Variance-Based Global Sensitivity Analysis: A Methodological Framework and Case Study for Microkinetic Modeling

Bart F.H. van den Boorn[1]
b.f.h.vandenboorn@differ.nl

Matthijs van Berkel[2]
m.vanberkel@differ.nl

*Anja Bieberle-Hütter[1]
a.bieberle@differ.nl
+31 (0)40 3334 801

[1]DIFFER - Dutch Institute for Fundamental Energy Research, Electrochemical Materials and Interfaces, De Zaale 20, 5612 AJ Eindhoven, The Netherlands
[2]DIFFER - Dutch Institute for Fundamental Energy Research, Energy Systems and Control, De Zaale 20, 5612 AJ Eindhoven, The Netherlands

October 2022

Keywords — *Global Sensitivity Analysis, Electrochemistry, Oxygen Evolution Reaction, System Identification*

Abstract

Chemical models are built up from chemical reactions and parameters. Each of these parameters has a degree of uncertainty. Sensitivity analysis has proven to be an important tool to quantify and trace this uncertainty to specific input parameters. In this study, the methodology of a prominent global sensitivity analysis method, i.e. Sobol's variance-based method, is presented for chemical modeling with a focus on microkinetic modeling. Sobol's method is developed to be used as an analysis framework, which - once set-up for microkinetic modeling - can easily be used for different models.

This analysis framework is successfully demonstrated by means of two case studies from the field of microkinetic modeling: (1) CO oxidation, (2) oxygen evolution reaction (OER) at the photoanode in a photo-electrochemical cell. The results give insight into the influence of each input parameter on the output uncertainty. For CO oxidation, it is found that the temperature and chemisorption energies have most impact on the output. For the OER model, the valence band energy and solvent reorganization energy are most influential. Based on this, a workflow is proposed incorporating the sensitivity analysis into the modeling process, aimed at reducing the output uncertainty and at validating and optimizing the model.

Symbol list

$\mathbf{A}$	Left half of sampling sequence; $\mathbf{A} \in \mathbb{R}^{N \times k}$
$\mathbf{B}$	Right half of sampling sequence; $\mathbf{B} \in \mathbb{R}^{N \times k}$
X_i	Input parameter i; $X_i \in \mathbb{R}$
Y	Output parameter which equals $Y = f(X_1, X_2, \ldots, X_k)$; $Y \in \mathbb{R}$
$\mathbf{X}$	Matrix of all input parameters; $\mathbf{X} \in \mathbb{R}^{N \times k}$

1 Introduction

Chemical models are built up from chemical reactions and parameters. Each of these parameter has a certain degree of uncertainty. This uncertainty is reflected in the modeling results. Sensitivity analysis has proven to be an important tool to quantify and trace the effect of these uncertainties to specific model parameters and to validate models.[1] Over the years, with ongoing increase of computational power, sensitivity analyses have become more popular.[2] Within the field of electrochemistry, and particularly in microkinetic modeling, sensitivity analyses have been used, however infrequently. Examples include: The use of the Morris method on models of photochemical degradation of ethene in the atmosphere and the combustion of methane identifying the most sensitive reactions;[3] Application of global sensitivity analysis based on Sobol's method to identify the pertinent reaction steps in a model describing the dehydrogenation reaction kinetics of methanol from various metal surfaces;[4] Use of local sensitivity analysis to develop ethanol steam reforming models at reduced computational costs;[5] Implementation of global sensitivity analysis to highlight the correlation in energies of species and reactions and the uncertainty of current Density Functional Theory (DFT) methods.[6,7] These examples show the benefits of sensitivity analysis for their respective applications: It indicates which parameters or reactions have the largest impact on the model, and as such highlight where improvements can be made. Furthermore, it can identify how uncertainty propagates in the model, and provide a confidence bound to the model results.

In this study, we demonstrate the advantages of global sensitivity analysis in the broader context of microkinetic modeling and present the procedure of a prime global sensitivity analysis method, the Sobol's method, for chemical modeling by means of two case studies. In Sobol's method, the variance of the model output parameter, i.e. the parameter obtained by execution of model simulations, which could for example be the current density, is partitioned in Sensitivity Indices (SIs) according to fractional contribution of the input parameters.[8,9] Here, the set of input parameters includes all parameters chosen to be analyzed and can refer to both (1) tuneable parameters, which are parameters that can be set and fixed in an experiment, e.g. illumination intensity, and (2) fixed parameters that are typically constants within the system, but due to a lack of knowledge have a range of uncertainty, e.g. rate constants. We aim to illustrate that performing sensitivity analysis is not necessarily a complex procedure and can provide beneficial model insights. To this end, we introduce an easily adaptable framework for the implementation of Sobol's method, which is disconnected from the model except for a number of inputs and outputs, see Figure 1. The components of this figure are discussed in Section 2. By developing a framework, the method can be applied to any chemical model. We demonstrate the analysis implementation and highlight typical formats of the analysis results, by means of two case studies from the field of microkinetic modeling, a state-space model of CO oxidation and the oxygen evolution reaction (OER) in water-splitting electrolysis. Additionally, we introduce visualizations that can be generated from data resulting from the sensitivity analysis.

The structure of this work is as follows: Section 2 describes our motivation for use of Sobol's method as well as its implementation and the sampling method using the Sobol' sequence. Section 3 describes the model and results of the first case study, the CO oxidation model. Subsequently, Section 4 describes the

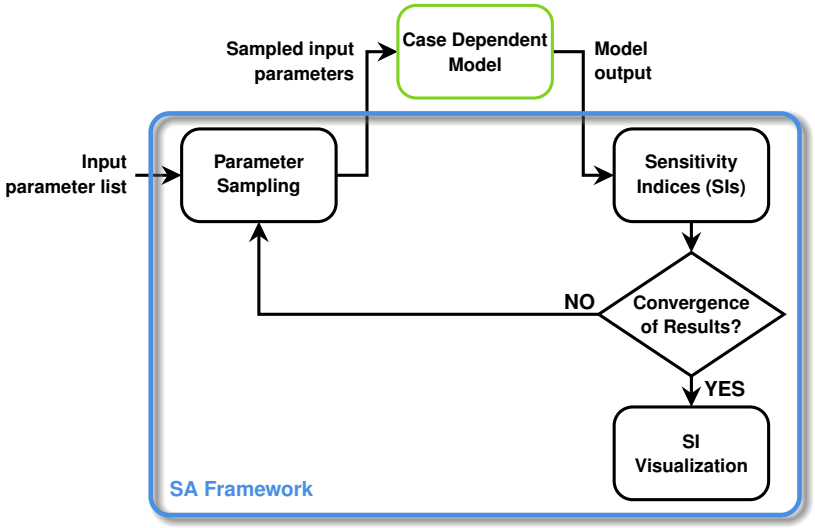


Figure 1: Sensitivity Analysis (SA) process chart with the generally applicable SA framework and the case dependent model.

second case study, the model of the OER in water-splitting electrolysis. A concluding discussion is given in Section 5.

2 Sobol's method

In this section, we motivate the use of Sobol's sensitivity analysis method and explain its implementation. The explanation is divided into three parts, represented by the components of the framework in Figure 1: (1) '*Sensitivity Indices (SIs)*': calculation of the SIs based on the contributions of input parameters to output parameter variance, (2) '*Parameter Sampling*': determination of the input parameters from a specified sampling sequence, and (3) '*Convergence of results?*': discussion of three different criteria for convergence of the SIs. These three parts are detailed in subsections of the same name. The visualization of the SIs is discussed by means of the case studies in Sections 3 and 4.

2.1 Motivation

Sobol's method is a global variance-based sensitivity analysis method.[8,9] Hence, the measure for sensitivity of the model is the variance of the model output parameters as a function of the input parameters. Variance is defined as the square of the standard deviation. The motivation for a variance-based method is twofold:

- Variance-based methods are global sensitivity analysis methods (GSA), which analyze the response of the model to variation of multiple inputs simultaneously.[1] Hence, these are able to detect interaction effects between multiple parameters. Additionally, GSA methods are suited to handle non-linear input/output relations and are independent of model monoticity.[1] On the other hand, in local methods, parameters are varied one-at-a-time. This is computationally less intensive, but leaves a large part of the input-space unexplored. This leads to an incomplete view of model sensitivity, as the parameter interactions are not identified.[10]

- The direct relation between the SIs and the

model output parameter variance is advantageous when the goal of the analysis is to reduce the output uncertainty by decreasing its variance, an approach commonly referred to as variance cutting.[11]

Sobol's method and FAST[12] (primarily the extended method eFAST)[13] are the most widely used global variance-based methods. Both methods generally give similar results.[14] We focus on Sobol's method in this study, as the computation of the higher order SIs that take the interactions into account is more straightforward and intuitive to implement compared to eFAST.[15,16] Sobol' method has been used earlier in fields ranging from environmental and hydrologic modeling, e.g. surface water runoff models,[11,17,18] crop growth models[19] and water distribution systems,[20] to biological/biomedical modeling, e.g. epidemic models[21] and models describing stem cell differentiation in the colon and colorectal cancer.[22] It has not been frequently used for microkinetic modeling so far.

