3}$$

with $Y = (y_1, \ldots, y_n)^T$, $1_n = (1, \ldots, 1)^T$, $\alpha = (\alpha_1, \ldots, \alpha_n)^T$ and $\Omega_{kl} = \varphi(u_k)^T \varphi(u_l) = K(u_k, u_l)$, with $K(u_k, u_l)$ a positive definite kernel $(k, l = 1, \ldots, n)$. According to Mercer's theorem (Mercer, 1909), the resulting LS-SVM model for new inputs becomes:

$$\hat{m}(u^\star) = \sum_{k=1}^{n} \hat{\alpha}_k K(u^\star, u_k) + \hat{b}, \tag{4}$$

where K is any positive definite kernel. In this paper we choose $K = \exp(-\|u_k - u_l\|_2^2/\sigma^2)$. The training of the LS-SVM model involves an optimal selection of the tuning parameters e.g. σ and γ, which are tuned via 10-fold cross-validation.

3. DISTILLATION COLUMN SET-UP

The experimental set-up involves a computer-controlled packed distillation column (see Fig. 1 and 2). The column is about 6 m high and has an internal diameter of 6 cm. The column works under atmospheric conditions and contains three sections of about 1.5 m with Sulzer CY packing (Sulzer, Winterthur) responsible for the separation. This packing has a contact surface of 700 m^2/m^3 and each meter packing is equivalent to 3 theoretical trays. The feed stream containing a mixture of methanol and isopropanol is fed into the column between packed sections 2 and 3. The temperature of the feed can be adjusted by an electric heater of maximum 250 W. At the bottom of the column a reboiler is present containing two electric heaters of maximum 3000 W each. In the reboiler, a part of the liquid is vaporised while the rest is extracted as bottom stream. At the column top a total condenser allows the condensation of the entire overhead vapour stream, which is then collected in a reflux drum. A part of the condensed liquid is fed back to the column as reflux, while the remainder leaves the column as the distillate stream.

In this set-up the following four variables can be manipulated: the reboiler duty Qr (W), the feed rate Fv (g/min), the duty of the feed heater Qv (W) and the reflux flow rate Fr (g/min). The distillate flow Fd (g/min) is adjusted to maintain a constant reflux drum level. Measurements are available for the reflux flow rate Fr, the distillate flow rate Fd, the feed flow rate Fv and nine temperatures, i.e., the temperature at the top of the column Tt, the temperatures in the center of every packing section (i.e. $Ts1$, $Ts2$ and $Ts3$, respectively), the temperature $Tv1$ between section 1 and 2, the temperature $Tv2$ between section 2 and 3, the temperature Tb in the reboiler of the column, and the temperatures of the feed before and after heating (i.e. $Tv0$ and $Tv2$, respectively). All temperatures are measured in degrees Celsius. The actuators and sensors are connected to a Compact Fieldpoint (National Instruments, Austin)

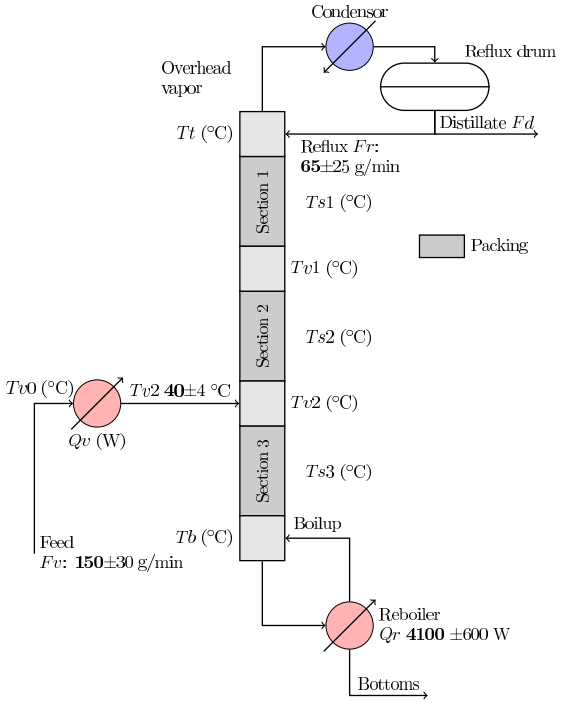

Fig. 1. Diagram of the pilot scale distillation column. Nominal set-points are printed in bold and are followed by the maximum admissible deviations.

with a controller interface cFP-2100 and I/O modules cFP-AIO-610, cFP-AIO-610 and cFP-AI-110. A Labview (National Instruments, Austin) program has been developed to control the actuators and to register the variables.

