Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Low-Complexity Polynomial Regression Framework for Environmental Drift Compensation in MOX Gas Sensing Systems

Abdulnasser Nabil Abdullah
Universiti Malaysia Perlis (UniMAP)
Kamarulzaman Kamarudin
kamarulzaman@unimap.edu.my

Universiti Malaysia Perlis (UniMAP)
Ahmad Syamil
Universiti Malaysia Perlis (UniMAP)
Md Ashequl Islam
Adelaide University
Syed Muhammad Mamduh
Universiti Malaysia Perlis (UniMAP)
Ammar Zakaria
Universiti Malaysia Perlis (UniMAP)

Research Article

Keywords: MOX gas sensors, environmental drift compensation, temperature and humidity cross-sensitivity, polynomial regression

Posted Date: August 31st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10228738/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Sensing and Imaging on November 13th, 2026. See the published version at https://doi.org/10.1007/s11220-026-00917-1.

Low-Complexity Polynomial Regression Framework for Environmental Drift Compensation in MOX Gas Sensing Systems

Abdulnasser Nabil Abdullah[1,2,*], Kamarulzaman Kamarudin[1,2,*], Ahmad Syamil[1], Md Ashequl Islam[3], Syed Muhammad Mamduh[1,2] and Ammar Zakaria[1,2]

[1] Faculty of Electrical Engineering & Technology, Universiti Malaysia Perlis (UniMAP), Arau, Malaysia
[2] Centre of Excellence for Advanced Sensor Technology (CEASTech), Universiti Malaysia Perlis, Arau, Malaysia
[3] School of Electrical and Mechanical Engineering, Adelaide University, South Australia, Australia

[*]Correspondence: nassernabil@studentmail.unimap.edu.my; kamarulzaman@unimap.edu.my

Keywords: MOX gas sensors, environmental drift compensation, temperature and humidity cross-sensitivity, polynomial regression.

Abstract

Metal-oxide semiconductor (MOX) gas sensors are critical components in low-cost environmental monitoring architectures; however, their cross-sensitivity to ambient temperature and relative humidity induces severe baseline drift, compromising measurement accuracy and long-term stability. Therefore, accurately modelling the nonlinear influence of ambient temperature and relative humidity on sensor resistance remains a significant challenge for environmental drift compensation. This study proposes a structured polynomial mathematical framework to optimize and implement drift compensation across three distinct commercial MOX sensors (MQ4, MQ136, and TGS2442). The proposed approach systematically evaluates and compares the performance of six higher-order polynomial regression formulations (M1-M6) incorporating varying degrees of interaction and non-linear terms. Model structures were comprehensively evaluated using 10-fold cross-validation to establish the optimal balance between mathematical complexity and prediction accuracy. Cross-validation results show that increasing model complexity yields significant error reduction initially, leading to the selection of Polynomial Model M4 as representative models that provide the best trade-off between prediction accuracy and mathematical complexity. Compared with the selected linear regression model, M4 reduced the cross-validation mean squared error (MSE) by 18.5% for the MQ4 sensor, 46.5% for the MQ136 sensor, and 25.0% for the pulse-heated TGS2442 sensor. External verification further demonstrated the effectiveness of the selected polynomial model on an independent dataset, where the standard deviation of the corrected sensor responses was reduced by 71.1% for MQ4, 57.2% for MQ136, and 15.5% for TGS2442, indicating improved measurement stability under previously unseen operating conditions. These findings demonstrate that the proposed polynomial regression framework provides a mathematically transparent, computationally efficient, and low-overhead drift compensation solution suitable for gas sensing applications.

1. Introduction

Gas sensing technologies play a crucial role in modern industrial and environmental monitoring systems, where accurate detection of hazardous and combustible gases is essential for safety, process control, and regulatory compliance [1]. Industries such as petrochemical processing, wastewater treatment, agriculture, and smart building management increasingly rely on gas sensors for continuous real-time monitoring. In these applications, sensor reliability and measurement stability are critical, particularly under dynamic environmental conditions [2]. However, practical deployment environments often experience significant fluctuations in ambient temperature and relative humidity, particularly in tropical and industrial settings. These environmental variations introduce measurement instability and response drift even when the target gas concentration remains constant. Previous studies have shown

that changes in temperature and humidity can alter sensor baseline resistance and sensitivity, leading to substantial measurement errors and long-term instability if not properly compensated [3]. Consequently, achieving robust and accurate gas detection under varying environmental conditions remains a significant challenge in real-world sensing applications.

