

Combined Estimation and Optimal Control of Batch Membrane Processes

Authors:

Martin Jelemenský, Daniela Pakšiová, Radoslav Paulen, Abderrazak Latifi, Miroslav Fikar

Date Submitted: 2018-07-30

Keywords: fouling estimation, time-optimal operation, membrane fouling, batch diafiltration

Abstract:

In this paper, we deal with the model-based time-optimal operation of a batch diafiltration process in the presence of membrane fouling. Membrane fouling poses one of the major problems in the field of membrane processes. We model the fouling behavior and estimate its parameters using various methods. Least-squares, least-squares with a moving horizon, recursive least-squares methods and the extended Kalman filter are applied and discussed for the estimation of the fouling behavior on-line during the process run. Model-based optimal non-linear control coupled with parameter estimation is applied in a simulation case study to show the benefits of the proposed approach.

Record Type: Published Article

Submitted To: LAPSE (Living Archive for Process Systems Engineering)

Citation (overall record, always the latest version):

LAPSE:2018.0194

Citation (this specific file, latest version):

LAPSE:2018.0194-1

Citation (this specific file, this version):

LAPSE:2018.0194-1v1

DOI of Published Version: <https://doi.org/10.3390/pr4040043>

License: Creative Commons Attribution 4.0 International (CC BY 4.0)

Article

Combined Estimation and Optimal Control of Batch Membrane Processes

Martin Jelemenský ^{1,*}, Daniela Pakšiová ¹, Radoslav Paulen ², Abderrazak Latifi ³
and Miroslav Fikar ¹

¹ Faculty of Chemical and Food Technology, Slovak University of Technology in Bratislava, Radlinskeho 9, 81237 Bratislava, Slovakia; daniela.paksiova@stuba.sk (D.P.); miroslav.fikar@stuba.sk (M.F.)

² Department of Chemical and Biochemical Engineering, Technische Universität Dortmund, Emil-Figge-Strasse 70, 44221 Dortmund, Germany; radoslav.paulen@bci.tu-dortmund.de

³ Laboratoire Réactions et Génie des Procédés, CNRS-ENSIC, Université de Lorraine, 1 rue Grandville, 54001 Nancy, France; Abderrazak.Latifi@ensic.inpl-nancy.fr

* Correspondence: martin.jelemensky@stuba.sk; Tel.: +421-259-325-730

Academic Editor: Dominique Bonvin

Received: 13 October 2016; Accepted: 10 November 2016; Published: 18 November 2016

Abstract: In this paper, we deal with the model-based time-optimal operation of a batch diafiltration process in the presence of membrane fouling. Membrane fouling poses one of the major problems in the field of membrane processes. We model the fouling behavior and estimate its parameters using various methods. Least-squares, least-squares with a moving horizon, recursive least-squares methods and the extended Kalman filter are applied and discussed for the estimation of the fouling behavior on-line during the process run. Model-based optimal non-linear control coupled with parameter estimation is applied in a simulation case study to show the benefits of the proposed approach.

Keywords: batch diafiltration; membrane fouling; time-optimal operation; fouling estimation

1. Introduction

In the last few decades, an increased focus has been given to the membrane technology, which is used both in production and down-stream processing. Membrane technologies are exploited in a variety of processes, such as in membrane reactors, membrane distillation, diafiltration, dialysis, electrolysis, etc. All of these employ principles of membrane separation [1]. There are, however, certain limitations for the application. The major obstacle is the membrane fouling, which refers to the blockage of membrane pores during filtration, causing a decrease in the production rate. Although membrane fouling is an inevitable phenomenon occurring during the filtration process, it can be controlled and its influence alleviated [2].

There are several operational approaches to deal with membrane fouling. These mainly include scheduling of membrane cleaning cycles and on-line fouling control strategies [2]. These techniques usually require mathematical models for predicting the fouling behavior. The first attempt to describe the fouling behavior in a unified model was presented by Hermia [3] who developed four types of fouling behavior. Few works have then been devoted to model-based optimal control of membrane-assisted processes [4,5] and membrane fouling and cleaning [6,7]. However, optimal control requires knowledge of process model parameters. Thus, in the case of unknown parameters in the model, on-line parameter estimation may be performed in order to improve the process control performance. For this purpose, several methods can be applied, such as common least-squares or some more advanced methods. In Charfi et al. [8], an estimation of the fouling mechanism model proposed by Hermia was conducted for microfiltration and ultrafiltration in a membrane bioreactor using experimental data reported in the literature.

Parameter identifiability has to be addressed before the actual estimation step. If a model is non-identifiable, the estimated parameters will lead to errors in subsequent model predictions [9]. The identifiability test also provides a guideline for how to simplify the model structure or indicates when more measurements are needed to allow for unique identifiability [10].

Recently, the optimal control of a diafiltration process in the presence of fouling was proposed [11]. The obtained control law depends mainly on the parameters of membrane flux and fouling. In the subsequent work, an on-line control strategy was considered that incorporated parameter estimation using the extended Kalman filter (EKF) [12]. In this paper, we focus more broadly on the optimal operation of a diafiltration process under fouling where parameter estimation is used to determine the parameters of the fouling model. Compared to the work [12], in this paper, we provide a comparison of the estimation of the unknown parameters using several optimization-based (least-squares and moving-horizon estimation), as well as recursive (recursive least-squares estimation and extended Kalman filter) methods by considering also measurement and process noise, and we assess their performance on the highly non-linear estimation problem that arises from the nature of the employed process model. In addition, the identifiability test is also performed in order to reveal whether the model parameters can be estimated based on available measurements.

The paper is organized as follows. In the next section, the general formulation of a batch diafiltration process is introduced. In Section 3, we provide a detailed analysis of the membrane fouling phenomena, and we also discuss the modeling of the membrane fouling. Section 4 provides the definition of the problem of the optimal operation of the diafiltration process. In Section 5, a review of identifiability detection methods is given with the main focus on non-linear models described by ordinary differential equations (ODEs). The results and implications of the performed identifiability tests are presented. An introduction is then given to the considered parameter estimation problem, and four different estimation methods are briefly presented. Finally, a simulation study is provided in Section 6 with a detailed discussion and analysis of the obtained results.

2. Process Description

In this paper, we consider the batch diafiltration process shown in Figure 1. The batch process operates under constant pressure and temperature. The overall batch process consists of a feed tank and a membrane module. We consider that the process solution consist of a solvent (water), macro-solute (high molecular weight component) and micro-solute (low molecular weight component). The process solution is brought from the feed tank to the membrane by means of mechanical energy (e.g., pump). The membrane is designed to retain the macro-solute, while the micro-solute can easily pass through the membrane pores. The stream rejected by the membrane is called the retentate and is taken back into the feed tank. The stream that leaves the system is called the permeate, and its flow-rate is defined as $q = AJ$, where A is the membrane area and J is the permeate flux per unit area of the membrane. The flow rate of the permeate stream is a function of both concentrations of the individual solutes and of time, in the case of the occurrence of fouling phenomena.

The control of the batch diafiltration process can be achieved by adjusting the flow-rate of the solvent (diluent) into the feed tank. Then, the control variable α is defined as the ratio between the inflow into the feed tank and the outflow from the membrane module. Traditional approaches to the operation of the diafiltration process consider sequences of operating regimes, which differ in the rate of diluent addition:

1. Concentration mode: During this mode, no diluent is added into the feed tank, i.e., $\alpha = 0$.
2. Constant-volume diafiltration mode: Here, the rate of inflow of the diluent is kept the same as the rate of permeate outflow, i.e., $\alpha = 1$.
3. (Pure) Dilution mode: In this mode, a certain amount of diluent can be added instantaneously into the feed tank. This can be represented as $\alpha \rightarrow \infty$.

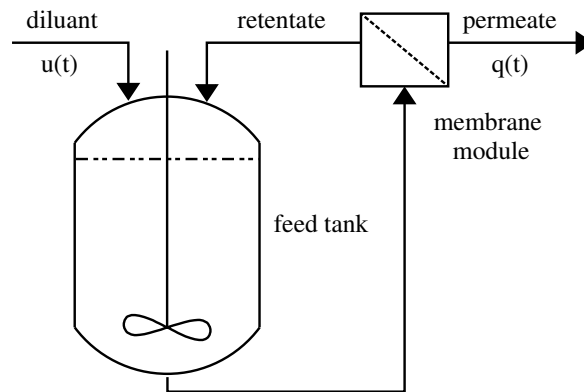


Figure 1. Schematic representation of a generalized ultrafiltration/diafiltration process.

Regarding the practical applicability of the pure dilution mode, as will be shown further, the optimal operation would involve the pure dilution mode as an initial or a final step. In both of these cases, the pure dilution step can be performed out of the batch as a pre-/post-treatment of the separated solution.

The process model [13] is constituted by the solutes mass balance equations, which have the following form:

$$\frac{dc_i}{dt} = \frac{c_i q}{V} (R_i - \alpha), \quad c_i(0) = c_{i,0}, \quad i = 1, 2, \quad (1)$$

where V is the volume in the feed tank and c_i denotes the concentration of the i -th solute; $i = 1$ for macro-solute; and $i = 2$ for micro-solute. R_i is the observed rejection coefficient of the i -th solute, which describes the ability of the membrane to retain the solute. The total mass balance is defined as follows:

$$\frac{dV}{dt} = \alpha q - q = (\alpha - 1)AJ, \quad V(0) = V_0, \quad (2)$$

where V_0 is the initial volume of the process solution.