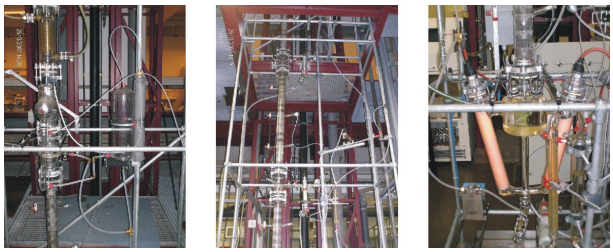

Fig. 2. Pictures of the pilot-scale distillation column: condenser (left), packed section and feed introduction (center), and reboiler (right).

4. MODEL IDENTIFICATION

In order to construct the LS-SVM model, the following steps are performed: (*i*) Experiment, (*ii*) Data preparation, (*iii*) Parameter estimation as described in Section 2 performed with the LS-SVMlab Toolbox (De Brabanter et al., 2010), and (*iv*) Validation. For the linear models, (*iii*) is replaced by model selection and parameter estimation performed with the Matlab System Identification Toolbox (Ljung, 2009).

4.1 Experiments

In order to generate estimation and validation data for system identification, an experiment is performed. The excitation signal is build up with Pseudo Random Binary (PRB) signals for the different manipulated variables.

Before the excitation signals are applied, the column is kept at a constant operating point for two hours to ensure the column is in steady-state. The nominal steady-state values of the different manipulated variables are: a reflux flow rate Fd of 65 g/min, a feed flow rate Fv of 150 g/min, a feed heater duty Qv to maintain a feed temperature $Tv2$ of 40 °C and a reboiler power Qr of 4100 W. These nominal values are known to yield an appropriate operating point for the column. All manipulated variables are controlled by PI controllers except the reboiler power. When the column has reached steady-state, the experiment is started.

When the excitation signals are applied, all manipulated variables stay between two values. The reflux flow rate Fr fluctuates between 40 and 90 g/min, while the feed flow rate Fv changes between 120 and 180 g/min. The feed heater duty Qv is manipulated to obtain feed temperatures Tv between 38 and 42 °C and the reboiler power Qr switches between 3500 and 4700 W. The distillate flow rate Fr is manipulated in order to keep the content of the reflux drum at 40% of its maximum content. All data are recorded with a sampling period of 100 ms.

The PRB input signal is constructed in the following way. The reboiler duty Qr is a repeated periodic signal of 6000 s. The clock period, i.e. the minimum time before the signal is allowed to switch, is 300 s. From former experiments (Logist et al., 2009), it is known that the dynamics of the system are faster at the top of the column. Therefore, the clock period of the other inputs is slightly smaller. For the feed flow rate Fv and feed temperature $Tv2$ a clock period of 120 s is taken by a period of length 3720 s and for the reflux flow rate Fr the clock period is 20 s with a period length of 5100 s. These input signals are combined into one experiment with a time span of 25000 seconds.

4.2 Model in- and outputs

The considered outputs of the system are two temperatures along the column, i.e., the top temperature Tt and the reboiler temperature Tb. The inputs are feed flow rate Fr, feed duty Qv, reboiler duty Qr and reflux flow rate Fr. See Fig. 3 for an overview.

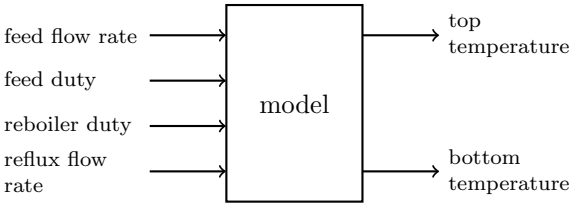

Fig. 3. Overview of in- & outputs of the column model.