2.2 Sensitivity indices (SIs)

The primary step of the sensitivity analysis procedure is the determination of the SIs. We describe this for a model of the form $Y = f(X_1, X_2, \ldots, X_k)$, with a single output parameter Y and input parameters X_i, where it is assumed that the inputs are independently distributed. All input parameters are combined in the array $\mathbf{X}$. The description given in this section follows the method in Saltelli et al.[23] An SI represents the degree to which an input parameter affects the variance of the output parameter, $V(Y)$. This output variance can be deconstructed into the variance contributions of each input separately, as well as their interactions, i.e. the variances that depend on a subset of the parameters,[23]

$$V(Y) = \sum_{i=1}^{k} V_i + \sum_{i=1}^{k} \sum_{j>i}^{k} V_{ij} + ... + V_{12...k}, \qquad (1)$$

where V_i is the direct contribution of parameter X_i to the output parameter variance and V_{ij} are interaction terms that depend on contributions of both input parameter X_i and X_j, etc. The direct variance contribution V_i is defined as,[23]

$$V_i \coloneqq V_{X_i}\left(\mathbb{E}_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)\right), \qquad (2)$$

where the term $\mathbb{E}_{\mathbf{X}_{\sim i}}\left(Y \mid X_i\right)$ denotes the mean value of output Y under variation of every input parameter except X_i, which is set to a fixed value. (Note that $\mathbf{X}_{\sim i}$ refers to all parameters except X_i.) The operation $V_{X_i}(\ldots)$ determines the variance of this mean when changing the given value for X_i. The interaction term V_{ij} is defined as

$$V_{ij} \coloneqq V_{X_i,X_j}\left(\mathbb{E}_{\mathbf{X}_{\sim i,j}}\left(Y \mid X_i, X_j\right)\right) - V_i - V_j, \qquad (3)$$

which performs the same operation on the combination of parameters X_i and X_j, minus their direct effects.

The SI which represents the direct contribution of each parameter, the **first-order (FO)** SI, is determined by dividing the direct contribution by the output variance,[23]

$$S_{Y,i}^{\mathrm{FO}} = \frac{V_i}{V(Y)}. \qquad (4)$$

Since V_i is a part of $V(Y)$, this sensitivity measure is a fraction, $S_{Y,i}^{\mathrm{FO}} \in [0, 1]$. The larger the fractional contribution of a particular input, the more sensitive the output parameter is to this input. Instead of fractions, SIs are commonly represented as percentage. The sensitivities of the model depend on which parameter is taken as output. Therefore, it is important to specify the output parameter for which the sensitivities are determined.

As seen in Equation (1), parameter X_i can also contribute to the output variance through interactions with other parameters. Therefore, the **total-effect (TE)** SI is used, which includes the direct effect of X_i, *plus* the effect of interactions of X_i with all other inputs X_j(,...,X_k).[23] It is given as

$$S_{Y,i}^{\mathrm{TE}} = \frac{V_i + V_{ij} + \cdots + V_{i..k}}{V(Y)} \coloneqq \frac{V_{Ti}}{V(Y)}. \qquad (5)$$

The two types of SIs represent respectively the highest ($S_{Y,i}^{\mathrm{TE}}$) and lowest ($S_{Y,i}^{\mathrm{FO}}$) measures for the con-

tribution of parameter X_i to Y, as the former includes the latter plus possible interaction terms between various inputs.[24] Consequently,

$$S_{Y,i}^{\mathrm{TE}} \geq S_{Y,i}^{\mathrm{FO}}. \tag{6}$$

2.3 Parameter sampling

The variance terms for the first-order and total-effect SIs are determined using a Monte-Carlo simulation, in which the input parameters are sampled from a range of values. These ranges of the different input parameters are defined prior to the analysis and can for example be obtained from experimental results or literature. A simple sampling method is taking samples randomly from the input parameter space. However, a quasi-random sampling sequence is preferred, as it explores the input parameter space more efficiently, thus requiring fewer samples, see Burhenne et al.[25] for comparisons. Furthermore, it allows for sequential addition of samples. As such, the number of samples can efficiently be extended until the parameter space is sufficiently explored and convergence of the analysis results is achieved, see Section 2.4 for more details. A number of quasi-random sequences have been developed, such as by Halton, Niederreiter, Faure and Sobol', of which some have been reviewed in Bratley et al.[26] In this work, we restrict to the Sobol' sequence as quasi-random sequence of choice.

A set of values is generated on the unit hypercube, in dimension equal to twice the amount of input parameters ($2k$). Initially, this set is generated N times, i.e. for N (initial) samples. A next set of N samples is generated from the Sobol' sequence at each additional sampling iteration. We then derive the sample variance terms following Saltelli et al.[23] As small example, we take $k = 2$ input parameters and $N = 2$ samples. The generated Sobol' sequence matrix of this sampling iteration yields

$$SOB_{\mathrm{seq}} = {}_{N}\downarrow \overbrace{\begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & x_{1,4} \\ x_{2,1} & x_{2,2} & x_{2,3} & x_{2,4} \end{bmatrix}}^{2k\ \longrightarrow} . \tag{7}$$

$$\underbrace{\phantom{x_{1,1}\ x_{1,2}}}_{\mathbf{A}} \quad \underbrace{\phantom{x_{1,3}\ x_{1,4}}}_{\mathbf{B}}$$

The first (left) and second (right) half of the sampled dimensions are used separately in the estimators of the variance terms. Hence, we define these as matrix $\mathbf{A}$ and $\mathbf{B}$, respectively, see Equation (7). From these matrices, a new set of matrices $\mathbf{A_B} = \{\mathbf{A}_\mathbf{B}^{(i)}\}$ is created, where $\mathbf{A}_\mathbf{B}^{(i)} = \mathbf{A}$ except column i is the i-th column of $\mathbf{B}$. In case of our example, this yields

$$\mathbf{A}_\mathbf{B}^{(1)} = \begin{bmatrix} x_{1,3} & x_{1,2} \\ x_{2,3} & x_{2,2} \end{bmatrix}, \tag{8}$$

$$\mathbf{A}_\mathbf{B}^{(2)} = \begin{bmatrix} x_{1,1} & x_{1,4} \\ x_{2,1} & x_{2,4} \end{bmatrix}, \tag{9}$$

where the terms are taken from $\mathbf{A}$ and from $\mathbf{B}$. The general matrix representation of the Sobol' sequence of Equation (7) and $\mathbf{A_B}$ (8, 9) for any number of input parameters is given in Supporting Information Section S1. The expressions of the variance terms for the first-order, V_i, and total-effect SIs, V_{Ti}, in Equations (4) and (5) respectively, are derived as[23,24]

$$V_i = V(Y) - \frac{1}{2N}\sum_{j=1}^{N}\left(f\left((\mathbf{B})_j\right) - f\left(\left(\mathbf{A}_\mathbf{B}^{(i)}\right)_j\right)\right)^2, \tag{10}$$

$$V_{Ti} = \frac{1}{2N}\sum_{j=1}^{N}\left(f\left((\mathbf{A})_j\right) - f\left(\left(\mathbf{A}_\mathbf{B}^{(i)}\right)_j\right)\right)^2, \tag{11}$$

where $(\mathbf{A})_j$ denotes the j-th row of $\mathbf{A}$.

2.4 Convergence of results

As the number of samples is increased and the parameter input space is further explored, the SIs converge towards a stable value. However, there are different definitions of what constitutes sufficient convergence. Three different types of convergence can be used:[11]

(a) Convergence of **screening**, which groups input parameters based on influence on output variance: influential or (approximately) non-influential parameters;

(b) Convergence of **ranking**, which sorts all parameters based on their influence;

(c) **Full convergence** of the SIs, which provides numerical expressions of the influence of each parameter.

Generally, the minimum number of samples increases with the three types of convergence. It is therefore useful to assess prior to analysis what kind of convergence is necessary. For example, in case of a complex model with significant run time, computational limitations could be an argument for the use of screening or ranking convergence. On the other hand, while convergence of ranking gives a good indication of which parameters are more influential than others, it does not encompass information on the difference in influence between subsequently ranked parameters. This information is only available with full convergence of the SIs. As the prominence of these advantages/disadvantages depends on the model and the use-case, such trade-offs have to be made on a case-to-case basis (refer to Section 5).

Each of the three convergence types can be quantified with different numerical convergence criteria, which we evaluate in the following.