In many cases, the rejection coefficient R_i is a function of concentrations. In this paper, we will consider that the rejections are ideal, given and constant for both solutes ($R_1 = 1$ and $R_2 = 0$). This means that the membrane is perfectly impermeable to macro-solute and that only the micro-solute can pass through the membrane. Moreover, since only the micro-solute can pass through the membrane, the mass of macro-solute in the system will not change and remains constant, such that $c_1(t)V(t) = c_{1,0}V_0$. Based on these assumptions, we can eliminate the differential equation for the volume (2), and the model has now the following form:

$$\frac{dc_1}{dt} = c_1^2 \frac{AJ}{c_{1,0}V_0} (1 - \alpha), \quad c_1(0) = c_{1,0}, \quad (3)$$

$$\frac{dc_2}{dt} = -c_1 c_2 \frac{AJ}{c_{1,0}V_0} \alpha, \quad c_2(0) = c_{2,0}. \quad (4)$$

3. Membrane Fouling

One of the major obstacles in the field of membrane-assisted processing is the membrane fouling. The membrane fouling is caused by the deposit of the solutes in/on the membrane pores. The main consequence of the membrane fouling is the decrease in the permeate outflow due to the blockage of the pores. The most important factors that cause the membrane fouling are the feed properties and membrane material. Moreover, membrane fouling can increase when the concentration polarization effect takes place. In [14], the authors have shown that during the membrane filtration, the retained macro-solute can form a gel layer over the membrane surface. The formation of such a gel layer can increase the interactions between the solutes and the membrane surface, which eventually lead

to membrane fouling. Further, the process variables, such as temperature and pressure, may also contribute to membrane fouling [15].

As mentioned above, the main consequence of the membrane fouling is the decrease in the permeate flow rate, which results in the increase of the overall processing time. If the membrane becomes heavily fouled, cleaning must be performed. In some cases, even replacement of the membrane module must be carried out, which leads to the increase of the production costs.

In the past three decades, the modeling of membrane fouling became very important. In 1982, Hermia [3] derived a unified fouling model for dead-end filtration in terms of total permeate flux and time, which has the following form:

$$\frac{d^2t}{dV_p^2} = K \left(\frac{dt}{dV_p} \right)^n, \quad (5)$$

where V_p is the permeate volume, t is time and K represents the fouling rate constant. From the above unified fouling model, four classical fouling models can be derived for different values of the fouling parameter n . The four models are cake filtration ($n = 0$), intermediate blocking ($n = 1$), internal blocking ($n = 3/2$) and the complete blocking model ($n = 2$). By integration of (5), we obtain the differential equation for the permeate flux [16,17], which is expressed as:

$$\frac{dJ}{dt} = -KA^{2-n}J^{3-n}. \quad (6)$$

This equation can be solved to give an explicit solution $J(t, K, n, A, J_0)$ if the parameters A , K and n are constant and J_0 is the initial flux at time $t = 0$. However, this model holds only for dead-end filtration mode. In order to account for the dynamics of the cross-flow system, which is considered in this paper, we propose to substitute the initial flux $J_0(t = 0)$ with the flux of an unfouled membrane $J_0(c_1, c_2)$.

In Figure 2, we show the graphical representation of the four standard fouling models. The models differ in the way that the solutes deposit on/in the membrane surface. In the following subsections, we will briefly discuss the individual fouling models.

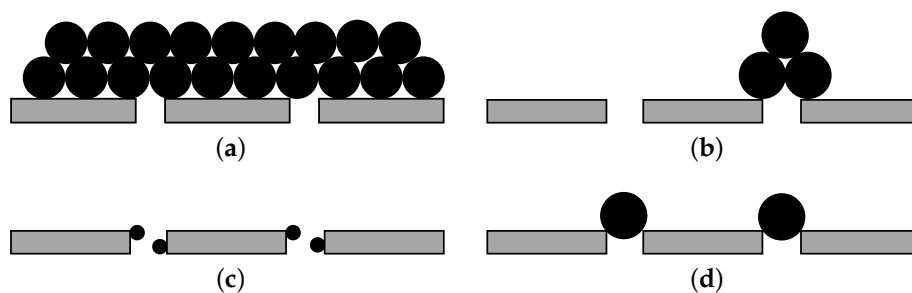


Figure 2. Graphical representation of the four classical fouling models developed by Hermia. (a) Cake filtration model ($n = 0$); (b) intermediate blocking model ($n = 1$); (c) internal blocking model ($n = 3/2$); (d) complete blocking model ($n = 2$).

3.1. Cake Filtration Model

The cake filtration model considers that the solutes brought to the membrane surface will form a multi-layered cake formation shown in Figure 2a. The multi-layered formation is caused by the repeating deposit of the solutes on the membrane surface. The parameter n is set to zero, and the permeate flux has the following form:

$$\frac{1}{J^2} = \frac{1}{J_0^2} + 2KA^2t. \quad (7)$$

where J in $\text{m}^3/\text{s}/\text{m}^2$ is the permeate flow per unit area of the membrane and J_0 is the flux of an unfouled membrane.

3.2. Intermediate Blocking Model

The intermediate fouling model also considers that the solute will block all of the pores. However, in this case, the solutes can also deposit on each other, as illustrated in Figure 2b. To derive the model, the parameter n is equal to one, and the permeate flux is of the form:

$$\frac{1}{J} = \frac{1}{J_0} + KA t. \quad (8)$$

3.3. Internal Blocking Model

The aforementioned fouling models do not consider the fouling to take place inside the membrane pores. This is not the case with an internal blocking model (Figure 2c). This type of fouling results in the decrease of the diameter of the membrane pores, which leads to the decrease of the permeate flow. For the internal blocking model, the parameter n is equal to $3/2$, and the permeate flux is derived as:

$$\frac{1}{\sqrt{J}} = \frac{1}{\sqrt{J_0}} + \frac{1}{2}KA^{1/2}t. \quad (9)$$

3.4. Complete Pore Blocking Model

The complete pore blocking model considers that all of the solutes brought to the membrane surface will block the membrane pores (Figure 2d). Therefore, the permeate flow will be reduced. Moreover, the blockage of the pores is caused only if the molecules of the solutes are larger than the membrane pores. The complete blocking model can be derived from (5) if $n = 2$ and is of the following form:

$$\ln J = \ln J_0 - Kt, \quad (10)$$

Note that unlike in the previous cases, the permeate flux does not depend on the membrane area.

4. Optimization Problem

The main operational objective of the diafiltration plant is to process as much solution as possible. This can only be achieved by pursuing the time-optimal batch recipes with tight scheduling. Consequently, as scheduling is here usually a rather trivial task, we concentrate on the time-optimal control problem of a batch membrane process in the presence of membrane fouling.

4.1. Problem Definition

The objective is to identify such a time-dependent function $\alpha(t)$ that will drive the system from the initial concentrations to the required final concentrations in the minimum time. The optimization problem reads as:

$$\mathcal{J}^* = \min_{\alpha(t)} \int_0^{t_f} 1 \, dt, \quad (11)$$

$$\text{s.t. } \dot{c}_1 = \frac{c_1^2 A J}{c_{1,0} V_0} (1 - \alpha), \quad c_1(0) = c_{1,0}, \quad c_1(t_f) = c_{1,f}, \quad (12)$$

$$\dot{c}_2 = -\frac{c_1 c_2 A J}{c_{1,0} V_0} \alpha, \quad c_2(0) = c_{2,0}, \quad c_2(t_f) = c_{2,f}, \quad (13)$$

$$J = J(t, J_0(c_1, c_2), K, n), \quad (14)$$

$$\alpha \in [0, \infty). \quad (15)$$

We assume that this optimization problem is feasible, i.e., the process design ensures that the final conditions can be met within admissible input trajectories. This optimization problem can be approximately solved using various numerical methods [18] of dynamic optimization (e.g., control vector parametrization or complete discretization). However, an analytical solution is available. This was shown in [19] and is briefly reviewed in the following.

4.2. Nominal Optimal Operation

In our recent study [19], we derived analytically the complete optimal operation for a diafiltration process in the presence of membrane fouling using Pontryagin's minimum principle [20,21]. The optimal operation is an explicit non-linear control strategy defined on three time intervals as follows:

1. In the first interval, the control is kept on the boundaries (e.g., minimum or maximum) until the singular curve is reached:

$$S(t, c_1, c_2, K, n) = J + c_1 \frac{\partial J}{\partial c_1} + c_2 \frac{\partial J}{\partial c_2} = 0. \quad (16)$$

The minimum control action ($\alpha = 0$) is applied if the initial concentrations lie to the left side of the singular surface ($S(0, c_{1,0}, c_{2,0}, K, n) < 0$) in the state diagram. Conversely, the maximum ($\alpha \rightarrow \infty$) control action is applied when the initial concentrations are on the right of the singular curve ($S(0, c_{1,0}, c_{2,0}, K, n) > 0$).

2. Once the singular curve is reached, the singular control is applied, which forces the states to stay on the singular curve:

$$\alpha(t, c_1, c_2, K, n) = \frac{\frac{\partial S}{\partial c_1} c_1}{\frac{\partial S}{\partial c_1} c_1 + \frac{\partial S}{\partial c_2} c_2} + \frac{\frac{\partial S}{\partial t}}{\frac{c_1 A J}{c_{10} V_0} \left(\frac{\partial S}{\partial c_1} c_1 + \frac{\partial S}{\partial c_2} c_2 \right)}. \quad (17)$$

This step is terminated once the ratio of the concentrations is equal to the ratio of the final concentrations or once the final concentration of the micro-solute is reached.

