Identification of kinetic models of methanol oxidation on silver in the presence of uncertain catalyst behavior

Marco Quaglio1 | Fabrizio Bezzo2 | Asterios Gavriilidis1 | Enhong Cao1
Noor Al-Rifai1 | Federico Galvanin1

1Department of Chemical Engineering, University College London (UCL), London, UK
2CAPE-Lab (Computer-Aided Process Engineering Laboratory), Department of Industrial Engineering, University of Padova (UNIPD), Padova, Italy

Abstract
Catalytic oxidation of methanol to formaldehyde is an important industrial process due to the value of formaldehyde either as a final product or as a precursor of numerous chemicals. The study of kinetics in this system is hindered by sources of uncertainty that are inherently associated to the nature and state of the catalyst (e.g., uncertain reactivity level, deactivation phenomena), the measurement system and the structure of the kinetic model equations. In this work, a simplified kinetic model is identified from data collected from continuous flow microreactor systems where catalysts with assorted levels of reactivity are employed. Tailored model-based data mining methods are proposed and applied for the effective estimation of the kinetic parameters and for identifying robust experimental conditions to be exploited for the kinetic characterization of catalysts with different reactivity, whose kinetic behavior is yet to be investigated.

KEYWORDS
data mining, information, parameter estimation, reactivity, uncertainty

1 | INTRODUCTION

The oxidation of methanol to formaldehyde over silver catalyst is an industrial process of great significance due to the importance of formaldehyde as a precursor to the production of valuable chemicals. The versatility demonstrated by this organic compound led to its usage in numerous production chains. Agriculture, medicine, cosmetics industry, pesticides and fertilizers production, and dyes and explosive synthesis are only some of the areas in which formaldehyde has found successful application.1 Under industrial conditions, the process is usually carried out at atmospheric pressure and relatively high temperature (T = 850–923 K) on silver catalyst.1 The overall mechanism involves the conversion of methanol CH3OH to formaldehyde CH2O, through both oxidation and dehydrogenation reactions:

\[
\text{CH}_3\text{OH} + \frac{1}{2}\text{O}_2 \rightarrow \text{CH}_2\text{O} + \text{H}_2\text{O} \quad (1)
\]

\[
\text{CH}_3\text{OH} \rightarrow \text{CH}_2\text{O} + \text{H}_2 \quad (2)
\]

The principal by-products are: hydrogen H2, water H2O, carbon dioxide CO2, formic acid HCOOH, and traces of carbon monoxide CO. The industrial process is carried out adopting rich inlet methanol/oxygen mixtures, introducing steam for achieving high selectivity. The main undesired reactions occurring in the system, affecting the selectivity of formaldehyde, involve the complete...
oxidation of methanol and the oxidation of formaldehyde into carbon dioxide:

\[
\begin{align*}
\text{CH}_3\text{OH} + 3/2\text{O}_2 &\rightarrow 2\text{H}_2\text{O} + \text{CO}_2 \\
\text{CH}_2\text{O} + \text{O}_2 &\rightarrow \text{H}_2\text{O} + \text{CO}_2
\end{align*}
\]  

(3)  

(4)

Despite the great industrial importance of this process, the key phenomena occurring on the catalyst surface are yet to be completely understood. Many researchers investigated the catalytic role of silver in the reaction and the possible mechanisms occurring on the surface of the catalyst film. In 2003, a micro-kinetic model for methanol oxidation on silver was proposed by Andreesen et al\(^6\) to explain the surface phenomena in experimental conditions of interest for industrial application. However, the complexity of the proposed mechanism made unrealistic its use for engineering purposes and, for this reason, a simplified model derived from the micro-kinetic one was proposed by the same authors.\(^7\)

One of the main issues making this system so complicated to be unraveled is related to the highly uncertain nature of the catalyst. Indeed, the reactivity is strongly dependent on the density and distribution of the active sites on the catalyst surface. This crucial aspect has a strong impact on the catalyst behavior and can be influenced by a multitude of factors, from the production of the catalyst to the operating conditions the catalyst has to withstand.\(^8\) The identification of a kinetic model in a reacting system affected by such uncertainty relies on the collection of valuable information from the experiments. Microreactor platforms are suitable for collecting information with the aim of identifying kinetic models for fast exothermic and endothermic reactions.\(^9\) The small dimensions characterizing the channel of these devices permit the execution of the reaction in isothermal conditions and in the absence of mass transfer limitations. However, the information content of an experiment does not rely exclusively on the experimental setup, but also on the thoughtful choice of the experimental conditions to investigate. Model-based design of experiment (MBDoE) techniques have been proposed in literature to identify the best experimental conditions depending on the purpose of the experimental investigation: selecting a model out of a set of proposed candidates (i.e., MBDoE for model discrimination)\(^10-12\); improving the parameter estimates of an already selected model (i.e., MBDoE for parameter precision)\(^12-15\); or both (i.e., joint-MBDoE)\(^16-19\). However, only few works have been proposed in the literature to address the problem of model identification in the presence of high system uncertainty (i.e., uncertain behavior of the physical system and structural uncertainty in the model equations).\(^18-21\)

In this manuscript, data obtained from experiments performed on microreactors using silver catalysts for formaldehyde synthesis are analyzed for:

1. Estimating precisely the kinetic parameters in a simplified model of methanol oxidation on silver\(^7\) dealing with the sources of uncertainty occurring in the system and in the model;
2. Identifying robust experimental conditions for the collection of valuable information from the kinetic experiments regardless of the uncertainty on the catalyst reactivity.

Task 1 is fulfilled employing machine learning technologies.\(^22,23\) Statistical learning methods are increasingly being applied in surface and material sciences for the discovery and characterization of heterogeneous catalysts.\(^24,25\) Specifically, data mining techniques\(^26\) have been employed for supporting the recognition of patterns in kinetic datasets and suggest improved catalyst designs.\(^27,28\) In this work, tailored data mining methods based on robust regression techniques\(^29-31\) are used to identify automatically the experimental data that are not significant for the estimation of the kinetic parameters embracing the following sources of uncertainty: (a) measurement errors in the data; (b) catalyst deactivation; and (c) structural uncertainty in the candidate kinetic model. For addressing Task 2, MBDoE techniques based on the concept of Fisher information\(^32\) are applied to quantify the information associated with kinetic experiments in a convenient range of experimental conditions and for variable catalyst reactivity. This allows for the identification of experimental conditions whose information is insensitive to the catalyst behavior. These conditions represent the best compromise for starting the experimental activity with the aim of identifying kinetic models of methanol oxidation whenever the behavior of the catalyst is still highly uncertain.

