Optimisation of a Batch Distillation Process by Applying Surrogate Models

Laszlo Hegely, Marton Tamas Szucs, Ömer Faruk Karaman, Peter Lang

Department of Building Sevices and Process Engineering, Faculty of Mechanical Engineering, Budapest University of Technology and Economics, Műegyetem rkp. 3., H-1111 Budapest, Hungary
hegely.laszlo@gpk.bme.hu; szucs.marton1997@gmail.com; karaman@edu.bme.hu; lang.peter@gpk.bme.hu

Abstract. A surrogate model-based method is proposed for the optimisation of batch distillation processes and applied to the recovery of methanol from a five-component azeotropic waste solvent mixture, where pollutants are removed in two fore-cuts and an after-cut. The objective function is the profit of a single batch, while constraints are given on the purity of the main cut and composition of the second fore-cut. Simulations are performed by a flow-sheet simulator in a set of points (generated by Latin hypercube sampling) in the space of optimisation variables (reflux ratios of steps, stopping criteria of the fore-cuts). Algebraic surrogate models are fitted by ALAMO to the simulation results to describe the objective function and the constraints. The resulting optimisation problem is solved numerically. The profit obtained is 5 % higher than the one previously obtained by a genetic algorithm, while the number of simulations is reduced to its third.

Keywords Batch Distillation, Simulation, Optimisation, Surrogate Model, Waste Solvent

Introduction

The treatment of waste solvent mixtures is frequently performed by batch distillation (BD). These mixtures often form azeotropes. Components or azeotropes more volatile than the main component can be removed in fore-cut(s), after which the main component is obtained in high purity as main cut. An after-cut can also be taken to remove either pollutants or the main component from the still residue.

For optimizating BD processes, Mujtaba (2004) distinguished three optimisation problems: maximum distillate, minimum time and maximum profit. Minimising the time also decreases the energy demand. As BD is a dynamic process, a dynamic optimisation problem must be solved. By the commonly used feasible path approach, the objective function (OF) is evaluated by solving the model of the process repeatedly at different points in the space of optimisation variables. However, simulation of the process is time-consuming. If a flow-sheet simulator is applied, the optimisation is most frequently performed by an external tool using an evolutionary (usually a genetic) algorithm. These methods require a large number of evaluations of the OF, making the optimisation computationally very intensive.

To make optimisation faster, a new surrogate model-based optimisation (SMBO) method is proposed here. Surrogate models (SMs) or metamodels are reduced models constructed from the inputs and outputs of rigorous models, whose evaluation is considerably less computationally intensive yet mimic the behaviour of the rigorous models. If SMs of OF (and eventually of the constraints) are available, an estimation of the real optimum can be rapidly obtained by finding the optimum of the surrogate OF. Several surrogate modelling techniques were applied recently for the optimisation of continuous distillation columns, such as kriging (Quirante et al., 2015), support vector machines (Jia et al., 2017) or artificial neural networks (ANN; Ibrahim et al., 2017). However, SMBO of BD was only performed in a few number works. For such a dynamic optimisation problem, two different approaches can be distinguished. In the first one, SMs are used to describe the evolution of certain variables (e.g. concentration of the desired component in the distillate) in time. The optimisation, in this case, is still a dynamic optimisation problem, but the dynamic SMs are used to evaluate OF. Greaves et al. (2003) developed a dynamic SM to replace the rigorous one for the optimisation of a middle-vessel column. The behaviour of the real plant was reproduced with good accuracy. The amount of products was maximised using SQP with low computational effort. Khazraee et al. (2011) applied an adaptive neuro-fuzzy inference system to describe the evolution of the amount and composition of the distillate of a batch reactive distillation process. Optimisation was performed by differential evolution (DE); however, a questionable, dimensionally heterogeneous OF was used. The optimisation variables were reflux ratio and total batch time. In the second approach, SMs are fitted to the results of a large number of dynamic simulations, and the optimum of the surrogate OF is determined without a need for dynamic optimisation. Safe et al. (2013) studied a reactive distillation process in a batch dividing-wall column. A polynomial response surface was fitted to OF as a function of the only two