4.3 Data preparation

The sampling period of the recorded dataset is reduced to 10 s. Therefore every 10 s a sample is taken from the original recorded data without averaging or filtering. Before identification, an identification and a validation dataset has to be created. The identification dataset consists out of the first 2/3 of the recorded dataset. The remaining 1/3 is employed as validation data. Each of the following identification methods use the same estimation and validation dataset.

4.4 Model selection and parameter estimation

The aim is to construct a *Multiple Input - Single Output* (MISO) black-box model for each of the outputs of the distillation column. The following model structures will be explored:

Least Squares Support Vector Machines (LS-SVM) The general model structure is a NARX of the form:

$$y_t = f(y_{t-1}, \ldots, y_{t-R}; u_{t-1}, \ldots, u_{t-R}) + e_t, \qquad (5)$$

where R denotes the order of the NARX model (number of lags). The number of lags is determined via 10-fold cross-validation (CV). The algoritm for the LS-SVM estimation is summarized follows:

(1) Select the model order R.
(2) Select the regularisation parameter γ, the kernel function K and its bandwidth σ.
(3) Compute the kernel matrix Ω.
(4) Solve the dual linear system (3)

Transfer Functions As known from first principles, distillation column subsystems consist of low-order subsystems. To account for this factor, linear, low-order, continuous-time transfer functions will be fitted into the data. A first-order model (Eq. 6) with time delay is estimated (Ljung, 2009)):

$$G(s) = \frac{K_p}{1 + T_{p1}s} e^{T_d s} \qquad (6)$$

Subspace state-space identification A MISO state-space formulation for each of the outputs is identified:

$$\begin{aligned} x(kT+T) &= Ax(kT) + Bu(kT) + Ke(kT) \qquad (7) \\ y(kT) &= Cx(kT) + Du(kT) + e(kT) \end{aligned}$$

with parameter matrices A, B, C, D, and K. The measurements are sampled at time instances $t = kT$, with $k = 1, 2, \ldots$. The parameters in the general formulation (Eq. 7) are identified using the subspace identification method (Van Overschee and De Moor, 1996).

Box-Jenkins model structure Different polynomial models are examined, e.g., ARX, ARMAX, Output Error (OE) and Box-Jenkis (BJ), which can all be represented for n_u control variables and n_y output variables by the following general formulation (Ljung, 2009):

$$A(q)y(t) = \sum_{i=1}^{n_u} \frac{B_i(q)}{F_i(q)} u_i(t - n_{k_i}) + \frac{C(q)}{D(q)} e(t). \qquad (8)$$

Here, $A(q)$, $B_i(q)$, $C(q)$, $D(q)$ and $F_i(q)$ are matrix polynomial expressions in the shift operator q^{-1} which shifts samples back in time. The order of the polynomial expressions is indicated by respectively n_a, n_b, n_c, n_d and n_f. n_k introduces an additional shift back in time in order to incorporate system delays.

4.5 Validation

Validation is performed on the basis of three quality measures: the Akaike Information Criterion (AIC) (Akaike, 1974), a Fit measure and the Mean Squared Error (MSE).

Akaike Information Criterion (AIC) for linear models is defined as:

$$AIC = \log(V) + \frac{2d}{N} \tag{9}$$

where V is the loss function, d the number of estimated parameters, and N the number of values in the estimation data set. The loss function V is equal to the residual sum of squares: $V = \frac{1}{N}\sum_{i=1}^{N} \hat{\epsilon_i}^2$. For LS-SVM models, the number of estimated parameters is replaced by the degrees of freedom or effective number of parameters given by the trace of the smoother matrix. The smoother matrix L is defined as (De Brabanter et al., 2011):

$$L = \Omega(Z^{-1} - Z^{-1}\frac{J_n}{c}Z^{-1}) + \frac{J_n}{c}Z^{-1}. \tag{10}$$

with $Z = \Omega + \frac{I_n}{\gamma}$, $c = 1_n^T(\Omega + \frac{I_n}{\gamma})^{-1}1_n$ and J_n is a square matrix with all elements equal to 1. hence, for LS-SVM $d = \text{trace}(L)$.

Fit measure is defined as:

$$\text{fit} = 100\% \left(1 - \frac{\|\hat{y}(t) - y(t)\|_2}{\|y(t) - \bar{y}(t)\|_2}\right) \tag{11}$$

where $\hat{y}(t)$ is the simulated output, $y(t)$ the measured output and $\bar{y}(t)$ the mean of the measured output. A fit value of 100% means that the simulation is the same as the measure output. If the simulated value is equal to the mean value, the result is 0%.