## 2 | KINETIC MODEL IDENTIFICATION PROCEDURE

The identification of a reliable set of equations to model the kinetics of a catalytic reaction is not a straightforward process. Frequently, the process of kinetic modeling requires dealing with the impossibility of measuring directly certain physical quantities, or the difficulty of separating the key mechanisms because of an overlapping of their effects in the reacting system (e.g., uncertainty on the structure of complex reaction networks).\(^33\) Furthermore, there might be some phenomena occurring in the system (catalyst deactivation is a typical example) that are too complex to describe. For these reasons, proposed kinetic model structures frequently embody a certain degree of approximation and uncertainty. The precise identification of the parameters in these models requires the identification of highly informative experimental conditions whose evaluation may not be a trivial task because of the following reasons:

- The range of experimental conditions in which the simplifying (possibly inappropriate) modeling hypotheses are acceptable (i.e., the range of conditions in which the model gives accurate predictions), may not be known a priori;
- The model may be capable of computing reliable predictions only for a subset of its predicted quantities;
- Some measurements in the dataset available to fit the non-measurable model parameters may contain gross errors;
- The location of experimental conditions carrying valuable Fisher information (i.e., information for estimating nonmeasurable model parameters through data fitting)\(^32\) in the experimental design space may be very sensitive to the reactivity of the specific catalyst analyzed; reactivity is normally unknown before preliminary kinetic experiments are performed.
A framework is here proposed for the systematic identification of kinetic models in catalytic systems embracing the aforementioned sources of uncertainty. The scheme of the proposed procedure is given in Figure 1. The procedure starts from the availability of: (a) a highly reliable (benchmark) dataset, assumed as reference; (b) a number \( n \) of new type catalysts showing a highly heterogeneous range of behaviors (i.e., a wide range of different reactivities); and (c) a proposed model structure for describing the kinetics. We assume that kinetic experiments are performed on the \( n \) available catalysts obtaining \( n \) distinct datasets. In the proposed framework three main stages can be distinguished:

1. **A reference model identification step.** The kinetic parameters of the proposed model structure are estimated from the benchmark dataset obtaining an instance of kinetic parameters for the proposed model structure.

2. **A model-based data mining (MBDM) stage for parameter estimation.** In order to account for the intrinsic variability on the reaction system, the kinetic parameters that are correlated to the reactivity (e.g., pre-exponential factors of catalytic reactions), have to be adjusted to each of the \( n \) datasets through data fitting. MBDM is employed for the automated model-based identification of outliers, that is, data for which there is a significant discrepancy between model prediction and measurements. Provided that the \( n \) datasets are sufficiently rich, the application of MBDM generates two outputs: (a) \( n \) sets of kinetic parameters; (b) it labels experimental data either as compatible or incompatible with the candidate model. Knowledge of the distribution of outliers in the

![Figure 1](image-url)  
**Figure 1** Framework for systematic identification of kinetic models in the presence of uncertain catalyst behavior
model input space provides information on where the model is inaccurate, leading to the determination of the domain of model reliability.

3. A robust MBDoE stage. The n instances of the reference model identified at the previous stage capture the variability of behaviors across the analyzed catalysts. Information on the previously observed range of variability in the catalyst behavior is exploited at this stage to design highly informative experiments employing robust MBDoE methodologies.\(^{34,35}\) The employment of the robust MBDoE leads to the identification of robust experimental conditions that are insensitive to catalyst reactivity and guarantee the collection of high quality information for a wide range of catalyst behaviors. Knowledge on the geometry of the domain of model reliability can be included in the analysis to bound the design problem to regions of the design space where the model is expected to provide a good fitting.

The procedure takes full advantage of previously observed catalyst behaviors in order to minimize the experimental effort required for investigating the kinetics in new catalysts that are yet to be analyzed. In the following sections, the properties of the MBDM method will be discussed. Two tailored MBDM filters are then defined for the specific application presented in this paper. Robust MBDoE methods are then detailed.

### 2.1 MBDM for parameter estimation

A model structure is proposed to describe a certain physical system. A model in its standard reduced form can be expressed in terms of the following system of equations:

\[
\begin{align*}
\hat{y} = f(x, x, u, t, \theta)
\end{align*}
\]

where quantities \(x\) and \(u\) ∈ \(U\) represent vectors of state variables and control variables, respectively, \(t\) is time, \(\theta \in \Theta\) is a set of \(N_\theta\) non measurable model parameters and \(\hat{y}\) is a vector of \(N_o\) predictions associated to \(N_o\) measurable quantities of the physical system. The estimation of \(\theta\) is achieved fitting experimental data through the optimization of an opportune scalar function. A popular technique for parameter estimation is the maximum likelihood (ML) approach, which demonstrated to produce good estimates in a wide range of situations.\(^{36}\) Assume that \(N_o\) experiments were performed and that in each trial the values of the \(N_m\) measurable variables were collected. The set of data \(\Psi\), defined as in Equation (6), is then available to define a parameter estimation problem.

\[
\Psi = \{y_{ij} || j = 1,...,N_m^\ast \wedge j = 1,...,N_o^\ast\}
\]

In Equation (6) \(y_{ij}\) is the \(i\)-th measured response in the \(j\)-th experiment. Measurements are assumed to be affected by uncorrelated Gaussian noise with zero mean and known SDs \(\sigma_{y}\). If \(\hat{y}_{ij}(\theta)\) is the model prediction of \(y_{ij}\), the computed value of the parameters \(\theta\) is called maximum likelihood estimate if it maximizes the likelihood function \(L(\theta|\Psi)\). In case of known Gaussian measurement noise, however, maximizing \(L(\theta|\Psi)\) is equivalent to minimizing the sum of normalized squared residuals.

\[
\hat{\theta}_{\text{ML}} = \arg\min_{\theta \in \Theta} \sum_{j=1}^{N_o} \sum_{i=1}^{N_m} r_{ij}(\theta)^2
\]

\[
r_{ij}(\theta) = \frac{(\hat{y}_{ij}(\theta) - y_{ij})}{\sigma_{y_{ij}}} \forall i,j
\]

The least squares estimator represents the cornerstone of classical statistics.\(^{30}\) However, statisticians soon realized that traditional least squares methods were inappropriate in any situation where some measurements follow distributions that are significantly different from the postulated ones.\(^{29}\) The concept of breakdown point was developed to assess the sensitivity of a parameter estimator towards the presence of outliers in the dataset.\(^{29}\) Hampel defines the breakdown point as “the smallest fraction of free contamination that can carry the estimated value beyond any bound.”\(^{37}\) For least squares estimators, the breakdown point approaches 0 as the number of fitted measurements increases, meaning that one outlier in the dataset is sufficient to have dramatic effects on the parameter estimation. The acknowledgment of this weakness led to the development of alternative robust estimators and a whole subarea of statistics, namely, the field of robust regression.\(^{29,30,38}\)

In the field of kinetic modeling, the presence of gross measurement errors in the dataset and the presence of inappropriate modeling assumptions are frequent causes of discrepancy between postulated and true measurements distributions.\(^{31}\) This discrepancy makes the employment of least squares methodologies inappropriate for the identification of simplified kinetic models. In the following subsection, tailored MBDM techniques for parameter estimation are derived from robust regression methods to prompt the exclusion of the outliers from the fitting. It is shown that such robust estimators can be tailored to detect automatically whether excessively large model residuals are caused by the attempt of fitting certain experiments or certain model output variables.

#### 2.1.1 Data mining formulation

The weighted least squares method plays an important role in robust statistics.\(^{30,36}\) By weighting the contribution of each residual in the objective function, it is possible to give fitting priority to relevant data and to reduce the effect of outliers on the parameter estimation. Rousseauw and Leroy proposed the use of weights as switchers to activate or deactivate the presence of single residuals in the objective function to be minimised.\(^{30}\)

\[
\hat{\theta}_{\text{DM}} = \arg\min_{\theta \in \Theta} \sum_{j=1}^{N_o} \sum_{i=1}^{N_m} \lambda_{ij} r_{ij}(\theta)^2
\]

\[
\text{s.t. } \lambda_{ij} = \begin{cases} 1 & \text{if } r_{ij}(\theta)^2 < c^2 \\ 0 & \text{if } r_{ij}(\theta)^2 > c^2 \end{cases} \forall i,j
\]

