Technical University of Denmark

DTU

Modelling Framework for the Identification of Critical Variables and Parameters under Uncertainty in the Bioethanol Production from Lignocellulose

Morales Rodriguez, Ricardo; Meyer, Anne S.; Gernaey, Krist V.; Sin, Gürkan

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

MODELLING FRAMEWORK FOR THE IDENTIFICATION OF CRITICAL VARIABLES AND PARAMETERS UNDER UNCERTAINTY IN THE BIOETHANOL PRODUCTION FROM LIGNOCELLULOSE

R. Morales-Rodriguez[1*], A.S. Meyer[2], K.V. Gernaey[3], G. Sin[1]

[1]CAPEC, [2]BIOENG, [3]PROCESS, Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark. e-mail: rmr@kt.dtu.dk

Abstract

This study presents the development of a systematic modelling framework for identification of the most critical variables and parameters under uncertainty, evaluated on a lignocellulosic ethanol production case study. The systematic framework starts with: (1) definition of the objectives; (2) Collection of data and the implementation of dynamic models for each unit operation in the process; (3) Uncertainty and sensitivity analysis, performed to identify the critical operational variables and parameters in the process. The uncertainty analysis is carried out using the Monte-Carlo technique. Sensitivity analysis employs the standardized regression coefficient (SRC) method, which provides a global sensitivity measure, β_i, thereby showing how much each parameter contributes to the variance (uncertainty) of the model predictions. Thus, identifying the most critical parameters involved in the process, suitable for further analysis of the bioprocess. The uncertainty and sensitivity analysis identified the following most critical variables and parameters involved in the lignocellulosic ethanol production case study. For the operating cost, the enzyme loading showed the strongest impact, while reaction volume showed a significant impact on the ethanol/biomass ratio. The results showed also that it is possible to find a better alternative operation of the plant in comparison with the base case.

INTRODUCTION

Process optimization is an important area within process systems engineering (PSE), since it is actively used in the development, decision making, and subsequent improvement of chemical processes (e.g. for the design, synthesis and operation), aiming at maximizing the process performance while at the same time minimizing the processing costs.

What makes the process optimization a challenging task in bioprocess development are the uncertainties present in the system, which are a result of technological factors, operational conditions as well as economical factors. These uncertainties then lead to uncertainties in the predictions of key performance indices of bioprocesses such as cellulosic ethanol production yield and unit cost. To address these uncertainties, it is required to perform a formal uncertainty and sensitivity analysis previously to the optimization phase. Hence the objective of this paper is to develop a systematic framework for the identification of the most critical variables and parameter of bioprocesses subject to various sources of uncertainties. The framework is evaluated on a case study focusing on lignocellulosic bioethanol production. The problem statement in this case study is formulated as follows: given a process flowsheet, an operational configuration and a feed profile, how can an engineer predict and identify the uncertainty of the main performance criteria considered in plant design – e.g. bioethanol yield and concentration, water recovery, energy consumption, operational cost, etc. The process consists of four main operation steps: acid pre-treatment, enzymatic hydrolysis, fermentation and downstream processes (Morales-Rodriguez et al., 2011).

MODELLING FRAMEWORK FOR THE IDENTIFICATION OF CRITICAL VARIABLES AND PARAMETERS UNDER UNCERTAINTY

The systematic framework illustrated in Figure 1 starts with the definition of the objective of the study, followed by the collection of data, and the implementation of models (for each unit operation in the process) to finally develop an integrated dynamic model to describe the system. In the third step, the uncertainty and sensitivity analysis is performed to identify the critical process operational variables and parameters in the system. An important element here is the identification of sources of uncertainties in the system.
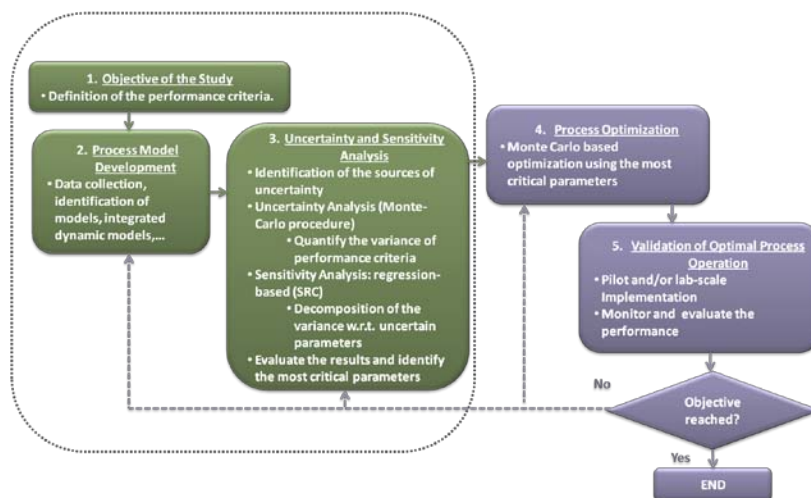