optimisation variables, the vapour and liquid split ratios. The optimum of the surface was then determined by DE. (It is not clear whether the presence of the dividing wall is advantageous.)
The goal of this work is to propose a SM-based method for the fast optimisation of BD processes. In the space of the optimisation variables, a set of points is selected with Latin hypercube sampling (LHS). At each point, OF is evaluated by dynamic simulation with a professional flow-sheet simulator, and algebraic SMs are fitted to the results by using ALAMO (Automatic Learning of Algebraic MOdel) software (Cozad et al., 2014). If necessary, the search space can be narrowed based on the results, and additional sampling can be performed. ALAMO is a machine learning tool that constructs algebraic models from predefined basis functions without the need to specify a function form a priori. The advantages of algebraic models are that they can be readily interpreted by humans and their sensitivity to the input parameters is easy to calculate. The novelty of the work is that non-dynamic SMs are applied to optimise a BD process by considering all operational parameters for the first time. Moreover, the SMs used can provide a better fit than the polynomials applied by Safe et al. (2013).
Recovery of methanol (B) from an azeotropic waste solvent mixture containing acetone (A), tetrahydrofuran (C), water (D) and toluene (E) by batch distillation is optimised. The profit of one batch is maximised. The results are compared with those obtained with a genetic algorithm (Hegely and Lang, 2016).

Process description

The waste solvent mixture to be treated contains 0.07 mass% acetone (A), 37.4 % methanol (B), 4.89 % tetrahydrofuran (C), 56.34 % water (D) and 1.56 % toluene (E). B must be recovered with a purity of 99.5 %. Five minimum-boiling azeotropes are formed, in increasing order of boiling points: A-B, B-C, B-E, C-D and D-E. The azeotropes (except D-E) and A have lower boiling points than B. The recovery of B is hindered by the B-C and B-E azeotropes. (The concentration of A is very low, while the azeotrope C-D does not present a problem since C leaves earlier in a mixture of B and C.) Therefore, C and E must be removed in fore-cuts, causing a considerable loss of B. VLE calculations were performed by using the UNIQUAC model. A more detailed description is given in Hegely and Lang (2016).
The separation is performed in a distillation column with 27 theoretical plates (including the reboiler and the total condenser) (Hegely and Lang, 2016). The top of the column is at atmospheric pressure, while the total pressure drop is 0.25 bar. The volume of the charge is 25 m^3 (at 20 °C). The liquid hold-up of the condenser is 0.45 m^3, that of the column is 0.05 m^3/plate. The reboiler is heated with a heat duty (Q_{st}=1800 MJ/h), provided by saturated steam with a pressure of 3 bar (its heat of condensation is r_{st}=2263.5 MJ/t).
The treatment of one batch consists of the following steps:
Step 0: heating-up of the column with total reflux in order to approach steady-state conditions. The step is finished after 360 min. At this point, the condensate contains mainly B and C with a composition close to the azeotropic one.
Step 1: taking of the first fore-cut with a finite reflux ratio R_1 to remove the bulk C and E with a considerable loss of B. Fore-cut 1 is incinerated. Step 1 is finished when $x_{d,C}<Cr_1$ where $x_{d,C}$ is the instantaneous mass fraction of C in the distillate, and Cr1 is the stopping criterion for Step 1.
Step 2: taking of the second fore-cut with reflux ratio R_2. This cut already contains B in a considerable concentration, but its pollutant (C and E) content is still too high. This cut is recycled to the next batch to limit the loss of B. This step is stopped when $x_{d,C}<Cr_2$.
Step 3: taking of the main cut (B product) with R_3. This step is finished (because of the increasing $x_{d,D}$) when $x_{mc,B}<0.9952$ where $x_{mc,B}$ is the mass fraction of B in the main cut.
Step 4: taking of the after-cut with R_4. The aim of the after-cut is to remove B from the still residue so that it can be sent to biological purification. The after-cut has a considerable B content, and it is recycled to the next batch. Taking of the cut is finished when the B content of the still residue ($x_{sr,B}$) becomes lower than 0.25 %.

