

## Article

# A Novel Approach for Prediction of Industrial Catalyst Deactivation Using Soft Sensor Modeling

Hamed Gharehbaghi and Jafar Sadeghi \*

Department of Chemical Engineering, University of Sistan and Baluchestan, P.O. Box 98164-161, Zahedan, Iran; hamed.gharehbaghi@gmail.com

\* Correspondence: sadeghi@eng.usb.ac.ir; Tel.: +98-915-549-4265

Academic Editor: Keith Hohn

Received: 13 January 2016; Accepted: 26 May 2016; Published: 23 June 2016

**Abstract:** Soft sensors are used for fault detection and prediction of the process variables in chemical processing units, for which the online measurement is difficult. The present study addresses soft sensor design and identification for deactivation of zeolite catalyst in an industrial-scale fixed bed reactor based on the process data. The two main reactions are disproportionation (DP) and transalkylation (TA), which change toluene and C<sub>9</sub> aromatics into xylenes and benzene. Two models are considered based on the mass conservation around the reactor. The model parameters are estimated by data-based modeling (DBM) philosophy and state dependent parameter (SDP) method. In the SDP method, the parameters are assumed to be a function of the system states. The results show that the catalyst activity during the period under study has approximately a monotonic trend. Identification of the system clearly shows that the xylene concentration has a determining role in the conversion of reactions. The activation energies for both DP and TA reactions are found to be 43.8 and 18 kJ/mol, respectively. The model prediction is in good agreement with the observed industrial data.

**Keywords:** soft sensor; catalyst deactivation; data-based modeling; state-dependent-parameter

## 1. Introduction

Due to large-scale application of catalysts in the petroleum and petrochemical processes, the main problem concerning these processes is the deactivation of catalyst. As the catalyst activity is related to the availability of catalyst active area for the reactants, deactivation refers to reducing the activated area of the catalyst or blocking the path of the moving reactants and the products. Mathematical modeling of catalyst deactivation is the main concern of these processes during operation. One of these processes with high potential of deactivation is the disproportionation (DP) of toluene and/or transalkylation (TA) of C<sub>9</sub> aromatics and toluene to C<sub>8</sub> aromatics over zeolites in fixed-bed reactors. DP and TA involve solid-acid catalytic processes which occur over medium-pore zeolites, such as ZSM-5, and large-pore zeolites, such as mordenite, Beta, and Y [1–4]. For H-mordenite, the temperature is a dominant factor that determines the activity and coking tendency, as the undesired reaction, due to its one-dimensional channel system [5].

The major problem for the commercialization of solid acid catalyzed processes is the rapid deactivation of the catalyst [4,6,7]. Therefore, modeling of the deactivation of the catalyst is a major problem that must be taken into account for the success of such processes. Some studies have already reported the deactivation behavior of the solid acid catalysts quantitatively [8–11]. These studies have focused on the solution of the set of differential and algebraic equations that describe catalyst deactivation in various reactor systems by experimental data. In these models, however, many simplified conditions have been assumed. Most of these models have been conducted at the laboratory scale under certain conditions with applying no dynamic data [1–7,12]. Additionally, online investigation of the industrial reactors' parameters (e.g., catalyst deactivation, conversion, product

quality, etc.) can be of considerable significance in the identification of these parameters. In the last two decades, researchers began to make use of the vast amounts of data being measured and collected in the industrial processes by constructing predictive models based on such data. In the context of the industrial processes, these predictive models are called soft sensors [13,14]. In chemical plants, soft sensors are widely used to estimate a process variable which is difficult to be measured online. In general, one can broadly classify soft sensors into two types, namely, model-based (white-box) and data-based (black-box). White-box models are dependent on a prior theoretical knowledge and thus often unavailable since industrial processes are commonly too complicated to be analyzed theoretically. These models are developed primarily for the planning and the design of the processing plants and, therefore, usually focus on the description of the ideal steady states of the processes. As a solution, the data-based soft sensors gained increasing popularity in the industrial processes. A wide variety of statistical inference techniques and machine learning techniques have been employed in data-based soft sensors, among which representative examples are principal component regression (PCR), partial least squares (PLS) regression, support vector machines (SVM), and artificial neural networks (ANN) [13–22].

Reactors, by their nature, are non-stationary because their complexity cannot be fully described using constant-parameter models. The modeling of such systems can be represented by a two-step procedure. First, the dynamic behavior of the system is described by a deterministic mathematical model. Second, the model is simplified to utilize in conventional control theories. The first assumption causes errors in modeling, while the second suffers from the complexity and dimensionality problem. Moreover, digitization in order to use computer facilities for application of the continuous control theories provides another source of errors in modeling and control strategies [23,24]. Hence, soft sensor modeling of these systems requires the use of powerful statistical inference techniques.

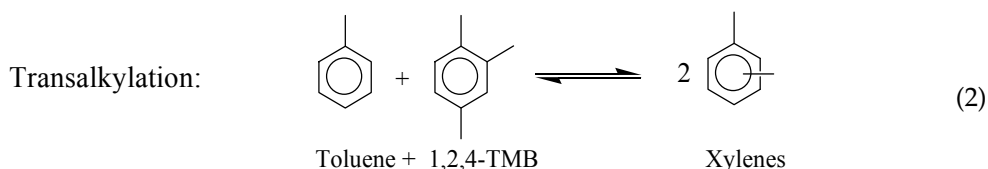
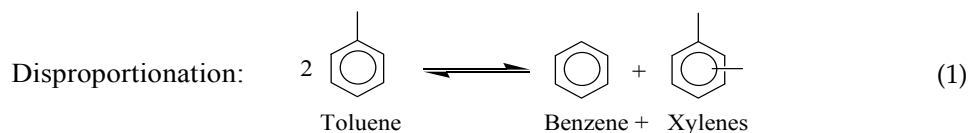
Data-based mechanistic (DBM) modeling philosophy, as introduced by Young [25], is an alternative to complicated modeling strategies. The major advantage of this method is balancing between order of the model (complexity) and the model efficiency while, at the same time, good explanation of data and statistical analysis is required for identification and estimation of such systems [25–30]. Since the resulting model is simple and directly in digital form, it avoids both the model complication and errors produced through either simplification or digitization of a complex continuous-time model. It is worth mentioning that the resulting model and its parameters are, according to DBM philosophy, meaningful and interpretable while in data-based modeling, which is not mechanistic, this is generally not the case. In this philosophy, the non-stationary aspects of the system are reflected by a simple structure-varying parameter model, whether its parameters are varying with time or as a function of some states of the system. This model is called a state-dependent parameter (SDP) model [26,30]. In this case, all nonlinearities of the model can be estimated in model parameters, while the structure of the model remains linear and simple. In other words, when a system is described by a nonlinear model, such as  $y_t = \alpha y_{t-1} + \beta y_{t-1}^2$ , the SDP model is written as  $y_t = a_t y_{t-1}$  which has a linear structure but its parameter is  $a_t = \alpha + \beta y_{t-1}$ . This model has the complexity of a nonlinear model while using a linear structure. The parameters are also identified in a non-parametric manner; i.e., when this approach is used, the functionality of the identified parameter can be seen in a graph rather than a mathematical relation.

In this study, a new SDP method for identification and modeling of the kinetics and catalyst activity has been proposed. The reactor modeling and identification have been considered for an industrial TA reactor using the input and output dynamic data. The obtained soft sensor model has a simple structure and appropriate for control purposes. This model also presents a good identification based on effective parameters of the catalyst activity.

## 2. Kinetic Model

In the present study, the process data of the fixed bed TA reactor of Bou-Ali-Sina Petrochemical Complex, Mahshahr, Iran, have been collected and analyzed daily for two years of reactor's operation.

In the TA reactor, benzene and xylenes are produced from toluene and C<sub>9</sub> aromatics through the two main reactions, DP (Equation (1)) and TA (Equation (2)). The DP produces benzene and xylenes from toluene, while the TA forms xylenes from toluene and C<sub>9</sub> aromatics:



Different kinetic models have been proposed to describe the catalytic toluene disproportionation and C<sub>9</sub> aromatics transalkylation. Table 1 summarizes the kinetic and deactivation models proposed in the literature, as well as the experimental conditions.