In Equation (9) the quantity \(\lambda_{ij}\) represents the weight associated to the residual \(r_{ij}\). In Equation (10) the switcher behavior of the weights
is defined where the hyperparameter \( c \) defines the maximum threshold of acceptability for an absolute normalized residual to be included in the fitting. Given a robust estimation of the SDs \( \sigma_i \) and a suitable choice of the hyperparameter \( c \) the estimator (9) subject to conditions (10) can attain a breakdown point up to 50% (i.e., the maximum breakdown point theoretically achievable), thus making it an extremely robust estimator that can deal with a high level of outlier contamination in the dataset.\(^{30}\) In the field of gross error detection, the recommended value for the hyperparameter \( c \) is in the range \( c = 2.5-3.0 \), which is equivalent to treating as an outlier any normalized residual exceeding the \( \pm 2.5 \sigma_i \) or \( \pm 3.0 \sigma_i \) confidence range.\(^{30}\) If the candidate model structure is particularly inaccurate, however, the mismatch between model predictions and measurements is not only caused by gross errors, but also by the inadequate model structure. In that case, setting \( c = 2.5-3.0 \) might be too restrictive, leading to the exclusion of too many data points and a significant loss of information. In those cases, if model accuracy is not a stringent model requirement, one may choose to relax the hyperparameter \( c \) adopting values larger than 3.0. Instead, if model accuracy is a fundamental model requirement, but the application of the weighted least squares method results in the exclusion of too many data points, one shall proceed collecting additional data performing experiments in the domain of model reliability (i.e., the range of conditions where the candidate model is expected to give accurate fitting and accurate predictions). This scenario is not considered in the present manuscript and it will be object of study in other works.

Rousseeuw and Leroy also propose an extension of the approach above to a multivariate case in which multiple residuals are grouped under the same switcher weight.\(^{30}\) In this work, the formulation introduced by Rousseeuw and Leroy is employed for obtaining MBDM estimators taking into account the possible limits of the reference model in representing certain experimental conditions or certain output variables. For this purpose, weights \( a_i \) with \( j = 1,...,N_{\text{exp}} \) are defined for the experiments and weights \( \beta_i \) with \( i = 1,...,N_m \) are defined for the output variables.\(^{36}\) A MBDM filter for neglecting the experiments whose associated residuals are not compatible with measurement errors is defined as in Equation (11), including switcher conditions on the residuals (12).

\[
\hat{\theta}_{\text{DM,exp}} = \arg \min_{\theta \in \Theta} \sum_{i=1}^{N_{\text{exp}}} \alpha_i \sum_{j=1}^{N_{\text{exp}}} r_i(\theta)^2 \tag{11}
\]

\[
s.t. \quad \alpha_i = \begin{cases} 
1 & \text{if } \sum_{j=1}^{N_{\text{exp}}} r_i(\theta)^2 \leq N_{\text{exp}} \sigma_i^2 \\
0 & \text{if } \sum_{j=1}^{N_{\text{exp}}} r_i(\theta)^2 > N_{\text{exp}} \sigma_i^2 
\end{cases} \forall j \tag{12}
\]

Similarly, an MBDM filter acting on model output variables whose associated residuals are incompatible with measurement noise is defined as in Equation (13) with the respective switcher conditions on the residuals (14).

\[
\hat{\theta}_{\text{DM,m}} = \arg \min_{\theta \in \Theta} \sum_{i=1}^{N_m} \sum_{j=1}^{N_{\text{exp}}} \beta_j r_i(\theta)^2 \tag{13}
\]

\[
s.t. \quad \beta_j = \begin{cases} 
1 & \text{if } \sum_{i=1}^{N_m} r_i(\theta)^2 \leq N_m \sigma_i^2 \\
0 & \text{if } \sum_{i=1}^{N_m} r_i(\theta)^2 > N_m \sigma_i^2 
\end{cases} \forall j \tag{14}
\]

If the fitting of the whole dataset \( \mathbf{y} \) results in unacceptably high residuals, the MBDM estimators defined above shall be employed to detect if the bad fitting is caused by specific experiments or by specific model output variables. The quantities \( \hat{\theta}_{\text{DM,exp}} \) and \( \hat{\theta}_{\text{DM,m}} \) represent robust maximum likelihood estimates obtained from the fitting of a possibly reduced dataset.

Notice that whenever an experiment or an output variable is labeled as an outlier by an MBDM estimator, no direct information is returned on whether the outlier is caused by gross measurement errors or whether the outlier is caused by an inappropriate mathematical structure of the model (5).

### 2.2 Robust model-based Design of Experiments

A parameter estimate \( \hat{\theta} \) is obtained maximizing a certain likelihood function. The values of these estimates depend on experimental data, which are affected by measurement errors. A complete characterization of the computed parameters shall also consider this aspect and the estimates shall be treated as random variables. The covariance matrix \( \mathbf{V}_\theta \) associated to the parameter estimates is well approximated by the inverse of the observed Fisher information matrix \( \mathbf{H}_\theta \):\(^{36}\)

\[
\mathbf{V}_\theta = \mathbf{H}_\theta^{-1} \tag{15}
\]

From the computed covariance matrix \( \mathbf{V}_\theta \) it is possible to assess the statistical significance of each parameter through a t-test. Whenever some parameter estimate is found to be statistically unsatisfactory, its precision can be improved collecting new experimental data. The quantity of information collected during an experiment, is strongly dependent on both the model equations and the investigated experimental conditions. MBDoE techniques for parameter precision have been proposed in the literature for designing highly informative experiments, bearing the richest information possible to achieve a certain goal with a limited experimental budget (e.g., maximum number of experiments allowed and/or maximum number of measurements permitted within an experiment).\(^{12}\) Central to MBDoE methods is the evaluation of the expected Fisher information matrix \( \mathbf{H}_\theta \). \( \mathbf{H}_\theta \) represents an approximation of the observed Fisher information matrix \( \mathbf{H}_\theta \) and does not require measured quantities to be computed. The quality of the approximation improves as the model becomes adequate to represent the process during the model identification task.\(^{36}\) Assuming that a single experiment has to be designed in which the values of \( N_m \) variables can be collected, then, under the assumption of normally
distributed measurement errors, the expected Fisher information $H_b$ is quantified as follows:

$$H_b = \sum_{i=1}^{N_0} \frac{1}{\sigma_i^2} \nabla \psi \nabla \psi^T$$

Equation (16) can be evaluated in the entire experimental design space adopting the best parameter estimate available. In order to summarize the multidimensional nature of $H_b$, the information is frequently evaluated with a convenient scalar measure $\psi$ of the matrix. Popular choices for $\psi$ are: the trace $\text{Tr}(H_b)$, the determinant $\text{Det}(H_b)$, and the largest eigenvalue.\(^{33}\) Whenever the model $f(\theta; X)$ is nonlinear in the parameters, the distribution of $\psi$ in the experimental design space is highly sensitive to the values of $\theta$ adopted for the evaluation of Equation (16). At the beginning of the experimental activity, some model parameters may be affected by a wide range of uncertainty (e.g., the values of parameters correlated to catalyst reactivity are typically affected by high uncertainty before any experiment is performed). In this situation, the aim of the scientist is planning preliminary experiments at robust conditions that can provide high Fisher information for all the possible values of the uncertain parameters.\(^{19,39}\) Optimal robust conditions $u_R$ are identified by solving the following min-max optimization problem:\(^{34}\)

$$u_R = \arg \max_{u \in U} \psi_{\text{MIN}}(u) \quad (17)$$

$$\psi_{\text{MIN}}(u) = \min_{\theta \in \Theta} \psi(u, \theta) \quad (18)$$

In Equation (17), $\psi_{\text{MIN}}$ is maximized acting on the experimental conditions $u$ for the worst case scenario, that is, considering the value for $\theta \in \Theta$ that minimizes $\psi$ across the experimental design space. In the proposed framework in Figure 1, the execution of kinetic experiments on catalysts with assorted levels of reactivity results in the estimation of different sets of parameters. The sets of kinetic parameters $\theta_{\text{CLR}}$ and $\theta_{\text{CHR}}$ identified for the catalyst with low reactivity (CLR) and for the catalyst with high reactivity (CHR) can then be used to define the expected hyperbox of variability for the kinetic parameters, that is, $\Theta = [\theta_{\text{CLR}}, \theta_{\text{CHR}}]$ in which the optimization problem (18) is solved.