Calculation method

The objective function (OF; Eq. 1) is the profit of a single batch. It is composed of the price of methanol in the main cut, the costs of incineration of the Fore-cut 1 and of steam consumption during the process (Hegely and Lang, 2016).

$$OF = p_B m_{mc} - c_{inc} m_{fc1} - c_{st} \frac{Q_{st}}{r_{st}} t \quad (1)$$

where: p_B: price of methanol, 0.46 US$/kg, m_{mc}: mass of the main cut, kg, c_{inc}: cost of incineration, 0.21 $/kg, m_{fc1}: mass of Fore-cut 1, kg, c_{st}: cost of steam, 57.6 $/t, t: duration of the process, h.
The optimization problem is subject to the inequality constraints: Constraint 1: $x_{mc,B}\geq0.9952$, Constraint 2: $x_{mc,C}/x_{mc,B}\leq0.107$, Constraint 3: $x_{mc,E}/x_{mc,B}\leq0.12$, where $x_{fc2,B}$, $x_{fc2,C}$ and $x_{fc2,E}$ are the concentration of B, C and E in Fore-cut 2, respectively.

Constraint 1 guarantees the required purity of the product. Constraints 2 and 3 are needed to ensure that the organic pollutants C and E are not accumulated in Fore-cut 2 so that it can be recycled to the next batch. The optimisation variables are: R_1, R_2, R_3, and Cr_1, Cr_2. Since previous calculations showed that the effect of R_4 on OF is negligible, its value is kept constant at 5.41.

In this work, a SMBO method is proposed (Fig. 1). First, a large number of test points are generated in the space of the optimization variables by Latin hypercube sampling (LHS). In each point, simulation is performed by using ChemCad, whose results are the values of dependent variables necessary to calculate OF and the left-hand side of the constraints. If the number of feasible points (that is, where the constraints are not violated) is deemed sufficient for model fitting, surrogate models are generated by ALAMO; otherwise, the generation of test points is repeated by using a narrower range of the optimisation variables. Optimisation of the surrogate OF is then performed in Maple. Finally, a simulation is performed with the values of the optimisation variables obtained to evaluate the difference of the OF values calculated by the SMs and by rigorous simulation. The smallest the difference, the better the fit of the SMs and the more likely that a good approximation of the true optimum is obtained.

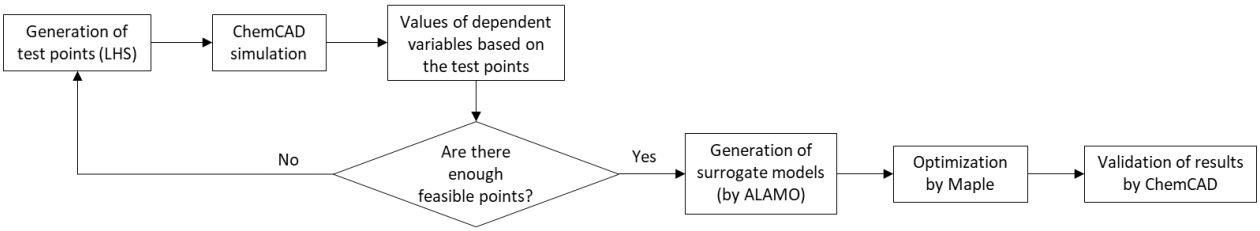

Fig. 1 Flow chart of the surrogate model-based optimisation method.

In order to obtain accurate SMs for the whole range of the optimisation variables, a more uniform sampling pattern than simple random sampling is preferred. In the space of the optimisation variables given in Table 1 as the original range, 500 points are generated by LHS. Inside the intervals selected by LHS, the values of the optimisation variables are randomly generated. These ranges are selected based on previous calculation experience.