**Table 1.** Kinetic studies of toluene DP and C<sub>9</sub> aromatics TA over mordenite zeolite catalyst.

| Reaction | P(bar) | T(°C)   | E (kJ/mol) | Kinetic Models <sup>a</sup> | Deactivation Function                                              | Ref. |
|----------|--------|---------|------------|-----------------------------|--------------------------------------------------------------------|------|
| DP       | 1      | 350–450 | 60.7       | SO                          | $a = a_0 e^{-0.003t}$                                              | [6]  |
| DP       | 1      | 350–450 | 51.9       | RSO                         | $-\frac{da}{dt} = \frac{k_d P_T^2}{(1 + K_B P_B + K_T P_T)^2} a^2$ | [31] |
| DP       | 28.59  | 372–352 | 104.6      | RSO                         | NA                                                                 | [5]  |
| DP       | 31.2   | 347–352 | 102        | RSO                         | $a = \exp(-0.0006.t) WHSV^{-1.3781}$                               | [32] |
| TA       | 31.2   | 347–352 | 96.77      | RSO                         | $a = \exp(-0.0006.t) WHSV^{-1.3781}$                               | [32] |
| DP       | NA     | 300–400 | 54.22      | RSO                         | $a = e^{-a_0 t}$                                                   | [1]  |
| TA       | NA     | 300–400 | 7.21       | RSO                         | $a = e^{-a_0 t}$                                                   | [1]  |
| DP       | 5–32   | 360–431 | 27.17      | RSO                         | NA                                                                 | [12] |
| TA       | 5–32   | 360–431 | 151        | RSO                         | NA                                                                 | [12] |
| TA       | NA     | 300–400 | 17.84      | RSO                         | $a = e^{-a_0 C_{Ar}}$                                              | [33] |
| DP       | NA     | 300–400 | 50.4       | RSO                         | $a = e^{-a_0 t}$                                                   | [34] |

<sup>a</sup> SO: Second Order; RSO: Reversible Second Order.

The reactor contains H-mordenite zeolite catalyst, which is manufactured by Sinopec Tech Company. Its characteristics are as follows (Table 2).

**Table 2.** Characteristics of mordenite zeolite catalyst.

| Appearance       | White, Cylindrical                                |
|------------------|---------------------------------------------------|
| Diameter/Length  | 1.6–2.0 mm/3–20 mm                                |
| Bulk density     | 0.72 ± 0.05 kg/L                                  |
| Crush strength   | >100 N/cm                                         |
| Main ingredients | mordenite and alumina                             |
| Main components  | SiO <sub>2</sub> , Al <sub>2</sub> O <sub>3</sub> |

### 3. Identification Methodology

DBM modeling which was initially acquired from the analysis of observational time-series data has only been considered reliable if it can be interpreted in physically meaningful terms. This model is identified from a general and widely-used class of linear/nonlinear models. In this method, the non-stationary aspects (data with variable mean and/or variance) of the system are shown by a simple-structure varying-parameter model, either its parameters are varying with time or as a function of some states of the system. One example of these models is the SDP model [25,26,35].

### 3.1. State Dependent Parameter (SDP) Model

The SDP models are representative of nonlinear systems which have similar structure to the linear time-invariant (LTI) systems, but their parameters are changing as a function of the state variables to accommodate the system's non-linearity [36].

Young [27] introduced the estimation of 'slowly changing' time-variable parameters (TVPs) in the various kinds of linear regression models. One of these TVP models is the Dynamic Auto-Regressive with eXogenous (DARX) variables model (the observed output is dependent on its past values and the past values of exogenous inputs). The DARX model can describe a real dynamic system and the output is a complex response. The model parameters are changing with time, when they are functions of the system variables. The results of the model can behave in a heavily nonlinear or even chaotic manner. Young [25,36] proposed an SDP model of this type as a state dependent parameter DARX (SDARX) model. The SDARX model relates a single input variable  $u_t$  to an output variable  $y_t$  that can be written conveniently as:

$$y_t = \mathbf{z}_t^T \mathbf{p}_t + e_t \quad (3)$$

More generally, it can be expressed as follows:

$$y_t = \sum_{i=1}^{n+m+1} p_i(\mathbf{x}_{1,t}^{(i)}, t) z_{i,t} + e_t \text{ with } \begin{cases} e_t = N(0, \sigma^2) \\ \mathbf{x}_{1,t}^{(i)} = [x_{1,t}^{(i)} \ x_{2,t}^{(i)} \ \dots \ x_{ns_i,t}^{(i)}] \end{cases} \quad (4)$$

$p_i(\cdot)$  is the  $i$ th state dependent parameter ( $p_i = \begin{cases} -a_i & i \leq n \\ b_{i-(m+1)} & i > n \end{cases}$ ),  $y_t$  and  $x_{j,t}^{(i)}$  are the observed output and the  $j$ th state corresponds to  $p_i(\cdot)$  at  $t$ , respectively. In addition,  $n + m + 1$  is the number of model parameters,  $ns_i$  is the number of states in which  $p_i(\cdot)$  depends on,  $z_{i,t}$  is the  $i$ th regressor ( $z_{i,t} = \begin{cases} y_{t-i} & i \leq n \\ u_{t-\delta+n+1-i} & i > n \end{cases}$ ), and  $e_t$  is a zero-mean white-noise input with Gaussian normal amplitude distribution and variance  $\sigma^2$ .

The TA reactor is described by mass balances for benzene and C<sub>9</sub> aromatics [37]:

$$\frac{\partial C_B}{\partial t} - D_A \frac{\partial^2 C_B}{\partial l^2} - D_R \left( \frac{\partial^2 C_B}{\partial r^2} + \frac{1}{r} \frac{\partial C_B}{\partial r} \right) + \frac{Q}{S} \frac{\partial C_B}{\partial l} - \Re'_B \rho_x = 0 \quad (5)$$

$$\frac{\partial C_{Ar}}{\partial t} - D_A \frac{\partial^2 C_{Ar}}{\partial l^2} - D_R \left( \frac{\partial^2 C_{Ar}}{\partial r^2} + \frac{1}{r} \frac{\partial C_{Ar}}{\partial r} \right) + \frac{Q}{S} \frac{\partial C_{Ar}}{\partial l} - \Re'_{Ar} \rho_x = 0 \quad (6)$$

Since the reactor has a large volume, axial and radial diffusion terms can be neglected and assumed to be a part of the internal system error. Additionally, since the data are collected daily, the system is supposed to be in steady-state condition; therefore, the accumulation term is ignored and Equations (5) and (6) can be rewritten by following simple forms:

$$\frac{Q}{S} \frac{\partial C_B}{\partial l} - \Re'_B \rho_x = \xi \quad (7)$$

$$\frac{Q}{S} \frac{\partial C_{Ar}}{\partial l} - \Re'_{Ar} \rho_x = \xi \quad (8)$$

where  $\xi$  is the error. Discretization of Equations (7) and (8) in backward difference leads to:

$$\frac{Q}{S} \frac{(C_{B,L} - C_{B,0})_{t-1}}{\Delta l} = \Re'_B \rho_x + \xi \quad (9)$$

$$\frac{Q}{S} \frac{(C_{Ar,L} - C_{Ar,0})_{t-1}}{\Delta l} = \Re'_{Ar} \rho_x + \xi \quad (10)$$

Since the terms related to diffusion (adsorption term) are considered as a part of catalyst deactivation, the reaction rate is considered to be power function of the concentration of the reaction components, temperature-dependent specific reaction rate, and time-dependent deactivation rate as follows:

$$\mathcal{R}'_B = C_{Tol}^{\lambda_1} \cdot C_B^{\lambda_2} \cdot C_X^{\lambda_3} \cdot k_0 \cdot \exp\left(\frac{-E_1}{RT}\right) \cdot \exp(-a_1 \cdot t) \quad (11)$$

$$\mathcal{R}'_{Ar} = C_{Tol}^{\mu_1} \cdot C_{Ar}^{\mu_2} \cdot C_X^{\mu_3} \cdot k_0 \cdot \exp\left(\frac{-E_2}{RT}\right) \cdot \exp(-a_2 \cdot t) \quad (12)$$

Substitution of Equations (11) and (12) into Equations (9) and (10) and rearrangement gives:

$$(C_{B,L} - C_{B,0}) = \frac{S \cdot L}{Q} k_{01} C_{Tol}^{\lambda_1} C_B^{\lambda_2} C_X^{\lambda_3} \exp\left(\frac{-E_1}{RT}\right) \exp(-a_1 t) \quad (13)$$

$$(C_{Ar,L} - C_{Ar,0}) = \frac{S \cdot L}{Q} k_{02} C_{Tol}^{\mu_1} C_{Ar}^{\mu_2} C_X^{\mu_3} \exp\left(\frac{-E_2}{RT}\right) \exp(-a_2 t) \quad (14)$$

Taking the logarithm of both sides of Equation (10) yields:

$$\log(C_{B,L} - C_{B,0}) = \lambda_1 \log(C_{Tol}) + \lambda_2 \log(C_B) + \lambda_3 \log(C_X) + \log\left(\frac{S \cdot L}{Q} k_{01} \rho_x\right) + \frac{-E_1}{R} \frac{1}{T} + (-a_1 t) \quad (15)$$

$$\log(C_{Ar,L} - C_{Ar,0}) = \mu_1 \log(C_{Tol}) + \mu_2 \log(C_{Ar}) + \mu_3 \log(C_X) + \log\left(\frac{S \cdot L}{Q} k_{02} \rho_x\right) + \frac{-E_2}{R} \frac{1}{T} + (-a_2 t) \quad (16)$$

Having time-series concentration data of the process, the identification method of Equations (15) and (16) will be discussed in the following section.

It should be noted that in Equations (15) and (16) the terms of the reaction kinetics have been considered based on the following assumptions:

- Reaction specific rate is a function of temperature according to the Arrhenius equation;
- Due to reversible reactions, the concentration of the reaction components is considered by a power law expression (since all the adsorption terms are included in the deactivation term); and
- Catalyst deactivation is a function of time and the concentrations of some other components. This function is assumed to be exponential (see Table 1).