An approximated definition of $\psi_{\text{MIN}}$ is used in this work to represent graphically the distribution of the information for the worst case scenario. For this purpose, the following assumptions are made: the input space $U \subset \mathbb{R}^2$ is assumed to be bi-dimensional for visualization purposes; it is assumed that the parameter values $\theta$ that are representative of the worst case scenarios (i.e., the parameters $\theta \in \Theta$ minimizing $\psi$ according to Equation (18)) are the kinetic parameters associated to the least reactive and to the most reactive catalyst, that is, the distribution of information $\psi_{\text{MIN}}$ representative of the worst case scenario is evaluated according to Equation (19):

$$\psi_{\text{MIN}}(u) = \min \left\{ \psi(u, \theta_{\text{CLR}}), \psi(u, \theta_{\text{CHR}}) \right\} \quad (19)$$

An example, derived from the case study presented in this manuscript, is reported in Figure 2 to illustrate the adopted procedure. In the top-left corner of Figure 2, a plot shows the information surface $\psi$ computed adopting the kinetic parameters resulting in the most divergent catalyst behavior: $\psi(u, \theta_{\text{CLR}})$ computed adopting the kinetic parameters $\theta_{\text{CLR}}$ associated with the catalyst with lowest reactivity and $\psi(u, \theta_{\text{CHR}})$, computed adopting the parameters $\theta_{\text{CHR}}$ related to the catalyst with highest reactivity observed in the experiments. The distribution of $\psi_{\text{MIN}}$ is then obtained according to Equation (19) and it is represented in the plot to the right of Figure 2. Experimental conditions along the line of intersection between the CLR and CHR information surfaces (i.e., red region in the plot on the right of Figure 2) represent a good compromise for beginning the experimental activity when the level of reactivity is still unknown, that is, the conditions with high information content for a wide range of catalyst behaviors.

3 | DATASETS—MATERIALS AND METHODS

Data collected with microreactors are analyzed for the purpose of identifying simplified kinetic models of methanol oxidation over silver. A description of the experimental setups is given in this section. The available datasets and the investigated experimental conditions are presented and a graphical assessment of the catalyst reactivity is proposed. The modeling assumptions are then illustrated and discussed. Eventually, the assumptions for the application of the methodology (Figure 1) to this specific case study are detailed.

3.1 | Experimental setups

Experimental data collected using five different silicon-glass microreactors were employed for the identification of kinetic models of methanol oxidation on silver. The silicon reactor chips were fabricated through photolithography and deep reactive ion etching. A thin film of silver catalyst was deposited on the bottom of the microchannels by sputtering. A schematic representation of the chips is given in Supporting Information Appendix A with a concise description of their geometry. Five datasets are distinguished according to the microreactor on which the experiments were performed. Three different geometries of reactors were employed in different years for the data collection: (a) wide-channel reactor C1 in 2008; (b) wide-channel reactor C2 in 2013; and (c) serpentine reactors R1, R2, and R3 in 2015. The catalysts employed in 2008, 2013, and 2015 went through different fabrication histories and were therefore expected to behave differently. Experiments were performed at steady-state conditions injecting the gaseous mixture containing the reactants (i.e., methanol and oxygen), water and helium (used as inert carrier), through mass flow controllers. The composition of the gaseous mixture at the outlet was detected using a ThermoQuest Trace and a Shimadzu gas chromatograph for experiments conducted on C1. Outlet compositions in experiments performed on reactors C2, R1, R2, and R3 were measured with an Agilent 7890A gas chromatograph. A more detailed
description of the reactors and the setups can be found in the literature.\textsuperscript{40}

3.2 Datasets description and comparison

The set of independent variables altered in the experiments were: inlet volumetric flowrate $F_{\text{IN}}^{\text{STC}}$ (STC: $T^* = 273 \, \text{K}; P^* = 101,325 \, \text{Pa}$); reactor temperature $T$; molar fractions of methanol, oxygen, and water at the inlet (i.e., $y_{\text{IN}}^{\text{CH}_3\text{OH}}, y_{\text{IN}}^{\text{O}_2}$, and $y_{\text{IN}}^{\text{H}_2\text{O}}$, respectively). The dependent output variables of the system were: outlet volumetric flowrate $F_{\text{OUT}}^{\text{STC}}$ (the values $F_{\text{IN}}^{\text{STC}}$ and $F_{\text{OUT}}^{\text{STC}}$ are different because of a change in the molarity of the mixture during the reaction); inlet and outlet pressure $P_{\text{IN}}^{\text{N}}$ and $P_{\text{OUT}}^{\text{N}}$; outlet composition of the gaseous mixture. Methanol, oxygen, water, formaldehyde, hydrogen, carbon dioxide, and carbon monoxide were the measured species (i.e., $y_{\text{OUT}}^{\text{CH}_3\text{OH}}, y_{\text{OUT}}^{\text{O}_2}, y_{\text{OUT}}^{\text{H}_2\text{O}}, y_{\text{OUT}}^{\text{CH}_2\text{O}}, y_{\text{OUT}}^{\text{H}_2}, y_{\text{OUT}}^{\text{CO}_2},$ and $y_{\text{OUT}}^{\text{CO}}$, respectively). Different ranges of experimental conditions were investigated in the setups. A summary of the values for the considered independent variables of the system is given in Table 1 (the detailed datasets including also the values measured for the output variables can be found in Supporting Information Appendix B). Methanol conversion at comparable conditions of temperature and inlet composition is plotted in Figure 3 for the various datasets as a function of the residence time. The plot in Figure 3 allows for a preliminary qualitative comparison of the datasets, demonstrating the presence of different levels of reactivity in the reactors. In this reacting system, catalyst reactivity is influenced by numerous independent factors that are complex to control and to model. Difficulties in controlling the deposition process of the silver film contribute to the final uncertainty on the catalyst behavior. Furthermore, reactivity is also sensitive to the experimental history, that is, during its lifetime, the catalyst undergoes an ageing process (typically a decay in the reactivity) with dynamics that depend on the experimental campaign conducted on the catalyst. Another source of uncertainty is due to the nonuniform deactivation along the catalyst film. The first part of the catalyst film is typically more stressed and undergoes a more rapid deactivation process than the rest of the film.\textsuperscript{41} Since different quantities of catalyst (i.e., different catalyst film lengths) were present in the five microreactor chips (see Supporting Information Appendix B).
Information Appendix A), the nonuniform deactivation may result in different observed dynamics of deactivation.