Among the various gas sensing technologies available, metal-oxide semiconductor (MOX) gas sensors are widely adopted in industrial and environmental monitoring systems due to their low cost, simple structure, high sensitivity, and ease of integration into embedded platforms [4]. These chemiresistive sensors operate based on changes in the electrical resistance of a metal oxide sensing layer, typically composed of materials such as SnO_2, ZnO, WO_3, or TiO_2, when exposed to target gases [5], [6]. The sensing mechanism relies on surface adsorption and redox reactions between gas molecules and oxygen species adsorbed on the sensor surface, which modifies the charge carrier concentration and consequently alters the sensor resistance. However, the same surface-based sensing mechanism that provides high sensitivity also makes MOX sensors inherently susceptible to environmental influences, particularly temperature and humidity variations [7]. This fundamental dependency on surface reactions underpins the environmental instability issues described earlier, making compensation strategies essential for reliable deployment in real-world applications.

Despite their extensive use, MOX gas sensors suffer from inherent limitations that affect measurement reliability, particularly cross-sensitivity and sensor drift [8]. Cross-sensitivity refers to the phenomenon where a sensor responds not only to the target gas but also to other interfering gases or environmental parameters due to the non-specific surface reaction mechanism of metal oxides [6]. Since the sensing principle is based on oxygen adsorption and redox reactions, various oxidizing and reducing gases can trigger similar resistance changes, thereby reducing selectivity and complicating signal interpretation [9], [10]. In addition to cross-sensitivity, sensor drift represents another critical challenge in long-term applications. Drift is characterized by a gradual change in sensor baseline and sensitivity over time, even under constant gas exposure conditions. Studies have shown that drift may arise from reversible processes such as humidity adsorption as well as irreversible mechanisms including material aging and microstructural changes in the sensing layer [11 - 12]. Without appropriate correction, these phenomena introduce systematic errors that can exceed actual gas concentration variations, significantly degrading measurement accuracy and reliability in continuous monitoring systems.

Among the environmental factors influencing MOX gas sensors, ambient temperature and relative humidity are the most dominant contributors to response instability [13], [14]. Variations in temperature directly affect the kinetics of surface reactions and charge carrier mobility within the metal oxide layer, resulting in significant changes in baseline resistance and sensor sensitivity [15], [16]. Even small temperature fluctuations have been reported to induce noticeable baseline shifts and alter calibration curves, thereby producing inconsistent sensor outputs under constant gas concentrations [17]. Similarly, humidity exerts a substantial influence on sensor performance through the adsorption of water molecules onto the sensing surface. The presence of water vapor modifies surface chemistry, competes with oxygen adsorption sites, and forms hydroxyl groups that interfere with gas surface interactions [18]. This leads to nonlinear variations in sensor resistance, sensitivity reduction at high humidity levels, and in some cases prolonged transient recovery behaviour [19]. The combined effects of temperature and humidity therefore create complex and interdependent disturbances in sensor output, reinforcing the need for robust compensation mechanisms capable of modelling these nonlinear environmental interactions. Consequently, accurately modelling these nonlinear environmental effects has become one of the primary challenges in MOX sensor calibration.