3. In the last step, the control is kept on one of the boundaries similar to the first step with the difference that the operation mode to be applied is determined by the final time constraints. The concentration mode ($\alpha = 0$) is applied if the final concentration of the micro-solute has been reached. The dilution mode with $\alpha \rightarrow \infty$ is applied once the ratio of the final concentrations is equal to the final one. Both steps are performed until the final concentrations are reached.

For a particular case of initial and final concentrations, any of the three steps can be missing. For instance, if initial (respectively final) concentrations lie on the singular case, then the first (respectively last) step will be skipped. Another case would occur if the singular case cannot be reached before the control has to switch from the first to the last step, so the optimal operation will consist entirely of saturated control arcs.

4.3. Optimal Operation with Imperfect Knowledge of Fouling Model Parameters

It can be easily shown that the optimal control structure does not depend on the unknown fouling parameters in the first and last step of the optimal operation. This comes from the fact that, on the one hand, the initial flux is not influenced by the fouling, as $J(0) = J_0$, and on the other hand, the last step is determined by the active constraints, which are independent of the fouling behavior.

These observations imply that one can distinguish two cases in the on-line optimal operation under imperfect knowledge of the fouling model:

1. If the operation starts with the concentration mode, the a priori knowledge of unknown parameters K and n should be improved during this step by parameter estimation, such that the switching time can be determined accurately. If this estimation is successful, the optimal singular surface is followed thereafter.
2. Should the optimal operation commence with a pure dilution step, the entry point to the singular surface would be known exactly, as it only depends on J_0 . In other words, the amount of the diluent that should be added at the beginning of the operation does not depend on the parameters of the fouling model. It would then be again necessary to improve the knowledge about the unknown parameters in order to maintain the singular-surface condition.

Clearly, the on-line parameter estimation will yield adequate results only if (a) a sufficient amount of data can be gathered and if (b) significant sensitivity is present between measured outputs and estimated parameters on the particular arc of the optimal solution. The first condition requires that the optimal concentration mode, respectively singular arc, lasts sufficiently long, which will likely result in the negligible suboptimality of the on-line control strategy. The latter condition suggests the provision of the persistence of excitation in the system. This might be critical should the parameter estimation be required to yield accurate parameter estimates when the system resides on the singular arc. For example, if the optimal singular flux is constant, then the set of obtained flux measurements will not bring enough new information about the fouling behavior. This case will further be addressed in Section 6 where the interesting question is to find out what the trade-off is between the insensitivity of the objective function to the control on the singular arc and the inaccuracy of parameter estimates.

5. Parameter Estimation Problem

In this section, we will discuss parameter identifiability and parameter estimation techniques. Several studies have employed standard estimation methods to obtain the fouling parameters necessary for the control of membrane processes. In [8], the authors have proposed to employ the least-squares method to identify the fouling rate constant (K) for the standard fouling models. However, this procedure was performed completely off-line with the necessity of initial probing runs to obtain the experimental measurements of flux based on which the fouling rate constant could be determined. Moreover, only the fouling rate constant was estimated with a fixed type of fouling model. Since the derived optimal control depends also on the information of the type of fouling model, the estimation of both the unknown fouling parameters (K, n) is followed in this work. Furthermore the estimation is performed on-line, which eliminates the necessity of any initial experiments.

The on-line estimation is coupled with the aforementioned optimal feedback control law, which is updated iteratively with the estimated values of the parameters of the fouling model. In our recent paper [12], we have introduced this strategy using EKF. In this paper, we provide an assessment of the performance of some other commonly-used model-based estimators.

5.1. Problem Definition

We will consider that the process dynamics is described by a set of ODEs with initial conditions and expressed in the following form:

$$\dot{c}_1 = \frac{c_1^2 A J}{c_{1,0} V_0} (1 - \alpha), \quad c_1(0) = c_{1,0}, \quad (18)$$

$$\dot{c}_2 = -\frac{c_1 c_2 A J}{c_{1,0} V_0} \alpha, \quad c_2(0) = c_{2,0}, \quad (19)$$

$$\dot{\tau} = 1, \quad \tau(0) = 0, \quad (20)$$

where τ replaces the explicit appearance of time with a new state, and the permeate flux J is defined as follows:

$$J(\tau, J_0(c_1, c_2), n, K, A) = J_0 \left(1 + K(2 - n)(AJ_0)^{2-n}\tau \right)^{1/(n-2)}, \quad (21)$$

We will rewrite the above differential equations into the following compact form:

$$\dot{x} = f(x, \theta, u), \quad (22)$$

where $x = (c_1, c_2, \tau)^T$ is the vector of state variables, $\theta = (K, n)^T$ is the vector of estimated parameters and u is the control action. For the estimation of the unknown parameters, we assume that the concentrations of the individual solutes (c_1, c_2) and also the permeate flux (J) can be measured.

5.2. Identifiability Problem

The problem of the parameter identifiability of dynamic models is a special case of the observability problem. It is often the case that the parameters of a model have unknown values. These values can be estimated from observing the input-output behavior.

Before the model parameters can be estimated, the identifiability analysis must be performed. The property of identifiability is the ability to deduce the values of the model parameters uniquely in terms of known quantities based on input and output variables and their time derivatives. In order to investigate identifiability of (18) to (20), we will first introduce a definition of identifiability.

Definition 1. A dynamic system described by (22) and (25) is identifiable if the parameter vector θ can uniquely be determined from the given system input u and the measurable system output y . If no solution can be found for the given vector of unknown parameters θ , then the model is non-identifiable.

The identifiability analysis also distinguishes whether the parameter vector can uniquely be identified or not. In this regard, a system is said to be structurally globally identifiable if for any admissible input and any two parameter vectors θ_1 and θ_2 , $y(u, \theta_1) = y(u, \theta_2)$ holds if and only if $\theta_1 = \theta_2$. A system is said to be structurally locally identifiable if for any parameter vector θ within an open neighborhood of some point, $y(u, \theta_1) = y(u, \theta_2)$ holds if and only if $\theta_1 = \theta_2$.

There are many well-established methods for determining the identifiability of linear systems. However, fewer methods are applicable for non-linear systems. The most common methods for the identifiability test include: the direct test [22], the identifiability of the corresponding linearized system [22], the similarity transformation method [23,24], power series expansion [25], the generating series approach [26] and differential algebra [27]. Two of these methods are used for the identifiability analysis in this work.

5.2.1. Identifiability of the Corresponding Linearized System

Local linearization of the non-linear models can be considered for the identifiability test. However, the identifiability of the linearized model is only a necessary condition [22].

There are several methods for the identifiability of linear systems. One method is to consider parameters as dummy state variables with time derivatives equal to zero. The observability rank test can then be applied where the observability matrix O of a linear system $\dot{x} = A(\theta)x + B(\theta)u$, $y = C(\theta)x$,

$$O = \begin{bmatrix} C & CA & \cdots & CA^{n-1} \end{bmatrix}^T, \quad (23)$$

should be of full rank n , where n is the dimension of the vector of state variables. Rigorously speaking, this approach can only serve to detect the non-identifiability of the parameters.

5.2.2. Generating Series Approach

This approach is based on the computation of Lie derivatives, which allows one to expand the output variables with respect to input variables. By computing Lie derivatives, a set of resulting coefficients is formed, which is then used to compute the parametric solution. Based on this solution, the identifiability of the model is checked. A drawback of this method is that the sufficient number of Lie derivatives to be considered is unknown. In order to easily visualize the structural identifiability problem and to assist in the solution, identifiability tableau is proposed [26]. This tableau expresses the dependency between parameters by using the Jacobian matrix of exhaustive summary, i.e., a set of relations that depends only on the parameters. The tableau is represented by a “0 – 1” matrix, which represents whether there is a dependency between Lie derivatives and the parameters “1”, or not “0”. The identifiability is then checked by examining whether there is a unique “1” in a row of the tableau. If so, then there exists a unique dependency between the Lie derivative and a parameter. Thus, the corresponding parameter is structurally identifiable [9,28].

5.3. Software Tools for the Identifiability Test

There are few frameworks and software tools available that assist in identifiability analysis [29,30]. To the authors knowledge, there are three common software tools that use one of the above-mentioned methods for testing global identifiability of non-linear systems. These are DAISY [31], COMBOS [30] and GenSSI. The open-source tool GenSSI implements the generating series approach and is chosen here for the identifiability analysis of the parameters. This is due to the model structure, which must be strictly polynomial in order to be implemented in DAISY and in COMBOS. In our case, the model of the diafiltration process with fouling does not satisfy this requirement.

GenSSI module is written in MATLAB and can be used for any general linear, non-linear ODE or differential algebraic system. To run the identifiability test, the user needs to specify the model equations, input and output variables, initial conditions, parameters to be identified and the number of Lie derivatives. In the case of the detection of non-identifiability, GenSSI also provides information on individual parameters and their global or local identifiability. This information can guide the proper reformulation of the model [9,28].

Results and Discussion

The identifiability problem was solved for the membrane fouling process described by (3) and (4) using both the observability matrix test and generating series approach. The results showed that the parameters K , n in the membrane fouling model are not observable with the vector of measurement $\mathbf{y} = (c_1, c_2, J)^T$. This is due to the fact that parameters enter the process equations via J only, and therefore, there are infinitely many combinations that can lead to actual J . Therefore, a new measurement must be added, which is a different function of the estimated parameters. For this reason, the time derivative of the flux $\dot{J}$ is considered as a new measurement. Based on the results given by GenSSI, the augmented model of membrane fouling, which also considers the time-derivative of the flux as a measured variable, is globally identifiable. This means that the two parameters K , n can be identified based on the measured variables from any given initial conditions. The identifiability of the parameters was also confirmed by using the observability matrix.