### 3.2. Time-Varying Parameter (TVP) Estimation

SDP estimation is a form of TVP estimation. Normally, when little is known about the nature of the time variability, the model needs to be both simple and flexible. One of the simplest and most-useful class of stochastic, state-space models involves the assumption that the  $i$ th parameter ( $p_i$ ) is defined by a two-dimensional stochastic state vector  $\mathbf{x}_{i,t} = [l_{i,t} d_{i,t}]^T$ , where  $l_{i,t}$  and  $d_{i,t}$  are the changing level and slope of the associated TVP, respectively. The recursive SDP estimation algorithms are an extension of the stochastic approach to TVP estimation, where the time variations in the parameters are assumed to evolve as one of the generalized random walk (GRW) class of the Gauss-Markov process [25,27]. Early research in the 1960s [38,39] tended to use a simple scalar random walk model for the parameter variations when the parameters are supposed to change the TVP models. The stochastic evaluation of each parameter is described by a GRW process [40]:

$$\mathbf{x}_{i,t} = \mathbf{F}_i \mathbf{x}_{i,t-1} + \mathbf{G}_i \boldsymbol{\eta}_{i,t} \quad i = 1, 2, \dots, m + n + 1 \quad (17)$$

where:

$$\mathbf{F}_i = \begin{bmatrix} \alpha & \beta \\ \gamma & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{G}_i = \begin{bmatrix} \delta & 0 \\ 0 & \varepsilon \end{bmatrix} \quad (18)$$

$\boldsymbol{\eta}_{i,t} = [\eta_{1i,t} \ \eta_{2i,t}]^T$  is a  $2 \times 1$ , zero mean, white-noise vector that allows for stochastic variability in the parameters and is assumed to be characterized by a covariance diagonal matrix  $\mathbf{Q}_{\eta,t}$  that each diagonal element  $\mathbf{Q}_i$  corresponds to  $\boldsymbol{\eta}_{i,t}$ . The various, normally constant, parameters in the GRW model ( $\alpha, \beta, \gamma, \delta, \varepsilon$ , and the elements of  $\mathbf{Q}_{\eta,t}$ ) are often referred to as hyper-parameters. This is to differentiate them from the TVPs that are the main object of the estimation analysis. Having introduced the GRW models for the parameter variations, an overall state space model can be constructed through the aggregation of the subsystem matrices defined by Equation (17), with the ‘observation’ equation defined by Equation (3):

$$\left. \begin{array}{l} \text{Observation Equation} \quad y_t = \mathbf{H}_t \mathbf{x}_t + e_t \\ \text{State Equation} \quad \mathbf{x}_t = \mathbf{F} \mathbf{x}_{t-1} + \mathbf{G} \boldsymbol{\eta}_t \end{array} \right\} \quad (19)$$

where  $\mathbf{x}_t = [\mathbf{x}_{1,t}^T \ \mathbf{x}_{2,t}^T \ \dots \ \mathbf{x}_{n+m+1,t}^T]^T$  if  $p = 2(n+m+1)$ ,  $\mathbf{F}$  is a  $p \times p$  diagonal matrix with blocks defined by the  $\mathbf{F}_i$  matrices in Equation (17),  $\mathbf{G}$  is a  $p \times p$  block diagonal matrix with blocks defined by the corresponding subsystem matrices  $\mathbf{G}_i$  in Equation (17), and  $\boldsymbol{\eta}_t$  is a  $(n+m+1) \times 1$  vector containing the white-noise inputs which is assumed to be independent of the observation noise  $e_t$  and has a covariance diagonal  $(n+m+1) \times (n+m+1)$  matrix  $\mathbf{Q}$  that each diagonal element is  $\mathbf{Q}_i$  corresponding to  $\boldsymbol{\eta}_{i,t}$ . Additionally,  $\mathbf{H}_t$  is a row vector of the following form:

$$\mathbf{H}_t = [y_{t-1} \ 0_1 \ y_{t-2} \ 0_2 \ \dots \ y_{t-n} \ 0_n \ u_{t-\delta} \ 0_{n+1} \ \dots \ u_{t-\delta-m} \ 0_{n+m+1}] \quad (20)$$

where  $0_i, i = 1, 2, \dots, n+m+1$  are defined as  $1 \times q_i$  vectors of zero and the term  $\delta$  is a pure time delay, measured in sampling intervals. Now the relation between scalar observation  $y_t$  and the state variables defined by Equation (17) can be defined by Equation (19), which represents the model Equation (17), where each parameter is defined as a GRW process. The state space formulation of Equation (17) is particularly well-suited for optimal recursive estimation in which the time variable parameters are estimated sequentially while working through the data in temporal order through two separate passes, “forward pass filtering” and “backward pass smoothing” [25,41].

Forward Pass Recursive LS equations:

$$\text{Prediction : } \begin{array}{l} \hat{\mathbf{x}}_{t|t-1} = \mathbf{F} \hat{\mathbf{x}}_{t-1} \\ \mathbf{P}_{t|t-1} = \mathbf{F} \mathbf{P}_{t-1} \mathbf{F}^T + \mathbf{G} \mathbf{Q}_r \mathbf{G}^T \end{array} \quad (21)$$

The estimation is based on all the times before  $t$ . To include time  $t$  the correction is needed:

$$\text{Correction : } \begin{array}{l} \hat{\mathbf{x}}_t = \hat{\mathbf{x}}_{t|t-1} + \mathbf{P}_{t|t-1} \mathbf{H}_t^T \left[ 1 + \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^T \right]^{-1} \left( y_t - \mathbf{H}_t \hat{\mathbf{x}}_{t|t-1} \right) \\ \mathbf{P}_t = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} \mathbf{H}_t^T \left[ 1 + \mathbf{H}_t \mathbf{P}_{t|t-1} \mathbf{H}_t^T \right]^{-1} \mathbf{H}_t \mathbf{P}_{t|t-1} \end{array} \quad (22)$$

where  $t = 1, 2, \dots, N$  and  $\hat{\mathbf{x}}_0$  and  $\mathbf{P}_0$  are the initial value and its covariance matrix, respectively.

### 3.3. Backward Pass Smoothing Equations

Estimated values that are obtained from the forward-pass filtering algorithm are based on the initial conditions,  $\hat{\mathbf{x}}_0$  and  $\mathbf{P}_0$ , and the values are prior to and include sample  $t$ . Therefore, they need to be updated sequentially, while working through the data, in reverse temporal order (backward pass smoothing) using a backward-recursive procedure termed fixed interval smoothing (FIS) algorithm,

Equation (21), to obtain a smooth estimate based on all the elements of the time series and to remove any lag effect.

$$\begin{aligned}\hat{\mathbf{x}}_{t|N} &= \mathbf{F}^{-1} \left[ \hat{\mathbf{x}}_{t+1|N} + \mathbf{G}\mathbf{Q}_r\mathbf{G}^T\mathbf{L}_t \right] \\ \mathbf{L}_t &= \begin{cases} [\mathbf{I} - \mathbf{P}_{t+1}\mathbf{H}_{t+1}^T\mathbf{H}_{t+1}]^T \times \{\mathbf{F}^T\mathbf{L}_{t+1} - \mathbf{H}_{t+1}^T(y_{t+1} - \mathbf{H}_{t+1}\hat{\mathbf{x}}_{t+1})\} & t < N \\ 0 & t = N \end{cases} \\ \mathbf{P}_{t|N} &= \begin{cases} \hat{\mathbf{P}}_t + \hat{\mathbf{P}}_t\mathbf{F}^T\mathbf{P}_{t+1|t}^{-1} [\mathbf{P}_{t+1|N} - \mathbf{P}_{t+1|t}] \mathbf{P}_{t+1|t}^{-1}\mathbf{F}^T\hat{\mathbf{P}}_t & t < N \\ \hat{\mathbf{P}}_t & t = N \end{cases}\end{aligned}\quad (23)$$

In these algorithms the  $p \times p$  noise variance ratio (NVR) matrix,  $\mathbf{Q}_r$ , and the  $p \times p$  matrix  $\hat{\mathbf{P}}_t$  are defined as:

$$\mathbf{Q}_r = \frac{\mathbf{Q}}{\sigma^2} \quad \text{and} \quad \hat{\mathbf{P}}_t = \frac{\mathbf{P}_t^*}{\sigma^2} \quad (24)$$

$\mathbf{P}_t^*$  is the error covariance matrix associated with the state estimates which defines the estimated uncertainty in the parameters. For convenience, it is usually assumed that the NVR matrix  $\mathbf{Q}_r$  is diagonal, although it is not necessary. The NVR parameters that characterize  $\mathbf{Q}_r$  are unknown prior to the analysis and clearly need to be estimated based on of the time-series data  $y_t$  and  $u_t$  before the filtering and smoothing algorithms can be utilized. The usual way of dealing with this estimation problem is to formulate it in maximum likelihood (ML) methods of estimation based on Gaussian assumptions for the stochastic disturbances and the application of prediction error decomposition [25,26,39,41]:

$$e_t = y_t - \mathbf{H}_t\hat{\mathbf{x}}_{t|t-1} \quad t = 1, 2, \dots, N \quad (25)$$