To mitigate these effects, various mathematical methods have been investigated. Among these, regression-based modelling approaches have been widely used for environmental drift compensation because of their simplicity and suitability for embedded sensing platforms. Multivariate linear regression models that incorporate environmental variables such as temperature and humidity have demonstrated the ability to reduce response variability and improve drift stability under controlled conditions [20]. However, several studies have reported that the relationship between sensor resistance and environmental parameters is inherently nonlinear due to complex surface adsorption and reaction kinetics within MOX materials [21 - 24]. Consequently, linear models may fail to accurately capture curvature effects and higher-order interactions between environmental variables and sensor response, particularly over wider operating ranges. Recently, machine learning techniques have been increasingly applied to gas sensor calibration and drift compensation problems. In particular, supervised learning

models can learn complex relationships between sensor measurements and environmental variables directly from experimental data. Regression-based machine learning approaches have therefore emerged as effective tools for modelling the nonlinear influence of temperature and humidity on MOX sensor behaviour. By treating sensor resistance and environmental parameters as input features and the reference resistance as the target output, these models can automatically learn compensation functions that improve measurement stability under varying environmental conditions.

This study proposes a polynomial regression framework for compensating for the temperature and humidity induced drift in MOX gas sensors by modelling the nonlinear relationship between sensor resistance, ambient temperature, and relative humidity. Accurate environmental drift compensation is essential because variations in temperature and humidity significantly influence sensor responses, leading to reduced measurement accuracy and stability even under constant gas concentrations. Therefore, an effective mathematical compensation model is required to improve measurement reliability while maintaining the computational efficiency needed for practical implementation in low-cost environmental monitoring systems.

Despite the promising capability of polynomial regression for modelling nonlinear sensor behaviour, existing studies have primarily focused on a single sensor type or a single regression formulation, with limited investigation into the influence of model complexity on drift compensation performance [20], [25], [26]. Consequently, the applicability of polynomial regression across different commercial MOX gas sensors operating under varying environmental conditions remains insufficiently understood. Building upon this research gap, the present study systematically develops and evaluates six polynomial regression models with progressively increasing complexity to identify an appropriate balance between prediction accuracy and mathematical simplicity. The developed models are assessed using three commercial MOX gas sensors (MQ4, MQ136, and TGS2442) to investigate their effectiveness in compensating temperature and humidity induced drift across different sensing characteristics. Finally, the selected polynomial model is validated using an independent dataset to verify its generalization capability, thereby providing a practical and computationally efficient framework for environmental drift compensation in low-cost MOX gas sensing applications.

2. Challenges and limitations of MOX gas sensor

In this study, the MOX sensors used as sensing elements are MQ-4 for CH_4, MQ-136 for hydrogen sulfide H_2S, and TGS2442 for CO. These sensors are often used in environmental monitoring applications and microcontroller-based sensing systems because their output signals based on electrical resistance changes can be easily obtained and processed. The wide availability and low cost make MOX sensors suitable for multi-gas sensing systems in real environments.

Although MOX gas sensors have many advantages, their performance in practical applications is often hampered by several limitations stemming from the surface-based sensing mechanism. The response of MOX sensors depends on the adsorption and desorption processes of gases on the surface of metal oxide materials, which are very sensitive to external operating conditions. In real environments, sensor operation rarely occurs under ideal or controlled conditions, and variations in ambient temperature, relative humidity, and gas composition are unavoidable. This can cause instability in the sensor response and make it difficult to directly interpret the sensor readings, especially for continuous or long-term monitoring applications.

The MOX gas sensors using sensing materials such as tin oxide (SnO_2) and zinc oxide (ZnO) have been widely used to detect gases such as methane, carbon monoxide, hydrogen sulphide, and volatile organic compounds (VOCs). Based on the research obtained, several practical problems have been reported regarding the use of these sensors. Among them are baseline value fluctuations, sensitivity changes under different operating conditions, and inconsistency of sensor responses to the same gas concentration over different periods of time. This behavior is attributed to changes in sensor surface properties, repeated exposure to gases, and the influence of environmental factors [27].

Additionally, the main challenges of semiconductor gas sensors, particularly metal oxide-based sensors. MOX sensors using materials such as SnO_2 are reported to exhibit performance that is highly dependent on surface conditions and operating temperature, where small changes in the environment can cause significant variations in sensor response. It is also reported that lack of selectivity is a

fundamental problem of MOX sensors because the sensing mechanism that relies on oxygen redox reactions causes the sensor to respond to a wide range of gases. Furthermore, long-term operation at high temperatures can cause changes in the microstructure of the sensor material, which in turn contributes to response instability and changes in sensor characteristics over time [28], [29].