Figure 1 Systematic framework for bioprocess identification under uncertainty.

The uncertainty analysis is then carried out using the Monte-Carlo technique, which involves four steps (see Figure 2): (i) specification of input uncertainty, (ii)sampling of (uncertain) parameters (Latin Hypercube Sampling, LHS), (iii) Monte-Carlo simulations with the sampled parameter values and (iv) representation of uncertainty (e.g. mean, standard deviation, variance (Helton and Davis, 2003). For the sensitivity analysis, the standardized regression coefficient (SRC) method is chosen as it provides a good approximation to global sensitivity measure with an affordable computational demand compared with more computational exhaustive global sensitivity analysis methods such as FAST or *Sobol's* sensitivity indices (Sin et al., 2009). The SRC method involves building linear regression models on the output of the Monte-Carlo simulations (Helton and Davis, 2003). The SRC method provides a global sensitivity measure, β_i, which is a quantitative measure of how much each parameter contributes to the variance (uncertainty) of the model predictions. This sensitivity measure is then used as basis to identify the most critical parameters involved in the process. Further analysis can be performed by using the obtained results for a process optimization study (using for instance Monte-Carlo based technique), which uses the most critical parameters identified above. A subsequence step can be the evaluation of the performance of the optimized process operation via comparison to data obtained in lab or pilot-scale experiments. If the validation results are satisfactory, then the systematic procedure will be terminated. Otherwise the procedure needs to be iterated, either by reviewing the models used for the optimization or by evaluating a different set of critical system parameters. Thus study presents the results that have covered the first three steps.
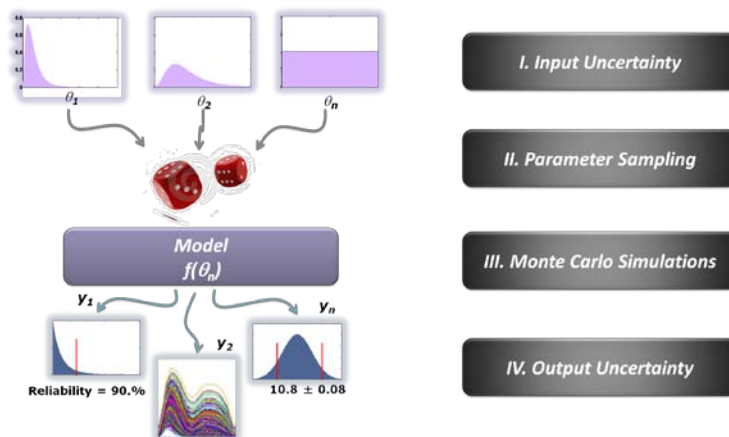
Figure 2 Monte-Carlo technique.

PROCESS CHARACTERISTICS DATA AND SIMULATION PLATFORM

The model implementation, the simulations and the uncertainty and sensitivity analysis have all been performed in Matlab (The Mathworks, Natick, Massachusetts). The basic process characteristics and information regarding conversion rates and dimensions of key unit operations are from Aden et al. (2002), but expanded for dynamic modelling with specific rate equations for enzyme and co-fermentations kinetics as outlined in Morales-Rodriguez et al. (2011).

CASE STUDY: LIGNOCELLULOSIC ETHANOL PRODUCTION.