It is here assumed that the stochastic white noise inputs to the system have a normal distribution. Now, if  $p$  is the number of TVPs being estimated and the first  $p$  observations are regarded as fixed, then the log-likelihood function of the series  $y_{p+1}, \dots, y_N$  can be defined as follows in terms of the prediction error decomposition:

$$\log L = \frac{-(N-p)}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2) - \frac{1}{2} \sum_{t=p+1}^N \log(1 + \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T) - \frac{1}{2\sigma^2} \sum_{t=p+1}^N \frac{e_t^2}{1 + \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T} \quad (26)$$

It can be shown that  $\sigma^2(1 + \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T)$  is the variance of  $e_t$ , so the last term in Equation (26) is based on the sum of the squares of the normalized one-step-ahead prediction errors. Now, the ML estimate of  $\sigma^2$ , conditional on the hyper-parameters, is given by:

$$\hat{\sigma}^2 = \frac{1}{N-p} \sum_{t=p+1}^N \frac{e_t^2}{1 + \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T} \quad (27)$$

By substituting Equation (27) into Equation(26) and removing the constant term, the ‘concentrated likelihood’ can be expressed as:

$$\log(L_c) = -\frac{1}{2} \sum_{t=p+1}^N \log(1 + \mathbf{H}_t\mathbf{P}_{t|t-1}\mathbf{H}_t^T) + \frac{N-p}{2} \log(\hat{\sigma}^2) \quad (28)$$

which can then be minimized if it is multiplied by  $-1$ . This minimization is accomplished by initiating the optimization with the hyper-parameter estimates either selected by the user or set to some default values. Since  $p_t$  is potentially state-dependent, it cannot be assumed that the simple GRW model Equation (17) is appropriate to describe the parametric variation over time, which could be very fast. One obvious requirement with this new approach is that the sorting of data, prior to FIS estimation, must be common to all of the variables in Equation (3). The simplest strategy is to sort according to the ascending order of the ‘dependent’ variable  $y_t$ . In back-fitting algorithm, each parameter is estimated,

in turn, based on the ‘modified dependent variable’ (MDV) series obtained by subtracting all the other terms on the right-hand side of Equation (3) from  $y_t$ , using the values of the other parameter estimates from the previous iteration. At each such back-fitting iteration, the sorting can then be based on the single variable associated with the current SDP being estimated.

### 3.4. Back-Fitting Algorithm for SDP Models

1. Assume FIS estimation has yielded prior TVP estimates  $\hat{p}_{i,t|N}^0$ ,  $i = 1, 2, \dots, m + n + 1$  of the SDPs.
2. Iterate:  $i = 1, 2, \dots, m + n + 1$ ;  $k = 1, 2, \dots, k_c$
3. Form the MDV:  $y_t^i = y_t - \sum_{j \neq i} z_{i,t} \cdot \hat{p}_{j,t|N}^k$ .
4. Sort both  $y_t^i$  and  $z_{i,t}$  according to the ascending order of  $z_{i,t}$ .
5. Obtain an FIS estimate  $\hat{p}_{i,t|N}^k$  of  $p_{i,t}$  in the MDV relationship  $y_t^i = p_{i,t} z_{i,t}$ .
6. Repeat steps 2–5 until convergence criterion occurs.

The smoothing hyper parameters required for FIS estimation at each stage are optimized by ML.

Sadeghi et al. [42,43] extended this approach to accommodate the multidimensionality of parameters vs. its states. In this modeling approach, the model is optimized and is regarded as non-parametric by ML.

## 4. Results and Discussion

To verify the effectiveness of the SDP modeling, data series of fixed bed industrial TA reactor of the Bou-Ali-Sina Petrochemical Complex are analyzed. A record of 623 data as daily means for two years of reactor operation are used. Overall process data for the input and output concentrations of toluene, C<sub>9</sub> aromatics, olefin, hydrogen, benzene, xylene, and temperature  $T$  versus time are shown in Figure 1.

As the figure shows, the data are recorded with high input and output oscillation, without interfering with the working conditions and the dynamics form.

By designing the soft sensor, it is expected that while all of the states or variables are changing simultaneously, the system identification is performed. Moreover, model parameters are expected to be functions of the system state variables. When a complex nonlinear model with constant parameters is replaced by a linear structure with varying parameters, it is clear that the model with linear variations of the state variables gives a better control. The following assumptions are considered:

- Modeling is based on the continuity equation;
- Grey-box model is considered;
- Data are registered as daily average, that is, for each 24 h one data is recorded;
- Since the residence time of the reactor is half an hour and the data are recorded per day, the system is considered stable and the term for the accumulation in the continuity equation is vanishes;
- Due to the high volumetric flow rate of the reactor, the terms for the axial and radial diffusions are considered as part of the system error and hence removed; and
- The hydrogen molar percentage in the feed is about 90%. Therefore, the internal porous mass transfer resistance is disregarded.

Using the SDP modeling and through the investigation of the considered reactor model parameters Equations (15) and (16), it has been shown that most parameters have slight changes. Therefore, the parameters are supposed to be constant and are represented in Table 3.

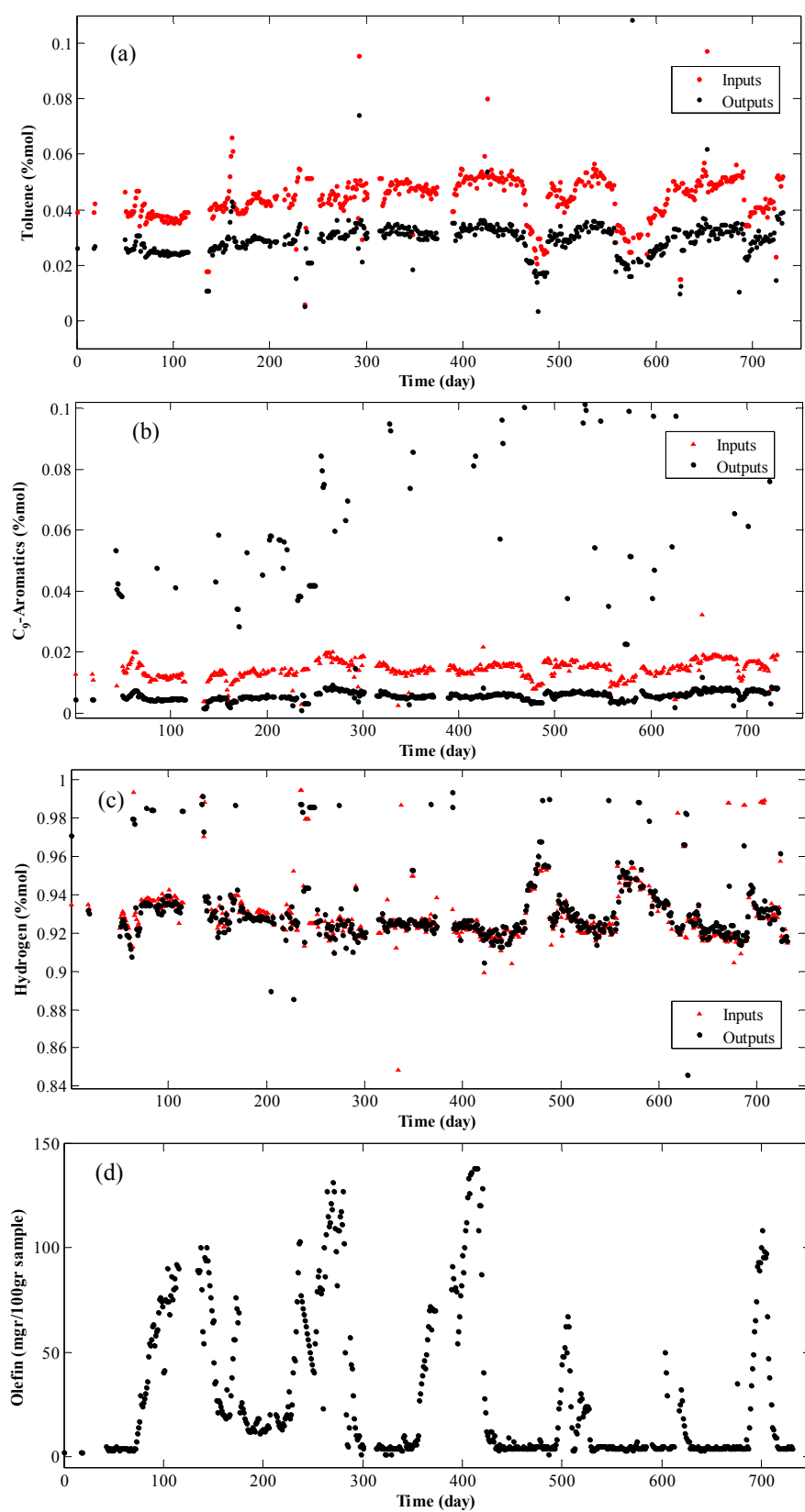
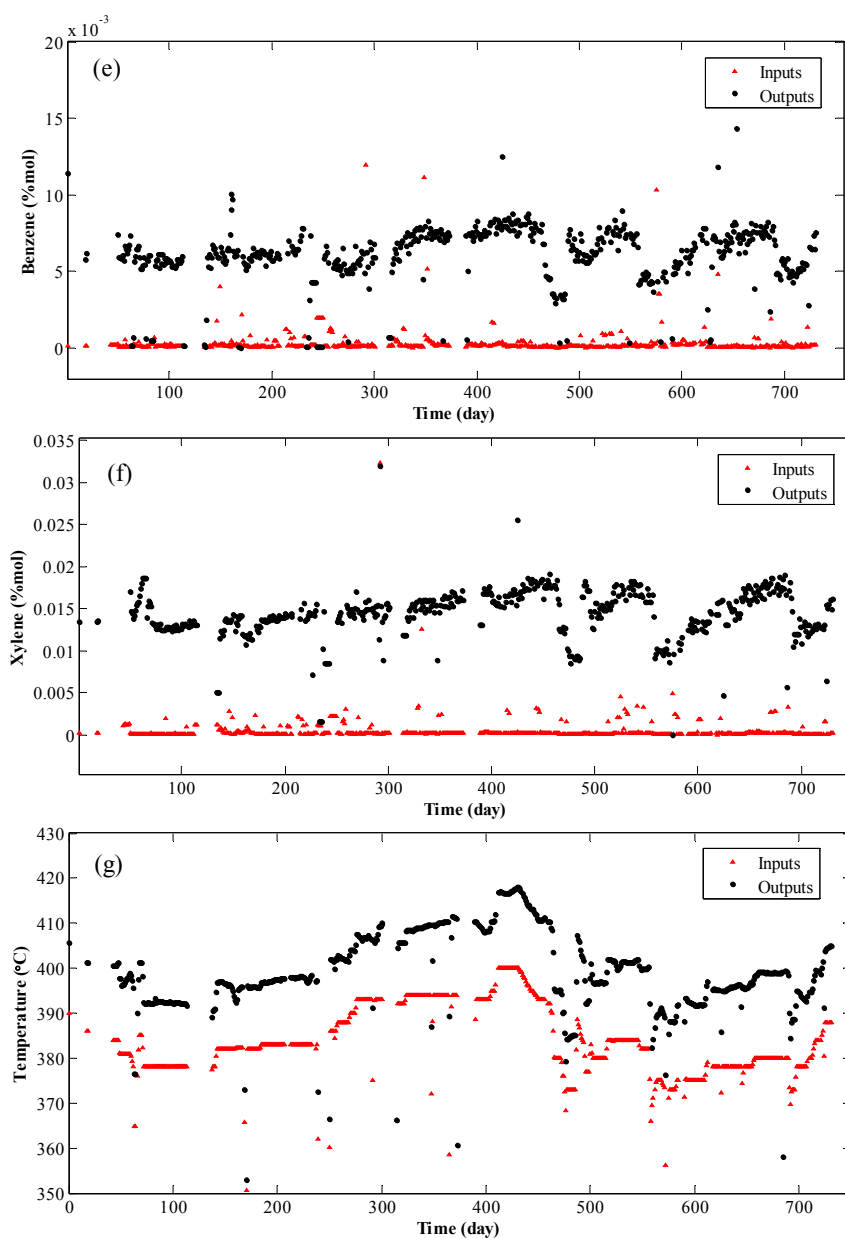