It is usually not possible to measure the derivative of the flux directly. Therefore, an approximation of the third order is considered, which reads as:

$$\frac{dJ}{dt} \approx \frac{J_{k-3} - 6J_{k-2} + 3J_{k-1} + 2J_k}{6T_s} \quad (24)$$

where T_s represents the sampling time. Different approximations were tested; the third order is a good compromise between accuracy and complexity.

Then, the considered measured outputs in discrete-time samples (k) are as follows:

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \boldsymbol{\theta}) = (c_1, c_2, J, \dot{J})^T. \quad (25)$$

This form of output equation is further considered for parameter estimation.

5.4. Parameter Estimation Methods

Different parameter estimation methods are discussed in the following subsections. They mainly differ in the consideration of the parameter estimation problem as non-linear (least-square method), mildly non-linear (least-squares method with moving horizon) or almost linear (recursive least-squares method and EKF). Note that this consideration implies whether the non-linear programming problem (NLP) needs to be solved with each newly-gathered measurement or whether a recursive approach is sufficient, where only a few mathematical operation (e.g., matrix-vector multiplications) are performed. Here, a critical point is presented by the tendency of the recursive methods to rapidly diverge should the process (and its model) exhibit strong non-linear behavior.

5.4.1. Weighted Least-Squares Method

One of the most used methods for parameter estimation is the weighted least-squares (WLS) method. Compared to the standard least-squares method, the WLS method considers that different measured outputs can contribute differently to the estimation of unknown parameters based on their variability and statistical significance [32]. For the on-line parameter estimation, it must be, however, considered that with each new measurement, a new NLP problem must be solved, and moreover, if a large set of experimental points is available, numerical problems can occur due to the gradually increasing complexity of the optimization problem. The optimization problem that needs to be solved in each sampling time is as follows:

$$\hat{\boldsymbol{\theta}}_k = \arg \min_{\boldsymbol{\theta}} \sum_{j=0}^k (\mathbf{y}_j - \hat{\mathbf{y}}_j)^T \mathbf{R}^{-1} (\mathbf{y}_j - \hat{\mathbf{y}}_j) \quad (26)$$

$$\text{s.t. (18) to (20),} \quad (27)$$

$$\hat{\mathbf{y}}_j = \mathbf{h}(\mathbf{x}_j, \boldsymbol{\theta}), \quad (28)$$

$$\boldsymbol{\theta}_{\min} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\max}, \quad (29)$$

where $\mathbf{y}$ represents the vector of obtained measurements, $\mathbf{R}$ is the covariance matrix for the measurement noise and $\boldsymbol{\theta}_{\min}$, $\boldsymbol{\theta}_{\max}$ are the minimum and maximum values of the parameters, respectively.

5.4.2. Weighted Least-Squares Method with Moving Horizon

The weighted least-squares method with moving horizon [33] is very similar to the previously-discussed method. However, in this case, not the whole measurement set is considered, but only recent measurements on a constant moving horizon (MH) contribute actively to the estimation. The influence of the past measurements, gathered prior to the moving horizon, is not neglected, but rather considered through the parametric covariance matrix $\mathbf{P}$ as a priori knowledge of the parameters, similarly to Bayesian estimation. The covariance matrix then penalizes a drift of the new estimates from their a priori values in the objective function [34]. In [35], it has also been shown that the moving-horizon estimation is equivalent to the extended Kalman filter if the first-order approximation of dynamics and no constraints are considered. A distinct advantage of this method is that since a fixed-length horizon of measured outputs is considered, the NLP problem that needs to be solved, when new measurements are gathered, is of approximately fixed complexity and might be easier to solve than the optimization problem in WLS estimation. The resulting non-linear problem reads as:

$$\hat{\theta}_k = \arg \min_{\theta} (\theta - \hat{\theta}_{k-MH})^T P_{k-MH}^{-1} (\theta - \hat{\theta}_{k-MH}) + \sum_{j=k-MH}^k (y_j - \hat{y}_j)^T R^{-1} (y_j - \hat{y}_j) \quad (30)$$

$$\text{s.t. (18) to (20),} \quad (31)$$

$$\hat{y}_j = h(x_j, \theta), \quad (32)$$

$$\theta_{\min} \leq \theta \leq \theta_{\max}. \quad (33)$$

It is worth mentioning that the matrix P can be either selected as a constant matrix or one can apply the extended Kalman filter method to update the matrix P in each sampling time to improve the convergence.

5.4.3. Recursive Least-Squares Method

The recursive least-squares method (RLS) iteratively finds parameter estimates that minimize the linearized least-squares objective function [36]. The parameter estimates are calculated at each sampling time based on the a priori knowledge of the parameters (through parametric covariance matrix) and the current measured outputs. The advantage of the method is that during the estimation, no NLP problem has to be solved. On the other hand, RLS does not explicitly consider the constraints on the estimated parameters in the updates of the covariance matrix, and thus, the a priori knowledge of parameters might be biased.

In order to apply the RLS method, we consider a Taylor expansion of the output function (25) to the first order:

$$\hat{y}_k = h(x_k, \theta) \approx h(x_k, \hat{\theta}_{k-1}) + \left(\frac{\partial h(x_k, \theta)}{\partial x_k} \bigg|_{\hat{\theta}_{k-1}} \frac{\partial x}{\partial \theta} \bigg|_{\hat{\theta}_{k-1}} + \frac{\partial h(x_k, \theta)}{\partial \theta} \bigg|_{\hat{\theta}_{k-1}} \right)^T (\hat{\theta}_k - \hat{\theta}_{k-1}), \quad (34)$$

$$\approx h(x_k, \hat{\theta}_{k-1}) + \varphi_k^T (\hat{\theta}_k - \hat{\theta}_{k-1}), \quad (35)$$

where $(\partial x / \partial \theta)$ is the matrix of parametric sensitivities, which can be propagated in time as follows:

$$\dot{s} = \frac{d}{dt} \frac{\partial x}{\partial \theta} \bigg|_{\hat{\theta}_{k-1}} = \frac{\partial f(x, \theta, u)}{\partial x} \bigg|_{\hat{\theta}_{k-1}} \frac{\partial x}{\partial \theta} \bigg|_{\hat{\theta}_{k-1}} + \frac{\partial f}{\partial \theta} \bigg|_{\hat{\theta}_{k-1}}, \quad s(0) = \frac{\partial x_0}{\partial \theta} \bigg|_{\hat{\theta}_{k-1}}, \quad (36)$$

for the given $u(t)$. Here, we can observe that the sensitivity of the system needs to be ideally solved at each sampling time from the initial conditions at $t = 0$. This can impose significant computational burden and certain numerical problems, since many states and parameters are present in the problem. This does not, however, present a significant issue in the present study.

The estimation error (ϵ) can be defined via the difference between the measured and calculated outputs as:

$$\epsilon = y_k - \hat{y}_k = y_k - h(x_k, \hat{\theta}_k) - \varphi_k^T \hat{\theta}_k + \varphi_k^T \hat{\theta}_{k-1}, \quad (37)$$

The overall RLS algorithm [37] is then, using (34), defined as follows:

$$\hat{\theta}_k = \left(I + L_k \varphi_k^T \right)^{-1} \left(\hat{\theta}_{k-1} + L_k \left[y_k - h(x_k, \hat{\theta}_{k-1}) + \varphi_k^T \hat{\theta}_{k-1} \right] \right), \quad (38)$$

$$L_k = P_{k-1} \varphi_k \left(\varphi_k^T P_{k-1} \varphi_k + R \right)^{-1}, \quad (39)$$

$$P_k = P_{k-1} - P_{k-1} \varphi_k \left(\varphi_k^T P_{k-1} \varphi_k + R \right)^{-1} \varphi_k^T P_{k-1}, \quad (40)$$

where I is the identity matrix and L is the so-called gain matrix.

5.4.4. Extended Kalman Filter

The Extended Kalman Filter is one of the most used methods for joint state and parameter estimation of nonlinear systems [38,39]. The main idea is the use of the prediction-correction principle where the non-linear model is used for the prediction (denoted by the superscript $-$), and its linearized counterpart around the current estimate is exploited in the propagation of the covariance matrix P and in the subsequent correction step (denoted by the superscript $+$) with the discrete system measurements. Compared to the recursive least-squares method, one can additionally account for the nonlinearity and the noise that affects the system dynamics using EKF. This can prove to be highly advantageous if parametric uncertainty influences the measurements strongly. Similarly to the recursive least-squares method, the absence of the treatment of the constraints and severe nonlinearities can lead to the divergence of the estimator. On the other hand, the advantage of EKF lies in its simpler implementation, where no parametric sensitivities are required, and in increased tuning capabilities, which might yield faster and more robust convergence.