Table 1.
Ranges of the values of the optimisation variables used for Latin hypercube sampling.

	R_1	R_2	R_3	Cr_1	Cr_2
Original range	1-10	1-10	1-5	0.10-0.30	0.015-0.050
Narrowed range	2-7	2-7	2-4.5	0.13-0.25	0.020-0.035

Simulation is then performed at each point. The batch distillation process is modelled in ChemCad Version 7.1 in dynamic mode. To automate the calculation, ChemCad is coupled to Excel. A VBA macro is used to control the simulation by detecting the fulfilment of termination criteria and updating the value of R and the setting of a divider at the end of the steps. The divider is used to switch between the accumulators. At each time step, the current values of R and divider setting are transferred to ChemCad, which then gives back selected results. To reduce the time requirement of the simulation, several measures are implemented. Since Step 0 has no optimisation variables, it is only simulated once; subsequent calculation starts from the end of Step 0. If, at the end of Step 2, Constraints 2 or 3 are violated, the simulation is terminated. In Step 3, the B content of the main cut has a maximum value ($x_{mc,B,max}$) in time. If $x_{mc,B,max}$ does not reach 0.9952, the simulation is also terminated since the purity of the product will not be acceptable. Performing the calculations on the original range results in a low number of points with acceptable purity. Model fitting to too few points might lead to low accuracy of the models. To avoid this, a second set of 500 points are generated by LHS on a narrower domain (Table 1) based on those points with acceptable purity. All the models are fitted by using the narrower domain.

Surrogate models are fitted by ALAMO to the results of the simulation necessary to calculate OF and the left-hand sides of Constraints 2 and 3: m_{fc1}, m_{mc}, t, $x_{fc2,B}$, $x_{fc2,C}$, $x_{fc2,E}$ and additionally to $x_{mc,B,max}$. The latter value shows not only if Constraint 1 is violated (the product purity is unacceptable) but also the level of the violation. (The value of $x_{mc,B}$ is not suitable for this purpose.) Alternatively, it would also be possible to fit a SM directly to OF; however, this would reduce the level of insight that can be obtained by analysing the models and would likely lead to a less accurate model.

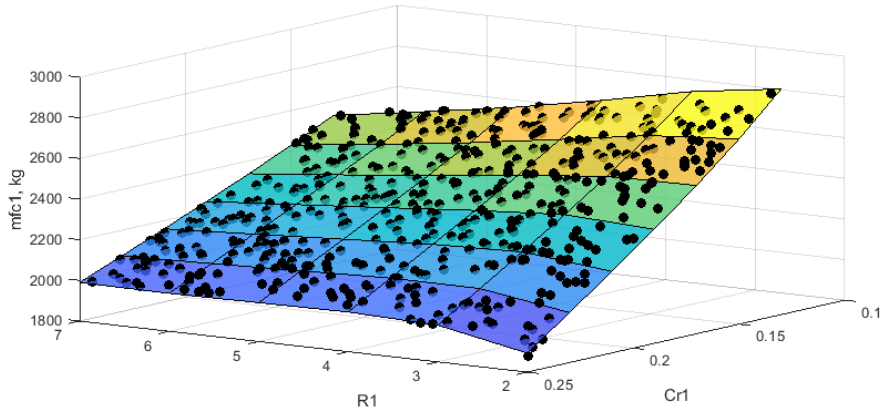

Fig. 2 Mass of Fore-cut 1 calculated by simulation (dots) and by the surrogate model fitted (surface).