Figure 1. Cont.



**Figure 1.** Input and output data of process. (a) Toluene; (b) C<sub>9</sub> aromatics; (c) hydrogen; (d) olefin; (e) benzene; (f) xylene; and (g) temperature.

**Table 3.** Initial estimates for fixed parameters of models.

| Parameter            | Mean (95% CL)          |
|----------------------|------------------------|
| $\lambda_1$          | −0.19 ( $\pm 0.30$ )   |
| $\lambda_2$          | −0.04 ( $\pm 0.20$ )   |
| $\lambda_3$          | 1.85 ( $\pm 0.28$ )    |
| $\mu_1$              | 0.10 ( $\pm 0.27$ )    |
| $\mu_2$              | 0.60 ( $\pm 0.24$ )    |
| $\mu_3$              | 1.17 ( $\pm 0.15$ )    |
| $E_1$ (J/kmol·K)     | 43388 ( $\pm 921.15$ ) |
| $k_{01}$ (kg·h/kmol) | 2.90 ( $\pm 1.01$ )    |
| $E_2$ (J/kmol·K)     | 18201 ( $\pm 490.40$ ) |
| $k_{02}$ (kg·h/kmol) | 0.32 ( $\pm 0.43$ )    |

The only exception is due to the variations of the term for the deactivation of catalyst *a*, which is considered to be a function of different variables of the system states. The changes have been found to be significant in terms of the time and hydrogen and olefin concentrations. An important issue is the high sensitivity of the catalyst activity to the olefin concentration (lower than 200ppm). Olefin causes coke formation (as the undesired reaction) and reduces the catalyst efficiency; hence, the conversion decreases. In order to compensate for this effect, the system operation needs to increase the temperature; however, increasing the reactor temperature will also increase the deactivation rate of the catalyst. Furthermore, high molar fraction of hydrogen is required to avoid undesired coking reaction and dilution of other components in order to reduce the diffusion resistance and increased temperature result in the recovery of the catalysts. However, the process is then able to perform the required conversion in a lower temperature. Thus, the effect of olefin on the catalyst poisoning, along with the high molar fraction of hydrogen and the increased temperature for the catalyst recovery, cause the catalyst activity to have a roughly monotonic trend over the two-year period of operation. One main advantage of the SDP method to design the soft sensor is that the modeling is exclusively performed by process data and, hence, exact determination of the process information (e.g., reaction mechanism, kinetics, etc.) is not needed. Therefore, in the present study, the parameters related to such problems as the order of concentration in the reaction are found to be different from those recorded in the literature (Table 3). As can be seen, some parameters have values close to zero with high uncertainty. Thus, the related parameters are omitted and the equation is again recovered and identified as follows:

$$\begin{aligned} \log(C_{B,L} - C_{B,0}) - \log\left(\frac{S \cdot L}{Q} \rho_x\right) &= \lambda_3 \log(C_X) + \frac{-E_1}{R} \frac{1}{T} + \phi_1 \\ \log(C_{Ar,L} - C_{Ar,0}) - \log\left(\frac{S \cdot L}{Q} \rho_x\right) &= \mu_2 \log(C_{Ar}) + \mu_3 \log(C_X) + \frac{-E_2}{R} \frac{1}{T} + \phi_2 \end{aligned} \quad (29)$$

In these equations, the term for the catalyst deactivation multiplied by the pre-exponential factor is replaced by  $\phi$ . The results of the new identification show a reduction in the uncertainty of the estimated parameters as represented in Table 4.

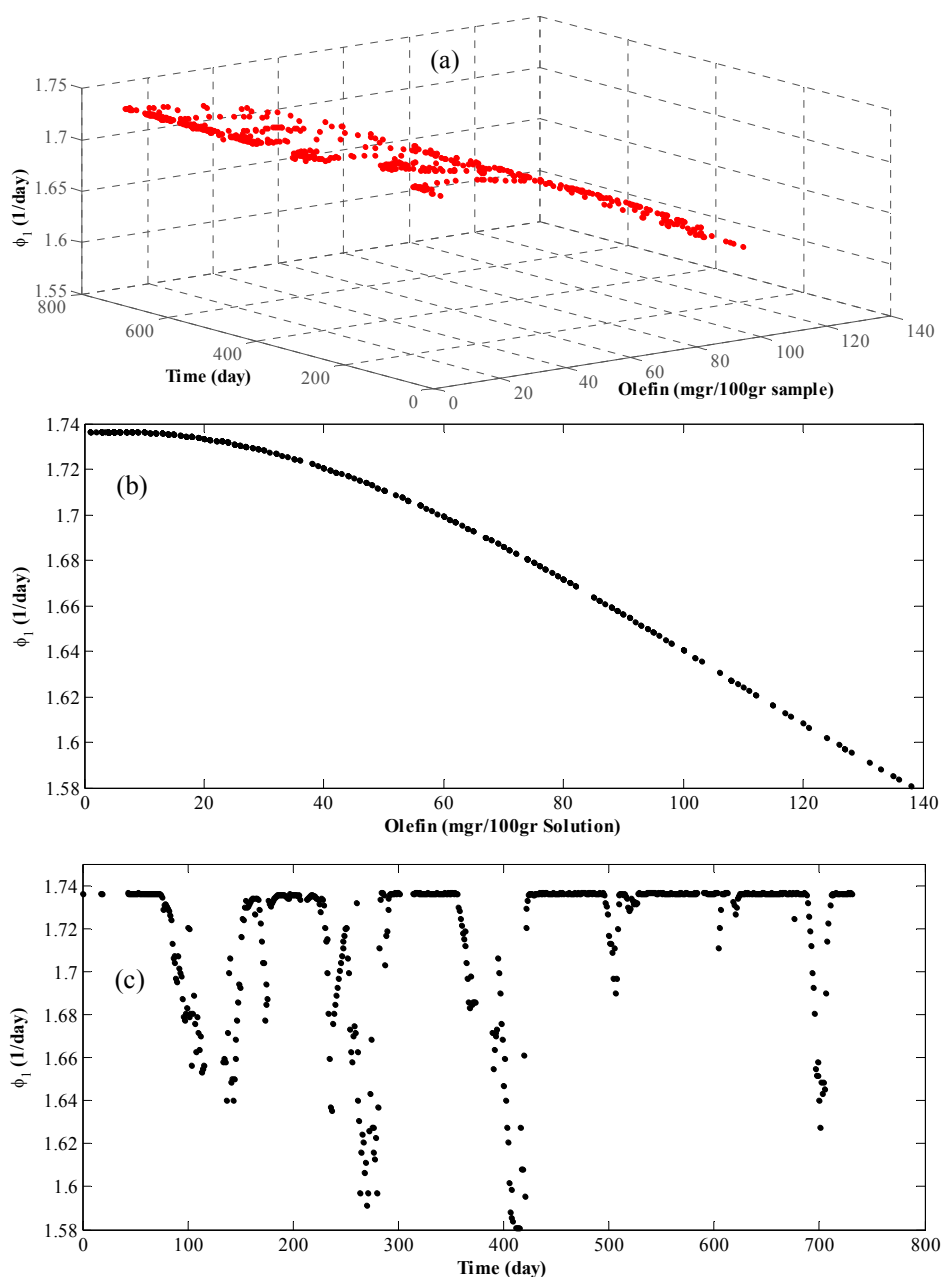
**Table 4.** Final estimates for fixed parameters of models.