Overall, the major challenges affecting MOX gas sensors include cross-sensitivity, sensor drift, and strong dependence on ambient temperature and relative humidity, as summarized in Table 1. These factors cause uncertainty in the sensor response and limit the accuracy of gas concentration estimation when raw data is used directly. Therefore, a thorough understanding of the factors affecting the performance of MOX sensors is essential before further analysis and correction methods are implemented. These key factors will be discussed in more detail in the next section.

Table 1: Summary of Challenges in Using MOX Gas Sensors in Practical Applications.

Challenge	Causes	Impact on Sensor Response
Cross-sensitivity	Oxygen redox-based sensing mechanism.	Response overlap between target gas and interferer gas.
Sensor drift	Material aging, microstructural changes, repeated exposure.	Baseline shifts and sensitivity changes over time
Ambient Temperature	Surface reaction rate depends on operating temperature.	Response magnitude variation and signal instability
Ambient Humidity	Adsorption of water molecules on the sensor surface.	Baseline changes and spurious responses

3. Methodology

This section presents the proposed polynomial regression framework developed to compensate for environmentally induced drift in MOX gas sensors. The framework consists of dataset preparation, polynomial model development, systematic model complexity evaluation, model selection using 10-fold cross-validation, and external validation using an independent dataset.

On the other hand, the linear regression framework used as the baseline comparison was previously developed and comprehensively described in [20]. Therefore, this paper focuses on the development and evaluation of the proposed polynomial regression models, while the linear regression results are presented only as the benchmark for performance comparison.

3.1 Dataset

This study employed the open-source gas sensing dataset published by Abdullah [30], which was developed for the characterization of MOX gas sensors under controlled temperature and humidity conditions for confined-space atmospheric monitoring applications. The dataset comprises synchronized measurements collected from three commercially available MOX gas sensors, namely MQ4, MQ136, and TGS2442, which are primarily sensitive to methane (CH_4), hydrogen sulfide (H_2S), and carbon monoxide (CO), respectively. In addition to sensor resistance (Rs), the dataset includes the corresponding ambient temperature and relative humidity measurements, providing a comprehensive representation of the environmental conditions affecting sensor responses.

The dataset was acquired under multiple target gas concentration levels and four controlled relative humidity conditions (45%, 55%, 65%, and 75% RH), with temperature systematically varied throughout the experiments (16°C, 18°C, 20°C, 22°C, 24°C, 26°C, 28°C, 30°C) as described in [30]. Such controlled environmental variations provide sufficient diversity for investigating the influence of temperature and humidity on MOX sensor behaviour. The availability of synchronized sensor resistance, temperature, and humidity measurements across different operating conditions enables the proposed regression models to effectively learn environmental dependencies, thereby improving the robustness and generalization capability of the developed compensation model.

Based on this dataset, a systematic polynomial regression framework was developed to model and compensate for the environmental drift of the investigated MOX gas sensors, as illustrated in Figure 1. The framework begins with data preparation and dataset partitioning, followed by polynomial model development and performance evaluation using 10-fold cross-validation. The optimal model is subsequently selected and validated using an independent unseen dataset to verify its generalization capability. The validated model is then used to generate the drift-compensated sensor response for subsequent performance assessment.

The proposed framework consists of four main stages: Data preparation and preprocessing, Model input and training, model selection, and External validation. Each stage is described in detail in the following