2.4.a Convergence of screening

Parameter screening is achieved when the input parameters are categorized in a set of influential and a set of non-influential parameters with sufficient certainty. Generally, the group of non-influential parameters also includes parameters with a small influence, that are still categorized as negligible.[11,17,27] This requires a predefined influence threshold which splits the parameters into these groups. A possible measure for parameter screening convergence is that the sensitivity indices of the parameters within the group of non-influential parameters should have converged.[11] To this end, the same index convergence criterion and bootstrapping method can be employed as for the convergence of indices described in Section 2.4.c. To make sure that the obtained parameters are indeed non-influential, the screening results have to be validated, for instance through use of the Kolmogorov–Smirnov (KS) test, further detailed in Sarrazin et al.[11]

2.4.b Convergence of ranking

Convergence of ranking sorts parameters based on their influence on output variability. Akin to screening, ranking requires a measure which signals sufficient convergence of the ranking order. Various measures exist, such as Spearman's rank correlation coefficient.[28] However, these do not differentiate between rank reversals of high-influence and low-influence parameters, even though the latter are often of little interest. Hence, in this work we use a weighted rank correlation coefficient ρ,[11,29]

$$\rho = \sum_{i=1}^{k} \left(|R_i^m - R_i^n| \frac{\max_{m,n}\left(S_{Y,i}^m, S_{Y,i}^n\right)^2}{\sum_{i=1}^{k} \max_{m,n}\left(S_{Y,i}^m, S_{Y,i}^n\right)^2} \right), \quad (12)$$

where $S_{Y,i}^m$ and $S_{Y,i}^n$ are the SIs of X_i at subsequent sampling iterations m and n and R_i^m and R_i^n are the ranks of these SIs. The fractional term of the SIs in Equation (12) is the weighting term. This term is introduced such that the parameter influence on the correlation term is in proportion to the parameter sensitivity, increasing the importance of higher SI parameters, while downplaying the the inputs with smaller SI values.

Based on the correlation coefficient, a criterion for ranking convergence can be defined. A reasonable measure is that the 95-percentile of correlation coefficients of the sequentially increasing sample sets should fall below a critical value ρ_{crit},[11]

$$Q_{95\%}(\rho) < \rho_{\text{crit}}. \quad (13)$$

Setting $\rho_{\text{crit}} = 1$, ensures that the ranking of each parameter varies on average by at most 1 position.

2.4.c Full convergence of indices

To analyze full numerical convergence of the sensitivity indices, a confidence bound on the SI values is defined based on bootstrapping. Bootstrapping is a commonly used method to obtain a confidence measure of the degree in which the calculated SIs approach the 'true' SIs (i.e., when the SIs are determined from infinite samples).[30,31] In short, in bootstrapping, the samples taken so far, N_{T}, are resampled repeatedly with replacement: Every time, a set of N_{T} samples is randomly drawn from the original set with probability $1/N_{\text{T}}$, with the possibility that samples are repeated. The SIs calculated from these

resampled sets can be used to determine an approximation of the SI distribution. In this work, we use the 95-percentile interval of this bootstrap distribution as the confidence bound (c_x), i.e. under assumption that the bootstrap curve is normal, the 95-percentile lies within [$-c_x$,$+c_x$] from the bootstrap mean. In this case,

$$c_x = Q_{95\%}(s_x) \tag{14}$$

where s_x is the bootstrap SI sample standard deviation, x refers to either the first order or total effect indices calculated for each bootstrap resample $x = S_Y^{\mathrm{FO}},\ S_Y^{\mathrm{TE}}$. Since the error of both first order and total effect indices should fall within margin, we determine the convergence based on the maximum error,

$$c_{\max} = \max\{c_{S_{\mathrm{FO}}}, c_{S_{\mathrm{TE}}}\} \tag{15}$$

This value is compared to a threshold value c_{crit}. When the error value falls below the desired threshold, this is deemed sufficient convergence,

$$c_{\max} < c_{\mathrm{crit}}. \tag{16}$$

3 Case Study 1: CO oxidation

The first application case of Sobol's method in this study is a microkinetic model of the oxidation of carbon monoxide, which is a typical prototype reaction of heterogeneous catalysis.

3.1 Microkinetic model

The four-step reaction mechanism used in this study follows Falsig et al.[32] and consists of CO and O_2 adsorption, dissociative adsorption of O_2 and a Langmuir-Henshelwood type surface reaction to form CO_2.

$$\mathrm{CO\,(g)} + * \rightleftharpoons \mathrm{CO}^*, \tag{17}$$
$$\mathrm{O_2(g)} + * \rightleftharpoons \mathrm{O_2}^*, \tag{18}$$
$$\mathrm{O_2}^* + * \rightleftharpoons 2\,\mathrm{O}^*, \tag{19}$$
$$\mathrm{CO}^* + \mathrm{O}^* \rightleftharpoons \mathrm{CO_2(g)} + 2\,*, \tag{20}$$

where $*$ represents a free surface site.

The microkinetic model is based on Grabow et al.[33] and describes the evolution of the coverages of the reaction intermediates (CO, O_2, O) in a state-space system, which is generally represented as

$$\frac{\partial x(t)}{\partial t} = f(x(t),\, u(t)), \tag{21}$$
$$y = g(x(t),\, u(t)). \tag{22}$$

In this case, each state describes the evolution of one intermediate species, $x = [\theta_{\mathrm{CO}}, \theta_{\mathrm{O_2}}, \theta_{\mathrm{O}}]$, where θ_γ represents the surface coverage of intermediate γ ($\theta_\gamma \in [0, 1]$). The system is described by the following ordinary differential equations (ODEs),

$$\frac{\mathrm{d}\theta_{\mathrm{CO}}}{\mathrm{d}t} = r_1 - r_4, \tag{23}$$
$$\frac{\mathrm{d}\theta_{\mathrm{O_2}}}{\mathrm{d}t} = r_2 - r_3, \tag{24}$$
$$\frac{\mathrm{d}\theta_{\mathrm{O}}}{\mathrm{d}t} = 2r_3 - r_4, \tag{25}$$

where r_1, r_2, r_3 and r_4 represent the net rates of the four elementary steps in Equations (17), (18), (19) and (20), respectively. These rates are given as follows

$$r_1 = k_1^+ P_{\mathrm{CO}}\theta_* - k_1^- \theta_{\mathrm{CO}}, \tag{26}$$
$$r_2 = k_2^+ P_{\mathrm{O_2}}\theta_* - k_2^- \theta_{\mathrm{O_2}}, \tag{27}$$
$$r_3 = k_3^+ \theta_{\mathrm{O_2}}\theta_* - k_3^- \theta_{\mathrm{O}}^2, \tag{28}$$
$$r_4 = k_4^+ \theta_{\mathrm{CO}}\theta_{\mathrm{O}} - k_4^- P_{\mathrm{CO_2}}\theta_*^2, \tag{29}$$

where k_ι^+ and k_ι^- for $\iota = 1, 2, 3, 4$ refer to the forward and reverse rate constants, respectively. P_{CO}, $P_{\mathrm{O_2}}$ and $P_{\mathrm{CO_2}}$ represent partial pressures, which are predefined. The fraction of empty surface sites is represented by

$$\theta_* = 1 - \theta_{\mathrm{CO}} - \theta_{\mathrm{O_2}} - \theta_{\mathrm{O}}. \tag{30}$$

The reaction rate constants are determined using the standard Arrhenius equation, which relies on the pre-exponential factor ν_ι and activation energy E_{a}

$$k_\iota = \nu_\iota \exp\left(\frac{-E_{\mathrm{a},\iota}}{k_{\mathrm{B}}T}\right). \tag{31}$$

We can determine the pre-exponential factor from the change in entropy, $\Delta S_\iota^\ddagger$, between the initial state of the reaction and the transition state, [33]

$$\nu_\iota = \frac{k_\mathrm{B} T}{h} \exp\left(\frac{\Delta S_\iota^\ddagger}{k_\mathrm{B}}\right). \tag{32}$$

It is assumed here that this entropy change is zero for the surface reactions in Equations (19) and (20), as the reactants and transition states are adsorbed species. It is further assumed that the entropy change for reactions in Equations (17) and (18) amounts to a loss of 25% from the initial state, in line with Grabow et al. [33]

Following Falsig et al. [32], we consider non-activated and fast adsorption of CO and O_2. Furthermore, we assume that these reactions are quasi-equilibrated: The temperature range is high enough such that (also) the desorption is fast, and does not vary with temperature. Therefore, $\Delta E_{\mathrm{a},1} = 0$ and $\Delta E_{\mathrm{a},2} = 0$. The activation energy of reaction steps in Equations (19) and (20) are determined from scaling relations identified in Falsig et al. [32] and can be expressed in the chemisorption energies of CO and O exclusively,

$$\Delta E_{\mathrm{a},3} = 0.5 \cdot \Delta E_\mathrm{O} + 1.39\,\mathrm{eV} \tag{33}$$

$$\Delta E_{\mathrm{a},4} = -0.3 \cdot (\Delta E_\mathrm{O} + \Delta E_\mathrm{CO}) + 0.02\,\mathrm{eV}. \tag{34}$$

The values of these scaling relations are kept constant in this analysis. A more detailed SA, which is beyond the focus of this paper, could also include the influence of varying the scaling relation terms to assess the sensitivity towards the scaling relations.

