Accepted version on Author's Personal Website: Armin Norouzi

Citation:

Aliramezani, Masoud, Armin Norouzi, and Charles Robert Koch. "A grey-box machine learning based model of an electrochemical gas sensor." Sensors and Actuators B: Chemical 321 (2020): 128414.

See also:

https://arminnorouzi.github.io/files/pdf/NOxsensorGreyboxSVM_R1-wfp.pdf

Article accepted version starts on the next page ⟶
Or link: to Author's Website

A grey-box machine learning based model of an electrochemical gas sensor

Masoud Aliramezani[a], Armin Norouzi[a], Charles Robert Koch[a]

[a]*Mechanical Engineering Department, University of Alberta, Edmonton, Canada T6G 1H9*

Abstract

A grey-box machine learning based model of an electrochemical O_2-NO_x sensor is developed using the physical understanding of the sensor working principles and a state-of-the-art machine learning technique: Support Vector Machine (SVM). The model is used to predict the sensor response at a wide range of sensor operating conditions in the presence of different concentrations of NO_x and ammonia. To prepare a comprehensive training and test data set, the production sensor is first mounted on the exhaust system of a spark ignition, a diesel engine, and then on a fully controlled sensor test rig. The sensor is not modified, rather the sensor working temperature, all of the sensor cell potentials, and the pumping current of the O_2 sensing cell are the model inputs that can be varied while the pumping current of the NO_x sensing cell is considered as the model output. A 9-feature low order model (LOM) and a 45-feature high order model (HOM) are developed with linear and Gaussian kernels. The model performance and generalizability are then verified by conducting input-output trend analysis. The LOM with Gaussian kernel and the HOM with linear kernel have shown the highest accuracy and the best response trend prediction.

Keywords: Electrochemical sensor; Machine learning; Grey-box model; On-board diagnostics; Combustion engines; NOx emission; Cross sensitivity.

1. Introduction

Electrochemical gas sensors are ideal for measuring emissions of combustion systems due to their small size, fast response, and long life span [1–5]. Amperometric NO_x sensors are widely used in commercial combustion engines to meet the stringent emission regulations [6–9]. Amperometric NO_x sensors simultaneously measure the O_2 and NO_x concentration in the exhaust gas [10, 11]. Urea-based selective catalytic reduction (SCR) is an effective technique to minimize diesel engine NO_x emissions to satisfy current and upcoming stringent emission regulations [12, 13]. Measuring the concentration of NO_x emission upstream and downstream of the SCR system is essential for its closed-loop control [14, 15]. However, the commercial electrochemical NO_x sensors used in the automotive industry are cross-sensitive to ammonia (NH_3) [16, 17]. This cross sensitivity causes large deviations in the NO_x sensor reading from the actual NO_x concentration [18]. Determining the actual NO_x concentration has been a challenge for controlling urea injection of SCR systems as the cross sensitivity of commercial NO_x sensors to NH_3, makes it difficult to achieve maximum NO_x conversion in SCR control. This can cause NH_3 slip downstream of the SCR which not only reduces the efficiency of an SCR system, but increases the hazardous NH_3 emission in the exhaust gas [19, 20]. The cross-sensitivity factor of a NO_x sensor to NH_3 is taken as a constant [21, 22], a function of time [18] or a function of ammonia concentration [23] but none of the these studies capture all the cross sensitivity factors of commercial NO_x sensors [16]. More accurate methods were developed to remove the cross sensitivity of NO_x sensors to ammonia by estimating ammonia slip downstream of the SCR [24–26]. However, all these methods require an SCR model which increases the estimation uncertainty due to possible model errors. In addition, these cross sensitivity removal methods are quite dependant on the SCR system. This means the sensor cross sensitivity estimation has to change if the SCR system changes and estimation error increases as the SCR ages [27, 28]. In addition, all these studies are carried out on a sensor operating at the normal design conditions i.e. typical sensor temperature, sensing cell potentials, reference cell potentials etc. The effect of changed sensor operating parameters on sensor performance and sensor cross sensitivity to other species particularly to NH_3 has not been considered.

Developing a comprehensive NO_x sensor model that includes all of the sensor operating conditions and is also capable of predicting the sensor cross sensitivity to NH_3 at a wide range of sensor inputs is the main objective of this work. The motivation is that it is more cost effective to change the sensor operating conditions via control electronics and use existing production sensors than to build a more complex sensor. Accurate physics-based models have been developed to simulate the sensor performance by modeling diffusion of species over the sensor diffusion barriers and the electrochemical reactions that take place over the sensing electrodes [29–31]. The physics-based understanding provided by these studies has been used to also developed new fast response electrochemical sensors [32]. However, to predict the sensor output variables over a wide range of operating conditions, a more complex model is required.

2

Developing such a complex model also facilitates exploring the sensor operating conditions that lead to a lower sensor cross sensitivity to other species. This model can also be used for future on-board diagnostics and for developing complex emission reduction control strategies. To develop the more complex model, a machine learning approach is combined with the physical understanding of the sensor working principles.

Machine learning is rapidly expanding to solve a wide range of engineering problems [33, 34]. Support Vector Machine (SVM) is a machine learning approach which is capable of generating high-precision decision boundaries based on a small subset of training data points with complex and non-linear relations [35, 36]. The use of machine learning facilitates modeling systems with a complicated nonlinear input-output relation especially for a high number of inputs [37]. This approach does not require a physical understanding of the system working principles and so is called a data-driven or black-box model [38, 39]. Although, it can be accurately trained if sufficient training data is available [40, 41], the main disadvantage of using data-driven black box models is the risk of overfitting since physical understanding of the system is lacking [42]. To address this issue, a grey-box NO_x sensor model is developed. A grey-box model combines a machine learning-based model with physical understanding of the system to make sure the the most appropriate features are used.

A grey-box SVM model of the sensor response at a wide range of operating conditions in the presence of different concentrations of NO_x and ammonia will be developed in this paper. To prepare a comprehensive training and test data set the sensor is first mounted on the exhaust system of a spark ignition (SI) and diesel engine and then on a fully controlled sensor test rig. The experimental points are used to train the model (89 %) and then the remaining points (11 %) are used to validate the model performance and test the model accuracy. The grey-box model features are selected using the physical understanding of the sensor working principles. The model performance is then verified by using input-output trend analysis.

2. Experimental setup

2.1. Spark Ignition (SI) and Diesel engines

To study the sensor behavior to engine exhaust gas the sensor was mounted on a four cylinder Cummins QSB-4.5-160 diesel engine and a four cylinder GM-Vortec-3000 spark ignition (SI) engine. The sensor was mounted to directly measure the engine-out emissions. The engine operating condition has been changed to expose the sensor to a wide range of concentration of species in the exhaust gas.

2.2. Gas mixture test rig

A fully-controlled sensor test rig was used to test the sensor at different concentrations and combinations of species from gas cylinders filled to known concentrations of the desired species. The test rig consists of six fast response fully controlled mass flow controllers (MKS-GE50A), six externally controlled 2-way valves, three 3-way valves, and four humidifying tanks. All of the setup actuators are controlled externally by the test rig control computer [31].

3

The sensor is tested at 126 different operating conditions exposed to different concentrations of species in the engine exhaust gas and the sensor test rig throughout the operating range illustrated later in Table 1.

2.3. Solid state O_2-NO_x sensor

The sensor used in the experiments was a production Bosch sensors with ECM electronics [43]. The sensor is calibrated and controlled using the corresponding control module (*ECM-NOxCANt* P/N: 02-07). The sensor control module is connected to the electrochemical NO_x sensor mounted on the engine exhaust pipe or on the sensor test rig. The ECM sensor control module facilitates measuring the sensor output current, for O_2 and NO_x and controls all the main sensor parameters such as the sensor temperature and the sensing cell voltages. The sensor and the control module are shown in Figure 1.

Figure 1: ECM production NO_x sensor and the corresponding control module

2.4. Fourier-Transform Infrared Spectroscopy (FTIR)

To measure the actual concentrations of all species in the gas stream, a *MultiGas 2030* FTIR analyser was used. To avoid exposing the FTIR to soot particles produced by the diesel and SI engines, the sample exhaust gas has passed through two Heated Filters (*Flexotherm*). To minimize water vapor condensation in the sample gas, the sample lines were heated to $191^{o}C$ [44].