To formulate the EKF algorithm, the estimated parameters are first represented by new dummy state variables with zero dynamics:

$$\frac{d\hat{\theta}}{dt} = \mathbf{0}, \quad \hat{\theta}(0) = \hat{\theta}_0, \quad (41)$$

and the dynamics of a new state vector is constructed using the dynamics of state and parameter estimates:

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u) = (f^T(\hat{x}, \hat{\theta}, u), \mathbf{0}^T)^T, \quad \text{with } \tilde{x} = (\hat{x}^T, \hat{\theta}^T)^T. \quad (42)$$

The observer dynamics then reads in the following form:

$$\dot{\tilde{x}}^- = \tilde{f}(\tilde{x}^-, u), \quad \tilde{x}^-(t_{k-1}) = \tilde{x}_{k-1}^+, \quad (43)$$

$$\dot{\tilde{P}}^- = F\tilde{P}^- + \tilde{P}^- F^T + Q, \quad \tilde{P}^-(t_{k-1}) = \tilde{P}_{k-1}^+, \quad (44)$$

with initial estimates $\tilde{x}^-(t_0) = \tilde{x}_0$ and $\tilde{P}^-(t_0) = \tilde{P}_0$ and with the update at measurement instants k defined as:

$$\tilde{x}_k^+ = \tilde{x}^-(t_k) + L_k[y_k - \tilde{h}(\tilde{x}^-(t_k))], \quad (45)$$

$$L_k = \tilde{P}^-(t_k)C_k^T [C_k\tilde{P}^-(t_k)C_k^T + R]^{-1}, \quad (46)$$

$$\tilde{P}_k^+ = (I - L_k C_k)\tilde{P}^-(t_k). \quad (47)$$

The state transition and observation matrices are defined by the following Jacobians:

$$F = \left. \frac{\partial \tilde{f}(\tilde{x}, u)}{\partial \tilde{x}} \right|_{\tilde{x}^-(t), u(t)}, \quad C_k = \left. \frac{\partial \tilde{h}(\tilde{x})}{\partial \tilde{x}} \right|_{\tilde{x}^-(t_k)}. \quad (48)$$

Matrices Q, P denote, respectively, the covariance matrix of the noise affecting the state dynamics and the covariance of the estimated states and parameters. These matrices can also be thought of as tuning knobs of the estimation algorithm affecting its estimation performance and convergence.

5.4.5. General Assessment of the Presented Estimation Methods

The presented estimation methods differ in several aspects. One can distinguish them based on the level of accounting for the nonlinearity of the process model and the number of tuning parameters. The WLS method is the only method here that employs the fully-nonlinear model and does not approximate the estimation problem. It does not require tuning of any parameters if we assume that

the information on the variance of the measurement noise is available based on the employed sensor equipment. A price to pay for this rigorousness is usually reflected in the computational time when measurements are obtained on some large time horizon. The WLSMH method uses an approximation to the full-horizon estimation problem and employs an approximation here that is parametrized by the tuning factor P . The RLS method works with the assumption that a linearized process model represents the behavior of the process within the subsequent sampling instants. This method can be regarded as residing on the other side of the spectrum of the estimation methods compared to WLS method, as it provides a computationally-efficient estimator, which only approximates the behavior of the optimal estimator. A similar behavior can be expected from the EKF algorithm. This can additionally account for the nonlinearity of the process model (via tuning parameter Q), so even further simplifications can be made for the linearization of the estimation problem (see the previous section).

6. Case Study

Here, we present the simulation results obtained with different estimation techniques presented above. We will consider the batch diafiltration process that operates under limiting flux conditions, and the permeate flux for the unfouled membrane reads as:

$$J_0(c_1) = k \ln \left(\frac{c_{\text{lim}}}{c_1} \right), \quad (49)$$

where k is the mass transfer coefficient and c_{lim} is the limiting concentration for the macro-solute. Note that the permeate flux depends only on the concentration of the macro-solute. The overall separation goal is to drive the system from the initial $[c_{1,0}, c_{2,0}]$ to the final point $[c_{1,f}, c_{2,f}]$ in the minimum time described by (11) to (15). The initial process volume is considered to be $V_0 = 0.1 \text{ m}^3$. The parameters for the limiting flux model are $k = 4.79 \text{ m/s}$, $c_{\text{lim}} = 319 \text{ mol/m}^3$ and the membrane area 1 m^2 . The a priori unknown fouling rate is $K = 2$ units. As the degree of nonlinearity of the model is strongly dependent on the nature of the fouling behavior, represented by the a priori unknown parameter n (see Section 3), we will study the cases when $n = \{0, 1, 1.5\}$.

We first study the case where $[c_{1,0}, c_{2,0}] = [10 \text{ mol/m}^3, 100 \text{ mol/m}^3]$ and $[c_{1,f}, c_{2,f}] = [100 \text{ mol/m}^3, 1 \text{ mol/m}^3]$. The time-optimal operation, as stated above, follows a three-step strategy. In the first step, the concentration mode is applied, followed with a singular arc:

$$S(\tau, J_0(c_1), K, n) = J \left(1 + k + k(n-2) J^{\frac{5-2n}{n-2}} K \tau (A J_0)^n \right) = 0, \quad (50)$$

and in the last step, pure dilution mode is performed. The switching times between the individual control arcs are determined by the fine precision of the numerical integration around the roots of the singular surface equation. The on-line estimation of the unknown parameters is performed where the samples of the measured process outputs (25) are assumed to be available with the sampling time ($T_s = 0.01 \text{ s}$). This means that the optimal control is updated at each sampling time based on the considered measured outputs. Based on our observations, the chosen sampling time did not pose any computational challenges in the estimation of the parameters. However, in the case of the application on a real process, this sampling time can pose computational difficulties for optimization-based estimation (WLS and WLSMH) since an optimization problem needs to be solved in each sampling time. In the case of the WLSMH method, the initial computing time for one estimation was observed to be approximately 10 s, and once the true values were reached, the computational time decreased to 2 s. For this reason, such a low sampling time is more adequate for recursive methods (RLS and EKF) where no non-linear optimization problem is needed to be solved.

For each studied estimation method (WLS, WLSMH, RLS and EKF), three simulations were performed, with $n = \{0, 1, 1.5\}$. For the WLSMH method, the moving horizon was set to $MH = 2$.

The length of the horizon is traditionally defined using the sampling instants (steps) as its units. The covariance matrices were tuned for all of the considered simulations as follows

$$P_0 = \text{diag}(0.1, 0.01), \quad (51)$$

$$\tilde{P}_0 = \text{diag}[\text{diag}(0.001, 0.001, 0.001), P_0], \quad (52)$$

$$R = \text{diag}(10^{-4}, 10^{-4}, 10^{-8}, 10^{-4}), \quad (53)$$

$$Q = \text{diag}(0.001, 0.001, 0.001, 0.3, 0.1). \quad (54)$$

For the WLSMH method, the covariance matrix P was updated using the EKF run in parallel. In our simulations with the RLS method, a different matrix P_0 had to be chosen for a different value of n ; otherwise, a divergent behavior was observed. We attribute this behavior to the strong non-linearities of the process model. Further, due to the non-linear behavior, the choice of P was very sensitive to the overall estimation. The noise in the measured data was simulated as a random normally-distributed Gaussian with the covariance matrix R (the noise covariance of the approximation of the flux derivative is determined empirically from the simulation data). The same evolution of the random noise is used in all presented simulations. According to our observations, the estimation performance is significantly influenced by the choice of R . A small change in the covariance matrix R can cause a big difference in the rate of estimation convergence. The chosen variances represent 1% standard deviations of the measurement noise. The actual magnitudes depend on the employed units of the process variables. Hence, the flux varies in the range $[0, 0.06]$ m/h.

The parameter estimation methods discussed in Section 5.4 were implemented in the MATLAB R2016a environment. For the optimization-based (WLS, WLSMH) methods, the built-in NLP solver `fmincon` was chosen. The tolerances for the optimized variables and the objective function were both set to 1×10^{-4} . The ordinary differential equations describing the process model and sensitivity equations were solved using the MATLAB subroutine `ode45`. The values for the relative and absolute tolerances were set to 1×10^{-8} and 1×10^{-6} , respectively. The precision on detecting the switching times by numerical integration was determined with the same tolerances as set by the subroutine `ode45`. All of the reported results were obtained on the workstation Intel Xeon CPU X5660 with 2.80 GHz and 16 GB RAM.

Table 1 presents a comparison of the rate of suboptimality of the processing time (δ_{tr}^*) achieved for the on-line control strategy with different estimation methods for different fouling models. We also show the values of normalized root mean squared error (NRMSE) for the estimated parameters ($\mathcal{R}_{\hat{\theta}}$) and for the concentrations and flux ($\mathcal{R}_{\hat{x}}$) and the cumulative computational time (t_c) needed for the used estimators. We can observe that the optimality loss for the different estimation techniques is negligible, and the same loss is achieved with all employed estimators. The highest optimality loss occurs in the case of $n = 0$. The increased optimality loss could be caused here because the value $n = 0$ is used as the lower bound for the estimated parameters, and thus, at least, local convergence problems could occur due to this hard constraint. We can observe that the NRMSE for the estimated parameters for the WLS and WLSMH is higher compared to the recursive methods (RLS and EKF). This was caused due to the high oscillations in the beginning of the estimation of the parameters. The NRMSEs for the concentrations and flux ($\mathcal{R}_{\hat{x}}$) indicate that the difference in the trajectories' profile in the case of ideal and estimated concentrations and flux is small. The highest difference in all cases was observed only in the EKF method. A possible explanation is that the estimated parameters using the rest of the methods (WLS, WLSMH and RLS) converged exactly or reasonably close with negligible difference to the true values of the parameters on the second interval. If we compare the WLS and WLSMH methods, we can notice a significant decrease in the computational time. This was expected since in the case of the WLS method, every new measurement increases the complexity of the optimization problem that needs to be solved, whereas the complexity remains the same for WLSMH, since only a constant amount of measurements is considered. Simulations with a larger horizon were also performed for WLSMH

($MH = 5, 10, 20$). Although the convergence of the estimated parameters was slightly faster and smoother when compared to the case with $MH = 2$, the computational time increased significantly.