We note furthermore that the model does not take into account the effect of lateral interactions of the adsorbates. Temperature and pressure affect the adsorbate configurations on a mesoscopic scale. [34] As such, these parameters have an additional effect on the adsorption energies. The lateral interactions could be studied by implementing different reaction rates dependent on the lateral interaction energies for all possible adsorbate configurations. Such microscopic lateral interaction energies have previously been determined using DFT for CO oxidation on Pt(111) surfaces with preadsorbed oxygen. [35] Such analysis can be implemented in a follow-up study.

3.2 Input parameters

For the sensitivity analysis of the CO oxidation model we vary six input parameters: The temperature T at which the oxidation takes place, the oxygen, carbon monoxide, and carbon dioxide pressures (P_{O_2}, P_CO and P_{CO_2}, respectively), and the chemisorption energies of CO and O (ΔE_CO and ΔE_O). The ranges of the values of the different parameters are summarized in Table 1. These values are substantiated by their occurrence in theoretical work, experiments or simulations, which are denoted in the reference column in Table 1. The amount of data points available for these parameters is limited. Therefore, we choose to make no presumptions about the probabilities of specific values and regard the probability distribution of the values as a continuous uniform distribution, bounded by the values defined in Table 1. Instead of, for instance, defining the probability as Gaussian distribution around a nominal value, or assuming a log-uniform distribution where the distribution is uniform along the logarithmic of the parameter. Some important notes on the parameter values and references are given in the following:

- We focus on high temperature CO oxidation, hence, the temperature ranges between $T = 300 - 800$ K.
- The partial pressures are assumed to be at industrial conditions, i.e. around the order of 1 bar, which is several orders higher than typical experimental conditions. [36,37]
- The chemisorption energies of CO and O have been varied around the values for platinum, which is found to be one of the most active catalysts for high temperature CO oxidation, $\Delta E_{\mathrm{O,Pt}} = -1.25\,\mathrm{eV}$ and $\Delta E_{\mathrm{CO,Pt}} = -1.22\,\mathrm{eV}$. [33]

3.3 Sensitivity analysis

The sensitivity analysis is conducted using MATLAB. The stiff ODE solver 'ode15s' is used at default settings to obtain a numerical solution of the model.

Table 1: List of parameters, descriptions and value ranges of the CO oxidation model sensitivity analysis.

Parameters	Symbol	Parameter description	Unit	Minimum	Maximum	References
Reaction conditions	T	Temperature	[K]	3.00×10^{2}	8.00×10^{2}	[38]
	P_{CO}	Pressure of CO	[bar]	1.00×10^{-2}	1.00×10^{1}	[36,37]
	$P_{\mathrm{O_2}}$	Pressure of O_2	[bar]	1.00×10^{-2}	1.00×10^{1}	[36,37]
	$P_{\mathrm{CO_2}}$	Pressure of CO_2	[bar]	1.00×10^{-2}	1.00×10^{1}	[36,37]
Descriptors	ΔE_{CO}	Chemisorption energy of CO	[eV]	-1.50	-1.00	[33]
	ΔE_{O}	Chemisorption energy of O	[eV]	-1.50	-1.00	[33]

The model is solved in the time domain until equilibrium is reached (~ 3 ms CPU-time per Quasi-Monte-Carlo point). Alternatively, the equilibrium can be found by setting the ODE derivatives to zero and solving the algebraic equations. Due to the limited complexity of the CO oxidation model and limited number of parameters, achieving full convergence of the sensitivity indices is achieved within a reasonable time frame. Hence, we skip the convergence of parameter screening and ranking and look directly at the fully converged indices. The sensitivity of a model is always determined with respect to a specific output parameter. Therefore, by choosing different output parameters for the analysis, different sensitivities of the model can be explored. We show this for the CO oxidation model by performing two sensitivity analyses, with respect to the surface coverage values of CO, θ_{CO}, and, O_2, $\theta_{\mathrm{O_2}}$, respectively. Another important model output parameter is the turnover frequency (TOF), which represents the rate at which chemical conversions occur, and is a measure for the efficiency of the catalyst.[39] The sensitivity analysis conducted with respect to the TOF is given in Supporting Information Section S3.

Regarding the model output θ_{CO}, the first-order (FO) and total-effect (TE) SIs are shown in Figure 2. Here, we have ensured convergence of the SIs to a confidence bound critical threshold of $c_{\mathrm{crit}} = 1\%$. We find that the temperature is the primary contribution to the output sensitivity, both directly (FO) and in interaction with other parameters (TE). The second greatest contribution is the chemisorption energy of CO, ΔE_{CO}. This is related to the fact that ΔE_{CO} is involved in two reactions, i.e. Equations (17) and (20), and is therefore a prominent term to determine the evolution of the surface coverage of CO in Equation (23).

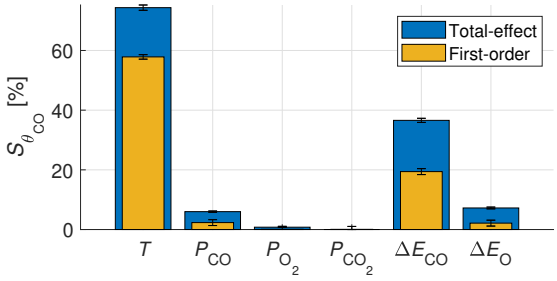


Figure 2: First-order and total-effect sensitivity indices $S_{\theta_{\mathrm{CO}}}$ of the CO oxidation model with respect to output θ_{CO} and the input parameters denoted in Table 1. Convergence is ensured to 95-percentile interval below $c_{\mathrm{crit}} = 1\%$.

Furthermore, we find that the partial pressure parameters $P_{\mathrm{O_2}}$ and $P_{\mathrm{CO_2}}$ have a minimal influence on θ_{CO}. This is primarily striking for $P_{\mathrm{CO_2}}$, since θ_{CO} depends directly on this partial pressure in Equation (26). However, the reverse rate constant of the surface reaction k_4^- is generally small, which diminishes the effect of $P_{\mathrm{CO_2}}$. Due to the negligible sensitivities to these pressures, any value within the respective range can be used without significant impact on the analysis. This is referred to as Factor Fixing.[1] To illustrate this, the two marginal partial pressures are resolved to their highest range boundary value and the sensitivity analysis is repeated against the four remaining parameters. The results are shown in Figure 3. Compared to Figure 2, only minor differences can be detected: The influence of the remaining

parameters is consistent with the aforementioned results (similar SI values), and the output variance is minimally impacted: 1.69×10^{-1} for the smaller parameter set (Figure 3) versus 1.66×10^{-1} for the larger parameter set (Figure 2). Note that the larger value of the variance for the smaller parameter set can be explained by the marginally small, *but not completely zero*, influence of P_{O_2} and P_{CO_2}: Depending on the chosen value, the variance can either slightly increase or slightly decrease. Resolving the partial pressures affects the analysis runtime, which has decreased by approximately 20% compared to the original analysis.

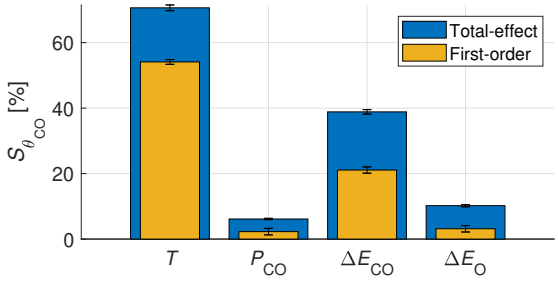


Figure 3: First-order and total-effect sensitivity indices $S_{\theta_{CO}}$ of the CO oxidation model with respect to output θ_{CO} and the input parameters denoted in Table 1, excluding P_{O_2} and P_{CO_2}. Convergence is ensured to 95-percentile interval below $c_{crit} = 1\%$.

Regarding the model output θ_{O_2}, the corresponding SIs are shown in Figure 4. The parameters T, ΔE_{CO} and ΔE_O have similar SIs, with a total effect of 60-70%. As in Figure 2, there is relatively small influence of the pressure parameters. However, in this case P_{O_2} is not negligible, likely due to the direct dependence in Equation (27). Furthermore, we find that in the analysis for θ_{O_2}, there is a large difference between first order and total effect SIs. This indicates significant interactions between the input parameters. Therefore, the uncertainty in the θ_{O_2} output cannot be attributed to a single input in particular but rather to a combination of inputs.