3. Grey-box SVM model

3.1. Support Vector Machine

Support Vector Machine (SVM) is one of the supervised machine learning methods which was introduced for a classification problem in 1964 [45, 46]. Application of SVM in the regression or so-called Support Vector Regression (SVR), was introduced by [47] in 1995. Generally, SVM is used for labeled data classification and function approximation (regression) by generating a set of hyperplanes in an infinite-dimensional space [48]. In this paper, SVR is

4

used for function approximation to find a correlation between labeled data. For a given learning data, $\{\mathbf{u_i}, \mathbf{z_i}\}$, where $\mathbf{u_i}$ is feature set and $\mathbf{z_i}$ is target output, SVR generates a function $\mathbf{y}(\mathbf{u_i})$ which is as smooth as possible and it is able to approximate training data in the maximum error of $|\epsilon|$. In this case, $\mathbf{y}(\mathbf{u_i})$ is a model to predict steady-state values. The approximate function for a given data set can be defined as

$$\mathbf{y}(\mathbf{u_i}) = \mathbf{w}^T \mathbf{u_i} + \mathbf{b} \tag{1}$$

where $\mathbf{w}$ and $\mathbf{b}$ are found by solving the SVR algorithm [47]. The smoothness of approximated function is achieved by minimizing the second norm of $\mathbf{w}$. Thus, a convex optimization problem is defined as

$$
\begin{aligned}
&\textbf{Minimize:} \quad \frac{1}{2}\|\mathbf{w}\|_2^2 \\
&\textbf{Subject to:} \quad
\begin{cases}
\mathbf{z_i} - \mathbf{w}^T \mathbf{u_i} - \mathbf{b} \le \epsilon \\[2mm]
\mathbf{w}^T \mathbf{u_i} + \mathbf{b} - \mathbf{z_i} \le -\epsilon
\end{cases}
\quad i = 1, ..., n
\end{aligned}
\tag{2}
$$

If such a function, $\mathbf{y}(\mathbf{u_i})$, which approximates all pairs of labeled learning data within a defined ϵ margin ($-\epsilon \le \mathbf{z_i} - \mathbf{y_i} \le \epsilon$), and is as smooth as possible, is found by solving Eq. (1), the convex optimization problem is feasible. Otherwise, the convex optimization problem of Eq. (1) is infeasible.

Slack variables (ζ_i^-, ζ_i^+) are added as a penalty variable which regulates error tolerate of optimization as

$$-\epsilon - \zeta_i^- \le \mathbf{z_i} - \mathbf{y_i} \le \epsilon + \zeta_i^+ \tag{3}$$

To overcome convex optimization problem infeasibility, the Soft Margin Loss Function (SMLF) with definition of slack variables (ζ_i^-, ζ_i^+) is used as [47]

$$
L_\epsilon(\mathbf{z_i}, \mathbf{y_i}) =
\begin{cases}
0 & |\mathbf{z_i} - \mathbf{y_i}| \le \epsilon \\[2mm]
|\mathbf{z_i} - \mathbf{y_i}| - \epsilon & otherwise
\end{cases}
\tag{4}
$$

The SMLF is equal to summation of slack variables as [47]

$$L_\epsilon(\zeta_i^-, \zeta_i^+) = \zeta_i^- + \zeta_i^+ \tag{5}$$

The SMLF with slacks variables is shown schematically in Figure. 2. Therefore, the convex problem is now defined to optimize smoothness of approximate function and minimize the SMLF, and associated constraint as [50]:

$$
\begin{aligned}
&\textbf{Minimize:} \quad \frac{1}{2}\|\mathbf{w}\|_2^2 + C \sum_{i=1}^{n} (\zeta_{\mathbf{i}}^+ + \zeta_{\mathbf{i}}^-) \\
&\textbf{Subject to:} \quad
\begin{cases}
\mathbf{z_i} - \mathbf{w}^T \mathbf{u_i} - \mathbf{b} \le \epsilon + \zeta_{\mathbf{i}}^+ \\[2mm]
\mathbf{w}^T \mathbf{u_i} + \mathbf{b} - \mathbf{z_i} \le -\epsilon + \zeta_{\mathbf{i}}^- \\[2mm]
\zeta_{\mathbf{i}}^-, \zeta_{\mathbf{i}}^+ \le 0
\end{cases}
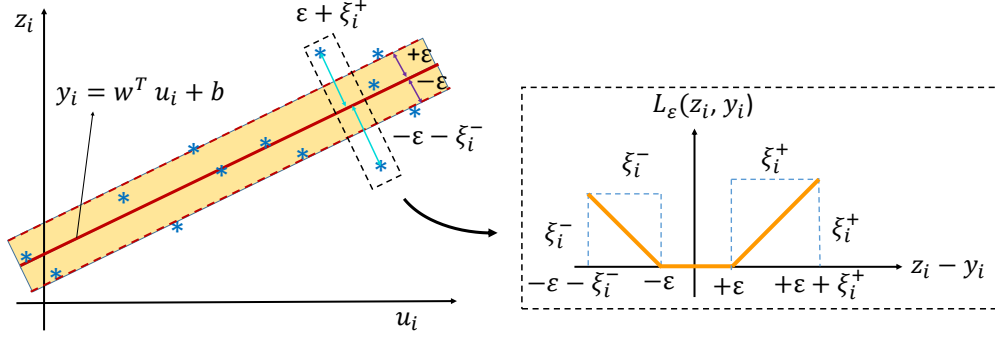\quad i = 1, ..., n
\end{aligned}
\tag{6}
$$

5

Figure 2: SVM regression and support vectors, and the soft margin loss function for a linear SVM (based on [49])

where C is a positive parameter which is used to adjust trade-off between model tolerated error and model smoothness. To solve the convex problem, the Lagrangian function (primal objective function) is calculated as [50]

$$
\begin{aligned}
L = &\frac{1}{2}\|\mathbf{w}\|_2^2 + C \sum_{i=1}^{N} (\zeta_i^- + \zeta_i^+) \\
&- \sum_{i=1}^{N} \alpha_i^+ (-\mathbf{z_i} + \mathbf{y_i} + \epsilon + \zeta_i^+) - \sum_{i=1}^{N} \mu_i^+ \zeta_i^+ \\
&- \sum_{i=1}^{N} \alpha_i^- (\mathbf{z_i} - \mathbf{y_i} + \epsilon + \zeta_i^-) - \sum_{i=1}^{N} \mu_i^- \zeta_i^-
\end{aligned}
\tag{7}
$$

where α_i^+, α_i^-, μ_i^+, and μ_i^- are Lagrangian multipliers [50]. The optimization problem is solved by calculating the partial differential of Lagrangian function with respect to the primal variables equal as [50]

$$
\frac{\partial L}{\partial \mathbf{w}} = 0 \;\rightarrow\; w = \sum_{i=1}^{N} (\alpha_i^+ - \alpha_i^-)\mathbf{u_i}
\tag{8a}
$$

$$
\frac{\partial L}{\partial \mathbf{b}} = 0 \;\rightarrow\; \sum_{i=1}^{N} (\alpha_i^+ - \alpha_i^-) = 0
\tag{8b}
$$

$$
\frac{\partial L}{\partial \zeta_i^+} = 0 \;\rightarrow\; \alpha_i^+ + \mu_i^+ = C
\tag{8c}
$$

$$
\frac{\partial L}{\partial \zeta_i^-} = 0 \;\rightarrow\; \alpha_i^- + \mu_i^- = C
\tag{8d}
$$

Eq. 8b is called support vector expansion. Substituting Eq. (8) into Eq. (7) generates the dual optimization problem as:

6

$$\textbf{Minimize: } L = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\alpha_i^+ - \alpha_i^-)(\alpha_j^+ - \alpha_j^-)\mathbf{u_i}^T\mathbf{u_j}$$

$$+ \sum_{i=1}^{N} (\alpha_i^+ - \alpha_i^-)\mathbf{z_i} - \epsilon \sum_{i=1}^{N} (\alpha_i^+ + \alpha_i^-)$$

$$\textbf{Subject to: } \begin{cases} \sum_{i=1}^{N}(\alpha_i^+ - \alpha_i^-) = 0 \\[2mm] 0 \leq \alpha_i^+ \leq C \\[2mm] 0 \leq \alpha_i^- \leq C \end{cases}$$

(9)