Mean Squared Error (MSE) is defined as:

$$\text{MSE} = \frac{\|\hat{y}(t) - y(t)\|_2^2}{N}. \tag{12}$$

5. RESULTS

This section describes the results for the different model types descibed in Section 4.4. Both outputs, i.e. the top temperature and the bottom temperature are treated separately.

5.1 Top Temperature

Least Squares Support Vector Machines (LS-SVM) On the estimation dataset, 10-fold crossvalidation is performed. The MSE of cross-validation (CV-MSE) is plotted in Fig. 4 (top). The MSE value evolves very fast to a value around 8×10^{-5}. Already at lag three this is achieved. Augmenting the lags does not decrease the CV-MSE value anymore. The observation that there is no clear minimum, indicates the possibility of the model class not being entirely correct. Although one cannot define a clear minimum, one can choose a lag value higher than three to represent a model in this model class. Based on the MSE on the validation dataset, depicted in Fig. 4 (bottom), a value $R = 35$ can be chosen. The MSE value is 0.0244. The main trend, seen in this figure, is a decrease of the MSE. For lags lower than 30 a large variation is seen. Between 30 and 40 lags, a clear valley is observed and the large variation between successive lags disappears. From lags 40 on, successive small peaks and valleys are seen. A low MSE, combined with a low lag number is situated between lags 30 and 40. Hence, the model with $R = 35$ is selected. The error bar drawn on $R = 35$, 59 and 99 suggest that the error on the MSE decreases for higher lags. On the other hand, the error on the CV-MSE value does not decrease anymore. This justifies the choice of the model with $R = 35$.

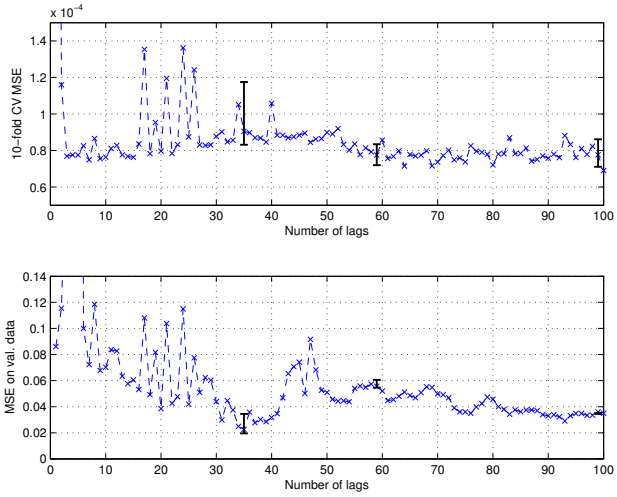

Fig. 4. Validation of the LS-SVM model for Tt.

Table 1. Model quality parameters for *transfer function*, *subspace state-space* and *LS-SVM* models on Tt validation data.

	Fit	AIC	MSE
P1D_t	68.71	-3.64	0.0499
$\text{N4SID}_{t,(16\text{ states})}$	55.18	-3.20	0.1024
$\text{LS-SVM}_{t,(R=35)}$	78.1	-3.28	0.0244

Transfer Functions and Subspace state-space identification The first-order transfer function, abbreviated as P1D_t is estimated and the model quality parameters are calculated on the validation dataset. In Table 1 the AIC, the MSE and Fit value are indicated. Subspace state-space identification picks a 16th order model as the *best* model. Six states are needed to incorporate the delays of the inputs. The delays for the state-space model are those estimated in the first-order transfer function rounded to the closest decade. It is clear that, according to all selection criteria, the transfer function model should be preferred to the subspace state-space model. In Fig. 5 (top), a validation plot is displayed for both model types together with the simulation on validation data for the LS-SVM.

Box-Jenkins model structure Table 2 summarizes the model quality parameters for the ARX, ARMAX, OE and BJ model structures. Two selection criteria are highlighted: the AIC and the minimum MSE. This last criterion is presented to indicate the possibilities in a particular model class. Based on the MSE, all model classes are comparable. AIC makes clear that only an OE model can be selected as this model structure can obtain a high Fit value toghether with a low number of parameters. Hence, following the Box-Jenkins model structure, OE models are to be preferred. According to AIC, ARX and ARMAX models are not suited for this measured data. In Fig. 5 OE an BJ models selected with AIC are displayed.