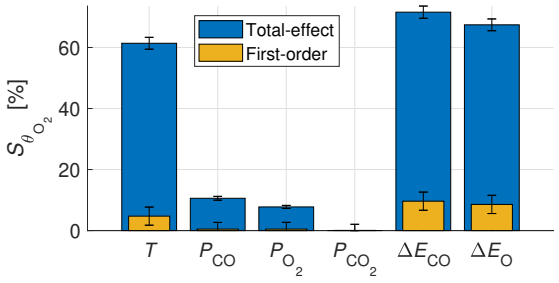


Figure 4: First-order and total-effect sensitivity indices of the CO oxidation model with respect to θ_{O_2} as output and the input parameters described in Table 1. Convergence is ensured to 95-percentile interval below $s_{crit} = 3\%$.

3.4 Sample visualization

The sensitivity indices are a quantitative expression of the influence of the input parameters. Hence, it does not show the impact of a specific input value on the output (whether it increases or decreases). This can be visualized using model samples as generated for Sobol's sensitivity analysis.[40] This type of qualitative analysis can indicate what parameter values lead to desirable system output.

A useful visualization method plots the sample input/output values for each input parameter. The distribution of these samples indicates the sensitivity to that input. Often, these patterns can instantly be recognized. However, in cases when the sample density is high overall, the number of data points might obstruct detecting patterns. We therefore propose a visualization method in which the samples are incorporated in a three-dimensional histogram. The height or color of the bars represent the amount of samples within the range of the bar (the number of samples per grid cell). Using the CO oxidation model as example, we visualize output samples of θ_{CO} against the sampled values of inputs T and P_{CO} in a three-dimensional histogram, see a) and b) in Figure 5, respectively. The absolute density values depend on the number of samples taken and is therefore of little relevance. Instead, the focus should be on the relative distribution.

In both examples in Figure 5, most samples are located around $\theta_{CO} \approx 0$ or $\theta_{CO} \approx 1$. However, for the variation of temperature, T (Figure 5a), a clear non-uniformity in the distribution on the Y-axis with respect to the X-axis can be observed. At low temperature, the surface is in most cases completely covered by CO ($\theta_{CO} \approx 1$), whereas at high temperatures, almost no CO is adsorbed on the surface ($\theta_{CO} \approx 0$). Similar sample visualizations for θ_{O_2} and θ_O are given in Supporting Information Section S2 and show relatively low coverage values over the entire temperature range, indicating that these species do not replace CO at high temperature. Instead, at high temperatures the surface sites are predominantly empty (Supporting Information Section S2). This can be expected: Following the Arrhenius equation, the oxidation reaction rate generally increases with the temperature. Hence, at a higher temperature, the CO would be completely converted into CO_2. The large temperature dependence shown here is directly related to the large SI of the temperature in Figure 2. Vice versa, a more uniform sample distribution such as observed for P_{CO} (Figure 5b), can indicate that a parameter is non-influential [1] (Note the small SI for P_{CO} in Figure 2).

4 Case Study 2: Oxygen evolution reaction (OER)

The second application case for sensitivity analysis using Sobol's method, is a model of water splitting in a photo-electrochemical cell (PEC). In this cell, the formation of hydrogen occurs at a metallic cathode, whereas the oxygen formation occurs at a semiconductor photoanode. The sustainable generation of hydrogen as energy carrier is a promising development in the shift towards renewable energy, as a way to deal with the intermittency of renewables, such as solar and wind energy.

4.1 Microkinetic model

The model for this case is derived from George et al. [41] and describes the OER occurring at the photoanode under alkaline conditions. The reaction steps are as follows,

$$* + OH^- + h^+ \underset{K_{b1}}{\overset{K_{f1}}{\rightleftharpoons}} OH^*, \tag{35}$$

$$OH^* + OH^- + h^+ \underset{K_{b2}}{\overset{K_{f2}}{\rightleftharpoons}} O^* + H_2O, \tag{36}$$

$$O^* + OH^- + h^+ \underset{K_{b3}}{\overset{K_{f3}}{\rightleftharpoons}} OOH^*, \tag{37}$$

$$OOH^* + OH^- + h^+ \underset{K_{b4}}{\overset{K_{f4}}{\rightleftharpoons}} {O_2}^* + H_2O, \tag{38}$$

where $*$ describes an adsorption site on the anode surface, and $*$ used as superscript, i.e. X^*, indicates a surface adsorbed species. These processes are captured in a 4^{th}-order state-space system of ordinary differential equations (ODEs), generally represented by Equations (21) and (22). The '4^{th}-order' terminology refers to the presence of four states in the system, which in this case are the surface coverages of the four intermediate species, $x = [\theta_{OH}, \theta_O, \theta_{OOH}, \theta_{O_2}]$. The evolution of each state is described by an ODE. In the analysis, the current density at equilibrium is chosen as the model output to which the sensitivity is determined, at fixed values of the applied potential u_{app}. The explicit expressions of the microkinetic model follow George et al. [41] and are denoted in Supporting Information Section S4.

4.2 Input parameters

The SA of the OER model investigates 14 input parameters, which are subject to varying degrees of uncertainty. The parameters, their ranges and corresponding literature references are denoted in Table 2. As explained for case study 1, we analyze the parameters based on a uniform probability distribution.

4.3 Sensitivity analysis

The OER model is more complex and contains more parameters than the CO oxidation model in Section 3. Therefore, we only omit the parameter screening step and first analyze the ranking of the parameters, before continuing with the full convergence of the SIs. The sensitivity analysis is conducted using the

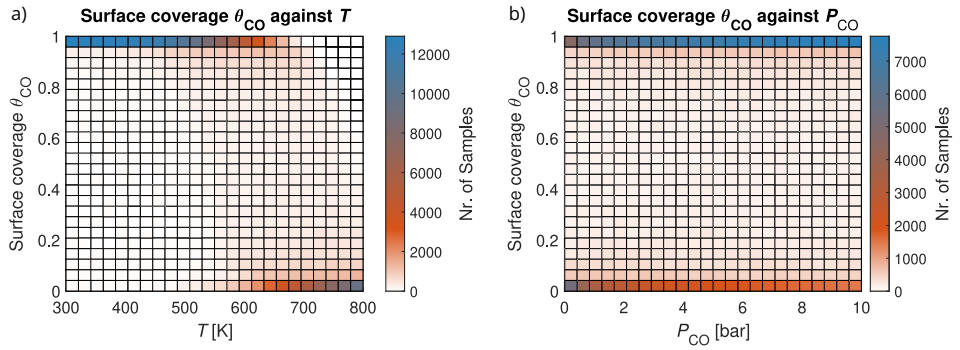


Figure 5: CO oxidation model input/output sample visualization using three-dimensional histogram. The output θ_{CO} is visualized against the parameter a) T and b) P_{CO}. The color represents the number of samples within the corresponding box. The sample distribution with respect to the parameter T is non-uniform, whereas the distribution against P_{CO} is approximately uniform.

Table 2: List of input parameters of the 4$^{\mathrm{th}}$ order oxygen evolution reaction model with lower and upper boundaries of their value ranges.

Parameters	**Symbol**	**Name**	**Unit**	**Minimum**	**Maximum**	**References**
Thermodynamic	pOH	Electrolyte pOH	-	0	1	[42–44]
	T	Temperature	[K]	293	303	[45]
	N_{V}	VB density of states	[cm^{-3}]	4.60×10^{19}	1.00×10^{22}	[46,47]
	n_{s0}	El. density in the dark	[cm^{-3}]	2.70×10^{17}	6.00×10^{20}	[46,48,49]
	p_{s0}	Hole density in the dark	[cm^{-3}]	1.00	1.00×10^{2}	[46,50]
	N_0	Adsorption site no. under r-SS	[cm^{-2}]	2.39×10^{14}	2.43×10^{15}	[48]
	λ	Solvent reorganization energy	[eV]	2.30	2.65	[51]
	ΔG_1	Gibbs free energy of OER step 1	[eV]	1.82	1.92	[48,52]
	ΔG_2	Gibbs free energy of OER step 2	[eV]	1.92	2.02	[48,52]
	ΔG_3	Gibbs free energy of OER step 3	[eV]	0.92	1.02	[48,52]
	ΔG_4	Gibbs free energy of OER step 4	[eV]	0.06	0.16	[48,52]
	E_{V}	VB energy level (hematite)	[eV]	2.40	2.70	[53–56]
Kinetic	$k_{\mathrm{vf,max}}$	VB max rate constant (forward)	[cm^4 s^{-1}]	1.00×10^{-17}	1.00×10^{-16}	[57]
	k_{f5}	Oxygen desorption rate	[s^{-1}]	9×10^{-12}	98	[58–61]

'ode15s' solver in MATLAB, with default settings. Using the method and the convergence criterion detailed in Section 2.4.b ($\rho_{\mathrm{crit}} = 1$), and at an applied voltage value of $u_{\mathrm{app}} = 1.3\,\mathrm{V}$, the parameter ranking is denoted in Table 3. A higher rank indicates a larger influence of on the output variability. Hence, the output has the strongest sensitivity to E_{V} and λ and the weakest sensitivity to N_{V} and n_{s0}. This ordering can be used as a basis for further parameter research: By decreasing uncertainty of parameters that drive the variability of the output the most, this variability can be decreased most efficiently. This is referred to as Factor Prioritization. [1,62]

Table 3: Parameter ranking of OER model at ranking convergence.