A standard Quadratic programming form (QP) is obtained by rewriting Eq. (9) as [51]:

$$\textbf{Minimize: } \frac{1}{2}\alpha^T \mathcal{H}\alpha + f^T\alpha$$

$$\textbf{Subject to: } A_{eq}\alpha = B_{eq}$$

(10)

where

$$\alpha = \begin{bmatrix} \alpha^+ \\ \alpha^- \end{bmatrix}, \quad \mathcal{H} = \begin{bmatrix} H & -H \\ -H & H \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} -\mathbf{z_i} + \epsilon \\ \mathbf{z_i} + \epsilon \end{bmatrix},$$

$$H = \begin{bmatrix} \mathbf{u_i}^T \mathbf{u_j} \end{bmatrix}, \quad \mathbf{A_{eq}} = [1...1 \quad -1... -1], \quad \mathbf{B_{eq}} = [0]$$

(11)

Now, α is found by solving Eq. (11) and then substituting α into Eq. (8a), $\mathbf{w}$ is obtained. To find $\mathbf{b}$ in Eq. 1, Karush-Kuhn-Tucker (KKT) conditions [52, 53] is applied to dual optimization problem. Based on KKT conditions, the following equations must be fulfilled at the optimum point

$$\alpha_i^+(-\mathbf{z_i} + \mathbf{y_i} + \epsilon + \zeta_i^+) = 0 \tag{12a}$$

$$\alpha_i^-(\mathbf{z_i} - \mathbf{y_i} + \epsilon + \zeta_i^-) = 0 \tag{12b}$$

$$\mu_i^+ \zeta_i^+ = (C - \alpha_i^+)\zeta_i^+ \tag{12c}$$

$$\mu_i^- \zeta_i^- = (C - \alpha_i^-)\zeta_i^- \tag{12d}$$

Based on our optimization problem constraints, Eq. (12) is reduced to:

$$\alpha_i^+ = \alpha_i^- = 0 \tag{13a}$$

$$0 < \alpha_i^+ < C, \; \alpha_i^- = 0 \tag{13b}$$

$$0 < \alpha_i^- < C, \; \alpha_i^+ = 0 \tag{13c}$$

$$\alpha_i^+ = C, \; \alpha_i^- = 0 \tag{13d}$$

$$\alpha_i^- = C, \; \alpha_i^+ = 0 \tag{13e}$$

If the Eqs. (13b) and (13c) are fulfilled, the exact value of $|\mathbf{z_i} - \mathbf{y_i}| = \epsilon$ will be obtained and the set of data which the error between prediction and actual equals to ϵ, can be calculated as

$$S = \{ \, i \, | \, 0 < \alpha_i^+ + \alpha_i^+ < C \} \tag{14}$$

7

where S is called support vectors index. As the support vector set $\mathbf{z_i} = \mathbf{y_i} + sign(\alpha_i^+ - \alpha_i^-)\epsilon$ is true, $\mathbf{b}$ can be calculated as

$$\mathbf{b} = \frac{1}{|S|} \sum_{i \in S}^{S} (\mathbf{z_i} - \mathbf{w}^T \mathbf{u_i} - sign(\alpha_i^+ - \alpha_i^-)\epsilon) \tag{15}$$

3.2. Kernel Based SVM

To improve the performance of conventional SVM in providing distribution predictions an implicit feature space is used instead of simply computing the inner product in Eq. (9), $\mathbf{u_i}^T \mathbf{u_j}$ (a high-dimensional map). In this case the optimization problem is to find the smoothest function in the feature space instead of the input space [50]. The nonlinear kernel in Eq. (9) is now used by mapping the training pattern into the feature space instead of using the term $\mathbf{u_i}^T \mathbf{u_j}$. The difference between non-linear SVM and linear SVM is that for non-linear SVM, a low-dimensional input space with a high-dimensional feature space is used, while for a linear SVM a high-dimensional linear input space is used [54]. Thus, the kernel is defined as $K(\mathbf{u_i}, \mathbf{u_j})$, and used instead of $\mathbf{u_i}^T \mathbf{u_j}$ in the Eq. (9). Therefore, in Eq. (11), H is changed to $H = \left[K(\mathbf{u_i}, \mathbf{u_j}) \right]$.

From Eq. (8a), $\mathbf{w}$ is calculated as $w = \sum_{i=1}^{N} (\alpha_i^+ - \alpha_i^-)\mathbf{u_i}$, which is linear combination of training data. Substituting Eq. (8a) into Eq. (1) and using kernel function instead of inner product ($K(\mathbf{u_i}, \mathbf{u_j})$ instead of $\mathbf{u_i}\mathbf{u}$), the approximate function is calculated as

$$\mathbf{y} = \sum_{i=1}^{n} (\alpha_i^+ - \alpha_i^-)K(\mathbf{u_i}, \mathbf{u}) + \mathbf{b} \tag{16}$$

as a special case, if radial basis function (RBF) kernel is used in kernel based SVM, it can be represent as Radial Basis Function Network (RBFN) which schematically shown in Figure. 3. The RBF kernel function is defined as

$$K(\mathbf{u_i}, \mathbf{u_j}) = exp\left(\frac{\|\mathbf{u_i} - \mathbf{u_j}\|_2^2}{2\sigma^2} \right) \tag{17}$$

where σ is free variables act like Gaussian variance in probability density function of a normally distributed random variable and $\|.\|_2$ is the Euclidean norm.

3.3. Physics-based feature selection

To capture the effect of sensor operating conditions on the sensor response and sensor cross sensitivity to ammonia, it is crucial to test the sensor at a wide range of operating conditions. Selecting appropriate features - that are based on the physical understanding of the sensor working principles - are essential for the SVM model. Physics-based models of an electrochemical NO_x sensor were previously developed in [29] and [31] and the physical insights provided by these references are used here as a guideline to select the physics-based features of the grey-box model. Employing physics-based features in any data-driven model potentially reduces the number of independent model features and also reduces the risk of model over-fitting [55]. To develop the grey-box model, the model features are selected based on the procedure described next.
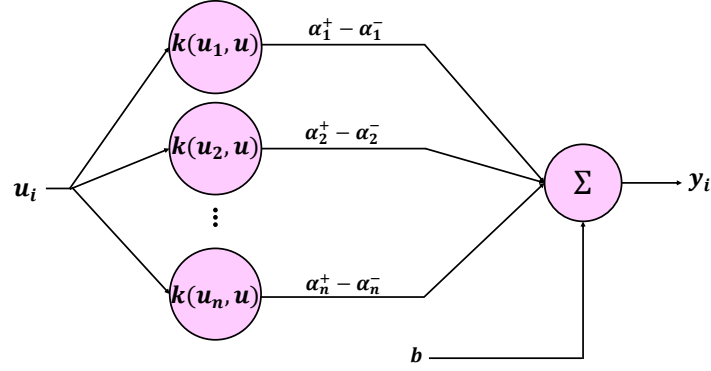
Figure 3: Radial Basis Function Network (RBFN) representation of kernel based SVR

3.3.1. NO$_x$ concentration

An amperometric NO$_x$ sensor typically operates at diffusion rate determining condition, where the electrochemical reactions are fast enough that the pumping current of the NO$_x$ sensing cell sensor only depends on the diffusion rate of species through the sensor diffusion barrier [29, 56] as schematically shown in Figure 4.
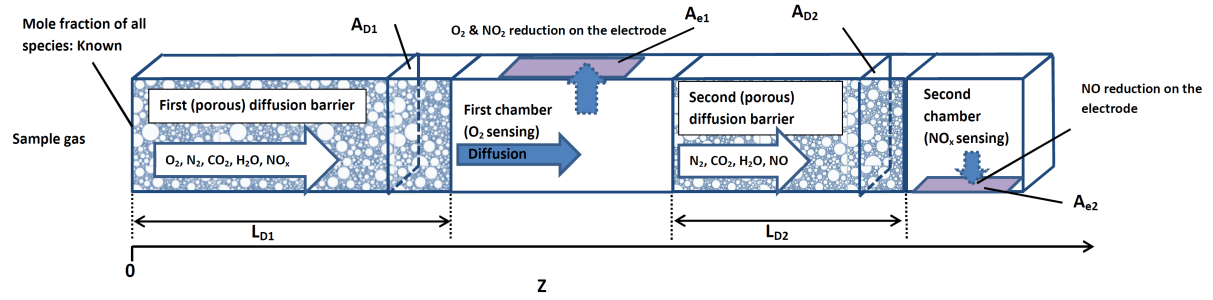

Figure 4: Amperometric NO$_x$ sensor working principle: Diffusion of species through the sensor diffusion barriers into the sensing chambers [57]

For the diffusion-rate-determining operating condition, the sensor response linearly changes with NO$_x$ concentration according to Fick's law [29]:

$$I_{p2} = 2F \times N_{NO,sc} \tag{18}$$

where, $N_{NO,sc}[\frac{mol}{s}]$ is the diffusive molar flow of NO through the second diffusion barrier into the second chamber.