Table 1. Comparison of the rate of the suboptimality of final processing times, the normalized root mean squared error for unknown parameters and concentrations and the computational time for different estimation techniques. MH, moving horizon; RLS, recursive least-squares.

Estimation	$n = 0$				$n = 1$				$n = 1.5$			
method	δ_{tf}^* (h)	$\mathcal{R}_{\hat{\theta}}$	$\mathcal{R}_{\hat{x}}$	t_c (h)	δ_{tf}^* (h)	$\mathcal{R}_{\hat{\theta}}$	$\mathcal{R}_{\hat{x}}$	t_c (h)	δ_{tf}^* (h)	$\mathcal{R}_{\hat{\theta}}$	$\mathcal{R}_{\hat{x}}$	t_c (h)
WLS	0.21	0.24	0.17	3.73	0.10	0.46	0.16	4.71	0.09	0.14	0.24	23.33
WLSMH	0.21	0.48	0.23	0.55	0.10	0.64	0.46	1.28	0.09	0.29	0.74	8.52
RLS	0.21	0.11	1.65	0.04	0.10	0.03	1.15	0.04	0.09	0.04	1.25	0.13
EKF	0.21	0.09	5.26	0.01	0.10	0.05	3.18	0.01	0.09	0.07	2.82	0.01

In the case of recursive methods, the computational time was low, since no NLP problem had to be solved and only the current measurements were considered. Overall, we can conclude that all of the estimation methods discussed in this paper were able to estimate the unknown fouling parameters either exactly or with only minor differences. This eventually resulted also in minor differences in the concentration and control trajectories. Further, based on the results, we can also conclude that the best method for on-line estimation is the EKF method. This is due to the fact that no NLP problem has to be solved at each sampling time, and the cumulative computational time is the lowest of all of the proposed methods. Moreover, compared to the RLS method, the EKF does not require different covariance matrices for the individual fouling models as discussed previously. Finally, the EKF method has a satisfactory convergence on the second interval, which leads to accurate information about the singular control.

The time-varying profiles of the unknown fouling parameters (K, n) are shown in Figure 3. We can observe that in the case of the WLS and WLSMH methods, the parameters converged quickly to a close neighborhood of the true values of the parameters ($K = 2$ and $n = 1.5$) prior to the switching time (vertical black dashed line) of the optimal control to the singular arc. The WLS method was able to estimate the true values of the parameters in the first few samples. However, in the case of WLSMH, the estimation needed more time, and we can observe that the parameter values were oscillating for the first two hours. This was due to the NLP solver falling into local minima. This can also be attributed to the chosen length of the moving horizon and the possibly imperfect choice of the initial covariance matrix P_0 . To overcome the issue of NLP solver falling into local minima, one can employ global optimizers. However, based on our observations, the local minima were only hit in the first samples of the estimation when only a few measured outputs were considered. When the recursive methods (RLS, EKF) were used, we can observe that the parameters converged almost to the true values of the parameters. The difference in the estimated and true values of the parameters is attributed to the strong nonlinearity of the process model. Moreover, the performance of these methods also strongly depends on the choice of the covariance matrices. Overall, we can conclude that all of the estimation methods were able to estimate the unknown fouling parameters or converge almost to true values of the parameters even before switching to the singular surface where singular control is applied. As a result, the theoretical optimality of the diafiltration process was almost restored when using the proposed estimators and coupling them within the feedback control law. Similar behavior for parameter estimation was also observed for other fouling models.

Figure 4 shows the concentration, flux and control trajectories for the ideal and estimated cases together with the considered measurements of concentrations and flux. The figures show only one of the estimation method (EKF) for $n = 1.5$ as the worst case scenario since in all other cases, the trajectories are closer to the ideal ones. In Figure 4a, we show the considered measured outputs (denoted by the circle and cross) and the ideal and estimated concentration trajectories. However, due to the small sampling time, we only display some of the measured points. The considered

measured flux with ideal and estimated parameters is shown in Figure 4b. Moreover, in Figure 4c,d, the optimal concentration state diagram and the optimal control profile are shown. In all figures, the solid lines represent the optimal scenario with a perfect knowledge of parameters. Further, we can observe that the estimated and ideal trajectories in the first step when the control is constant ($\alpha = 0$) are identical. This behavior was expected since the control in the first step does not depend on the unknown fouling parameters. The difference in the second interval is negligible. This was due to the convergence of the parameters during the first step. For this reason, the singular control in the second step is calculated with almost the true value of the parameters. The measured concentration of the macro-solute (blue dashed line in Figure 4a) shows the increasing difference to the ideal one at the singular step. However, this has only a minor impact on the overall optimal operation. Based on the results, we can conclude that by using the estimated parameters obtained by all estimation techniques, the overall operation is very close to optimal one with only minor differences in the processing and switching times. Moreover, it should be also mentioned that even if satisfactory convergence was obtained with the EKF method, different Kalman filter methods, like the unscented Kalman filter [40], could be also employed for the parameter estimation.

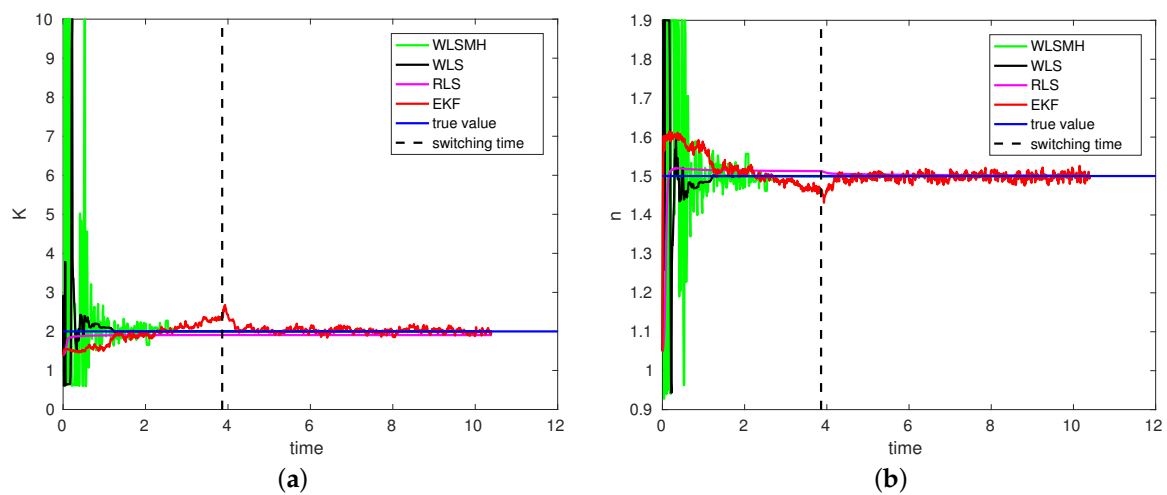


Figure 3. Estimation of the fouling parameters K, n for the three chosen cases together with the optimal switching time. (a) Estimation of the fouling parameter K ; (b) estimation of the fouling parameter n .

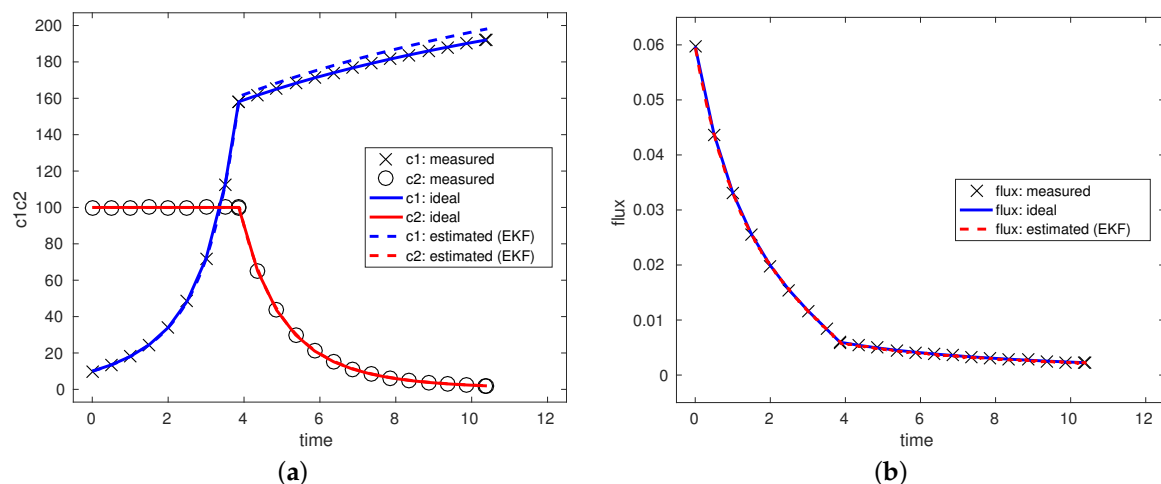


Figure 4. Cont.