Rank	Parameter
1	E_{V}
2	λ
3	pOH
4	$k_{\mathrm{vf,max}}$
5	p_{s0}
6	N_0
7	T
8	ΔG_2
9	ΔG_1
10	K_{f5}
11	ΔG_3
12	ΔG_4
13	N_{V}
14	n_{s0}

The parameter ranking only shows the relative influence on the model output. Hence, it does not indicate the significance of the influence of each parameter. Hence, for further analysis, we look at the fully converged SIs. The SIs describing the sensitivity of the current density towards the 14 inputs, for an applied potential of $u_{\mathrm{app}} = 1.3\,\mathrm{V}$ are represented in Figure 6. We confirm that the parameters E_{V} and λ have both the largest first-order and total-effect SIs. Furthermore, we find in Figure 6, that six parameters have an approximately negligible influence, i.e. K_{f5}, N_V, n_{s0}, ΔG_1, ΔG_3 and ΔG_4. We find that performing the SA excluding these parameters does not significantly affect the output uncertainty (4.8×10^{-4} to 5.5×10^{-4}) nor the SIs (see Supporting Information Section S5), but it is advantageous because it decreases the runtime. Moreover, we find that some of the parameters that are not well-known from the literature, such as the oxygen desorption rate K_{f5}, have minimal influence on the model output. SA can thus indicate at an early stage that further research on a certain parameter is not of the highest priority to decrease output variance.

Additionally, we note a significant difference between first-order and total-effect indices, pointing towards prominent interactions between input parameters. To further the analysis, it could be of interest to derive second-order effect SIs to study which parameters have the strongest combined effects. These low-order effects have been shown to typically encompass most of the interaction effects.[63] This information can be used to combine subsets of the input parameters to obtain a smaller set of primitive parameters that have proportionately larger first-order effects. Such analysis is outside the scope of this work. In a following study, analysis of the second-order effects will be incorporated.

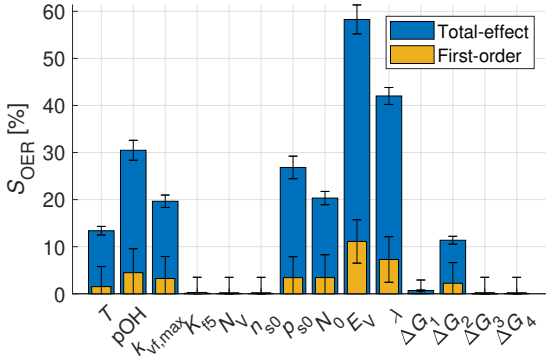


Figure 6: First-order and total-effect SIs S_{OER} of the parameters in the 4$^{\mathrm{th}}$ order OER model at applied potential of $u_{\mathrm{app}} = 1.3\,\mathrm{V}$. Convergence is ensured to a 95-percentile interval below $c_{\mathrm{crit}} = 5\%$.

Finally, as validation of the SA results, we look at the statistics of the convergence over the sampling iterations. The three convergence measures defined in Section 2.4, (a) screening, (b) ranking and (c) full convergence, are plotted against the sampling iterations, see Figure 7, with c_{max} the 95-percentile interval of the bootstrapped SIs, (Section 2.4.c) and $Q_{95\%}(\rho)$ the 95-percentile of rank correlation coefficients (Section 2.4.b). Convergence is achieved when the measures cross over a threshold (black dash-dot line). At the cross-over (red-dashed line) the required iterations N_{it} are indicated. For (a) and (c), this threshold is $c_{\mathrm{crit}} = 5\%$, i.e. the 95-percentile of bootstrapped samples deviate at most 5% from the original sample set. For (b), the threshold is $\rho_{\mathrm{crit}} = 1$, as discussed in Section 2.4.b. As expected,

we find that the full convergence of the indices is more costly, in a computational sense, than the convergence of ranking. Furthermore, we find that all three measures generally decrease with the sampling iterations. However, for (a) and (c), significant upward jumps can be observed, at which the error increases. This is likely due to parameter interactions: Specific combinations of samples, that might only appear at later sampling iterations, strongly affect the sensitivity indices. To substantiate this, the interactions ($S^{\mathrm{TE}} - S^{\mathrm{FO}}$) of each of the parameters have been plotted against the sampling iterations (Supporting Information Section S6). The jump in (c) at $N_{\mathrm{it}=338}$ corresponds with an approximate doubling of the interaction term of parameter p_{s0} while the other parameters are largely unaffected. For (b), we observe that the 95-percentile of the correlation coefficient decreases step-wise with the sampling iterations.

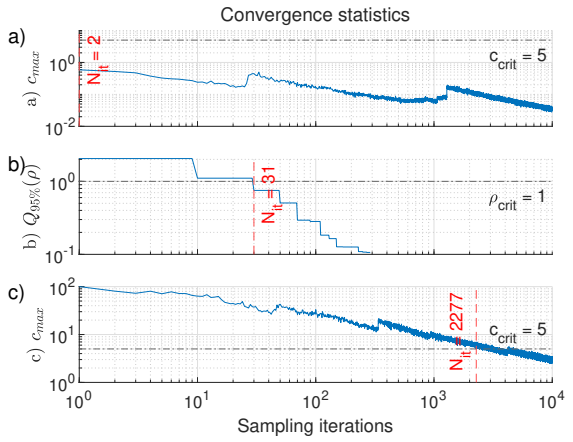


Figure 7: Screening (a), ranking (b) and full convergence (c) of the indices for the OER analysis, using their respective convergence measures. The convergence thresholds are indicated (black dash-dot), as well as the iteration number N_{it} at which the threshold is crossed (red dashed). For (a) and (c) c_{max} is the 95-percentile interval of the bootstrapped SIs. For (b) $Q_{95\%}(\rho)$ is the 95-percentile value of the weighted rank correlation coefficient. The three measures require increasing sampling iterations to convergence: (a) $N_{\mathrm{it}} = 2$; (b) $N_{\mathrm{it}} = 31$; (c) $N_{\mathrm{it}} = 2277$.

4.4 Sample visualization

In order to get an idea of the tendency of the current density with respect to the input parameters, the input/output samples are visualized. This is commonly referred to as Factor Mapping.[1] The output samples are plotted against the sampled input parameters at an applied voltage value of $u_{\mathrm{app}} = 1.3\,\mathrm{V}$, see Figure 8. This direct visualization of the samples shows for some parameters a strong non-uniformity of the sample distribution over different parameter values, e.g. for E_{V} and λ. Looking at E_{V} in detail, we see for example that the maximum value attainable value of the current density appears to increase exponentially with the value of E_{V}. From this, we can conclude that higher values of the current density can be achieved when the valence band energy level, E_{V}, is larger. Contrarily, there are other parameters for which the possible output values are not dependent on the specific parameter values, e.g. N_{V}. For this less-influential parameter, the distribution of samples across the current density range is approximately equal for all N_{V} values. These observations corroborate to the obtained sensitivity indices of the parameters. Such visualizations can be used to detect what parameter values lead to desirable or undesirable model output behavior.