When the sensor operating parameters change so that the sensor operates outside the diffusion rate determining condition, then the relation between the NO$_x$ concentration and sensor output follows the non-linear equation [31]:

$$i_{P2} = i_{P2}^o \left(\frac{x_{NO,sc}}{x_{NO,sc}^o}\right)^\gamma \left[exp\left(\frac{\alpha_{a2}F}{\bar{R}T}\eta_{ac}\right) - exp\left(\frac{-\alpha_{c2}F}{\bar{R}T}\eta_{ac}\right)\right] \tag{19}$$

9

where, α_{c2} and α_{a2} are the charge transfer coefficient of cathode and anode respectively, while i_{P2}^o, i_{P2}, $x_{NO,sc}^o$, and $x_{NO,sc}$ are the reference exchange current density; the NO_x sensing cell current density; the reference mole fraction of NO in the second chamber; and mole fraction of NO in the second chamber. The overpotential η_{ac} is calculated from $\eta_{ac} = V_{P2} - E_{OC2}$, where E_{OC2} and V_{P2} are the open circuit (zero-current) potential and the potential of the NO_x sensing cell. The NO_x concentration ratio is to the power of γ which represents the effect of concentration losses on the sensing cell current.

Although the sensor typically works at diffusion rate determining operating condition, to include the whole range of sensor operation, NO_x concentration and (NO_x concentration)$^\gamma$ are considered as features of the SVM. The value of γ is found to be equal to 0.1 for an amperometric NO_x sensor [31].

3.3.2. NH$_3$ concentration

The typical operating temperature of the electrochemical NO_x sensors is high enough to cause undesired oxidation of NH_3 inside the O_2 sensing chamber [16]. This oxidation produces an additional NO_x inside the sensor and therefore affects the pumping current of the NO_x sensing cell. Therefore, the same features that are allocated for NO_x should be considered for NH_3 (NH_3 concentration and (NH_3 concentration)$^\gamma$). To reduce the risk of overfitting, only the dominant feature, (NH_3 concentration), is included in the main SVM features.
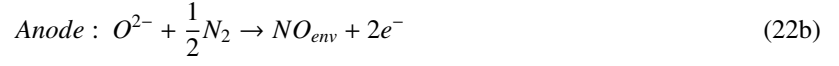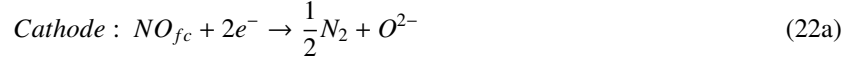
3.3.3. Sensor temperature

The sensor temperature has a complex effect on the NO_x sensor output. The multi-component molecular diffusion mechanism is found to be the dominant diffusion mechanism of species through the sensor diffusion barriers [29]. According to the molecular diffusion mechanism, the diffusion coefficient, D_n, increases with the sensor temperature as [30]:

$$D_n \propto T^{1.75} \tag{20}$$

Assuming that the sensor output current is limited by the diffusion rate, in the absence of any other electrochemical reactions, the sensor output, I_P, changes with temperature as [30]:

$$I_P \propto T^{0.75} \tag{21}$$

However, it is also shown in [31] that the NO_x species (NO_2 and NO) are completely or partly reduced in the first chamber before they reach to the NO_x sensing chamber. In the first chamber, NO is reduced through the following reactions [31]:

$$Cathode: \ NO_{fc} + 2e^- \rightarrow \frac{1}{2}N_2 + O^{2-} \tag{22a}$$

$$Anode: \ O^{2-} + \frac{1}{2}N_2 \rightarrow NO_{env} + 2e^- \tag{22b}$$

$$Overall: \ NO_{fc} \rightarrow NO_{env} \tag{22c}$$

where NO_{env} is the NO molecule in the sample gas and NO_{fc} is the NO molecule inside the first chamber.

The relation between of NO reduction and the first chamber cell voltage (V_{P1}) can be expressed by:

$$V_{P1} = \underbrace{E_{fc}^o + \frac{\Delta S_{ox-red,fc}}{2F}(T - T_o) - \frac{\bar{R}T}{2F}ln\left(\frac{p_{NO_{env}}}{p_{NO_{fc}}}\right)}_{E_N} - \eta_a \tag{23}$$

where, $p_{NO_{fc}}$ and $p_{NO_{env}}$ are the partial pressure of NO in the first chamber and in the sample gas respectively and η_a is the activation overpotential which will be explained later. Here, E_{fc}^o and $\Delta S_{ox-red,sc}$ are both equal to zero [31].

Using Dalton's law of partial pressures [58], Eq.(23) is written in terms of molar fraction of species as:

$$V_{P1} = \underbrace{E_{fc}^o + \frac{\Delta S_{ox-red,fc}}{2F}(T - T_o) - \frac{\bar{R}T}{2F}ln\left(\frac{x_{NO,env}}{x_{NO,fc}}\right)}_{E_N} - \eta_a \tag{24}$$

where, $x_{NO,env}$ and $x_{NO,fc}$ are the molar fraction of NO in the sample gas and in the first chamber respectively.

The first chamber cell potential is typically kept at high values (>0.42 V). Therefore, the activation is calculated using a Tafel approximation which is valid at high cell voltages [59]:

$$\eta_a = \frac{\bar{R}T}{\alpha_1 F}ln\left(\frac{i_{P1,NO}}{i_{P1}^o}\right) \tag{25}$$

where, $i_{P1,NO}$ and i_{P1}^o are the current density and the exchange current density of the first cell respectively and α_1 is the charge transfer coefficient of the first cell.

According to Eq. (24), for a constant cell potential, as the temperature increases, molar fraction of NO in the first chamber decreases. Reducing the molar fraction of NO in the first chamber, reduces the pumping current of the NO_x sensing cell according to the Fick's law [29]:

$$I_{P2} \propto (x_{NO,fc} - x_{NO,sc}) \tag{26}$$

Therefore, to include the effect of temperature on diffusion and the electrochemical reaction rates, features $T^{0.75}$ and $exp(\frac{1}{T})$ are included in the grey-box model based on Eqs (21) and (24) respectively.

3.3.4. NO_x sensing cell potential (V_{p2})

The effect of NO_x sensing cell potential is comprehensively studied in [31] using a physics-based approach.

$$V_{P2} = \underbrace{E^o_{sc} + \frac{\Delta S_{ox-red,sc}}{2F}(T - T_o) - \frac{\bar{R}T}{2F}ln\left(\frac{x^{0.5}_{O_2,rc}x^{0.5}_{N_2,rc}}{x_{NO,sc}}\right)}_{E_N} - \eta_\Omega - \eta_{ac} \tag{27}$$

Here, this physical understanding is used to select a physics-based feature for the grey-box model that reflects the effect of NO_x sensing cell potential, V_{p2}, on the NO_x sensing cell current. The effect of NO_x sensing cell potential on sensor output is schematically shown in Figure 5. Regions I to VI shown in Figure 5 are oxidation, activation, ohmic, ohmic-diffusion transition, diffusion-rate-determining and electrolyte reduction operating regions respectively [31].
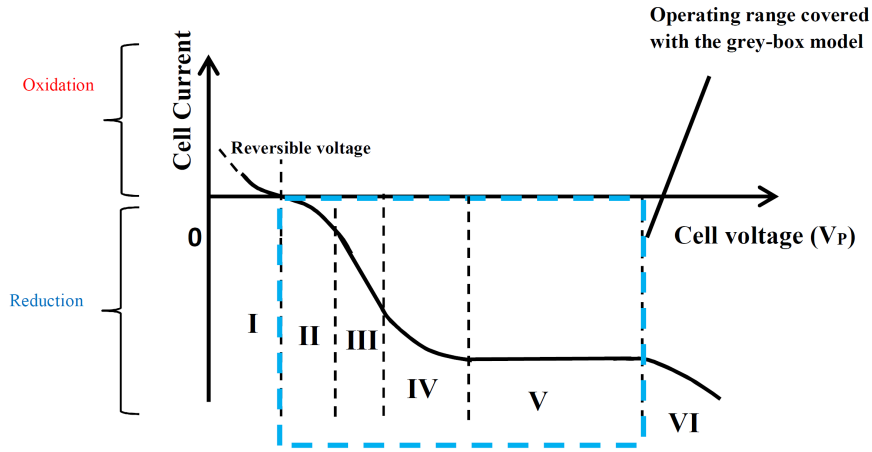

Figure 5: Typical electrode current vs potential relation at a constant NO_x concentration. Adapted from [31]

The full reduction operating range (regions II to V in Figure 5) is to be covered by the grey-box model. The following function is proposed and used to capture the affect of V_{p2} on I_{p2} over the whole reduction operating range:

$$I_{p2} = tanh(aV_{P2} + b) + c \tag{28}$$

where parameters a,b and c can be found based on the experimental data over the full range of V_{p2} when all other inputs are kept constant. The fitted values are found to be equal to 8.651, -0.2393 and 0.1595 for a, b and c respectively.