The independent (optimisation) variables influencing each dependent one are the following ones. m_{fc1} is a function of only R_1 and Cr_1. The composition of Fore-cut 2 ($x_{fc2,B}$, $x_{fc2,C}$ and $x_{fc2,E}$) depends on R_1, R_2, Cr_1 and Cr_2. m_{mc}, $x_{mc,B,max}$ and t can be influenced by all the independent variables (R_1, R_2, R_3, and Cr_1, Cr_2).
ALAMO fits algebraic models by optimising a selected criterion describing the goodness of the fit. The models are generated as combinations of previously chosen basis functions, not necessarily used in the final model. Here, Bayesian information criterion is selected as the measure of the goodness of the fit, which not only takes the model error into account but also penalises the model size to avoid overfitting. The basis functions allowed are constant terms, linear, logarithmic, and exponential functions, as well as polynomials of the variables and their binary and ternary products.
Optimisation is performed with SQP by using the NLPSolve function of Maple. The function takes as arguments the OF and (optionally) the optimisation constraints. Here, the bounds of the ranges of the independent variables were also given as constraints to avoid extrapolation. With the values of the independent variables obtained from the optimisation, a simulation is performed to verify the accuracy of SMs at the estimated optimum. Additionally, the gradient vector is calculated at the optimum determined by SMs. Simulations are performed following the direction of the gradient vector in order to verify whether it is possible to further increase OF by approaching the constraints more.

Results

From the original range, all the data points can be used for model fitting for m_{fc1}, $x_{fc2,B}$, $x_{fc2,C}$ and $x_{fc2,E}$. However, only 51 calculations do not violate Constraints 2 or 3 and thus can be used for $x_{mc,B,max}$. More importantly, there are only 18 feasible points where the product purity is acceptable, and that can be used for model fitting for m_{mc} and t. The highest OF value is 429.6 $ with R_1=5.77, R_2=3.23, R_3=2.62, Cr_1=0.1483 and Cr_2=0.0286.
By using the narrowed range, the number of feasible points increased from 18 to 46, which is deemed to be sufficient. The number of points that can be used for $x_{mc,B,max}$ increased slightly to 56. At the best point, OF equals 480.7 $ with R_1=5.80, R_2=2.13, R_3=3.22, Cr_1=0.2168 and Cr_2=0.0241.
The SMs are fitted on the narrowed ranges. The size of the models varies between 7 ($x_{fc2,B}$) and 25 ($x_{fc2,B}$). Interestingly, Cr_2 does not influence the distillation time:

$$t = 95.93 \cdot R_1 + 16.04 \cdot R_2 + 78.61 \cdot ln\, R_2 - 0.89 \cdot e^{R_3} + 939.52 \cdot e^{Cr_1} - 3{,}59 \cdot R_1^2 + 24.95 \cdot R_3^2 - 4531 \cdot Cr_1^2 \quad (2)$$

By increasing Cr_2, the duration of Step 2 decreases (less Fore-cut 2 is taken), but that of Step 3 is likely to increase to a very similar extent. As it is shown in Fig. 2, the surrogate model (surface) was able to predict the mass of Fore-cut 1 (dots) with good accuracy. On the increase of R_1, m_{fc1} decreases at low Cr_1 values but increases slightly at higher Cr_1 values. On the increase of Cr_1, m_{fc1}decreases since Step 1 is stopped earlier.
The results of SMBO are given in Table 2. In the optimum, both Constraints 2 and 3 are active. The results of the simulation and the SMs are very close to each other: the difference in OF is 2.6 $ (0.53 %). By the simulation, the constraints are fulfilled, but the concentration ratios are also very close to the constraints, with the higher deviation being only 0.57% for $x_{fc2,E}/x_{fc2,B}$.
Comparing the simulation results to the optimum found by Hegely and Lang (2016) by GA, SMBO gives a 5.0 % higher profit. Moreover, GA required 3000 simulations instead of 1000 by the present method. The suboptimality of the GA result is hinted at by the significant distance of $x_{fc2,C}/x_{fc2,B}$ from the limit of Constraint 2. R_1 decreased by 8.7

%, whereas R_3 and Cr_2 changed only slightly. R_2 decreased by 19 %. Cr_1 increased by 22 %, thereby reducing the cost of incineration. The mass of the main cut, and thus the income decreased slightly (by 2.5 %); however, the steam cost also decreased (by 3.1 %) due to the lower reflux ratios. It must also be noted that, in the narrowed ranges, a higher OF value than that of GA is already obtained by LHS only.