3.3 Reactor model

For the purpose of model identification, the microchannels were modeled as pseudo-homogeneous plug flow reactors. COMSOL simulations demonstrated that, at the tested experimental conditions, the silver catalyst film can be reasonably assumed to be isothermal and that transverse diffusion resistances are insignificant due to the small dimensions of the microchannel. Diffusion phenomena and energy balance are therefore neglected. Only the reactor part occupied by the catalyst film is modeled. The set of equations is given:

\[
\frac{d n_i(z)}{dz} = \sum_{j=1}^{N_C} a_j r_j(z) \quad \forall j = 1, \ldots, N_C \tag{20}
\]

\[
C_i(z) = \frac{\dot{n}_i(z)}{\sum_{j=1}^{N_C} \dot{n}_j(z)} \quad \forall j = 1, \ldots, N_C \tag{21}
\]

\[
r_j(z) = \alpha_k \prod_{i=1}^{N_C} C_i(z)^{\theta_i} \quad \forall j = 1, \ldots, N_R \tag{22}
\]

where \(C_i\) stands for the concentration of the \(i\)-th component in the mixture. \(\dot{n}_i\) is the molar flowrate of the \(i\)-th component in the mixture per unit of cross-sectional area expressed in mol/m²·s. \(N_C\) is the number of components considered in the gaseous mixture. The mass balance involves 7 species, that is, methanol, oxygen, water, formaldehyde, hydrogen, carbon dioxide and helium. \(N_R\) is the number of reactions included in the kinetic model. \(r_j\) is the reaction rate, expressed in mol/m²·s of the \(j\)-th reaction with kinetic constant \(k_j\) expressed per surface of catalyst and \(\alpha\) is the catalyst surface per unit of reactor.

### Table 1

Experimental conditions investigated in the experiments considered. The volumetric flowrate is expressed in STC. a Helium was used as balance.

| No. experiments | \(T\) (K) | \(F_\text{STC}^\text{IN}\) (mL/min) | \(y_\text{CH}_3\text{OH}^\text{IN}\) | \(y_\text{O}_2^\text{IN}\) | \(y_\text{H}_2\text{O}^\text{IN}\) |
|-----------------|----------|-------------------------------|-----------------|-----------------|-----------------|
| **C1—Wide-channel reactor—2008** | | | | | |
| 1–2 | 784–805 | 26.5 | 0.0982 | 0.0436 | 0.0744 |
| 3–6 | 764–826 | 26.6 | 0.1483 | 0.0606 | 0.1123 |
| 7–11 | 783 | 26.5 | 0.1407 | 0.0269–0.0962 | 0.0793 |
| 12–13 | 783 | 26.4 | 0.0684–0.1395 | 0.0445 | 0.0782 |
| 14–17 | 783 | 26.5 | 0.1042 | 0.0435 | 0.0184–0.2114 |
| **C2—Wide-channel reactor—2013** | | | | | |
| 1–4 | 783 | 13.7–58.8 | 0.099 | 0.0441 | 0.075 |
| **R1—Serpentine reactor—2015** | | | | | |
| 1–3 | 783 | 29.1–73.1 | 0.0996 | 0.0414 | 0.0754 |
| 4–7 | 733–826 | 50.9 | 0.1468 | 0.0975 | 0.2293 |
| 8–10 | 765–826 | 93.9 | 0.1469 | 0.098 | 0.2296 |
| 11–13 | 800–900 | 54.5 | 0.259 | 0.1064 | 0.2122 |
| **R2—Serpentine reactor—2015** | | | | | |
| 1–3 | 783 | 29.1–73.1 | 0.0996 | 0.0414 | 0.0754 |
| **R3—Serpentine reactor—2015** | | | | | |
| 1–3 | 783 | 29.1–73.1 | 0.0996 | 0.0414 | 0.0754 |
| 4–8 | 783–933 | 41.7 | 0.0997 | 0.0414 | 0.0755 |

aSTC are reference conditions of: temperature \(T^* = 273\) K; pressure \(P^* = 101,325\) Pa.

**Figure 3** Residence time impact on methanol conversion at 783 K for a fixed inlet composition. Conversion points are derived from experimental data collected on setups C1, C2, R1, R2, and R3.
3.4 | Assumptions for kinetic modeling

Andreasen et al formulated a micro-kinetic model of methanol oxidation on silver based on surface science studies. The authors provide numerical values for the parameters involved in the micro-kinetic model. Most kinetic parameters were either obtained from the literature or calculated using statistical thermodynamics. Only a small subset of parameters was estimated fitting the detailed micro-kinetic model to kinetic data. From the micro-kinetic model, Andreasen et al identified the presence of two limiting steps in the oxidation.

The first step is the intermediate methoxy decomposition that results in the generation of formaldehyde and hydrogen. The second step is the intermediate formate decomposition that results in the generation of carbon dioxide and hydrogen. Water is formed directly from methanol as a by-product of methoxy formation through an equilibrium reaction. From these considerations, the same authors also proposed a simplified power-law kinetic model, which involves only two lumped reactions. The simplified model was not validated directly against experimental data, but it is expected to be reasonably accurate at high temperature conditions (above 750 K), where the coverage of adsorbed species on the silver surface is negligible.

In this work, the two-reaction model derived by Andreasen et al is employed adding a third reaction of hydrogen oxidation to account for the low amounts of hydrogen detected at the outlet in the experiments. The stoichiometry and kinetics of the assumed reactions are given in Equation (23). The identification of this kinetic model requires the estimation of six kinetic parameters, that is, \( \theta = [A_1, A_2, E_{d1}, E_{d2}, E_{d3}] \).

\[
\begin{align*}
\text{CH}_3\text{OH} + \frac{1}{2}\text{O}_2 & \rightarrow \text{CH}_2\text{O} + \frac{1}{2}\text{H}_2 + \frac{1}{2}\text{H}_2\text{O} \\
\text{CH}_2\text{O} + \frac{1}{2}\text{O}_2 & \rightarrow \text{H}_2 + \text{CO}_2 \\
\text{H}_2 + \frac{1}{2}\text{O}_2 & \rightarrow \text{H}_2\text{O} \\
\end{align*}
\]

Different behaviors among catalysts are assumed to be a consequence of different active sites densities on the catalyst surface. In terms of model parameters, the active sites density is proportional to the pre-exponential factors of the catalytic reactions. The catalyst is primarily involved in promoting the partial methanol oxidation and formaldehyde oxidation, that is, reactions (1) and (2). Evidence reported in the literature suggests that the hydrogen oxidation reaction occurs slowly on the catalyst surface. Therefore, the catalytic effect of silver on reaction (3) is neglected and it is assumed that only parameters \( A_1 \) and \( A_2 \) are metrics of catalyst reactivity.

3.5 | Methods

Since different levels of reactivity were detected in the different microreactor chips, the five available datasets are treated independently for identifying five different sets of kinetic parameters. The steps detailed in this section follow directly from the general procedure presented in Figure 1. The criteria for selecting the reference dataset for identifying a reference set of parameters are presented and further detailed in Supporting Information Appendix D. The justification for the employment of MBDM methods for the identification of the kinetic parameters referring to the other datasets is then given. Eventually, the methods followed for evaluating the Fisher information across the design space and the impact of catalyst reactivity on its distribution will be illustrated. All the results presented in this work were obtained in the environment gPROMS ModelBuilder 4.1.

3.5.1 | Identification of the reference model

Dataset C1 is chosen as benchmark dataset for estimating the reference set of parameters \( \theta = [A_1, A_2, E_{d1}, E_{d2}, E_{d3}] \). The detailed explanation for the choice of C1 is reported in Supporting Information Appendix D. The model responses considered for data fitting are the outlet molar fractions of methanol, oxygen, water, formaldehyde, hydrogen, and carbon dioxide (i.e., \( N_m = 6 \)). A maximum likelihood method is employed assuming normally distributed measurement errors with a SD equal to \( \sigma = 3 \times 10^{-3} \) for \( i = 1, ..., N_m \) and \( j = 1, ..., N_{\text{exp}} \). Parameters are estimated employing the solver MAXLKHD implemented in gPROMS ModelBuilder 4.1. The quality of the solution achieved by MAXLKHD was ensured solving the parameter estimation problem multiple times adopting different initial guesses.