3.3.5. First sensing cell potential (V_{p1}) and the reference cell potential (V_S)

The reference cell potential, V_S, represents the partial pressure (and therefore molar fraction) of O_2 inside the O_2 sensing chamber [31]. A closed loop controller keeps V_S at the desired value by adjusting the O_2 sensing cell potential, V_{p1}. Unlike NO_x reduction in the NO_x sensing chamber, NO_x is always partially reduced in the O_2 sensing

chamber since the O_2 sensing electrode is an Au-Pt electrode with a higher activation energy than the Pt electrode used in the NO_x sensing chamber. Therefore the molar fraction of NO in the first chamber is proportional to exponential of VP1, according to Eq. (24). In addition, according to Eq. (26), the NO_x sensing pumping current is proportional to the molar fraction of NO in the first sensing chamber. Therefore, both $exp(V_S)$ and $exp(V_{p1})$ terms are considered as the features of the grey-box model.

3.3.6. First sensing cell current (I_{p1})

The last important sensor operating parameter that must be included in the sensor model is the pumping current of the O_2 sensing cell (I_{p1}) which represents O_2 concentration in the sample gas stream. Although the O_2 and NO_x pumping currents are not supposed to affect each other so that the sensor can measure both O_2 and NO_x concentration independently, this small effect still needs to be considered in the sensor model particularly when the sensor is operating away from diffusion-rate-determining condition. To consider this effect, feature I_{p1} is also included in the model.

4. Results and discussions

4.1. Training and test data

To provide a sufficient number of training and test data points for the model, the sensor is tested at 126 different operating conditions and the results were used to train and test the svm-based sensor model. Out of 126 data points, 112 points were used as training data and 14 points were used as the test data. To provide a better insight into the raw data used for model training and test, the range of all the main sensor input variables are listed in Table 1.

Table 1: The range of sensor operating condition used for model training and test (Total number of points: 126).

Variable	Maximum	Minimum	Mean
Sensor Temperature [K]	971.1	1079.2	1021.3
Second Cell potential (V_{P2}) [V]	0.272	0.450	0.417
Reference Cell potential (V_S) [V]	0.350	0.425	0.4212
First Cell potential (V_{P1}) [V]	3.154	4.414	4.084
First Cell Current (I_{P1}) [mA]	-0.0523	2.543	1.117
Second Cell Current (I_{P2}) [μA]	0.006	5.364	1.496
NO_x **concentration [ppm]**	2778	0	603
NH_3 **concentration [ppm]**	2000	0	249

4.2. Low Order SVM models

Two low order grey-box SVM models are developed using the physics-based features explained above and are listed in Table 2. One low order model is developed using a linear kernel and the other model is developed using a

Gaussian kernel as explained in section 3.2. The training and test results of the low order model with the linear and Gaussian kernel are shown in Figures 6 and 7 respectively. As shown in these figures, the Gaussian kernel based model has a significantly better performance on training and test data. The correlation coefficient (R^2) for LOM test and training are 0.875 and 0.970 for the linear kernel, while for the Gaussian kernel R^2 is equal to 0.999 and 0.921 for training and test data respectively.

Table 2: Features of the Low Order grey-box SVM

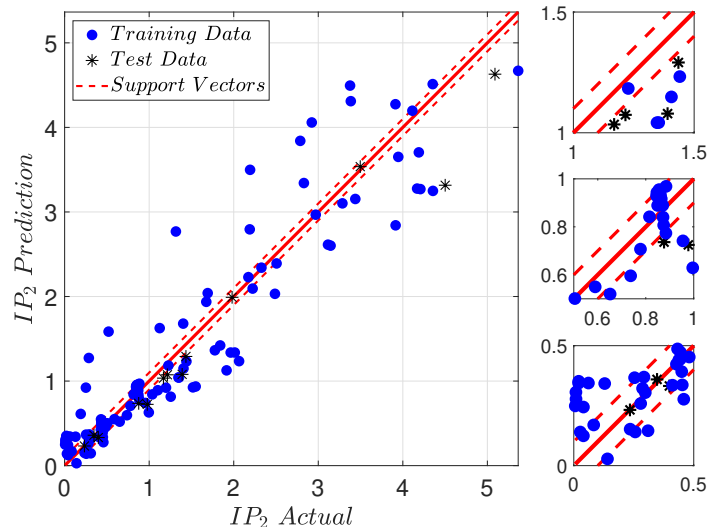|	
$U_1 = x_{NOx,\ env}$	$U_6 = exp(V_{p1})$
$U_2 = (x_{NOx,env})^{0.1}$	$U_7 = I_{p1}$
$U_3 = (T_{sen})^{0.75}$	$U_8 = exp(V_S)$
$U_4 = exp(T_{sen})$	$U_9 = x_{NH3,\ env}$
$U_5 = tanh(8.651(V_{p2} - 0.2393)) + 0.1595$	

Figure 6: Low order model with linear kernel. Training R^2=0.875. Test R^2=0.970.

4.3. High Order Models (HOM) SVM

To include the interactions between the features in the model and to improve the accuracy of the SVM training, the features of the low order model and their second order interactions are considered as the features of the high order model (HOM). Therefore, a 45-feature SVM model is obtained using all 9 grey-box features and all their second order interactions. Similar to the LOM, a linear and a Gaussian kernel were used to develop two high order SVM models.
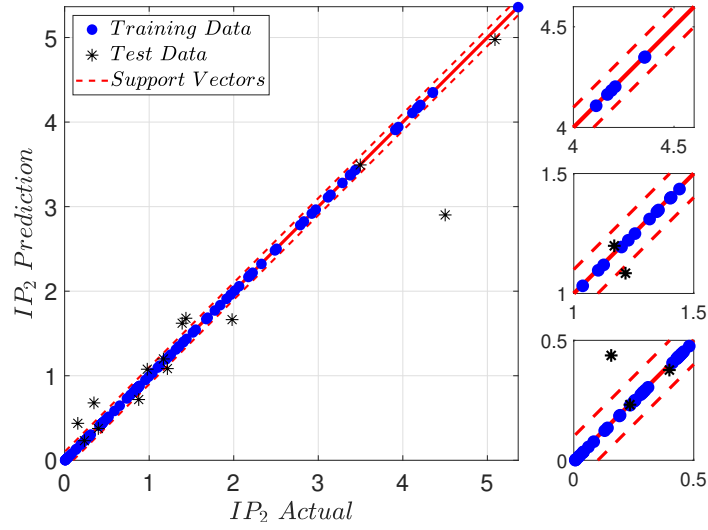
Figure 7: Low order model with Gaussian kernel. Training R^2=0.999. Test R^2=0.921.

The results of the HOM SVMs with linear and Gaussian kernels are shown in Figures 8 and 9 respectively. As shown in Figures 8 and 9, the accuracy of the SVM model significantly increases by increasing the number of features from 9 (LOM) to 45 (HOM) although a higher order is associated with a higher computational cost. In addition, similar to the LOM, the Gaussian kernel provides a better training and test performance compared to the linear kernel. The correlation coefficient (R^2) for HOM test and training are 0.984 and 0.986 for the linear kernel, while for the Gaussian kernel R^2 is equal to 0.999 and 0.993 for training and test respectively.
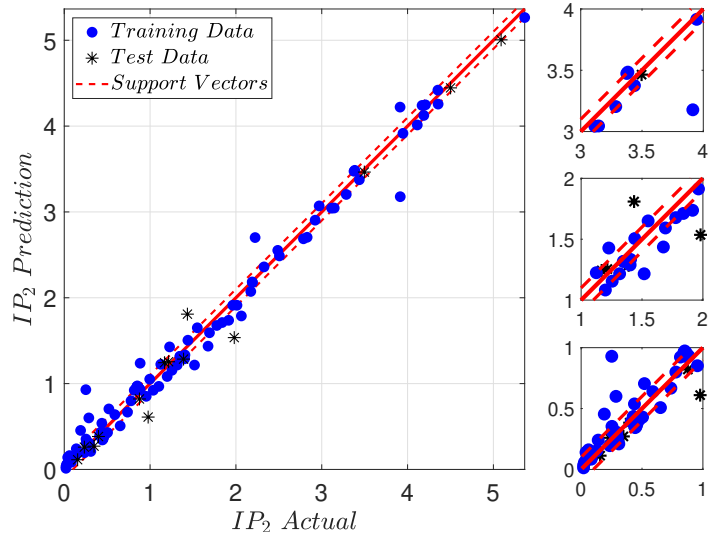