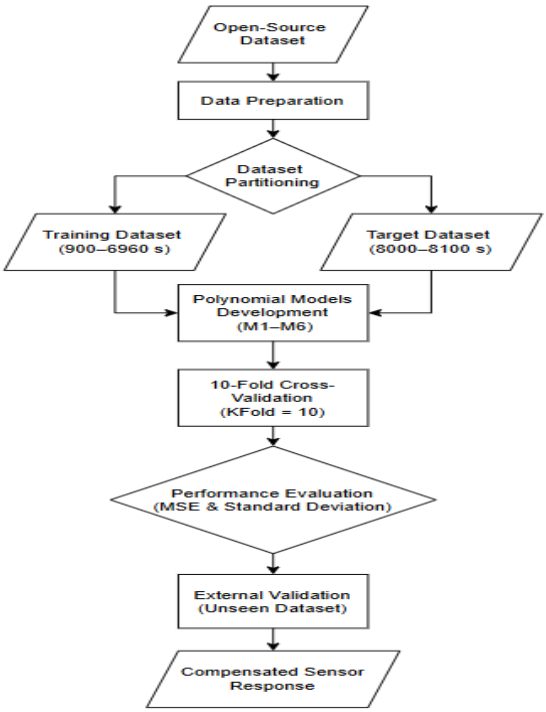


Figure 1. Overall workflow of the proposed polynomial regression framework for environmental drift compensation.

subsections.

3.2 Polynomial regression model

Polynomial regression extends conventional linear regression by incorporating higher-order terms to model nonlinear relationships between predictor variables and the target response. By introducing polynomial terms, the regression model can represent complex response patterns that cannot be adequately described using a purely linear relationship [27]. This capability makes polynomial regression particularly suitable for modelling the nonlinear influence of ambient temperature and relative humidity on the resistance of MOX gas sensors. The general equation for a polynomial model of degree (h) is expressed as:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \cdots + \beta_h x^h \quad (1)$$

where (y) is the predicted output, (x) is the predictor variable, β_0 is the intercept and $\beta_1, \beta_2, \ldots, \beta_h$ are the polynomial coefficients.

Although Equation (1) presents the general form of a polynomial regression model, the proposed compensation model employs multiple predictor variables, namely sensor resistance (R_s), ambient temperature (T), and relative humidity (RH), as expressed in Equation (2).

$$R_S^* = F(R_S, T, H) \tag{2}$$

where R_S^* represents the predicted drift-compensated resistance.

3.2.1 Data preparation and preprocessing

Prior to regression model development, the acquired sensor data was preprocessed to ensure that the input and target variables were prepared consistently for model training and evaluation. The raw dataset consisted of time-series measurements recorded over approximately 8460 s for each experimental file. Each file contained synchronized measurements of sensor resistance (Rs), temperature, and relative humidity. Before model development, the complete dataset was divided into two-time intervals based on the experimental timeline.

The first interval, from 900 s to 6960 s, was selected as the modelling dataset because it represented the period in which the sensors experienced environmental variations. The second interval, from 8000 s to 8100 s, represented the stable operating condition, and the average sensor resistance within this period was calculated as the reference value for each humidity level. These intervals were selected based on the experimental protocol reported in [26], where the initial transient response had stabilized and the sensor was fully exposed to the controlled environmental conditions. These datasets were subsequently used for regression model development and performance evaluation.

3.2.2 Model input and training

After the data preparation stage, the measured R_S, T, and H were used as the predictor variables, while the reference sensor resistance (R_S^*) obtained from the stable operating condition served as the target variable. Since the target values were available, the model development was formulated as a supervised regression problem.

To investigate the influence of polynomial and interaction terms on environmental drift compensation, six polynomial regression models with increasing complexity were developed. The models were constructed by gradually introducing second-order terms, third-order terms, and interaction terms involving R_S, T, and H. This progressive formulation enables the contribution of higher-order nonlinear effects and variable interactions to be systematically evaluated. The mathematical formulations of the proposed models are given in Equations (3) to (8).

$$M1: R_S^* = C_0 + C_1 R_S + C_2 T + C_3 H + C_4 {R_S}^2 + C_5 T^2 + C_6 H^2 \tag{3}$$

$$M2: R_S^* = C_0 + C_1 R_S + C_2 T + C_3 H + C_4 {R_S}^2 + C_5 T^2 + C_6 H^2 + C_7 R_S T + C_8 R_S H \tag{4}$$

$$M3: R_S^* = C_0 + C_1 R_S + C_2 T + C_3 H + C_4 {R_S}^2 + C_5 T^2 + C_6 H^2 + C_7 {R_S}^3 + C_8 T^3 + C_9 H^3 \tag{5}$$