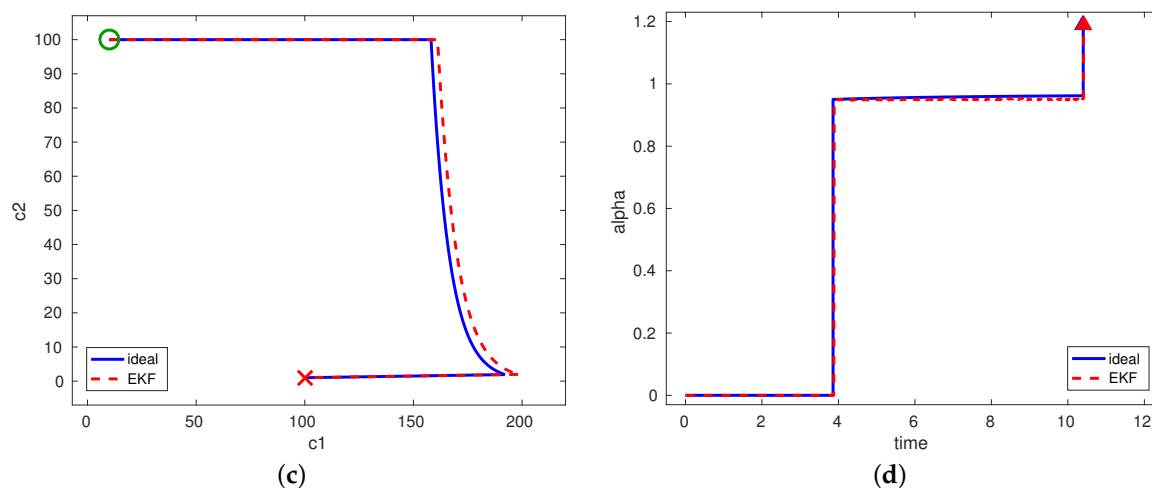


Figure 4. Concentration state diagram and optimal control profile for ideal and estimated fouling parameters ($K = 2$ and $n = 1.5$) by employing the EKF method. (a) Measured, ideal and estimated concentrations; (b) measured, ideal and estimated flux; (c) concentration state diagram; (d) optimal control profile.

Next, we study the case where the initial concentrations are $[c_{1,0}, c_{2,0}] = [130 \text{ mol/m}^3, 100 \text{ mol/m}^3]$, while the same final concentrations have to be met as in the previous cases. The optimal control strategy starts with pure dilution in the first step and in the last step, which follows after an operation on a singular arc. We use the same estimators as before, with the same tuning matrices.

In Figure 5, we show the results of the estimation of the unknown fouling parameters using the WLS method and the corresponding concentration state diagram and control profile of the on-line control strategy. As stated above (Section 4), the entry to the singular arc (50) can be determined without the knowledge of fouling parameters as $J = J_0$. On the singular arc, the estimation of the unknown fouling parameters commences. We can observe in Figure 5a,b that the unknown fouling parameters were estimated accurately in the first three hours of operation where the overall time-optimal operation was approximately 37 hours. The suboptimality of the on-line control strategy w.r.t. the ideal one is $\delta_{\text{tf}}^* = 0.025$. The oscillations in the first minutes of the operation were mainly caused by a small set of measured outputs. A similar behavior was observed in the previous case. Once the amount of measured outputs is increased, we can observe a quick convergence of the estimated parameters towards the true values. The comparison of ideal and estimated concentrations and control trajectories is shown in Figure 5c,d. As we can observe, the differences in the concentrations and control trajectories for ideal and estimated cases are negligible. This is due to the quick convergence of the estimated fouling parameters. We also performed simulations with the rest of the estimation methods.

In the case of the recursive methods (EKF and RLS), by using the same covariance matrices (51) to (54) as in the previous case, the estimated parameters diverged in the first hour of the batch and were not able to converge to the true values of the parameters. The same behavior was also observed in the case of the WLSMH method, since the covariance matrix P was updated based on the EKF method. The reasons for the divergence of the WLSMH, RLS and EKF methods was mainly caused due to the high sensitivity of the choice of the covariance matrices and the strong non-linearities of the process model. Other reasons for the possible divergence issues as discussed in Section 4.3 are the insufficient amount of measured outputs, since the estimation starts at the singular surface, and insignificant sensitivity between the measured outputs and the estimated parameters on the singular surface. This was caused by almost constant flux on the singular surface.

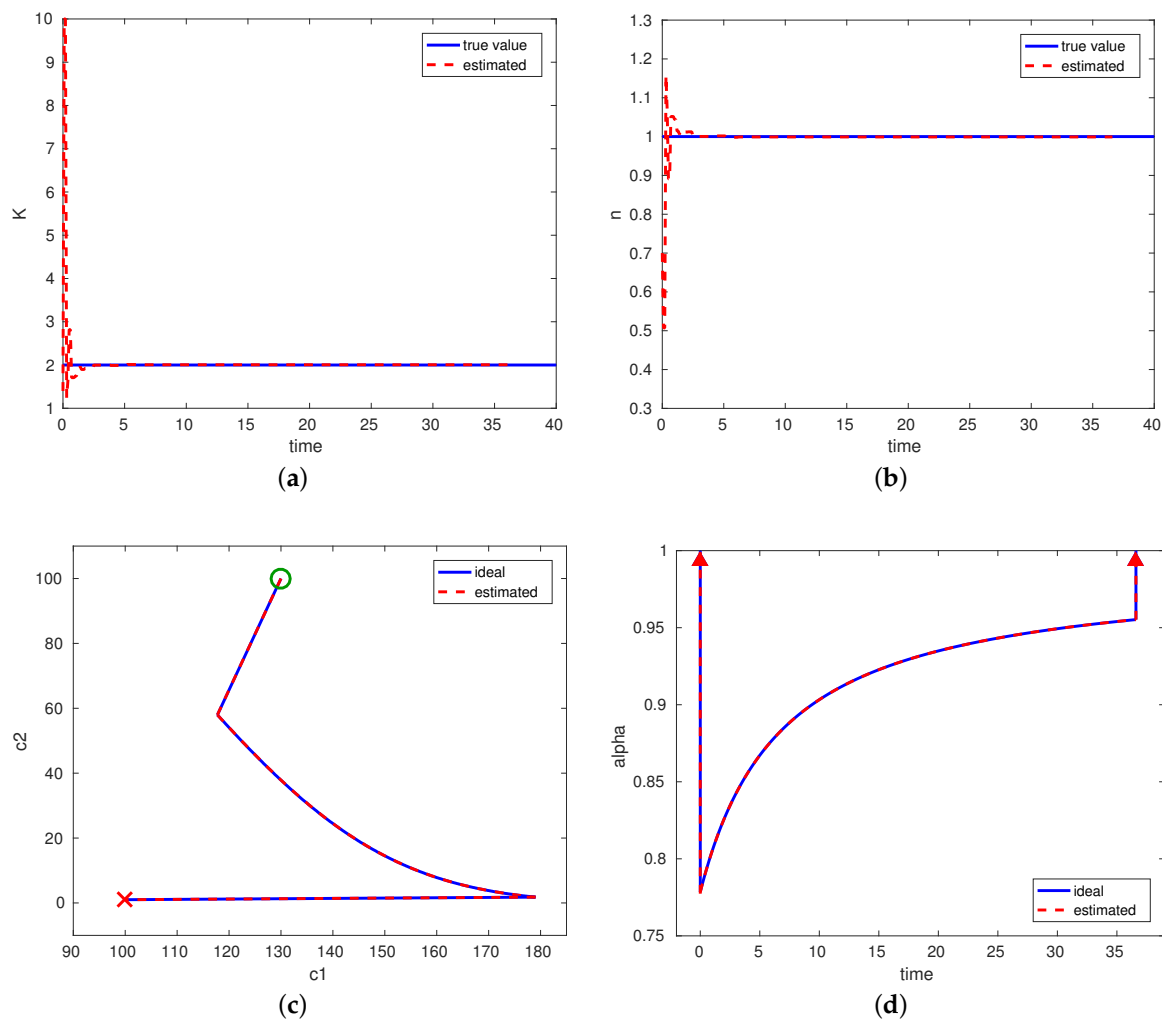


Figure 5. Estimation of fouling parameters, concentration state diagram and optimal control profile for ideal and estimated fouling parameters ($K = 2$ and $n = 1$) using the WLS estimation method. (a) Estimation of the fouling parameter K ; (b) estimation of the fouling parameter n ; (c) Concentration state diagram; (d) optimal control profile.

Overall, we can conclude that even if in the first step, pure dilution mode has to be applied and the estimation of unknown fouling parameters is only performed during the singular control, the differences between the ideal and estimated trajectories are negligible. The results indicate that the estimated parameters converged reasonably close to the true values of the parameters in the first hours of operation. For this reason, the overall operation is very close to the optimal one with only negligible differences in the case when the WLS method is applied. However, in the rest of the estimation methods, strong divergence of the estimated parameters was observed.

7. Conclusions

In this paper, we studied the time-optimal operation of a diafiltration process in the presence of membrane fouling. The optimal operation and control is an explicit non-linear control law defined over concentration regions. According to the theory, the optimal operation is defined as a three-step strategy with concentration or pure dilution mode in the first and the last step and with a singular arc in the middle step.

In the paper, we provided a detailed analysis on the problems related to the observation of the parameters based on the measured outputs. Based on the analysis, we conclude that by taking into

account the concentrations and flux as measured outputs, both parameters could not be estimated at the same time. Therefore, the derivative of flux w.r.t. to time had to be considered as a new measured outputs. Based on the analysis, four standard parameter estimation methods were discussed for the estimation of the unknown fouling parameters on-line during the separation. The crucial reason for estimating the parameters on-line is that the singular surface and control depend on the fouling parameters. The first two methods (WLS and WLSMH) were able to estimate the parameters exactly, whereas the recursive methods (RLS and EKF) converged almost to the true values of the parameters. However, the convergence was satisfactory and resulted in minor differences in the concentration and control profiles compared to the ideal case. Based on the results, we conclude that by using the EKF estimation method, we improve the convergence with the lowest computational time compared to other discussed methods. If the optimal operation starts with a pure dilution mode, only WLS and WLSMH were able to converge to the true values of the parameters, whereas the recursive methods (RLS and EKF) diverged completely from the true values. A possible remedy for the future could be to inject some dither signal into singular dilution rate to increase the excitation of the process.