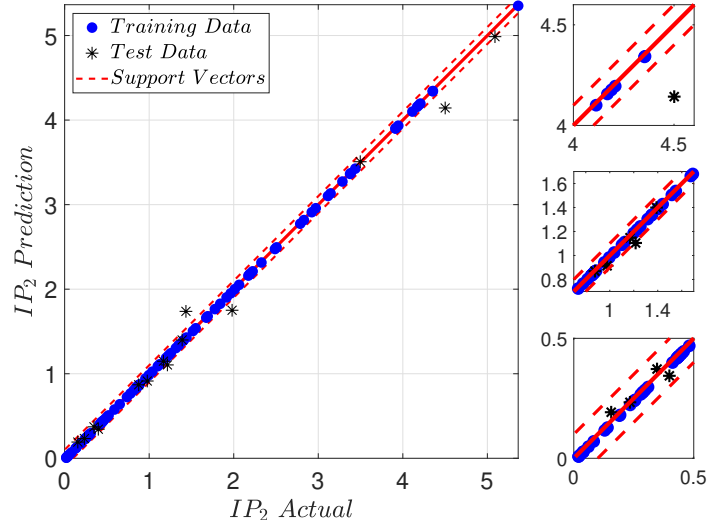Figure 8: High order model with linear kernel. Training R^2=0.984. Test R^2=0.986.

Figure 9: High order model (HOM) with Gaussian kernel. Training R^2=0.999. Test R^2=0.993

4.4. Trend Validation

One main concern with all types of data-driven models is the risk of overfitting the model, particularly when high number of features are included in the model. Allocating sufficient number of training and test data points can reduce the risk of overfitting by monitoring the model prediction performance for both the training and test data. To further investigate the model performance, its capability in correctly predicting the input-output relation without overfitting is evaluated by studying the effects of kernel type and model order on the input-output relation. To do so, the individual effect of NO_x and NH_3 concentration on the sensor output (IP_2) is calculated and shown in Figures 10 and 11 respectively. The effect of NO_x and NH_3 concentration on the sensor output (IP_2) is studied at typical potentials of the second and the reference cell of the sensor (V_{P2}= 0.44,V_S =0.423 V) which means that the sensor is in diffusion-rate-determining operation [31]. This means that the sensing current IP_2 should linearly change with NO_x and NH_3 concentration at this operating range. The model behavior is tested for both LOM and HOM with linear and Gaussian kernels and the models and performance are summarized in Table 3.

All but one of the models show an acceptably linear sensing current prediction. The HOM with Gaussian kernel has the highest order of magnitude RMS with respect to the linear response of I_{P2} as shown in Table 3. This means that although the HOM with Gaussian kernel has the best training performance amongst all the models, it is not capable of properly predicting the sensor output trend as it is overfitting the training data. Therefore, if we consider the training performance, the test performance and the output trend prediction, the LOM with Gaussian kernel and the HOM with linear kernel are the most appropriate models.
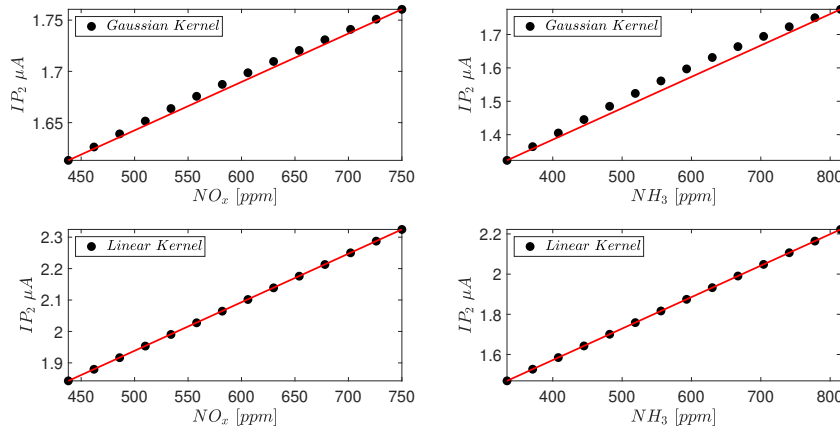
16

Figure 10: IP_2 vs NO_x and NH_3 prediction for the Low Order Model with linear and Gaussian kernels
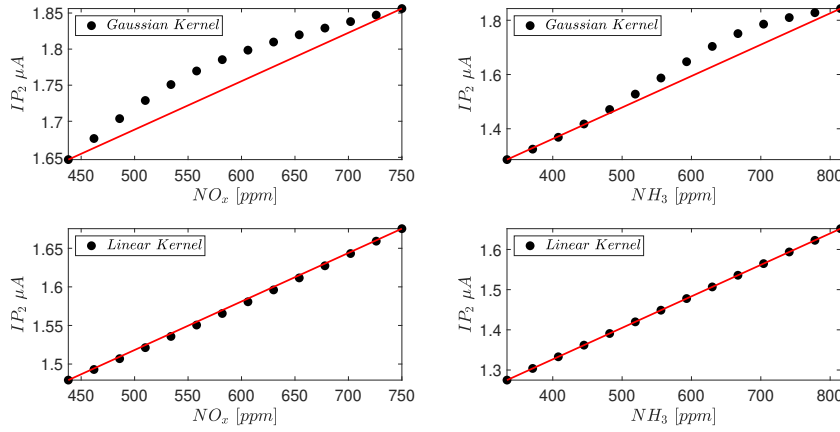

Figure 11: IP_2 vs NO_x and NH_3 prediction for the High Order Model with linear and Gaussian kernels

Table 3: Summary of LOM and HOM performance for training, test and trend analysis

Model Type	Number of the features	R^2 training	R^2 test	NO_x trend RMS	NH_3 trend RMS
Linear LOM	9	0.970	0.875	4.3571×10^{-4}	2.5867×10^{-16}
Kernel-based LOM	9	0.999	0.921	0.0043	0.0209
Linear HOM	45	0.985	0.984	0.003	1.0731×10^{-15}
Kernel-based HOM	45	0.999	0.993	0.0283	0.0439

As illustrated in Figures 10 and 11, the grey-box model is able to predict the effect of NO_x and NH_3 concentration on the sensor output. This provides a powerful tool to study the cross sensitivity of amperometric NO_x sensors to NH_3 which is essential for the automotive applications [16, 23]. This model can also predict the sensor output when

17

the sensor is simultaneously exposed to NO_x and NH_3.

The developed grey-box model can also be used for fault detection of production sensors. This can be done using sensor input perturbation technique which was originally presented in [30]. In this method, one of the sensor input variables (e.g. temperature) is changed and the effect of this perturbation on the sensor output is measured. The same sensor input perturbation is given to the model and the predicted output change is calculated and then compared with the measured value. If the difference between the estimated and the measured values are higher than a predefined value, it means the sensor is not working properly and must be changed. A simplified schematic of this strategy is shown in Figure 12. The model accuracy should be taken into account to define e_{MAX} in Figure 12. For instance, for NO_x concentration equal to 144 ppm, if sensor temperature, T_{sen}, and NO_x sensing potential, V_{P2} simultaneously change from 1009 K to 1025 K, and from 0.448V to 0.233V respectively resulted in .1911 change in I_{P2} in the experiments while this change is predicted to be 0.1581 μA, and 0.2417 μA by the linear HOM and the Kernel-based LOM respectively. Therefore if T_{sen} and V_{P2}, are both considered as the perturbed inputs, e_{MAX} must selected higher than approximately 0.04 and 0.05 for the linear HOM and the Kernel-based LOM, respectively to exclude the model error. In addition to the model-sensor mismatch, sensor aging can also increase the sensor and model deviation and therefore this factor should also be considered in selecting the e_{MAX} value to avoid undesired fault detection. Although it is highly recommended that a sensor/model recalibration process be added to the fault detection algorithm similar to the algorithm that was proposed in our previous work [30]. A more detailed study on this diagnostics technique will be carried out in the future works.
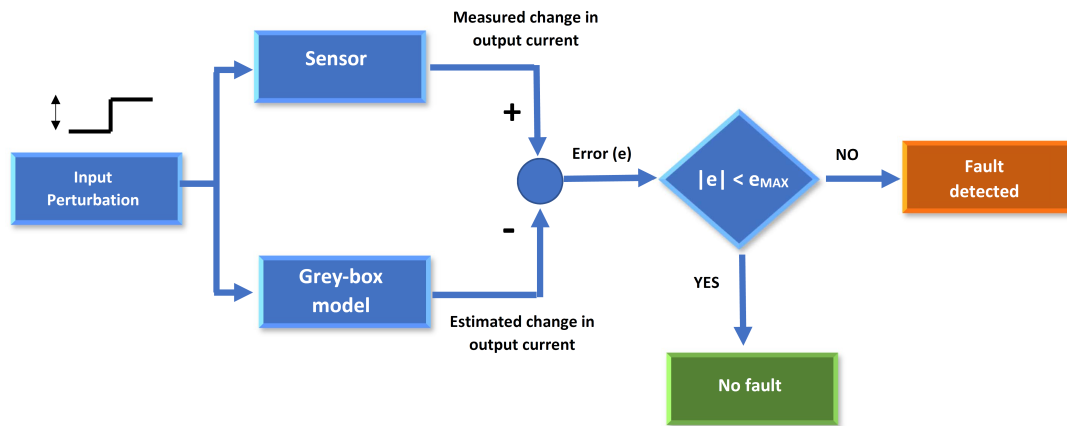