5 Concluding Discussion

This study presents an explanation and implementation of Sobol's global variance-based sensitivity analysis method for chemical modeling. We successfully demonstrated the application of this method on two cases from the field of microkinetic modeling. Additionally, we present informative visualizations of the model sensitivity and the response of the model to varying inputs. The presented analyses are able to uncover essential information from the microkinetic models, such as:

- The parameters with large direct influence terms and/or the largest interaction terms: These parameters represent the largest contribution to the output uncertainty, which could be a motivation for further investigation of these parameters

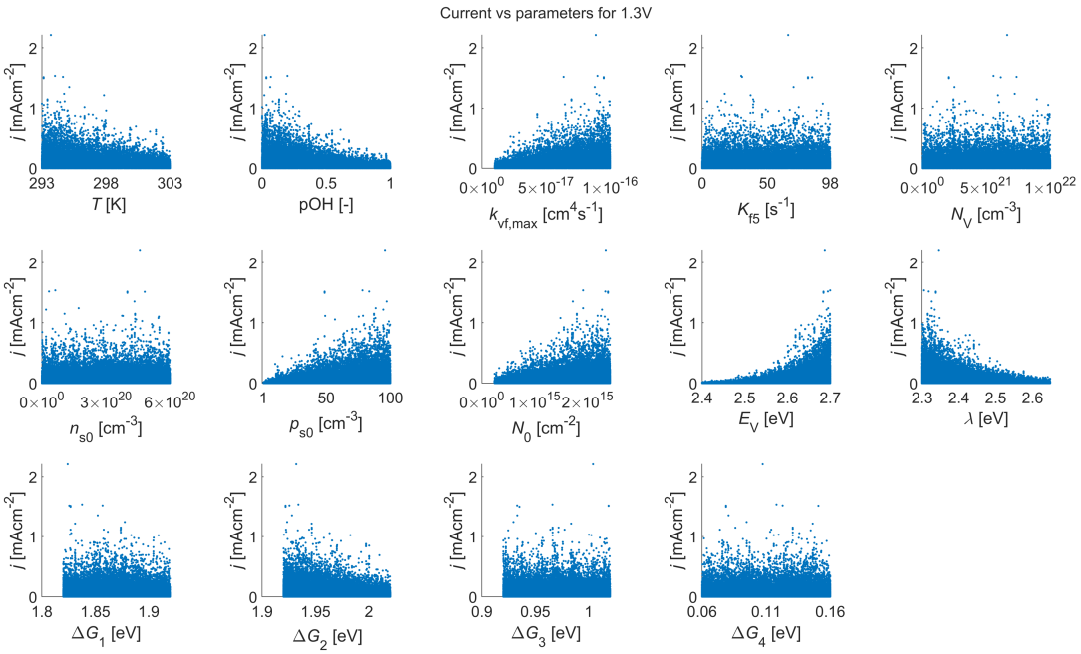


Figure 8: Monte-Carlo samples of model output, current density j, against input parameters at $u_{\mathrm{app}} = 1.3\,\mathrm{V}$.

both in modeling and in experiments.

- An overview of non-influential parameters: The effect of these parameters on the model are negligible, such that further efforts to decrease the uncertainty of these parameters do not affect the accuracy of the simulations. This could advocate for omitting these parameters when assigning often limited time resources for further study in the modeling process.

- The trend of the output with respect to each input parameter, if present: When aiming to improve the modeled system, this can be used to relate desirable or undesirable output behavior to parameter ranges of the input parameters.

Furthermore, in this study we highlight the impact of three commonly used criteria for the convergence of the analysis parameters: (a) parameter screening, (b) ranking and (c) full convergence of the sensitivity indices. For complex models, sensitivity analysis and, in particular, the global variance-based kind, is a computationally expensive and time costly operation. In order to perform sensitivity analysis in such cases, it could prove vital to pursue the computationally cheaper convergence of parameter screening or ranking, as well as resolving non-influential parameters (Factor Fixing). We propose the following workflow to decrease model output variance using sensitivity analysis in a computationally efficient manner (Figure 9):

1. Define initial parameter ranges of the input parameters to be included in the analysis.

2. Perform sensitivity analysis; Depending on the information required, choose between the convergence measures: (a) screening, (b) ranking, or (c) full convergence.

3. Check if the achieved output variance falls below the desired threshold for uncertainty.

4. *If no:* Refine the input parameters based on the choice in step 2; For (a) and (c): the identified non-influential parameters can be fixed to any value on their ranges, removing them from further analysis, and/or for (b) and (c): in-depth parameter research can be performed on the most influential parameters. Subsequently, repeat the analysis from step 2 until the output uncertainty is sufficiently small, i.e. until the uncertainty threshold is crossed.

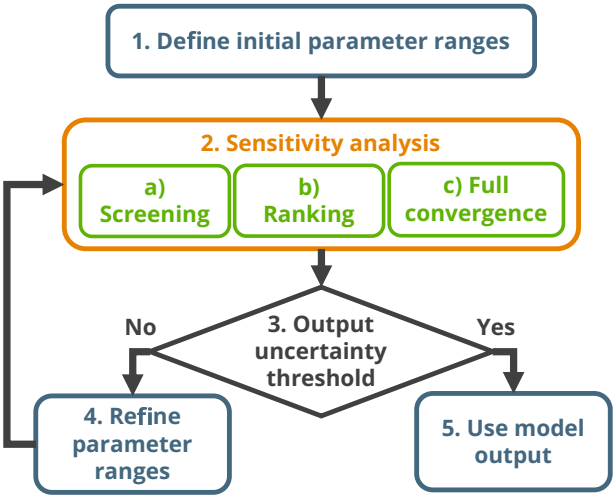


Figure 9: Analysis workflow to decrease model output variance using sensitivity analysis in a computationally efficient manner. The analysis is repeated when changes are made to the input parameter ranges.

5. *If yes:* The output uncertainty is below the desired threshold. The model output can be determined with a level of confidence based on this threshold.

In order to aid application of Sobol's method on other models within the field of chemical modeling, we have structured the method in this work as a separate framework, which does not inherently rely on a specific chemical model. The framework provides the model inputs for each Monte-Carlo simulation, and processes the model outputs into the sensitivity indices and visualizations. As such, other models can be analyzed in conjunction with this method, with only minor alterations to the modeling code.

6 Acknowledgement

Van den Boorn acknowledges funding from M-ERA.NET Call 2016 for the project "MuMo4PEC" with project number M-ERA.NET 4089. We would like to thank Dr. Oscar Diaz Morales for fruitful discussions.

References

[1] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, *Global Sensitivity Analysis. The Primer*, *Vol. 304*, Wiley, Chichester, England, **2008**.

[2] *Scopus - Analyze Search Results - Keyword: 'Senstivity analysis'*, **2022**.

[3] J. Zádor, I. Zsély, T. Turányi, *Reliab. Eng. Syst. Saf.* **2006**, *91*, 1232–1240.

[4] Z. C. Kramer, X.-K. Gu, D. D. Y. Zhou, W.-X. Li, R. T. Skodje, *J. Phys. Chem. C* **2014**, *118*, 12364–12383.

[5] J. E. Sutton, D. G. Vlachos, *Chem. Eng. Sci.* **2015**, *121*, 190–199.

[6] J. E. Sutton, W. Guo, M. A. Katsoulakis, D. G. Vlachos, *Nat. Chem.* **2016**, *8*, 331–337.

[7] S. Döpking, C. P. Plaisance, D. Strobusch, K. Reuter, C. Scheurer, S. Matera, *J. Chem. Phys.* **2018**, *148*, 034102.

[8] I. M. Sobol, *Math. Model. Comput. Exp.* **1993**, *1*, 407–414.

[9] I. M. Sobol, *Math. Comput. Simul* **2001**, *55*, 271–280.

[10] A. Saltelli, P. Annoni, *Environ. Modell. Software* **2010**, *25*, 1508–1517.

[11] F. Sarrazin, F. Pianosi, T. Wagener, *Environ. Modell. Software* **2016**, *79*, 135–152.

[12] R. I. Cukier, C. M. Fortuin, K. E. Shuler, A. G. Petschek, J. H. Schaibly, *J. Chem. Phys.* **1973**, *59*, 3873–3878.

[13] A. Saltelli, S. Tarantola, K. P.-S. Chan, *Technometrics* **1999**, *41*, 39–56.

[14] E. Ryan, O. Wild, A. Voulgarakis, L. Lee, *Geosci. Model Dev.* **2018**, *11*, 3131–3146.

[15] T. Homma, A. Saltelli, *Reliab. Eng. Syst. Saf.* **1996**, *52*, 1–17.

[16] A. Saltelli, M. Ratto, S. Tarantola, F. Campolongo, *Chem. Rev.* **2012**, *112*, PR1–PR21, PMID: 22515330.

[17] Y. Tang, P. Reed, T. Wagener, K. van Werkhoven, *Hydrol. Earth Syst. Sci.* **2007**, *11*, 793–817.

[18] Y. Tang, P. Reed, K. van Werkhoven, T. Wagener, *Water Resour. Res.* **2007**, *43*, 1–14.

[19] R. Confalonieri, G. Bellocchi, S. Bregaglio, M. Donatelli, M. Acutis, *Ecol. Modell.* **2010**, *221*, 1897–1906.

[20] G. Fu, Z. Kapelan, P. Reed, M. Asce, *J. Water Resour. Plann. Manage.* **2012**, *138*, 196–207.

[21] M. Tosin, A. M. A. Côrtes, A. Cunha Jr in *Networks in Systems Biology: Applications for Disease Modeling*, of *Mechanisms and Machine Science*, F. A. B. da Silva, N. Carels, M. Trindade dos Santos, F. J. P. Lopes (Eds.), Springer International Publishing, Switzerland, **2020**, chapter 6.