$$M4: R_S^* = C_0 + C_1 R_S + C_2 T + C_3 H + C_4 {R_S}^2 + C_5 T^2 + C_6 H^2 + C_7 {R_S}^3 + C_8 T^3 + C_9 H^3 + C_{10} R_S T + C_{11} R_S H \tag{6}$$

$$M5: R_S^* = C_0 + C_1 R_S + C_2 T + C_3 H + C_4 {R_S}^2 + C_5 T^2 + C_6 H^2 + C_7 {R_S}^3 + C_8 T^3 + C_9 H^3 + C_{10} R_S T + C_{11} R_S H + C_{12} TH \tag{7}$$

$$M6: R_S^* = C_0 + C_1 R_S + C_2 T + C_3 H + C_4 {R_S}^2 + C_5 T^2 + C_6 H^2 + C_7 {R_S}^3 + C_8 T^3 + C_9 H^3 + C_{10} R_S T + C_{11} R_S H + C_{12} TH + C_{13} R_S TH \tag{8}$$

The coefficients of each polynomial model were estimated using the least-squares method to minimize the prediction error between the measured and reference sensor resistance.

3.2.3 Model selection

The performance of the linear and polynomial regression models was evaluated using 10-fold cross-validation to assess their prediction accuracy and generalization capability. The complete dataset was randomly divided into ten mutually exclusive subsets of approximately equal size without overlapping. During each iteration, nine subsets were used for model training, while the remaining subset was used for validation. This process was repeated ten times such that each subset served once as the validation set, and the average mean squared error (MSE) across all folds was used to evaluate model performance.

The MSE was used as the primary evaluation metric for comparing the predictive performance of the developed regression models. MSE is defined as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|^2 \tag{9}$$

where y_i is the reference sensor resistance and $\hat{y}_i$ is the predicted drift-compensated sensor resistance. MSE quantifies the average squared difference between the predicted and reference values, assigning greater penalties to larger prediction errors. Therefore, a lower MSE indicates better prediction accuracy and more effective drift compensation.

The average MSE obtained from the 10-fold cross-validation was subsequently used to compare linear and polynomial regression models. Rather than selecting the model solely based on the lowest MSE, the trade-off between prediction accuracy and model complexity was also considered. Consequently, the regression model providing the best overall balance between these criteria was selected for the subsequent drift compensation analysis and external validation.

3.2.4 External validation

After the model selection process, the regression model that demonstrated the best overall performance was selected for external validation. The selected model was then applied to an independent dataset that was not included during model development or 10-fold cross-validation. This validation step was performed to assess the model's ability to generalize and maintain its drift compensation performance when applied to previously unseen data.

The unseen dataset was pre-processed using the same procedure described in Section 3.2.1 before being used for prediction. The predicted drift-compensated sensor resistance was subsequently compared with the measured sensor resistance to evaluate the effectiveness of the developed compensation model. This external validation provides an additional assessment of the model's robustness and practical applicability under new operating conditions.

4. Result and Discussion

4.1 Environmental effect on raw MOX sensor responses

Prior to compensation modelling, an exploratory analysis was performed to investigate the influence of temperature and relative humidity on the responses of the MQ4, MQ136, and TGS2442 sensors. Figures 2 to 4 present the three-dimensional response surfaces and corresponding mean response trends under different environmental conditions.

The results indicate that all sensors exhibited noticeable response variation with changes in ambient temperature and humidity. For the MQ4 sensor, the resistance generally decreased with increasing temperature across all investigated methane concentration levels. Similar trends were observed for MQ136 and TGS2442, although the magnitude of response variation differed among sensors. These

observations indicate that sensor response characteristics are not determined solely by target gas concentration but are also significantly affected by environmental conditions.