| Parameter        | Mean (95% CL)       |
|------------------|---------------------|
| $\lambda_3$      | 1.7 ( $\pm 0.15$ )  |
| $\mu_2$          | 0.67 ( $\pm 0.12$ ) |
| $\mu_3$          | 1.19 ( $\pm 0.13$ ) |
| $E_1$ (J/kmol·K) | 43842 ( $\pm 918$ ) |
| $E_2$ (J/kmol·K) | 18084 ( $\pm 489$ ) |

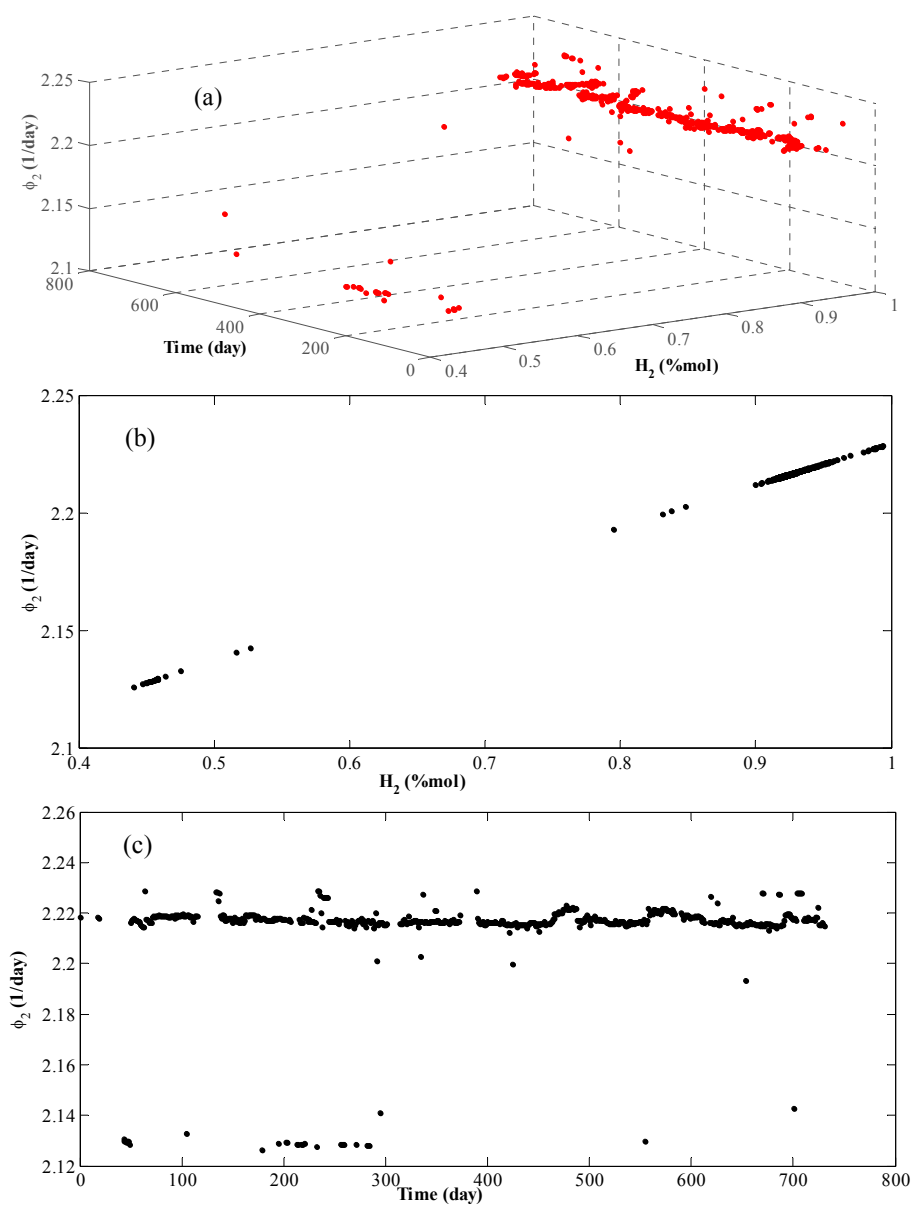
In both models, the xylene concentration plays a decisive role, compared to the other components. Therefore, it shows that with the increase in the xylene concentration in the feed, a higher conversion is obtained. As shown by Table 4, the activation energies of the DP and TA reactions are in an acceptable range and in compliance with the previous studies (see Table 1). Comparison of the activation energies of the DP and TA reactions shows that the TA reaction is faster. Concerning the catalyst deactivation term in the first model ( $\phi_1$ ), from the investigation of different states in the two dimensions of time and concentration of olefin, it can be observed that olefin causes catalyst toxicity and significantly decreases catalytic activity (see Figure 2). In this context, as shown in Figure 2a–c, it is suggested that for avoiding the catalyst toxicity, olefin concentration in the feed should be lower than 20ppm. However, during the time the catalyst activity is in an oscillatory form, and as mentioned before, it shows an approximately constant activity over the period of study. As shown in Figure 2, the catalyst deactivation term in the first model is changing by a function of olefin concentration and time. The catalyst deactivation term in the second model ( $\phi_2$ ) has been identified as a function of the two states of hydrogen concentration and time. Hydrogen concentration and the increased temperature have recovered the catalyst (Figure 3a–c).

Based on the changes in temperature and olefin concentration (Figure 4), it can be concluded that with an increase in the amount of olefin in the inlet flow, the reactor temperature is increased by the operator with a considerable time delay. The time delay causes a reduction in the conversion of the reactor. This can be improved by employing an online prediction method. In other words, by applying the soft sensor, the temperature can be adjusted based on the catalyst activity.

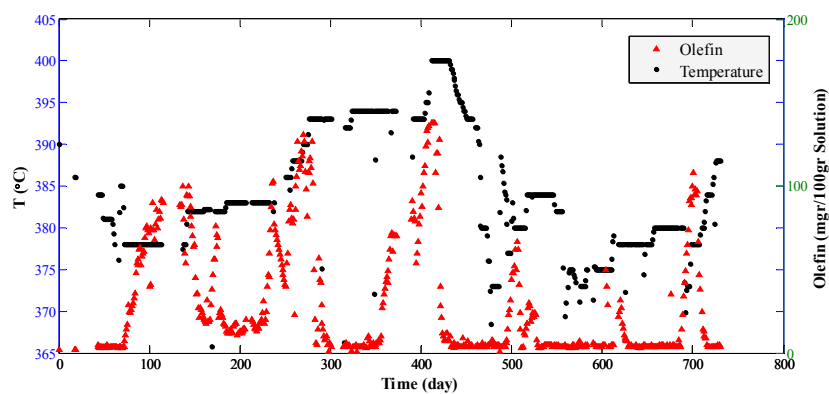
It would be a good progress for the industrial processes to be identified and fault-detected using the process data with no extra cost imposed. As shown in Figure 5, the predicted values are in good agreement with the process data which can be considered as an important result of the SDP method. It should be noted that the SDP method with reference to the reduction in the laboratory costs is a great success for petrochemical industries. As a result, it assists in the increase of prediction and fault detection ability.