Figure 12: An example for sensor diagnostics using the developed grey-box model: An input-perturbation-based sensor diagnostics strategy

Since the sensor model captures all the main factors which affect the sensor output, the sensor model can also be used as an accurate input to the plant (engine, combustion chamber) for both simulation and real-time model based observers and control. To use this sensor model, which is accurate in steady state, to simulate and control dynamic systems, the dynamic behavior of the sensor must be considered. Adding a first order lag to the base steady state

18

model produces accurate transient results for the cases tested [60].

5. Conclusions

A support vector machine (SVM) based grey-box model is developed to predict the sensing out put of an ampero-metric O_2-NO_x sensor that is highly cross sensitive to <mark>ammonia</mark>. The model features are selected based on the physics of the sensor working principles.

A 9-feature low order model (LOM) and a 45-feature high order model (HOM) were developed with linear and Gaussian kernels to predict the sensor response at a wide range of sensor operating conditions in the presence of a wide range of NO_x and ammonia concentrations.

The experimental data from diesel and SI engine tests as well as the tests carried out on a fully controlled sensor test rig running at constant operating condition were used to train and test the LOM and HOM machine learning based models. Changing the operating conditions of the sensor using the electronics, combined with a detailed understanding of the sensor, enables more functionality from the sensor without having to change the actual production sensor.

The sensor working temperature, all of the sensor cell potentials, and the pumping current of the O_2 sensing cell are used as the main model inputs while the pumping current of the NO_x sensing cell is considered as the model output. The model performance and generalizability are then verified by conducting input-output trend analysis. The LOM with Gaussian kernel and the HOM with linear kernel has shown the highest accuracy and the best generalizability.

The results also revealed that:

- The correlation coefficient (R^2) for LOM test and training are 0.875 and 0.970 for the linear kernel, while for the Gaussian kernel R^2 is equal to 0.999 and 0.921 for training and test respectively.

- The correlation coefficient (R^2) for the HOM test and training are 0.984 and 0.986 for the linear kernel, while for the Gaussian kernel R^2 is equal to 0.999 and 0.993 for training and test respectively.

- Considering the training performance, the test performance and the output trend prediction, the LOM with Gaussian kernel and the HOM with linear kernel are the most appropriate models.

The developed model can be used for on-board diagnostics and future control strategies development by considering the effect of sensor inputs and operating conditions such as sensor temperature, the sensing cell potentials and the reference cell potential on the sensor performance.

References

[1] T. Ritter, G. Hagen, J. Kita, S. Wiegärtner, F. Schubert, and R. Moos. Self-heated HTCC-based ceramic disc for mixed potential sensors and for direct conversion sensors for automotive catalysts. *Sensors and Actuators B: Chemical*, 248:793 – 802, 2017.

[2] Y. Guan, F. Liu, B. Wang, X. Yang, X. Liang, H. Suo, P. Sun, Y. Sun, J. Ma, J. Zheng, Y. Wang, and G. Lu. Highly sensitive amperometric nafion-based CO sensor using Pt/C electrodes with different kinds of carbon materials. *Sensors and Actuators B: Chemical*, 239:696 – 703, 2017.

[3] A. Kalyakin, J. Lyagaeva, D. Medvedev, A. Volkov, A. Demin, and P. Tsiakaras. Characterization of proton-conducting electrolyte based on La0.9Sr0.1YO3−δ and its application in a hydrogen amperometric sensor. *Sensors and Actuators B: Chemical*, 225:446 – 452, 2016.

[4] R. Sharan, M. Roy, A.K. Tyagi, and A. Dutta. Lanthanum gallate based amperometric electrochemical sensor for detecting ammonia in ppm level: Optimization of electrode compositions. *Sensors and Actuators B: Chemical*, 258:454 – 460, 2018.

[5] Thomas Ritter, Julia Lattus, Gunter Hagen, and Ralf Moos. A finite element model for mixed potential sensors. *Sensors and Actuators B: Chemical*, 287:476 – 485, 2019.

[6] T. Ritter, G. Hagen, J. Lattus, and R. Moos. Solid state mixed-potential sensors as direct conversion sensors for automotive catalysts. *Sensors and Actuators B: Chemical*, 2017.

[7] L. Wang, B. Han, Z. Wang, L. Dai, H. Zhou, Y. Li, and H. Wang. Effective improvement of sensing performance of amperometric NO2 sensor by Ag-modified nano-structured CuO sensing electrode. *Sensors and Actuators B: Chemical*, 207:791 – 800, 2015.

[8] F. Liu, B. Wang, X. Yang, Y. Guan, Q. Wang, X. Liang, P. Sun, Y. Wang, and G. Lu. High-temperature NO2 gas sensor based on stabilized zirconia and CoTa2O6 sensing electrode. *Sensors and Actuators B: Chemical*, 240:148 – 157, 2017.

[9] Thomas Ritter, Julia Lattus, Gunter Hagen, and Ralf Moos. On the influence of the NOx equilibrium reaction on mixed potential sensor signals: A comparison between FE modelling and experimental data. *Sensors and Actuators B: Chemical*, 296:126627, 2019.

[10] H. Sasaki, D. Scholl, M. Parsons, H. Inagaki, K. Shiotani, J. Visser, G. Zawacki, T. Kawai, S.Teramoto, and T. Kubinski. Development of an Al2O3/ZrO2-composite high-accuracy NOx sensor. In *SAE Technical Paper*. SAE International, 04 2010.

[11] T. Ritter, M. Seibel, F. Hofmann, M. Weibel, and R. Moos. Simulation of a NOx Sensor for Model-Based Control of Exhaust Aftertreatment Systems. *Topics in Catalysis*, Dec 2018.

[12] Y. Liu, J. Zhao, and J.M. Lee. Conventional and new materials for selective catalytic reduction (SCR) of NOx. *ChemCatChem*, 10(7):1499–1511, 2018.

[13] J.K. Lai and I.E. Wachs. A perspective on the selective catalytic reduction (SCR) of NO with NH3 by supported V2O5–WO3/TiO2 catalysts. *ACS Catalysis*, 8(7):6537–6551, 2018.

[14] A. Frobert, S. Raux, Y. Creff, and E. Jeudy. About cross-sensitivities of NOx sensors in SCR operation. SAE 2013-01-1512.

[15] Q. Lin and P. Chen. An NOx Sensor-Based Direct Algebraic Approach-Newton Observer for Urea Selective Catalytic Reduction System State Estimations. *Journal of Dynamic Systems, Measurement, and Control*, 140(11), 06 2018. 111004.

[16] M. Aliramezani, C.R. Koch, and R.E. Hayes. Estimating tailpipe NOx concentration using a dynamic NOx/ammonia cross sensitivity model coupled to a three state control oriented SCR model. *IFAC-PapersOnLine*, 49(11):8–13, 2016.

[17] J. Gong and V.R. Rajagopal. NH3 slip detection using NOx sensor, July 4 2019. US Patent App. 16/312,170.

[18] H. Zhang, J. Wang, and Y.Y. Wang. Removal of sensor ammonia cross sensitivity from contaminated measurements in Diesel-engine selective catalytic reduction systems. *Fuel*, 150(0):448 – 456, 2015.