Table 2.
Comparison of the results of the surrogate model-based optimisation with those of GA (Hegely and Lang, 2016).

Optimisation variable	GA	SMBO		Difference, %
R_1	6.22	5.68		-8.68
R_2	3.07	2.49		-18.8
R_3	3.05	3.09		+1.31
Cr_1	0.175	0.2138		+22.2
Cr_2	0.0262	0.0255		-2.67
Constraints		Model	Simulation	
$x_{fc2,C}/x_{fc2,B}$	0.0951	0.1070	0.1064	+11.9
$x_{fc2,E}/x_{fc2,B}$	0.1191	0.1200	0.1198	+0.588
Profit (OF) and its elements		Model	Simulation	
Income, $	2597	2534	2533	-2.46
Incineration cost, $	492	456	454	-7.72
Steam cost, $	1638	1585	1588	-3.05
Profit (OF), $	467	493.0	490.4	+5.01

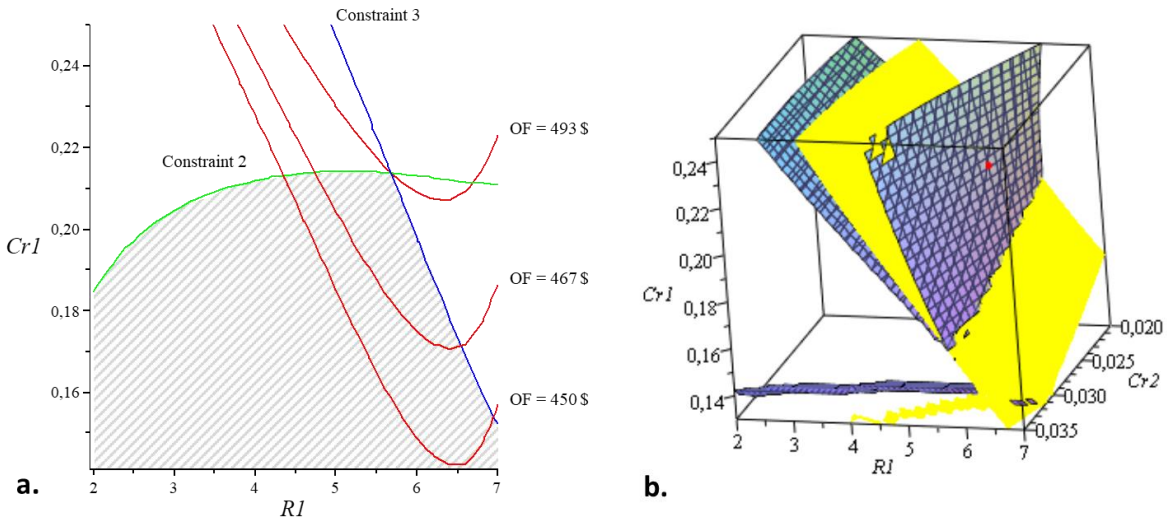

Fig. 3 a. Contour plot of OF (red lines) with R_1 and Cr_1 as independent variables. Constraint 2 is shown as a green line, Constraint 3 as a blue one, b. plot of the active constraints (multicolour surfaces), the limit of surrogate model validity (yellow surface) and the optimum obtained from the surrogate models (red dot).