Conclusions for Tt The discussion above illustrates that, according to AIC, ARX and ARMAX are not well suited model classes for the measured data. Both the

Table 2. Comparison between different Box-Jenkins model structures for the Top Temperature

	Selection Criterion:	Fit	AIC	MSE	n_a	n_b	n_c	n_d	n_f
ARX	minimum MSE	69.9	-0.69	0.0584	2	99	-	-	-
	Lowest AIC	69.8	-0.69	0.0589	2	93	-	-	-
ARMAX	minimum MSE	69.2	-0.74	0.0612	13	53	25	-	-
	Lowest AIC	69.2	-0.74	0.0612	13	53	25	-	-
OE	minimum MSE	72.3	-3.50	0.0496	-	20	-	-	1
	Lowest AIC	70.3	-3.74	0.0569	-	12	-	-	1
BJ	minimum MSE	71.4	-2.47	0.0527	-	5	1	1	4
	Lowest AIC	69.4	-3.43	0.0605	-	4	1	2	4

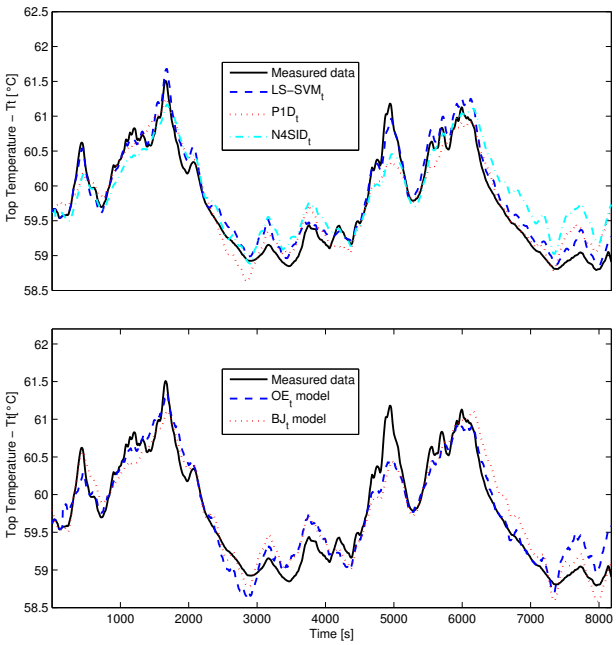

Fig. 5. Validation of the Top temperature Tt.

transfer models and the output error model perform well. The LS-SVM model, in fact a nonlinear ARX model, performs better than the linear variant but is only slightly better than the best performing linear model. In case one needs the best possible estimation of the temperature, the LS-SVM model can be preferred, but if speed and simplicity are important, it is better to choose a linear OE model or a transfer function model.

5.2 Bottom Temperature

Least Square Support Vector Machines (LS-SVM) Fig. 6 (top) displays the 10-fold CV-MSE for different lags of the estimation dataset representing the bottom temperature. A minimum is observed at lag 16. The corresponding MSE on validation data is 0.0130. The 10-fold CV-MSE is continually decreasing until lag 16 and slowly increasing for higher lags. From lag 27 on, the plot of successive CV-MSE values displays a less smooth behaviour. This is also seen in Fig. 6 (bottom). The MSE on validation data is highly fluctuating. Although the CV-MSE and MSE fluctuate rather strongly, a better description of the temperature is reached for lag 27 and 28 and some higher lags. The high corresponding CV-MSE and the higher AIC at e.g. lag 28 is an indication not to choose these models. So, based on the plots in Fig 6, a model with lag 16 is selected. Only when high accuracy is needed one can opt for the model with lag 28 as the MSE on validation data