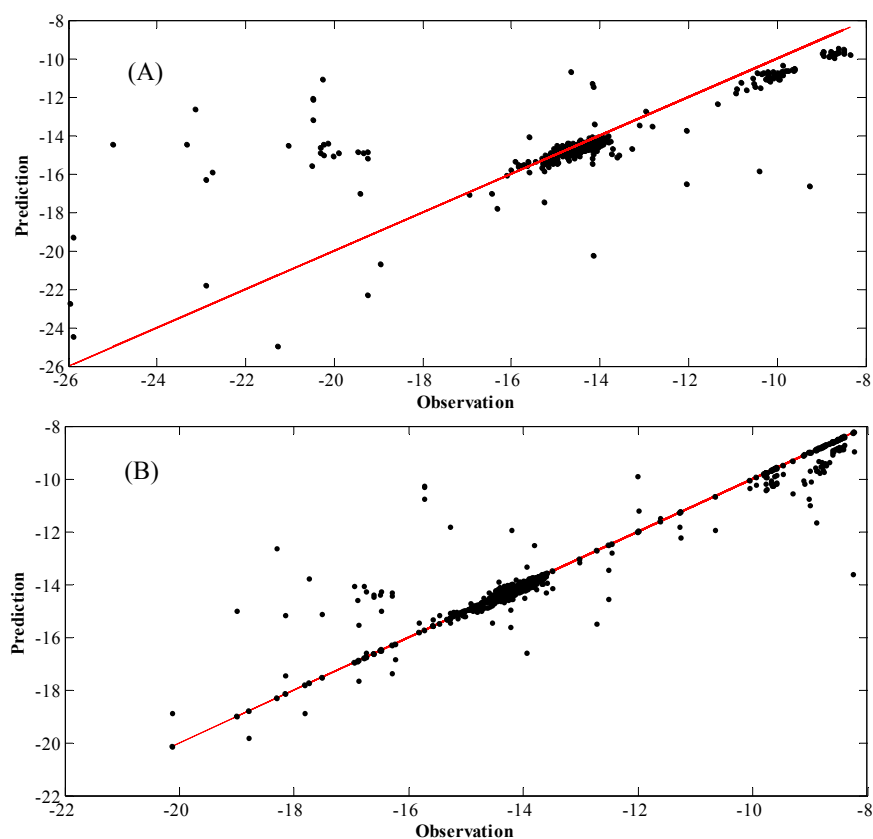
**Figure 2.** Catalyst deactivation term in the first model; (a) based on time and olefin states, simultaneously; (b) olefin state; and (c) time state.



**Figure 3.** Catalyst deactivation term in the second model, (a) based on time and hydrogen states; simultaneously; (b) hydrogen state; and (c) time state.



**Figure 4.** Temperature and olefin concentration vs. time, simultaneously.



**Figure 5.** The relationship between observed and predicted  $y$  in the SDP method. (A) First model; and (B) second model.

## 5. Conclusions

A soft sensor is designed for identification of the variations in the catalyst activity using process data of an industrial reactor. Based on the DBM philosophy, the present study is developed in an innovative method for soft sensor design, which demonstrates appropriate capability for analyzing the data with drastic changes. Using the SDP method, the catalyst activity has been estimated non-parametrically as a function of time, olefin concentration, and hydrogen concentration. Other parameters are considered to be constant. Based on the obtained results, it is suggested that the amount of olefin in the inlet flow should be lower than 20 ppm. In the corresponding models, for both reactions, the order of xylene concentration is greater than that of the other components and, thus, for the increase in the conversion it is suggested that the xylene concentration in the inlet flow should be increased through the recycle flow. However, the estimated order of concentration of the components does not show a good agreement with the previous results. The activation energies for the two DP and TA reactions have been found to be 43.8 and 18 kJ/mol, respectively. The model prediction is in good agreement with the process data. The present work can be considered as an advance towards the identification, modeling, and fault detection of a complex industrial process.

**Acknowledgments:** The authors acknowledge the technical support of Bou-Ali-Sina Petrochemical Company.

**Author Contributions:** Hamed Gharehbaghi wrote the first draft of the manuscript, which was then improved by Jafar Sadeghi.

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

## Symbols

$a_1$  the term of catalyst deactivation based on the DP reaction ( $\text{h}^{-1}$ )

|                                  |                                                                                                        |
|----------------------------------|--------------------------------------------------------------------------------------------------------|
| $a_2$                            | the term of catalyst deactivation based on the TA reaction ( $\text{h}^{-1}$ )                         |
| $a_i$                            | the $i$ th coefficient of model output (–)                                                             |
| $b_i$                            | the $i$ th coefficient of model input (–)                                                              |
| $C_i$                            | concentration of species $i$ in the fixed-bed reactor ( $\text{kmolm}^{-3}$ )                          |
| $D_A$                            | axial diffusion ( $\text{m}^2 \text{h}^{-1}$ )                                                         |
| $D_R$                            | radial diffusion ( $\text{m}^2 \text{h}^{-1}$ )                                                        |
| $E_i$                            | activation energy of $i$ th reaction ( $\text{J kmol}^{-1} \text{K}^{-1}$ )                            |
| $e_t$                            | observation error (–)                                                                                  |
| <b>F</b>                         | system coefficient matrix (–)                                                                          |
| <b>F<sub>i</sub></b>             | the $i$ th SDP's coefficient matrix (–)                                                                |
| <b>G</b>                         | input coefficient matrix (–)                                                                           |
| <b>G<sub>i</sub></b>             | the $i$ th SDP's coefficient matrix (–)                                                                |
| <b>H<sub>t</sub></b>             | SDP's coefficient matrix (–)                                                                           |
| <b>I</b>                         | identity matrix (–)                                                                                    |
| $k_{0i}$                         | pre-exponential factor for the $i$ th reaction ( $\text{kg} \cdot \text{h kmol}^{-1}$ )                |
| $L$                              | length of reactor bed (m)                                                                              |
| <b>L<sub>t</sub></b>             | lagrange multiplier in FIS smoothing (–)                                                               |
| $L_c$                            | concentrated likelihood function (–)                                                                   |
| $N$                              | Number of data points in time series (–)                                                               |
| $ns_i$                           | number of state that the $i$ th SDP is a function of them (–)                                          |
| $P_i$                            | partial pressure (Pa)                                                                                  |
| $p$                              | number of parameters (–)                                                                               |
| $p_i(\cdot)$                     | the $i$ th state dependent parameter (–)                                                               |
| <b>P<sub>t</sub></b>             | vector of model parameters (–)                                                                         |
| <b>P<sub>t</sub><sup>*</sup></b> | the error covariance matrix (–)                                                                        |
| $Q$                              | volumetric flow rate ( $\text{m}^3 \text{h}^{-1}$ )                                                    |
| <b>Q</b>                         | covariance matrix (–)                                                                                  |
| <b>Q<sub>r</sub></b>             | noise variance ratio matrix (–)                                                                        |
| $R$                              | gas constant ( $\text{J mol}^{-1} \text{K}^{-1}$ )                                                     |
| <b>R'<sub>i</sub></b>            | rate of generation or disappearance of component $i$ ( $\text{mol kg cat.}^{-1} \cdot \text{h}^{-1}$ ) |
| $r$                              | radius of reactor bed (m)                                                                              |
| $S$                              | cross-sectional area of reactor bed                                                                    |
| $T$                              | temperature (K)                                                                                        |
| $t$                              | time (h)                                                                                               |
| $u_t$                            | inputs (–)                                                                                             |
| WHSV                             | weight hourly space velocity ( $\text{h}^{-1}$ )                                                       |
| $x_t$                            | state vector (–)                                                                                       |
| $y_t$                            | output (–)                                                                                             |
| <b>y<sub>t</sub></b>             | Vector of outputs (–)                                                                                  |
| <b>z<sub>t</sub></b>             | vector of regressors (–)                                                                               |

### Greek Letters

|             |                                                    |
|-------------|----------------------------------------------------|
| $\eta_t$    | vector of parameters' noise (–)                    |
| $\lambda_i$ | order of component $i$ in the DP reaction (–)      |
| $\mu_i$     | order of component $i$ in the TA reaction (–)      |
| $\zeta$     | error of model (–)                                 |
| $\rho_x$    | catalyst density in the bed ( $\text{kg m}^{-3}$ ) |

|            |                                                                          |
|------------|--------------------------------------------------------------------------|
| $\sigma^2$ | variance of error (–)                                                    |
| $\phi_i$   | catalytic deactivation term multiplied by the pre-exponential factor (–) |