3.5.2 | Quantification of catalyst reactivity

It is assumed that variability in the catalytic behavior among the datasets can be explained by different active site densities. As a direct consequence of this assumption, only parameters \( A_1 \) and \( A_2 \) have to be re-estimated for adapting the reference set of parameters identified for C1 to the other datasets. Instead of estimating directly \( A_1 \) and \( A_2 \), two scaling parameters \( \kappa_1 \) and \( \kappa_2 \) are added to the reaction rates:

\[
\begin{align*}
A_1 & \rightarrow \kappa_1 A_1 \\
A_2 & \rightarrow \kappa_2 A_2 \\
\end{align*}
\]

In the present case study, since all microchannels are characterized by the same depth (see Supporting Information Appendix A for more details on the geometry of the microchannels), parameter \( \alpha \) is the same for all setups and will be therefore not considered explicitly in the following kinetic analysis.
At this stage, kinetic parameters $A_1$, $A_2$, $A_3$, $E_{\text{a,1}}$, $E_{\text{a,2}}$, and $E_{\text{a,3}}$ are fixed to the values estimated for C1 while parameters $k_1$ and $k_2$ are manipulated for fitting the kinetic model (i.e., $\theta = [k_1,k_2]$). As highlighted in Supporting Information Appendix D, high measurement errors and deactivation phenomena occurred in the latest datasets. Furthermore, in some datasets, the range of experimental conditions investigated far exceeds the range of conditions fitted for the identification of the reference model (e.g., the range of temperatures investigated in C1 was 764–826 K while in R3 the range was 783–933 K). Since the kinetic model is approximated, it may not provide accurate predictions outside the range of conditions used for its identification. For these reasons, MBDM methods for parameter estimation (11) and (13) were employed in series to estimate the kinetic parameters. For both MBDM estimators, the hyperparameter $c$ was set equal to 3, which is equivalent to treating as outliers experiments and measured variables for which the average residuals exceed the ±3$r_{ij}$ range of measurement error. The MBDM problems were solved employing the MINLP solver CVP_SS, implemented in the software gPROMS ModelBuilder 4.1, running the optimizations multiple times with different initial guesses. This is done to prevent the solver from getting stuck on local solutions.

### 3.5.3 Analysis of the distribution of information

The following analysis focuses only on the kinetic models identified for the most reactive and the least reactive catalysts towards reaction (1) (i.e., models with highest and lowest $k_1$ value in Equation (24)). These are selected as representatives of a CLR with kinetic parameters $\theta_{\text{CLR}}$ and a CHR with parameters $\theta_{\text{CHR}}$ respectively. The trace of the information matrix (16) is plotted in the experimental design space to evaluate the expected Fisher information associated to a kinetic experiment. The analysis aims at: 1) understanding the impact of the reactivity on the information obtainable from a kinetic experiment; 2) computing approximately the distribution of the information for the worst case scenario, that is, the distribution of $\psi_{\text{MIN}}$ in the special case of a bi-dimensional design space. With reference to aim 1, it is chosen to limit the investigation to three cases assuming different bi-dimensional design spaces and a fixed reactor geometry:

- Case 1: the information surface for CLR and CHR is evaluated in the design subspace identified by reactor temperature and residence time at fixed inlet composition;
- Case 2: the information surface for CLR and CHR is evaluated in the design subspace identified by reactor temperature and inlet ratio methanol/oxygen;
- Case 3: the impact of the inlet molar fraction of water on the information surface for CHR is assessed in the experimental design subspace identified by temperature and residence time.

In Table 2 values and ranges of variation for the input variables are given. Residence time $r$ is defined as the ratio between the catalyst-containing reactor volume and the inlet flowrate at standard conditions (i.e., at temperature $T^* = 273$ K and pressure $P^* = 101,325$ Pa). In all cases, a uniform pressure profile was assumed along the reactor with $P = 160,000$ Pa and the inlet molar fraction of oxygen was set at $y_{\text{O}_2}^\text{IN} = 0.044$.

### 4 RESULTS

#### 4.1 Reference model

The parameter estimates obtained fitting dataset C1 are given in Table 3 with related statistics, expressed in terms of $t$-values. A $t$-test performed on the estimates showed that the only parameter that was not estimated with satisfactory precision is the activation energy of the third reaction $E_{\text{a,3}}$. Other parameters passed the $t$-test. The parity plot in Figure 4 compares measurements against model predictions. The distribution of the points in proximity of the diagonal and within the error range demonstrates the good fitting achieved by the proposed kinetic model for the considered dataset.

#### 4.2 MBDM for catalyst reactivity quantification

In this section, the kinetic model identified for C1 is employed as reference, together with MBDM estimators (11) and (13) for the effective

| Case ID | $T$ (K) | $\tau$ (ms) | $y_{\text{H}_2\text{O}}^\text{IN}/y_{\text{O}_2}^\text{IN}$ | $y_{\text{H}_2\text{O}}^\text{IN}$ |
|---------|---------|-------------|----------------------------------|------------------|
| Case 1  | 764–834 | 1.5–12      | 2.25                             | 0.074            |
| Case 2  | 764–834 | 4.5         | 1.1–2.9                          | 0.074            |
| Case 3  | 764–834 | 1.5–12      | 2.25                             | 0.05–0.20        |

**TABLE 3** Maximum likelihood estimates obtained fitting dataset C1. The quantity $t_{\text{ref}}$ indicates the $t$-value of reference derived from a Student’s distribution with 95% of significance.

| Parameter | Estimation   | $t$-value$^a$ ($t_{\text{ref}} = 1.67$) |
|-----------|--------------|---------------------------------------|
| $A_1$ (mol/m$^3$·s$^{-1}$) | 5.33 · 10$^{11}$ | 158.9                                 |
| $A_2$ (s$^{-1}$) | 1.03 · 10$^{7}$ | 13.41                                 |
| $A_3$ (mol/m$^3$·s$^{-1}$) | 1.07 · 10$^{4}$ | 36.19                                 |
| $E_{\text{a,1}}$ (J/mol) | 1.42 · 10$^{5}$ | 7.39                                  |
| $E_{\text{a,2}}$ (J/mol) | 9.02 · 10$^{4}$ | 1.88                                  |
| $E_{\text{a,3}}$ (J/mol) | 1.83 · 10$^{4}$ | 0.43$^a$                              |

$^a$A $t$-value lower than $t_{\text{ref}}$ indicates poor parameter precision.
quantification of catalyst reactivity in datasets C2, R1, R2, and R3. The kinetic model (24) was initially fitted to the datasets adopting a conventional maximum likelihood method, thus fitting all the experimental data available. Parameter estimates computed for each dataset are given in Table 4. A comparison between measurements and model predictions is given in Figure 5, where residuals associated to all the datasets are shown. As one can see from the parity plot, a number of data points lay outside the error range. The plot highlights the impossibility of representing all the data points with the given model.