The Dynamic Lignocellulosic Bioethanol model version 1.0 (DLB1.0) (Morales-Rodriguez et al., 2011) was combined with rigorous dynamic downstream process models. The resulting dynamic plant-wide model was then employed to identify the operational window under uncertainty to assess operational cost of the conversion of lignocellulose to ethanol. A process configuration involving simultaneous saccharification and co-fermentation operating in a continuous process regime with recycle (SSCF-C_RECY) (see Figure 3), performed best with respect to producing maximal ethanol yield from the biomass, and was selected as the process regime for further investigation in the present case study. Indeed, SSCF-C_RECY showed a better ethanol yield compared to the process configuration proposed by NREL (Aden et al., 2002), which relies on fed-batch operation for the enzymatic hydrolysis and fermentation of sugars (FB-FB).
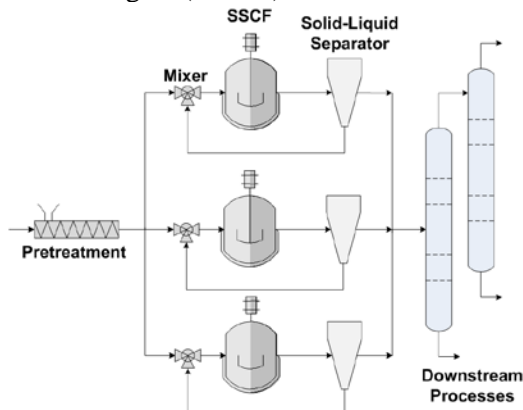

Figure 3 Bioethanol process configuration for Simultaneous Saccharification and Co-Fermentation operating in continuous regime with recycle (SSCF-C_RECY) (Morales-Rodriguez et al., 2011).

Implementation of the systematic framework for optimization under uncertainty
In the following, the case study will be used to illustrate the different steps of the procedure outlined in Figure 1.

Step 1: Objective of the optimization: The objective is to identify the most critical operational boundaries for the lignocellulosic ethanol production case study, with the intention to reduce the operational cost (for the additives such as, enzyme and acid loading) and improve the ethanol yield.

Step 2: Process model development: Collection of data, and implementation of models (for each unit in the process) to describe the integrated system is described in Morales-Rodriguez et al. (2011).

Step 3: Uncertainty and Sensitivity Analysis
3.1. The complete set of kinetic parameters for the SSCF mathematical model and operational variables (such as, acid concentration (C_{acid}) in the pretreatment step, temperature in the pretreatment (T_{PT}) and in the SSCF (T_{SSCF}) reactors, reactor volume of the SSCF unit (V_{SSCF}), enzyme loading of exo-β-1,4-cellobiohydrolase + endo-β-1,4-glucanase (EL_1) and β-glucosidase (EL_2), yeast loading (C_{yeast}), among others) were selected to identify the most critical parameters and variables under uncertainty in the lignocellulosic bioethanol production.

3.2. The sampling of uncertain variables and parameters was accomplished using LHS, which generated a sampling matrix with a dimension of 542x55, where rows corresponded to LHS samples and columns corresponded to variables and parameters. Subsequently, Monte-Carlo simulations were performed with the sampled parameter values, and the quantitative uncertainty results in the form of key performance indices (KPI) were plotted as histograms (see Figure 4). The uncertainty can be inferred from the variance of these histograms, which was particularly large for the operating cost (Figure 4a), meaning that the operating cost is likely to be the technological performance evaluation criterion which has a large deviation at the moment. The uncertainty on the ethanol/dry-biomass ratio is rather narrow with 0.13±0.1 kg/kg (Figure 4b)). Moreover, the simulation results show that among the 542 dynamic simulations (Monte-Carlo samples), 17% have resulted in higher ethanol yield compared with the base case. As far as the cost is concerned, 45% of these samples with a higher ethanol yield have a lower operating cost than the base case.
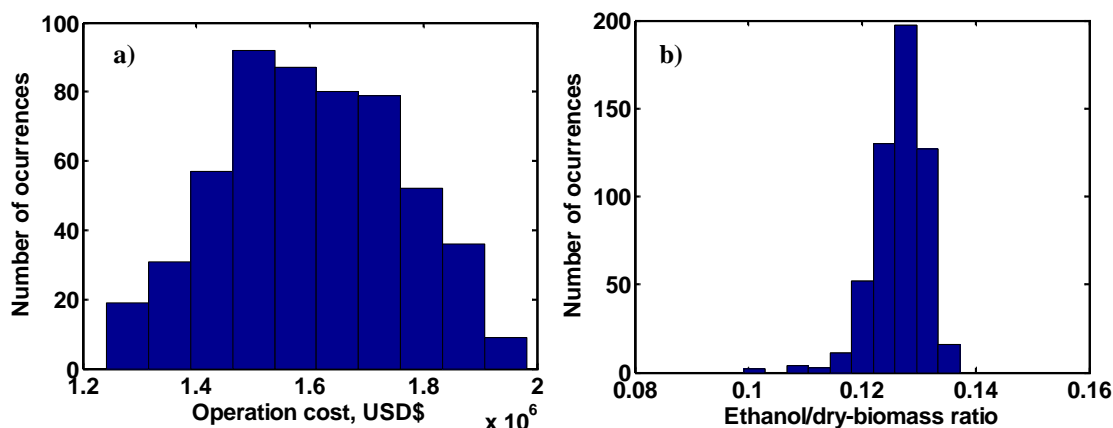