Since SMs are explicitly known, a more detailed analysis of the optimization problem is possible. To study the interaction between the variables related to Fore-cut 1, a contour plot of OF with R_1 and Cr_1 as independent variables is drawn (Fig. 3a; all other variables take their optimal value). Either Constraint 2 (green line) or Constraint 3 (blue line) are violated outside the shaded area. As in the optimum both constraints are active, the corresponding contour line and the constraint lines intersect in one point. By increasing R_1, Cr_1 must be decreased (more Fore-cut 1 must be taken) to keep OF constant (except at high R_1 values for low OF). The maximum possible Cr_1 value is determined by Constraint 2 below the optimal R_1 and by Constraint 3 above.
Although OF is a five-variable function, the visualisation of the optimum is still possible. An active Constraint 3 corresponds to the three-dimensional space shown in Fig. 3b. In each point of this space, R_2 is determined by the values of R_1, Cr_1 and Cr_2. The set of points where Constraint 2 is also active are the multicolour surfaces on which the optimum (red dot) is located. At lower R_1 values, Constraint 2 is violated. Points below the yellow surface correspond to R_2 values outside the range used for model fitting and thus represent an extrapolation of the models that should be avoided. The gradient of the objective function calculated from the SMs at the optimum is (17.2, 10.8, $-2.3 \cdot 10^{-6}$, 913, 4017), meaning that OF is most sensitive to Cr_2 and least sensitive to R_3. Since by the simulation, the constraints are not active, it is likely that OF can be further increased by approaching the constraints better. To follow the direction of the gradient, Cr_2 must be increased, which decreases the duration of Step 2. Even if the duration decreases by one time step (2 min), the constraints are violated, thus it was not possible to further increase OF.

Conclusions

A surrogate model-based method was proposed to reduce the computational intensity of the optimisation of batch distillation processes. The batch distillation treatment of a five-component azeotropic waste solvent mixture was optimized by Hegely and Lang (2016) using a genetic algorithm (GA). Methanol was obtained as main cut, while pollutants were removed in two fore-cuts and an after-cut. The objective function (OF) was the profit of a single batch, while constraints were given on the purity of the main cut and the composition of the second fore-cut.
By the method proposed, simulations were performed by a flow-sheet simulator in a set of points (generated by Latin hypercube sampling) in the space of optimisation variables (reflux ratios of the steps, stopping criteria of the fore-cuts).Algebraic surrogate models were then fitted to the simulation results to describe OF and the constraints by the ALAMO machine learning technique. The resulting optimisation problem was solved very easily by SQP.
The surrogate models accurately described the results of the simulation. The profit obtained by the surrogate model-based optimization was by 5 % higher than the one obtained by Hegely and Lang (2016), while the number of simulations was reduced from 3000 to 1000. By refining the iterative sampling step, further reduction in the number of simulations might be achieved.

Acknowledgements
The research reported in this paper and carried out at the BME has been supported by the National Research Development and Innovation Fund based on the charter of bolster issued by the National Research Development and Innovation Office under the auspices of the Ministry for Innovation and Technology, by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the ÚNKP-21-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund. Ömer Faruk Karaman wishes to express his gratitude to the SH program for the support in this research.

References

1. A. Cozad, N. V. Sahinidis, D. C. Miller, *AIChE Journal*. **60**, 2211–2227 (2014)
2. M. A. Greaves, I. M., Mujtaba, M. Barolo, A. Trotta, M. A. Hussain, *Chemical Engineering Research and Design*. **81**, 393-401 (2003)
3. L. Hegely and P. Lang, *Journal of Cleaner Production*, **136**, 99-110 (2016)
4. D. Ibrahim, M. Jobson, J. Li, G. Guillén-Gosálbez, *Computer Aided Chemical Engineering*. **40**, 481-486 (2017)
5. S. Jia, X. Qian, X. Yuan, *Chemical Engineering Research and Design*. **125**, 422-432 (2017)
6. S. M. Khazraee, A. H. Jahanmiri, S. A. Ghorayshi, *Neural Computing and Applications*. **20**, 239-248 (2011)
7. I. M. Mujtaba, *Batch distillation: design and operation*. London, UK: Imperial College Press, 2004.
8. N. Quirante, J. Javaloyes, J. A. Caballero, *AIChE Journal*. **61**, 2169-2187 (2015)
9. M. Safe, S. M. Khazraee, P. Setoodeh, A. H. Jahanmiri, *Mathematical and Computer Modelling of Dynamical Systems*. **19**, 29-50 (2013)