[19] C. Lee. Modeling urea-selective catalyst reduction with vanadium catalyst based on NH3 temperature programming desorption experiment. *Fuel*, 173:155 – 163, 2016.

[20] A. Ko, Y. Woo, J. Jang, Y. Jung, Y. Pyo, H. Jo, O. Lim, and Y. Jae Lee. Availability of NH3 adsorption in vanadium-based SCR for reducing NOx emission and NH3 slip. *Journal of Industrial and Engineering Chemistry*, 78:433 – 439, 2019.

[21] M. F. Hsieh and J. Wang. An extended Kalman filter for NOx sensor ammonia cross-sensitivity elimination in selective catalytic reduction applications. In *American Control Conference (ACC), 2010*, pages 3033–3038, June 2010.

[22] A. Bonfils, Y. Creff, O. Lepreux, and N. Petit. Closed-loop control of a SCR system using a NOx sensor cross-sensitive to NH3. *Journal of Process Control*, 24(2):368 – 378, 2014. ADCHEM 2012 Special Issue.

[23] M. Devarakonda, G. Parker, J.H. Johnson, and V. Strots. Model-based control system design in a urea-SCR aftertreatment system based on NH3 sensor feedback. *International Journal of Automotive Technology*, 10(6):653–662, 2009.

[24] K. Jiang, E. Cao, and L. Wei. NOx sensor ammonia cross-sensitivity estimation with adaptive unscented Kalman filter for Diesel-engine selective catalytic reduction systems. *Fuel*, 165:185 – 192, 2016.

[25] H. Zhang and J. Wang. NOx Sensor Ammonia-Cross-Sensitivity Factor Estimation in Diesel Engine Selective Catalytic Reduction Systems. *Journal of Dynamic Systems, Measurement, and Control*, 137(6), 06 2015. 061015.

[26] J. Figura, J. Pekar, P. Krejza, D. Mracek, D. Von Wissel, and T Zhang. NO2/NOx Ratio and NH3 Storage Estimation of Automotive SCR Multi-Brick Systems. In *WCX^{TM} 17: SAE World Congress Experience*. SAE International, mar 2017.

[27] S. Dosda, D. Berthout, G. Mauviot, and A. Nogre. Modeling of a DOC SCR-F SCR exhaust line for design optimization taking into account performance degradation due to hydrothermal aging. *SAE International Journal of Fuels and Lubricants*, 9(3):621–632, 2016.

[28] A. Wang, Y. Wang, E.D. Walter, N. M. Washton, Y. Guo, G. Lu, C.H.F. Peden, and F. Gao. NH3-SCR on Cu, Fe and Cu+Fe exchanged beta and SSZ-13 catalysts: Hydrothermal aging and propylene poisoning effects. *Catalysis Today*, 320:91–99, 2019.

[29] M. Aliramezani, C.R. Koch, R.E. Hayes, and R. Patrick. Amperometric solid electrolyte NOx sensors – the effect of temperature and diffusion mechanisms. *Solid State Ionics*, 313(Supplement C):7 – 13, 2017.

[30] M. Aliramezani, C.R. Koch, and R. Patrick. Phenomenological model of a solid electrolyte NOx and O2 sensor using temperature perturbation for on-board diagnostics. *Solid State Ionics*, 321:62 – 68, 2018.

[31] M. Aliramezani, C.R. Koch, M. Secanell, R.E. Hayes, and R. Patrick. An electrochemical model of an amperometric NOx sensor. *Sensors and Actuators B*, 290:302–311, 2019.

[32] M. Aliramezani, C.R. Koch, and R. Patrick. A variable-potential amperometric hydrocarbon sensor. *IEEE Sensors Journal*, 19(24):12003–12010, 2019.

[33] C.M. Bishop. *Pattern recognition and machine learning*. springer, 2006.

[34] I.H. Witten, E. Frank, M.A. Hall, and C.J. Pal. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.

[35] Y. Xu, R. Guo, and L. Wang. A twin multi-class classification support vector machine. *Cognitive computation*, 5(4):580–588, 2013.

[36] M. Tanveer. Robust and sparse linear programming twin support vector machines. *Cognitive Computation*, 7(1):137–149, 2015.

[37] J. Thangaraja and C. Kannan. Effect of exhaust gas recirculation on advanced Diesel combustion and alternate fuels - A review. *Applied Energy*, 180:169 – 184, 2016.

[38] I.M. Yusri, A.P.P. Abdul Majeed, R. Mamat, M.F. Ghazali, Omar I. Awad, and W.H. Azmi. A review on the application of response surface method and artificial neural network in engine performance and exhaust emissions characteristics in alternative fuel. *Renewable and Sustainable Energy Reviews*, 90:665 – 686, 2018.

[39] A.S. Silitonga, H.H. Masjuki, Hwai Chyuan Ong, A.H. Sebayang, S. Dharma, F. Kusumo, J. Siswantoro, Jassinnee Milano, Khairil Daud, T.M.I. Mahlia, Wei-Hsin Chen, and Bambang Sugiyanto. Evaluation of the engine performance and exhaust emissions of biodiesel-bioethanol-diesel blends using kernel-based extreme learning machine. *Energy*, 159:1075 – 1087, 2018.

[40] Pak Kin Wong, Xiang Hui Gao, Ka In Wong, and Chi Man Vong. Online extreme learning machine based modeling and optimization for point-by-point engine calibration. *Neurocomputing*, 277:187–197, 2018.

[41] Pak Kin Wong, Ka In Wong, Chi Man Vong, and Chun Shun Cheung. Modeling and optimization of biodiesel engine performance using kernel-based extreme learning machine and cuckoo search. *Renewable Energy*, 74:640 – 647, 2015.

[42] Ron Kohavi and Dan Sommerfield. Feature subset selection using the wrapper method: Overfitting and dynamic search space topology. In *KDD*, pages 192–197, 1995.

[43] http://www.ecm-co.com/.

[44] Robert Klikach. Investigation and Analysis of RCCI using NVO on a Converted Spark Ignition Engine. Master's thesis, 2018.

[45] Vladimir Vapnik and AJ Lerner. Generalized portrait method for pattern recognition. *Automation and Remote Control*, 24(6):774–780, 1963.

[46] Vladimir Vapnik and Alexey Chervonenkis. A note on class of perceptron. *Automation and Remote Control*, 24, 1964.

[47] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[48] Harris Drucker, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. Support vector regression machines. In *Advances in neural information processing systems*, pages 155–161, 1997.

[49] Bernhard Schölkopf, Alexander J Smola, Francis Bach, et al. *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2002.

[50] Alex J Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and computing*, 14(3):199–222, 2004.

[51] Richard Bellman et al. The theory of dynamic programming. *Bulletin of the American Mathematical Society*, 60(6):503–515, 1954.

[52] William Karush. Minima of functions of several variables with inequalities as side constraints. *M. Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago*, 1939.

[53] Harold W Kuhn and Albert W Tucker. Nonlinear programming, in (J. Neyman, ed.) Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1951.

[54] M Narasimha Murty and Rashmi Raghava. *Support Vector Machines and Perceptrons: Learning, Optimization, Classification, and Application to Social Networks*. Springer, 2016.

[55] Gregory A Landrum, Julie E Penzotti, and Santosh Putta. Machine-learning models for combinatorial catalyst discovery. *Measurement Science and Technology*, 16(1):270, 2004.

[56] T. Nakamura, Y. Sakamoto, K. Saji, and J. Sakata. NOx decomposition mechanism on the electrodes of a zirconia-based amperometric NOx sensor. *Sensors and Actuators B: Chemical*, 93(1):214–220, 2003.

[57] M. Aliramezani. *Production engine emission sensor modeling for in-use measurement and on-board diagnostics*. PhD thesis, University of Alberta, 2019.

[58] M.S. Silberberg, P. Amateis, S.. Lavieri, and R. Venkateswaran. *Chemistry: The molecular nature of matter and change*. Mosby St. Louis, Missouri, USA, 1996.

[59] J.C. Amphlett, R.M. Baumert, R.F. Mann, B.A. Peppley, P.R. Roberge, and T.J. Harris. Performance modeling of the Ballard Mark IV solid polymer electrolyte fuel cell I. Mechanistic model development. *Journal of the Electrochemical Society*, 142(1):1–8, 1995.

[60] M. Aliramezani, A. Norouzi, C.R. Koch, and R.E. Hayes. A control oriented diesel engine NOx emission model for on board diagnostics and engine control with sensor feedback. *Proceedings of Combustion Institute - Canadian Section*, 2019.