MBDM estimators (11) and (13) were employed in series for obtaining a more significant estimation of parameters $\kappa_1$ and $\kappa_2$. The parameter estimates obtained from the application of the two-stage MBDM method to datasets C2 to R3 are given in Table 5. No critical data were identified for the sets collected on C2 and R2 therefore leading to the same parameter estimates obtained with the conventional likelihood method. Regarding dataset R1, no weaknesses in the fitting of specific measured species were highlighted. However, as one can see from Table 5, it was not possible to fit the data collected in experiments 4, 8, 12, and 13. An interpretation for the exclusion of these experiments is now given. The operating temperatures set in these trials were 733, 765, 850, and 900 K, respectively. These values of temperature lay outside the extrema of the range of temperature investigated in C1 (see Table 1). Since the model represents an approximation of the comprehensive mechanism, it may not provide accurate predictions outside the range of data fitted for its identification. Concerning R3, only experiments 1–3 were retained, while trials 4–8 were excluded from the parameter estimation by the solver. The incompatibility between experiments 1–3 and experiments 4–8 (for which a graphical assessment is given in Supporting Information Appendix D), is interpreted as a consequence of a reduction in the catalyst reactivity that occurred during the campaign of experiments. Also notice that the reference model was found unable to describe the molar fraction of water at the outlet, which was therefore neglected for the purpose of parameter estimation. Estimates for $\kappa_1$ and $\kappa_2$ in R3 were obtained fitting only three experiments, that is, only three distinct experimental conditions. The fitting of the reduced dataset results in a negative t-test for parameter $\kappa_2$ in R3. Since it is known that the model is identifiable, the statistical quality of $\kappa_2$ in R3 can be effectively enhanced by the execution of additional experiments designed employing MBDoE methodologies for parameter precision.

In Figure 6, a parity plot shows the distribution of the residuals for all the fitted data. The application of MBDM led to the automated detection of the data causing the bad fitting and the identification of a model characterized by better descriptive capabilities. Comparing the parameter estimates obtained with the two methods in Tables 4 and 5, it can be noticed that MBDM allows for appreciating a regularity in the behavior of the catalyst used in 2015 in chips R1, R2, and R3. More specifically, if a conventional

| Reactor | Exp. excluded | Meas. excluded | $[\kappa_1, \kappa_2]$ | $t$-values | $t_{ref}$ |
|---------|---------------|----------------|------------------------|------------|----------|
| C1      | N/A           | N/A            | [1.00, 1.00]           | N/A        | N/A      |
| C2      | None          | None           | [2.83, 1.78]           | [6.52, 3.16]| 1.72     |
| R1      | None          | None           | [10.61, 7.12]          | [19.51, 15.39]| 1.66    |
| R2      | None          | None           | [8.82, 4.85]           | [6.50, 2.64]| 1.74     |
| R3      | None          | None           | [6.66, 2.00]           | [15.39, 3.12]| 1.68     |

*A t-value lower than $t_{ref}$ indicates poor parameter precision.
maximum likelihood estimator is employed, one may conclude that the catalysts in R1, R2, and R3 are characterized by a different level of reactivity, that is, significantly different estimates for $\kappa_1$ and $\kappa_2$ in R1, R2, and R3 (see Table 4). However, the silver catalysts sputtered on the microchannels of the serpentine reactors were synthesized in the same batch in 2015. Furthermore, the serpentine reactor chips were characterized by the same microchannel geometry (see Supporting Information Appendix A) and, although the catalyst film lengths were different, the same sputtering protocol was adopted. Comparable levels of reactivity were therefore expected in these setups. The application of MBDM allows for highlighting the fact that the catalyst sputtered in R1, R2, and R3 is actually the same, that is, similar values for $\kappa_1$ and $\kappa_2$ in R1, R2, and R3 (see Table 5), and that the different parameter estimates obtained with the maximum likelihood estimator are consequence of fitting data lying outside the domain of model reliability. Higher levels of reactivity with respect to C1 (i.e., $\kappa_1$ and $\kappa_2$ higher than 1), were detected in setups C2, R1, R2, and R3 (see Table 5).

### 4.3 Reactivity impact on the information content of the experiments

In this section, the impact of the catalyst reactivity on the information content of a kinetic experiment is assessed across the experimental design space. This is done evaluating the trace of the expected Fisher information matrix (16) with the kinetic parameters estimated for C1

| Reactor | Exp. excluded | Meas. excluded | $[\kappa_1, \kappa_2]$ | $t$-values | $t_{ref}$ |
|---------|---------------|----------------|----------------------|----------------|----------|
| C1      | N/A           | N/A            | [1.00, 1.00]         | N/A            | N/A      |
| C2      | None          | None           | [2.83, 1.78]         | [6.52, 3.16]   | 1.72     |
| R1      | 4, 8, 12, 13 | None           | [7.46, 5.98]         | [14.63, 11.26] | 1.67     |
| R2      | None          | None           | [8.82, 4.85]         | [6.50, 2.64]   | 1.74     |
| R3      | 4, 5, 6, 7, 8 | Water          | [8.42, 3.18]         | [9.85, 1.23$^*$] | 1.77     |

*A $t$-value lower than $t_{ref}$ indicates poor parameter precision.

Table 5: Results given by the application of the two-stage MBDM filtering method for each reactor. Excluded experiments and species are presented together with the computed parameters.

**FIGURE 6** Parity plot showing the quality of fitting associated to datasets C1, C2, R1, R2, and R3 if MBDM is employed. Only fitted data points are reported in the plot. MBDM, model-based data mining

**FIGURE 7** Information matrix trace distribution in the 2-dimensional experimental design subspace identified by the reactor temperature and residence time (Case 1), referring to a CLR (i.e., a C1-type catalyst) and a CHR (i.e., a R1-type catalyst). CHR, catalyst with high reactivity; CLR, catalyst with low reactivity [Color figure can be viewed at wileyonlinelibrary.com]

**FIGURE 8** Information matrix trace distribution in the 2-dimensional experimental design subspace identified by temperature and inlet molar ratio methanol/oxygen (Case 2). Information surfaces refer to a CLR (i.e., a C1-type catalyst) and a CHR (i.e., a R1-type catalyst). CHR, catalyst with high reactivity; CLR, catalyst with low reactivity [Color figure can be viewed at wileyonlinelibrary.com]
and R2, selected as representatives of a CLR and a CHR (θCLR and θCHR) respectively. These correspond to the parameter instances associated to the lowest and highest κ1 value. The same analysis considering the model associated to C1 and R1, that is, lowest and highest κ2 value, was found to give similar results and is not reported here.

4.3.1 | Case 1: Effect of reactor temperature and residence time

The information surfaces obtained in Case 1 are given in Figure 7. The optimal experimental conditions in the presence of a CLR and a CHR, identified by the crests of information in the plot, are located in different positions of the design space. The plot highlights the fact that information is particularly sensitive to catalyst reactivity and that robust MBDoE methodologies shall be employed to take this aspect into account for the design of informative kinetic experiments. The kinetic mechanism associated to a CLR shall be investigated adopting longer residence time and higher temperature with respect to a CHR. Being the peaks of information clearly distinct, the study of a CHR through the investigation of CLR optimal conditions (and vice versa), results in a dramatic reduction in the expected information. The intersection between the surfaces highlights the presence of an indifference curve (region of robust design) in the experimental design space, where conditions that are insensitive to catalyst reactivity are identified.

4.3.2 | Case 2: Effect of reactor temperature and methanol/oxygen inlet ratio

The effect of a variation on both temperature and inlet ratio methanol/oxygen was assessed for both a CLR and a CHR. Information surfaces are plotted in Figure 8. Distinct surfaces for the considered

![Figure 9](https://example.com/figure9.png)
values of reactivity were obtained also in this case. The analysis on information carried out for Case 2 still demonstrates the importance of a thoughtful choice of the experimental conditions and that this choice shall take into account the reactivity of the specific catalyst under analysis. For a fixed methanol/oxygen inlet ratio, higher temperature shall be chosen in the presence of a CLR with respect to a CHR. Both information surfaces are positively affected by an increase in the inlet molar ratio methanol/oxygen. Thus, rich mixtures shall always be adopted regardless of the reactivity.