[22] G. Qian, A. Mahdi, *Math. Biosci.* **2020**, *323*, 108306.

[23] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, *Comput. Phys. Commun.* **2010**, *181*, 259–270.

[24] M. J. Jansen, *Comput. Phys. Commun.* **1999**, *117*, 35–43.

[25] S. Burhenne, D. Jacob, G. P. Henze in *Proc. of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association*, Sydney, pp. 1816–1823.

[26] P. Bratley, B. L. Fox, *ACM Trans. Math. Softw.* **1988**, *14*, 88–100.

[27] P. A. Vanrolleghem, G. Mannina, A. Cosenza, M. B. Neumann, *J. Hydrol.* **2015**, *522*, 339–352.

[28] C. Spearman, *Am. J. Psychol.* **1904**, *15*, 72–101.

[29] M. Awad, T. Senga Kiesse, Z. Assaghir, A. Ventura, *Reliab. Eng. Syst. Saf.* **2019**, *189*, 109–122.

[30] B. Efron, R. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall/CRC, New York, **1994**.

[31] G. E. Archer, A. Saltelli, I. M. Sobol, *J. Stat. Comput. Simul.* **1997**, *58*, 99–120.

[32] H. Falsig, B. Hvolbæk, I. S. Kristensen, T. Jiang, T. Bligaard, C. H. Christensen, J. K. Nørskov, *Angew. Chem. Int. Ed.* **2008**, *47*, 4835–4839.

[33] L. C. Grabow in *Computational Catalysis*, *Vol. 14*, A. Asthagiri, M. J. Janik (Eds.), The Royal Society of Chemistry, Cambridge, UK, **2014**, chapter 1.

[34] S. Piccinin, M. Stamatakis, *ACS Catal.* **2014**, *4*, 2143–2152.

[35] M. Nagasaka, H. Kondoh, I. Nakai, T. Ohta, *J. Chem. Phys.* **2007**, *126*, 044704.

[36] S. Blomberg, U. Hejral, M. Shipilin, S. Albertin, H. Karlsson, C. Hulteberg, P. Lömker, C. Goodwin, D. Degerman, J. Gustafson, C. Schlueter, A. Nilsson, E. Lundgren, P. Amann, *ACS Catal.* **2021**, *11*, 9128–9135.

[37] M. A. van Spronsen, J. W. M. Frenken, I. M. N. Groot, *Chem. Soc. Rev.* **2017**, *46*, 4347–4374.

[38] R. Venderbosch, W. Prins, W. van Swaaij, *Chem. Eng. Sci.* **1998**, *53*, 3355–3366.

[39] S. Kozuch, S. Shaik, *Accounts of Chemical Research* **2011**, *44*, 101–110.

[40] F. Pianosi, K. Beven, J. Freer, J. W. Hall, J. Rougier, D. B. Stephenson, T. Wagener, *Environ. Modell. Software* **2016**, *79*, 214–232.

[41] K. George, M. van Berkel, X. Zhang, R. Sinha, A. Bieberle-Hütter, *J. Phys. Chem. C* **2019**, *123*, 9981–9992.

[42] C. Tung, Y. Hsu, Y. Shen, Y. Zheng, T. Chan, H. Sheu, *Nat. Commun.* **2015**, *6*, 1–9.

[43] C. Kronawitter, Y. Mao, B. Antoun, *Appl. Phys. Lett.* **2011**, *98*, 092108–092108.

[44] N. Iwasaki, Y. Sasaki, Y. Nishina, *J. Electrochem. Soc.* **1988**, *135*, 2531–2534.

[45] J. Rossmeisl, Z.-W. Qu, H. Zhu, G.-J. Kroes, J. Nørskov, *J. Electroanal. Chem.* **2007**, *607*, 83–89.

[46] P. Cendula, S. D. Tilley, S. Gimenez, J. Bisquert, M. Schmid, M. Grätzel, J. O. Schumacher, *J. Phys. Chem. C* **2014**, *118*, 29599–29607.

[47] O. Neufeld, M. Caspary Toroker, *J. Chem. Phys.* **2016**, *144*, 164704.

[48] K. George, T. Khachatrjan, M. van Berkel, V. Sinha, A. Bieberle-Hütter, *ACS Catal.* **2020**, *10*, 14649–14660.

[49] G. Wang, Y. Ling, D. A. Wheeler, K. E. George, K. Horsley, C. Heske, J. Z. Zhang, Y. Li, *Nano Lett.* **2011**, *11*.

[50] L. M. Peter, *J. Solid State Electrochem.* **2013**, *17*, 315–326.

[51] A. Govind Rajan, E. A. Carter, *Energy Environ. Sci.* **2020**, *13*, 4962–4976.

[52] X. Zhang, P. Klaver, R. van Santen, M. C. M. van de Sanden, A. Bieberle-Hütter, *J. Phys. Chem. C* **2016**, *120*.

[53] R. van de Krol, Y. Liang, J. Schoonman, *J. Mater. Chem.* **2008**, *18*, 2311–2320.

[54] M. Barroso, S. R. Pendlebury, A. J. Cowan, J. R. Durrant, *Chem. Sci.* **2013**, *4*, 2724–2734.

[55] A. Kudo, Y. Miseki, *Chem. Soc. Rev.* **2009**, *38*, 253–278.

[56] M. G. Walter, E. L. Warren, J. R. McKone, S. W. Boettcher, Q. Mi, E. A. Santori, N. S. Lewis, *Chem. Rev.* **2010**, *110*, 6446–6473.

[57] N. S. Lewis, *J. Phys. Chem. B* **1998**, *102*, 4843–4855.

[58] M. Iwamoto, Y. Yoda, N. Yamazoe, T. Seiyama, *Bull. Chem. Soc. Jpn.* **1978**, *51*, 2765–2770.

[59] Z. Wang, E. Seebauer, *Appl. Surf. Sci.* **2001**, *181*, 111–120.

[60] E. R. Blomiley, E. G. Seebauer, *Langmuir* **1999**, *15*, 5970–5976.

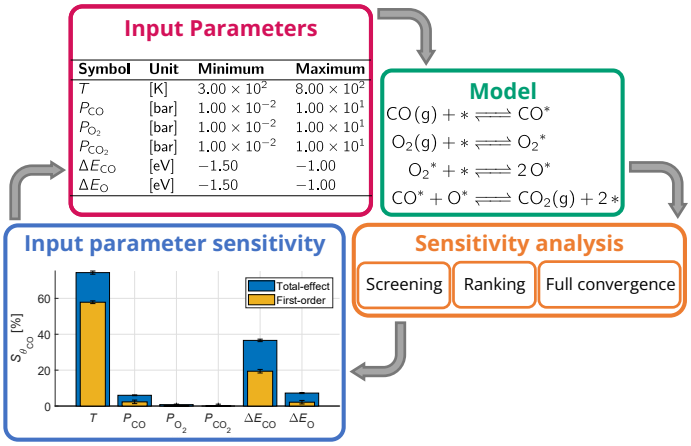
[61] G. W. Busser, O. Hinrichsen, M. Muhler, *Catal. Lett.* **2002**, *79*, 49–54.

[62] D. L. Allaire, K. E. Willcox, *Reliab. Eng. Syst. Saf.* **2012**, *107*, 107–114.

[63] S. Razavi, H. V. Gupta, *Water Resources Research* **2015**, *51*, 3070–3092.

7 Table of Contents

Sobol's sensitivity analysis method is presented for chemical modeling with a focus on microkinetic modeling. Its application is demonstrated with two case studies: (1) CO oxidation, (2) oxygen evolution reaction in a photo-electrochemical cell. The results show how each input parameter influences the output uncertainty. A workflow is proposed aimed at reducing output uncertainty and validating and optimizing the model.


Input Parameters
Symbol Unit Minimum Maximum
T [K] 3.00 × 10^2 8.00 × 10^2
P_CO [bar] 1.00 × 10^−2 1.00 × 10^1
P_O2 [bar] 1.00 × 10^−2 1.00 × 10^1
P_CO2 [bar] 1.00 × 10^−2 1.00 × 10^1
ΔE_CO [eV] −1.50 −1.00
ΔE_O [eV] −1.50 −1.00
Model
CO(g) + * ⇌ CO*
O2(g) + * ⇌ O2*
O2* + * ⇌ 2 O*
CO* + O* ⇌ CO2(g) + 2 *
Input parameter sensitivity
Total-effect
First-order
S_i,CO [%]
60
40
20
0
T
P_CO
P_O2
P_CO2
ΔE_CO
ΔE_O
Sensitivity analysis
Screening
Ranking
Full convergence

ToC Figure