All of the results presented in this paper are based on simulation results where the process noise was also considered in the simulations. Future research directions will be focused on experimental implementation of the proposed estimation techniques on a laboratory membrane process.

Acknowledgments: The authors gratefully acknowledge the contribution of the Scientific Grant Agency of the Slovak Republic under Grant 1/0053/13, the Slovak Research and Development Agency under the project APVV-15-0007 and also the international cooperation of the Slovak Research and Development Agency under the project APVV SK-FR-2015-0001. This contribution/publication is also the partial result of the Research & Development Operational Programme for the project University Scientific Park Slovak University of Technology in Bratislava, ITMS 26240220084, supported by the Research 7 Development Operational Programme funded by the ERDF.

Author Contributions: M.J. performed simulations and wrote most of the paper. D.P. performed identifiability tests and wrote the corresponding section. R.P. provided substantial background on the parameter estimation methods. A.L. and M.F. contributed to writing and correcting the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Paulen, R.; Fikar, M. *Optimal Operation of Batch Membrane Processes*; Springer: Cham, Switzerland, 2016; p. 158.
2. Nakajima, H. (Ed.) *Mass Transfer—Advances in Sustainable Energy and Environment Oriented Numerical Modeling*; InTech: Rijeka, Croatia, 2014; p. 522.
3. Hermia, J. Constant pressure blocking filtration laws-application to power-law non-Newtonian fluids. *Trans. Inst. Chem. Eng.* **1982**, *60*, 183–187.
4. Parvasi, P.; Mostafazadeh, A.K.; Rahimpour, M.R. Dynamic modeling and optimization of a novel methanol synthesis loop with hydrogen-permselective membrane reactor. *Int. J. Hydrog. Energy* **2009**, *34*, 29–37.
5. Bui, V.A.; Vu, L.T.T.; Nguyen, M.H. Simulation and optimisation of direct contact membrane distillation for energy efficiency. *Desalination* **2010**, *259*, 29–37.
6. Blankert, B.; Betlem, B.H.L.; Roffel, B. Dynamic optimization of a dead-end filtration trajectory: Blocking filtration laws. *J. Membr. Sci.* **2006**, *285*, 90–95.
7. Zondervan, E.; Roffel, B. Dynamic optimization of chemical cleaning in dead-end ultrafiltration. *J. Membr. Sci.* **2008**, *307*, 309–313.
8. Charfi, A.; Amar, N.B.; Harmand, J. Analysis of fouling mechanisms in anaerobic membrane bioreactors. *Water Res.* **2012**, *46*, 2637–2650.
9. Chis, O.; Banga, J.R.; Balsa-Canto, E. GenSSI: A software toolbox for structural identifiability analysis of biological models. *Bioinformatics* **2011**, *27*, 2610–2611.
10. Saccomani, M.P. Examples of testing global identifiability of biological and biomedical models with the DAISY software. *Comput. Methods Programs Biomed.* **2010**, *40*, 402–407.

11. Jelemenský, M.; Sharma, A.; Paulen, R.; Fikar, M. Time-optimal operation of diafiltration processes in the presence of fouling. In Proceedings of the 12th International Symposium on Process Systems Engineering and 25th European Symposium on Computer Aided Process Engineering, Copenhagen, Denmark, 31 May–4 June 2015; Gernaey, K.V., Huusom, J.K., Gani, R., Eds.; Elsevier: Copenhagen, Denmark, 2015; pp. 1577–1582.
12. Jelemenský, M.; Klaučo, M.; Paulen, R.; Lauwers, J.; Logist, F.; Van Impe, J.; Fikar, M. Time-optimal control and parameter estimation of diafiltration processes in the presence of membrane fouling. In Proceedings of the 11th IFAC Symposium on Dynamics and Control of Process Systems, Trondheim, Norway, 6–8 June 2016; Volume 11, pp. 242–247.
13. Kovács, Z.; Fikar, M.; Czermak, P. Mathematical modeling of diafiltration. In Proceedings of the International Conference of Chemical Engineering, San Francisco, CA, USA, 20–22 October 2009; Pannonia University: Veszprem, Hungary, 2009; pp. 135.
14. Baker, R.W. *Membrane Technology and Applications*, 3rd ed.; Wiley: Hoboken, NJ, USA, 2012.
15. Zhao, Y.; Wu, K.; Wang, Z.; Zhao, L.; Li, S.S. Fouling and cleaning of membrane—A literature review. *J. Environ. Sci.* **2000**, *12*, 241–251.
16. Bolton, G.; LaCasse, D.; Kuriyel, R. Combined models of membrane fouling: Development and application to microfiltration and ultrafiltration of biological fluids. *J. Membr. Sci.* **2006**, *277*, 75–84.
17. Vela, M.C.V.; Blanco, S.A.; García, J.L.; Rodríguez, E.B. Analysis of membrane pore blocking models applied to the ultrafiltration of PEG. *Sep. Purif. Technol.* **2008**, *62*, 489–498.
18. Goh, C.J.; Teo, K.L. Control parameterization: A unified approach to optimal control problems with general constraints. *Automatica* **1988**, *24*, 3–18.
19. Jelemenský, M.; Sharma, A.; Paulen, R.; Fikar, M. Time-optimal control of diafiltration processes in the presence of membrane fouling. *Comput. Chem. Eng.* **2016**, *91*, 343–351.
20. Pontryagin, L.S.; Boltyanskii, V.G.; Gamkrelidze, R.V.; Mishchenko, E.F. *The Mathematical Theory of Optimal Processes*; Wiley (Interscience): New York, NY, USA, 1962.
21. Bryson, A.E., Jr.; Ho, Y.C. *Applied Optimal Control*; Hemisphere Publishing Corporation: Washington, DC, USA, 1975.
22. Miao, H.; Xia, X.; Perelson, A.S.; Wu, H. On identifiability of nonlinear ODE models and applications in viral dynamics. *SIAM* **2011**, *53*, 3–39.
23. Godfrey, K.R.; Evans, N.D. Structural identifiability of nonlinear systems using linear/nonlinear splitting. *Int. J. Control* **2003**, *76*, 209–216.
24. Godfrey, K.R.; Evans, N.D. Identifiability of some nonlinear kinetics. In Proceedings of the 3rd Workshop on Modelling of Chemical Reaction Systems, Heidelberg, Germany, 24–26 July 1996.
25. Pohjanpalo, H. System identifiability based on the power series expansion of the solution. *Math. Biosci.* **1978**, *41*, 21–33.
26. Chis, O.; Banga, J.R.; Balsa-Canto, E. Structural identifiability of systems biology models: A critical comparison of methods. *PLoS ONE* **2011**, *6*, e27755.
27. DiStefano, J., III. *Dynamic Systems Biology Modeling and Simulation*; Academic Press: Cambridge, MA, USA, 2015; p. 884.
28. Chis, O.; Banga, J.R.; Balsa-Canto, E. *GenSSI: Generating Series Approach for Testing Structural Identifiability*; CSIC: Madrid, Spain, 2011.
29. Maiwald, T.; Timmer, J. Dynamical modeling and multi-experiment fitting with PottersWheel. *Bioinformatics* **2008**, *24*, 2037–2043.
30. Meshkat, N.; Kuo, C.; DiStefano, J., III. On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: A novel web implementation. *PLoS ONE* **2014**, *9*, e110261.
31. Bellu, G.; Saccomani, M.P.; Audoly, S.; D’Angio, L. DAISY: A new software tool to test global identifiability of biological and physiological systems. *Comput. Methods Programs Biomed.* **2007**, *88*, 52–61.
32. Galrinho, M.; Rojas, C.; Hjalmarsson, H. A weighted least-squares method for parameter estimation in structured models. In Proceedings of the 53rd IEEE Conference on Decision and Control, Los Angeles, CA, USA, 15–17 December 2014; pp. 3322–3327.
33. Gulán, M.; Salaj, M.; Rohal’-Ilkiv, B. Nonlinear model predictive control with moving horizon estimation of a pendubot system. In Proceedings of the 2015 International Conference on Process Control (PC), Strbske Pleso, Slovakia, 9–12 June 2015; pp. 226–231.

34. Kühn, P.; Diehl, M.; Kraus, T.; Schlöder, J.P.; Bock, H.G. A real-time algorithm for moving horizon state and parameter estimation. *Comput. Chem. Eng.* **2011**, *35*, 71–83.
35. Haverbeke, N. Efficient Numerical Methods for Moving Horizon Estimation. Ph.D. Thesis, Katholieke Universiteit Leuven, Leuven, Belgium, 2011.
36. Ding, F. Coupled-least-squares identification for multivariable systems. *IET Control Theory Appl.* **2013**, *7*, 68–79.
37. Ljung, L. *System Identification: Theory for the User*, 2nd ed.; Prentice Hall PTR: Upper Saddle River, NJ, USA, 1999.
38. Kalman, R.E. A new approach to linear filtering and prediction problems. *Trans. ASME J. Basic Eng.* **1960**, *82*, 35–45.
39. Bavdekar, V.A.; Deshpande, A.P.; Patwardhan, S.C. Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter. *J. Process Control* **2011**, *21*, 585–601.
40. Wan, E.A.; van der Merwe, R. The unscented Kalman filter for nonlinear estimation. In Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium, Lake Louise, AB, Canada, 1–4 October 2000.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).