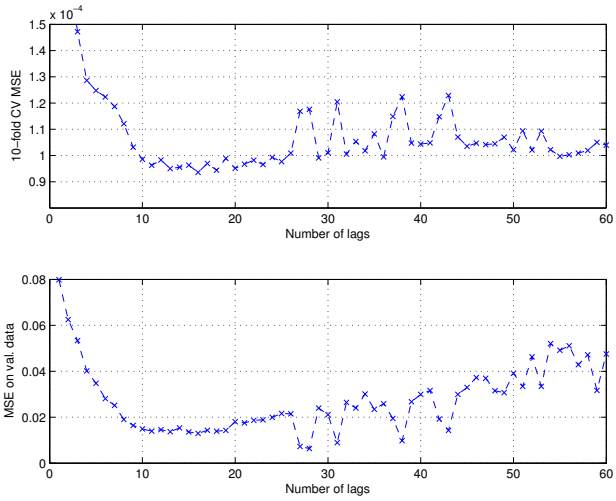

Fig. 6. Validation of the LS-SVM model for Tb.

decreased with a factor two compared to the selected LS-SVM model with lag 16.

Transfer Functions and Subspace state-space identification The model quality parameters for both the transfer function and subspace state-space models are given in Table 4. A first-order model is capable of simulating the temperature very well resulting in a MSE value of only 0.016. Subspace state-space identification fits the measured value less tightly. From the top plot in Fig. 7, it can be seen that the model deviates from the measured value when the temperature is away from the mean temperature (78.1 °C) between 2500 and 6500 seconds. This effect is less noticeable in the first-order transfer function. Compared to the LS-SVM model, transfer functions are only slightly worse, which is caused by the linear behaviour of this measured temperature.

Table 4. Model quality parameters for *transfer function* models, a *subspace state-space* model and *LS-SVM* for Tb on validation data.

	Fit	AIC	MSE
P1D$_b$	78.0	-3.94	0.0164
N4SID$_{b,11\ \text{states}}$	68.4	-1.97	0.0338
LS-SVM$_{b,(R=16)}$	80.4	-3.87	0.0130
LS-SVM$_{b,(R=28)}$	86.3	-3.59	0.0063

Box-Jenkins model structure Table 3 summarizes the model quality parameters for the ARX, ARMAX, OE and BJ model structures. Based on AIC, an ARMAX model has to be selected. The OE model performs better based on the Fit and MSE value, but AIC shows that the number of parameters is too high. For both ARX and BJ, all three

Table 3. Comparison between Box-Jenkins model structures for the Bottom Temperature

	Selection Criterion:	Fit	AIC	MSE	n_a	n_b	n_c	n_d	n_f
ARX	minimum MSE	72.8	-2.27	0.0617	1	33	-	-	-
	Lowest AIC	72.8	-2.27	0.0617	1	33	-	-	-
ARMAX	minimum MSE	81.8	-4.23	0.0274	56	51	16	-	-
	Lowest AIC	79.8	-4.44	0.0342	46	46	6	-	-
OE	minimum MSE	91.3	-0.54	0.0062	-	18	-	-	26
	Lowest AIC	88.1	-3.69	0.0117	-	11	-	-	29
BJ	minimum MSE	78.6	-3.92	0.0379	-	2	5	1	1
	Lowest AIC	78.6	-3.92	0.0379	-	2	5	1	1

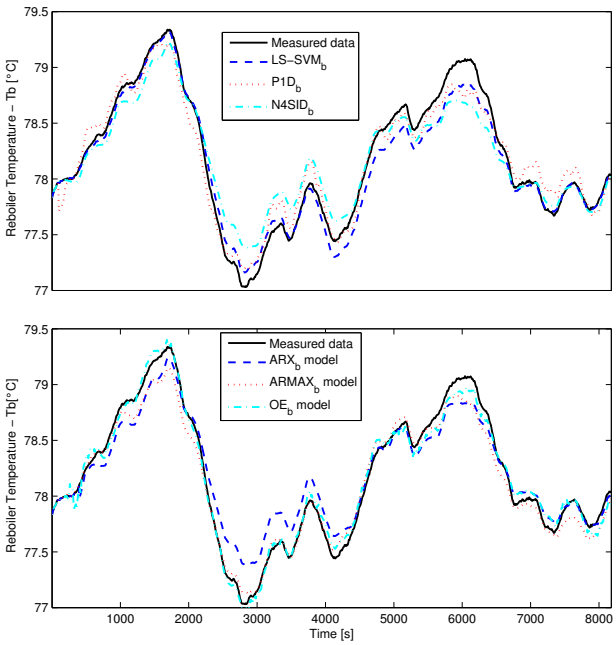

Fig. 7. Validation of the different models for the Bottom temperature Tb.

model selection criteria point to the same model. In Fig. 7 (bottom), the final results for ARX, ARMAX and OE-models are plotted.