4.3.3 | Case 3: Effect of inlet water concentration on information

In this case, the impact of different inlet fractions of water through the analysis of the effect produced on the information surface in the design space identified by temperature and residence time is assessed. Only the CHR case was considered. As one can see from Figure 9, an increase in the inlet fraction of water causes the information wave to shift towards the high temperature region of the design space. An interpretation for the shift of the information wave is now given. Reaction (1) is inhibited by the presence of water (see Equation (23)). Since reaction (1) generates the reactants for the other reactions included in the mechanism, adding water at the inlet slows down the overall process. If all the other experimental conditions are fixed and the amount of water in the system is increased, higher temperature is required to achieve the optimal level of reactant conversion and product generation required to maximize the expected information from the experiment.

4.3.4 | Robust design of kinetic experiments

In the previous section, it was highlighted that the distribution of information across the design space is highly sensitive to catalyst reactivity. This high sensitivity justifies the employment of robust experimental design methodologies. The existence of robust experimental conditions is demonstrated by the fact that information surfaces associated with CLR and CHR intersect in the design space (see Figures 7 and 8). The distribution of $\psi_{MIN}$ associated to Case 1 is shown in Figure 10. This represents the distribution across the design space of the minimum amount of information expected in the worst case scenario. The red-colored area in the plot of Figure 10 identifies robust experimental conditions where high information (i.e., high $\text{Tr}(\hat{H}_0)$) is expected for a wide range of catalyst reactivity if the inlet composition $[y_{\text{CH}_3\text{OH}}^\text{IN}, y_{\text{O}_2}^\text{IN}, y_{\text{H}_2\text{O}}^\text{IN}] = [0.098, 0.044, 0.074]$ is adopted and a uniform pressure profile along the reactor, with $P = 160,000$ Pa, is realized. Experimental conditions such as those highlighted by the red-colored area in Figure 10 represent the best compromise for collecting preliminary information whenever there is high uncertainty on the behavior of the catalyst.

FIGURE 10 Distribution of $\text{Tr}(\hat{H}_0)$ in the worst case scenario across the reduced experimental design space identified by: (1) temperature; (2) residence time. The colormap was obtained assuming that all the other experimental conditions are fixed. The red-colored area identifies robust experimental conditions where high Fisher information is expected for a wide range of catalyst behaviors, that is, in the range of reactivity delimited by a C1-type catalyst with low reactivity (CLR) and a R1-type catalyst with high reactivity (CHR) [Color figure can be viewed at wileyonlinelibrary.com]

5 | CONCLUSION

The analysis of datasets of methanol oxidation over silver presented in this work demonstrated that a number of significant sources of uncertainty may be present during the investigation of this system: uncertain level of catalyst reactivity, measurement errors, catalyst deactivation and uncertain structure of the equations for explaining the observations. These uncertainties are as much part of the physical reality as the kinetic mechanism under study and it is indispensable to embrace them in the data analysis stage. A MBDM framework derived from robust regression methodologies has been proposed in this manuscript for explicitly dealing with the aforementioned sources of uncertainty. MBDM methods were demonstrated to be powerful tools for effectively estimating kinetic parameters in simplified kinetic models through the automated identification of the outliers. This, paired with careful experimental design can lead to significant reduction of the experimental effort required for achieving an effective modeling of kinetics and characterization of the catalyst behavior. In this work, kinetic data were analyzed for identifying a simplified kinetic model of methanol oxidation on silver catalyst. The wide range of catalyst reactivity detected in the datasets advocated for considerations on the design of experiments for improving the precision of the kinetic parameter estimates in the considered model. The impact of the reactivity on the location of the optimal experimental conditions was studied plotting the trace of the
Fisher information matrix in the design space. The presence of conditions that are insensitive to catalyst reactivity was highlighted to support the design of preliminary trials, when the catalyst behavior is still characterized by a high degree of uncertainty. These represent robust conditions with high expected Fisher information for a wide variety of catalyst behaviors. Future research activities will focus on further improving the proposed MBDM method as a tool for diagnosing the model descriptive limits.

ACKNOWLEDGMENT

The authors want to express great appreciation to the anonymous reviewers for their valuable comments.

ORCID

Marco Quaglio https://orcid.org/0000-0002-0293-6174
Fabrizio Bezzo https://orcid.org/0000-0003-1561-0584
Asterios Gavriilidis https://orcid.org/0000-0003-3508-5043
Enhong Cao https://orcid.org/0000-0002-5156-1882
Federico Galvanin https://orcid.org/0000-0002-3296-1581

REFERENCES

1. Reuss G, Distelfeld W, Gamer AO, Hilt A. Formaldehyde in Ullmann’s Encyclopedia of Industrial Chemistry. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA; 2000.
2. Wachs IE, Madix RJ. The oxidation of methanol on a silver (110) catalyst. Surf Sci. 1978;76:531-558.
3. Robb DA, Harriott P. The kinetics of methanol oxidation on a supported silver catalyst. J Catal. 1974;35:176-183.
4. Montoya A, Haynes BS. Methanol and methoxide decomposition on silver. J Phys Chem C. 2007;111:9867-9876.
5. Schubert H, Tegtmeyer U, Schlogl R. On the mechanism of the selective oxidation of methanol over elemental silver. Catal Lett. 1994;28:383-395.
6. Andreassen A, Lynggaard H, Stegelmann C, Stolzke P. A microkinetic model of the methanol oxidation over silver. Surf Sci. 2003;544:5-23.
7. Andreassen A, Lynggaard H, Stegelmann C, Stolzke P. Simplified kinetic models of methanol oxidation on silver. Appl Catal Gen. 2005;289:267-273.
8. Bartholomew CH. Mechanisms of catalyst deactivation. Appl Catal Gen. 2001;212:17-60.
9. McMullen JP, Jensen KF. Rapid determination of reaction kinetics with an automated microfluidic system. Org Process Res Dev. 2011;15:398-407.
10. Hunter WG, Reiner AM. Designs for discriminating between two rival models. Dent Tech. 1965;7:307-323.
11. Box GEP, Hill WJ. Discrimination among mechanistic models. Dent Tech. 1967:9:57-71.
12. Espie D, Macchiottio S. The optimal design of dynamic experiments. AIChE J. 1989;35:223-229.
13. Galvanin F, Macchiottio S, Bezzo F. Model-based design of parallel experiments. Ind Eng Chem Res. 2007;46:871-882.
40. Cao E, Gavriilidis A. Oxidative dehydrogenation of methanol in a microstructured reactor. Catal Today. 2005;110:154-163.

41. Wilkinson SK. Reaction Kinetics in Formulated Industrial Catalysts. [PhD thesis]. Birmingham, University of Birmingham; 2014.

42. Schlichting H, Gersten K. Boundary-Layer Theory. 8th ed. Berlin; New York: Springer; 2000.

43. Reif F. Fundamentals of Statistical and Thermal Physics. 1st ed. Long Grove, IL: Waveland Press, Inc.; 2009.

44. Dokuchits EV, Khasin AV, Khasin AA. Mechanism and kinetics of hydrogen oxidation on silver. Russ Chem Bull. 2012;61:2225-2229.

45. Process Systems Enterprise, gPROMS, www.psenterprise.com/ gproms 1997–2017.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Quaglio M, Bezzo F, Gavriilidis A, Cao E, Al-Rifai N, Galvanin F. Identification of kinetic models of methanol oxidation on silver in the presence of uncertain catalyst behavior. AIChE J. 2019;e16707. https://doi.org/10.1002/aic.16707