Meanwhile, humidity affects sensor response through adsorption of water molecules on the sensor surface. Water molecules compete with oxygen species for active adsorption sites and alter surface chemistry through hydroxyl group formation. This behavior is consistent with the surface adsorption mechanism discussed in Section 2, where water molecules compete with oxygen species for active adsorption sites, thereby modifying electron exchange and altering sensor resistance. Consequently, variations in humidity may produce nonlinear baseline shifts and changes in sensor sensitivity

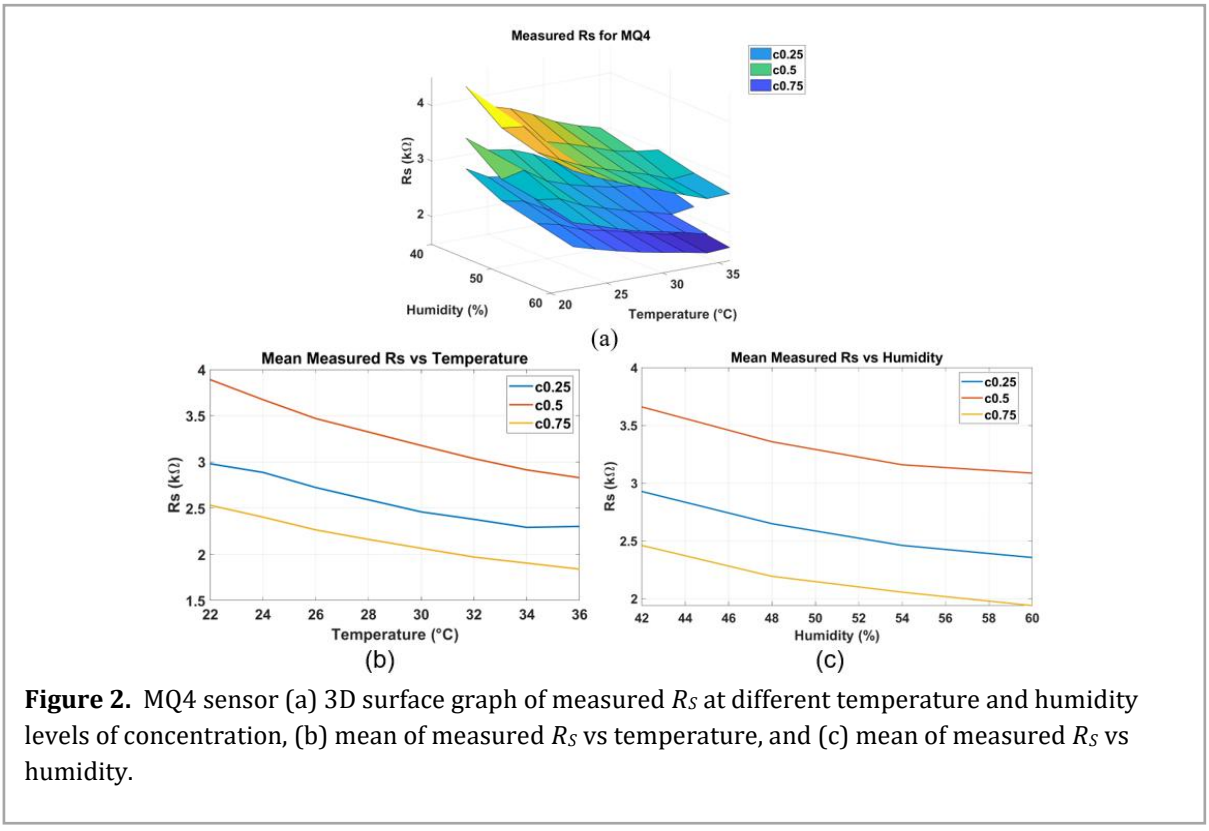


Figure 2. MQ4 sensor (a) 3D surface graph of measured R_S at different temperature and humidity levels of concentration, (b) mean of measured R_S vs temperature, and (c) mean of measured R_S vs humidity.