Conclusion for Tb The discussion above points to the first-order transfer function model and ARMAX model to be the best linear models describing the measured temperature. The selected LS-SVM model is able to describe this temperature only slightly better. This is caused by the linear nature of the measured signal. The additional computational effort to identify a nonlinear model can only be justified when by taking a higher lag e.g., R = 28 where the MSE value on validation data is devided by two compared to lag R = 16.

6. CONCLUSION

This paper discusses the use of LS-SVM to simulate a measured temperature compared to some well-known linear models types for a binary distillation column. In this real life example, there is a linear model describing the measured temperature very accurately for both the top as well as the bottom temperature. LS-SVM always competes with the best linear model, but is only slightly better. The use of a LS-SVM model for the bottom temperature cannot be justified. For simulation of the top temperature the benefit is sufficient to justify the use of an LS-SVM model.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6), 716–723.

De Brabanter, K., De Brabanter, J., Suykens, J., and De Moor, B. (2011). Approximate confidence and prediction intervals for least squares support vector regression. *IEEE Transactions on Neural Networks*, 22(1), 110–120.

De Brabanter, K., Karsmakers, P., Ojeda, F., Alzate, C., De Brabanter, J., Pelckmans, K., De Moor, B., Vandewalle, J., and Suykens, J. (2010). LS-SVMlab toolbox user's guide version 1.7. Technical report, ESAT-SISTA, K.U.Leuven (Leuven, Belgium).

Ljung, L. (1999). *System Identification: Theory for the User, Second Edition*. Prentice Hall, Upper Saddle River, New Jersey.

Ljung, L. (2009). *System Identification Toolbox Users Guide*. The MathWorks, Inc, Natick.

Logist, F., Huyck, B., Fabré, M., Verwerft, M., Pluymers, B., De Brabanter, J., De Moor, B., and Van Impe, J. (2009). Identification and control of a pilot scale binary distillation column. In *Proc. 10th European Control Conference (ECC '09). Budapest, Hungary.*

Mercer, J. (1909). Functions of positive and negative type and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society*, 209, 415–446.

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

Suykens, J., Van Gestel, T., De Brabanter, J., De Moor, B., and Vandewalle, J. (2002). *Least squares support vector machines*. World Scientific, Singapore.

Van Overschee, P. and De Moor, B. (1996). *Subspace Identification of Linear Systems: Theory, Implementation, Applications*. Kluwer Academic Publishers.

Vapnik, V. (1999). *Statistical Learning Theory*. John Wiley & Sons, Inc.

ACKNOWLEDGEMENTS

BDM is a full professor at the Katholieke Universiteit Leuven, Belgium. Research supported by: Research Council KUL: GOA/11/05 Ambiorics, GOA/10/09 MaNet, CoE EF/05/006 Optimization in Engineering (OPTEC) en PFV/10/002 (OPTEC), IOF-SCORES4CHEM, OT/09/025/TBA, OT/10/035; several PhD/postdoc & fellow grants; Flemish Government: FWO: PhD/postdoc grants, projects: G0226.06 (cooperative systems and optimization), G.0302.07 (SVM/Kernel), G.0320.08 (convex MPC), G.0558.08 (Robust MHE); G.0377.09 (Mechatronics MPC) IWT: PhD Grants, Eureka-Flite+, SBO LeCoPro, SBO Climaqs, SBO POM, O&O-Dsquare Belgian Federal Science Policy Office: IUAP P6/04 (DYSCO, Dynamical systems, control and optimization, 2007-2011) ; EU: ERNSI; FP7-HD-MPC (INFSO-ICT-223854), COST intelliCIS, FP7-EMBOCON (ICT-248940), FP7-SADCO (MC ITN-264735), ERC HIGHWIND (259 166). JVI holds the chair Safety Engineering sponsored by the Belgian chemistry and life sciences federation essenscia. The scientific responsibility is assumed by its authors.