Figure 4 Averaged plant performance criteria obtained from Monte-Carlo simulations plotted as histogram: a) operating cost, b) ethanol dry-biomass ratio.

3.3. For the sensitivity analysis, the standardized regression coefficient (SRC) method was used. Figure 5.a illustrates the linear model fit obtained when regressing the Monte-Carlo simulation outputs with respect to uncertainty parameters. Notice that the linear model determination coefficient (R^2) is equal to 1 which is quite high, meaning that the time-averaged model outputs could be linearized to a high degree, hence satisfying the requirement for β_i to be used as a reliable index of the sensitivity measure ($R^2 > 0.7$). As far as ethanol/dry-biomass ratio is concerned, R^2 was found to be lower than 0.7 ($R^2 = 0.271$), indicating that β_i cannot be used as a reliable index of sensitivity. Nevertheless, the correlation between reacting volume and ethanol/dry-biomass ratio was analyzed using scatter plots (see Figure 6.a), since it is qualitatively identifying the critical variables and parameters present in the process model. The ethanol/dry-biomass ratio increases when the reaction volume increases indicating a positive impact of reaction volume (within the studied range) on the ethanol yield. Another important point to highlight for some samples is that, even though the reacting volume in the SSCF section was reduced, the ethanol/dry-biomass ratio was not significantly compromised which requires further investigation in view of process optimization. Figure 6.b shows the correlation between enzyme loading type 1 with the ethanol/dry-biomass ratio, where a systematic trend was not found, thus implying that the impact of enzyme loading on ethanol yield was not significant within the studied range. For both operating variables (reaction volume and enzyme loading), a few outliers were observed in the scatter plots. This is probably due to a specific combination of the sampled input uncertainties which requires closer investigation.
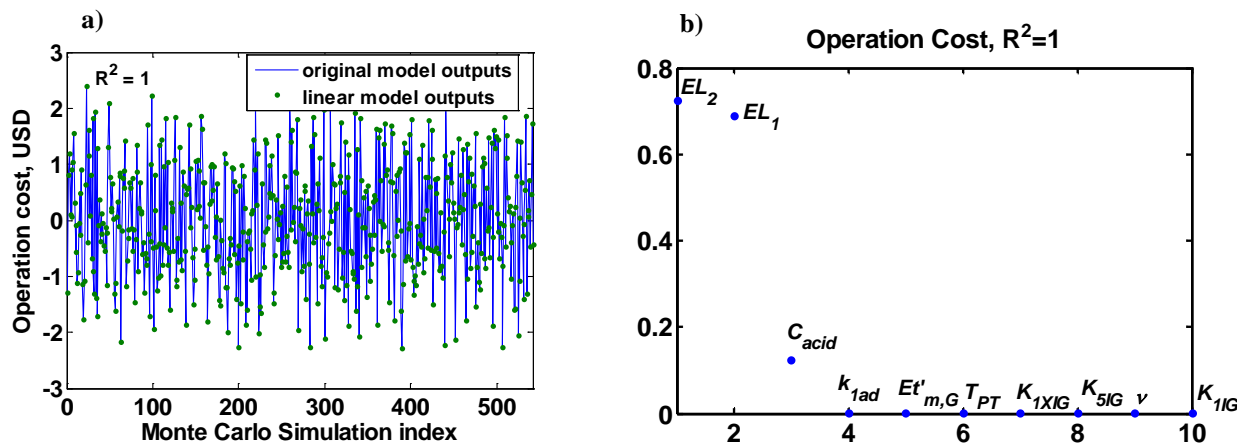

Figure 5 a) Linear model fit obtained from Monte-Carlo simulation for operating cost. b) Influence of variables and parameters on the operating cost.