## References

- Waziri, S.M.; Aitani, A.M.; Al-Khattaf, S. Transformation of Toluene and 1,2,4-Trimethylbenzene over ZSM-5 and Mordenite Catalysts: A Comprehensive Kinetic Model with Reversibility. *Ind. Eng. Chem. Res.* **2010**, *49*, 6376–6387. [[CrossRef](#)]
- Al-Khattaf, S. Catalytic transformation of toluene over a high-acidity Y-zeolite based catalyst. *Energy Fuels* **2006**, *20*, 946–954. [[CrossRef](#)]
- Wang, T.; Tsai, C.; Huang, S.T. Disproportionation of toluene and of trimethylbenzene and their transalkylation over zeolite beta. *Ind. Eng. Chem. Res.* **1990**, *29*, 2005–2012. [[CrossRef](#)]
- Tsai, T.-C.; Liu, S.-B.; Wang, I. Disproportionation and transalkylation of alkylbenzenes over zeolite catalysts. *Appl. Catal. A* **1999**, *181*, 355–398. [[CrossRef](#)]
- Tsai, T.C.; Chenb, W.H.; Lai, C.S.; Liub, S.B.; Wangc, I.; Ku, C.S. Kinetics of Toluene Disproportionation over Fresh and Coked H-mordenite. *Catal. Today* **2004**, *97*, 297–302. [[CrossRef](#)]
- Bhavkattl, S.S.; Patwardhan, S.R. Toluene Disproportionation over Nickel-Loaded Aluminum-Deficient Mordenite 2. *Kinetics. Ind. Eng. Chem. Prod. Res. Dev.* **1981**, *20*, 106–109.
- Al-Khattaf, S.; Rabiou, S.; Tukur, N.; Alnaizy, R. Kinetics of toluene methylation over USY-zeolite catalyst in a riser simulator. *Chem. Eng. J.* **2008**, *139*, 622–630. [[CrossRef](#)]
- De Jong, K.; Mesters, C.; Peferoen, D.; Van Brugge, P.; De Groot, C. Paraffin alkylation using zeolite catalysts in a slurry reactor: Chemical engineering principles to extend catalyst lifetime. *Chem. Eng. Sci.* **1996**, *51*, 2053–2060. [[CrossRef](#)]
- Taylor, R.J.; Sherwood, D.E. Effects of process parameters on isobutane/2-butene alkylation using a solid acid catalyst. *Appl. Catal. A* **1997**, *155*, 195–215. [[CrossRef](#)]
- Sahebdehfar, S.; Kazemeini, M.; Khorasheh, F.; Badakhshan, A. Deactivation behavior of the catalyst in solid acid catalyzed alkylation: Effect of pore mouth plugging. *Chem. Eng. Sci.* **2002**, *57*, 3611–3620. [[CrossRef](#)]
- Simpson, M.F.; Wei, J.; Sundaresan, S. Kinetic analysis of isobutane/butene alkylation over ultrastable HY zeolite. *Ind. Eng. Chem. Res.* **1996**, *35*, 3861–3873. [[CrossRef](#)]
- Lobao, M.W.N.; Alberton, A.L.; Melo, S.A.B.V.; Embirucu, M.; Monteiro, J.L.F.; Pinto, J.C. Kinetics of Toluene Disproportionation: Modeling and Experiments. *Ind. Eng. Chem. Res.* **2012**, *51*, 171–183. [[CrossRef](#)]
- Fortuna, L.; Graziani, S.A.; Rizzo, S.; Xibilia, M.G. *Soft Sensors for Monitoring and Control of Industrial Processes*; Springer: London, UK, 2007.
- Kadleca, P.; Gabrysa, B.; Strandt, S. Data-driven Soft Sensors in the process industry. *Comput. Chem. Eng.* **2009**, *33*, 795–814. [[CrossRef](#)]
- Kaneko, H.; Arakawa, M.; Funatsu, K. Development of a New Soft Sensor Method Using Independent Component Analysis and Partial Least Squares. *AIChE J.* **2009**, *55*, 87–98. [[CrossRef](#)]
- Lin, B.; Recke, B.; Schmidt, T.M.; Knudsen, J.K.H.; Jørgensen, S.B. Data-Driven Soft Sensor Design with Multiple-Rate Sampled Data: A Comparative Study. *Ind. Eng. Chem. Res.* **2009**, *48*, 5379–5387. [[CrossRef](#)]
- Liu, J.; Chen, D.-S.; Shen, J.-F. Development of Self-Validating Soft Sensors Using Fast Moving Window Partial Least Squares. *Ind. Eng. Chem. Res.* **2010**, *49*, 11530–11546. [[CrossRef](#)]
- Liu, Y.; Gao, Z.; Chen, J. Development of soft-sensors for online quality prediction of sequential-reactor-multi-grade industrial processes. *Chem. Eng. Sci.* **2013**, *102*, 602–612. [[CrossRef](#)]
- Shang, C.; Yang, F.; Huang, D.; Lyu, W. Data-driven soft sensor development based on deep learning technique. *J. Process Control* **2014**, *24*, 223–233. [[CrossRef](#)]
- Kaneko, H.; Funatsu, K. Database Monitoring Index for Adaptive Soft Sensors and the Application to Industrial Process. *AIChE J.* **2014**, *60*, 160–169. [[CrossRef](#)]
- Kaneko, H.; Funatsu, K. Adaptive soft sensor model using online support vector regression with time variable and discussion of appropriate hyperparameter settings and window size. *Comput. Chem. Eng.* **2013**, *58*, 288–297. [[CrossRef](#)]
- Kaneko, H.; Funatsu, K. Adaptive soft sensor based on online support vector regression and Bayesian ensemble learning for various states in chemical plants. *Chemom. Intell. Lab. Syst.* **2014**, *137*, 57–66. [[CrossRef](#)]

23. Elnashaie, S.S.E.H.; Elshishini, S.S. *Modelling, Simulation and Optimization of Industrial Fixed Bed Catalytic Reactors*; Gordon & Breach Science: Philadelphia, PA, USA, 1993.
24. Manenti, F.; Cieri, S.; Restelli, M.; Bozzano, G. Dynamic modeling of the methanol synthesis fixed-bed reactor. *Comput. Chem. Eng.* **2013**, *48*, 325–334. [[CrossRef](#)]
25. Young, P. *Recursive Estimation and Time-Series Analysis an Introduction for the Student and Practitioner*; Springer: Berlin, Germany, 2011.
26. Young, P. Data-based mechanistic modelling of environmental, ecological, economic and engineering systems. *Environ. Model. Softw.* **1998**, *13*, 105–122. [[CrossRef](#)]
27. Young, P. Data-based mechanistic modelling, generalised sensitivity and dominant mode analysis. *Comput. Phys. Commun.* **1999**, *117*, 113–129. [[CrossRef](#)]
28. Price, L.; Young, P.; Berckmans, D.; Janssens, K.; Taylor, J. Data-based mechanistic modelling (DBM) and control of mass and energy transfer in agricultural buildings. *Ann. Rev. Control* **1999**, *23*, 71–82. [[CrossRef](#)]
29. Young, P. The data-based mechanistic approach to the modelling, forecasting and control of environmental systems. *Ann. Rev. Control* **2006**, *30*, 169–182. [[CrossRef](#)]
30. Taylor, C.J.; Young, P.C.; Chotai, A. *True Digital Control: Statistical Modelling and Non-Minimal State Space Design*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
31. Bharati, S.P.; Bhatia, S. Deactivation Kinetics of Toluene Disproportionation over Hydrogen Mordenite Catalyst. *Ind. Eng. Chem. Res.* **1987**, *26*, 1854–1860. [[CrossRef](#)]
32. Xu, O.; Su, H.; Ji, J.; Jin, X.; Chu, J. Kinetic Model and Simulation Analysis for Toluene Disproportionation and C9-Aromatics Transalkylation. *Chin. J. Chem. Eng.* **2007**, *15*, 326–332. [[CrossRef](#)]
33. Al-Mubaiyedha, U.A.; Ali, S.A.; Al-Khattaf, S.S. Kinetic Modeling of Heavy Reformate Conversion into Xylenes over Mordenite-ZSM5 Based Catalysts. *Chem. Eng. Res. Des.* **2012**, *90*, 1943–1955. [[CrossRef](#)]
34. Odedairo, T.; Balasamy, R.J.S.; Al-Khattaf, R.J. Toluene Disproportionation and Methylation over Zeolites TNU-9, SSZ-33, ZSM-5, and Mordenite Using Different Reactor Systems. *Ind. Eng. Chem. Res.* **2011**, *50*, 3169–3183. [[CrossRef](#)]
35. Beven, K.J.; Leedal, D.T.; Smith, P.J.; Young, P.C. *Identification and Representation of State Dependent Non-Linearities in Flood Forecasting Using the DBM Methodology*; Springer: Berlin, Germany, 2012; pp. 341–366.
36. Toivonen, H.; Tötterman, S.; Åkesson, B. Identification of state-dependent parameter models with support vector regression. *Int. J. Control* **2007**, *80*, 1454–1470. [[CrossRef](#)]
37. Trambouze, P.; Landeghem, H.V.; Wauquier, J.P. *Chemical Reactors Design/Engineering/Operation*; Éditions Technip: Paris, France, 1988.
38. Lee, R.C. *Optimal Estimation, Identification, and Control*; MIT Press: Cambridge, MA, USA, 1964.
39. Young, P. An instrumental variable method for real-time identification of a noisy process. *Automatica* **1970**, *6*, 271–287. [[CrossRef](#)]
40. Young, P. *Recursive Estimation and Time-Series Analysis: An Introduction*; Springer-Verlag: New York, NY, USA, 1984.
41. Norton, J. *An Introduction to Identification*; Academic Press: London, UK, 1986.
42. Sadeghi, J.; Tych, W.; Chotai, A.A.; Young, P. Multi-state dependent parameter model identification and estimation for nonlinear dynamic systems. *Electron. Lett.* **2010**, *46*, 1265–1266. [[CrossRef](#)]
43. Tych, W.; Sadeghi, J.; Smith, P.J.; Chotai, A.; Taylor, C.J. *Multi-state Dependent Parameter Model Identification and Estimation*; Springer: London, UK, 2012; pp. 191–210.