3.4. In this sub-step, the sensitivity measure was then used as basis to identify the most critical operational variables and parameters involved in the process (Figure 5.b), which shows a significant impact on the operating cost. For the operating cost, the enzyme and sulfuric acid loadings were found the most significant sources of uncertainty. That is an obvious outcome since it is well known that enzymes for cellulose degradation represent one of the highest costs in the bioethanol production from lignocellulosic biomass (Alvarado-Morales et al, 2009).

Of course, it is important to highlight that formal optimization should be performed using the identified most critical variables and parameters in this work (step number four). To this end, a sophisticated Monte-Carlo based global optimization (GO) or Dynamic Optimization (DO) solvers can be used. This work is in progress.
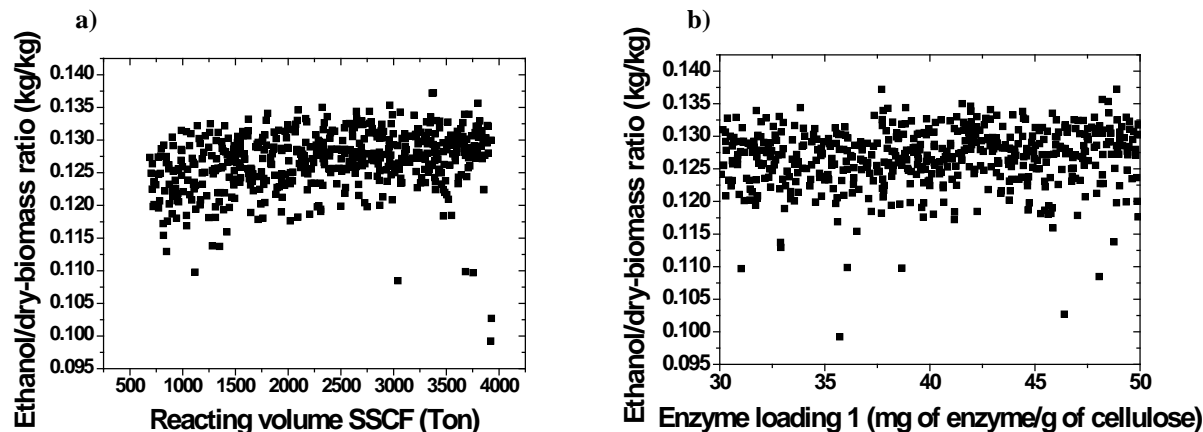
Figure 6 Correlation among parameters and performance criteria. a) reacting volume vs. ethanol/dry-biomass ratio, b) enzyme loading 1 vs. ethanol/dry-biomass ratio

CONCLUDING REMARKS

This study introduced a systematic framework for identification of the most critical parameters and operation variables under uncertainty, involved in the bioethanol production from lignocellulosic biomass. The uncertainty and sensitivity analysis identified the following most critical variables involved in the DLB1.0 model. For the operating cost, the enzyme loading showed the strongest impact, while reaction volume showed a significant impact on the ethanol/biomass ratio. The results showed also that it is possible to find a better alternative operation of the plant in comparison with the base case.

ACKNOWLEDGEMENTS

REFERENCES

1. Aden, A., Ruth, M., Ibsen, K., et al. (2002). Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover. *National Renewable Energy Laboratory Technical Report*. **NREL/TP-510-32438**.
2. Alvarado-Morales, M., Terra, J., Gernaey, K.V., Woodley, J.M., Gani, R., 2009. Biorefining: Computer aided tools for sustainable design and analysis of bioethanol production. *Chem. Eng. Res. Des.* **(87)** 1171–1183.
3. Helton, J.C., Davis, F.J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab. Eng. Syst. Saf.* **(81)** 23–69.
4. McKay, M.D., Beckman, R.J., Conover, W.J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **(21)** 239–245.
5. Morales-Rodriguez, R., Meyer, A.S., Gernaey, K.V., Sin, G. (2011).Dynamic Model-based Evaluation of Process Configurations for Integrated Operation of Hydrolysis and Co-Fermentation for Bioethanol Production from Lignicellulose. *Bioresour. Technol.* **(102)** 1174-1184.
6. Sin, G., Gernaey, K.V., Lantz, A.E. (2009). Good Modeling Practice for PAT Applications: Propagation of Input Uncertainty and Sensitivity Analysis. *Biotechnol. Prog.* **(25)** 1043-1053.