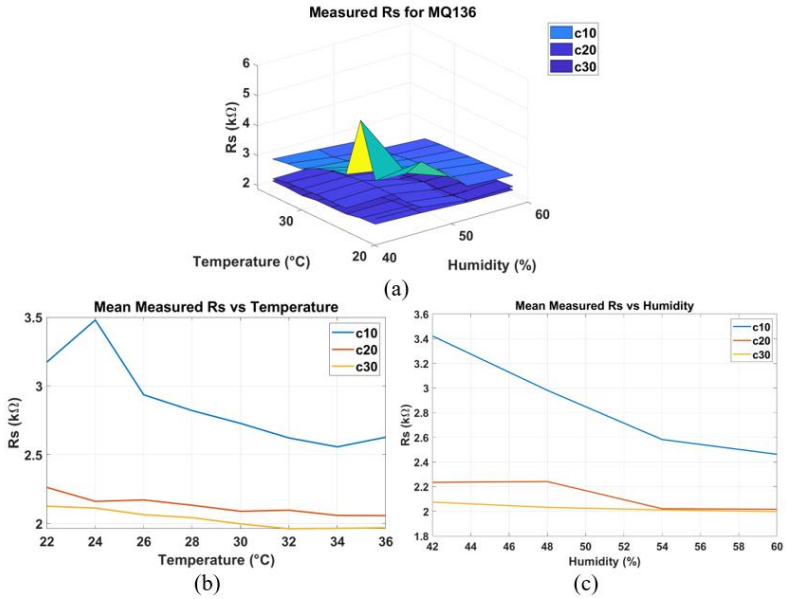


Figure 3. MQ136 sensor (a) 3D surface graph of measured R_S at different temperature and humidity levels of concentration, (b) mean of measured R_S versus temperature, and (c) mean of measured R_S versus humidity.

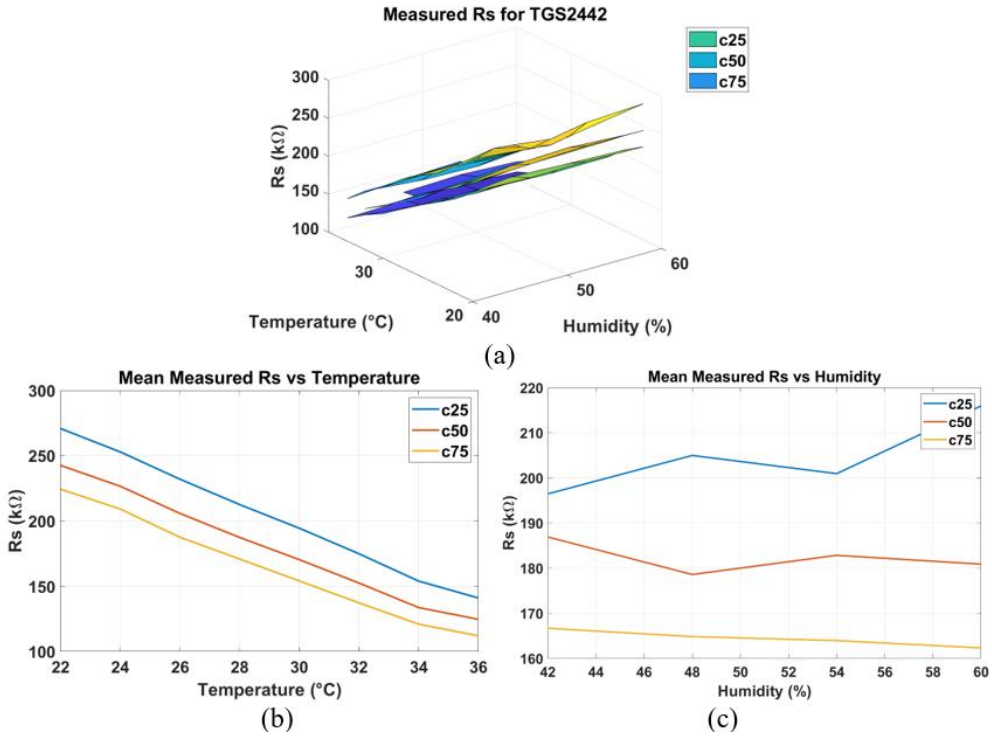


Figure 4. TGS 2442 sensor (a) 3D surface graph of measured R_S at different temperature and humidity levels of concentration, (b) mean of measured R_S versus temperature, and (c) mean of measured R_S versus humidity.

The observed environmental dependencies therefore justify the use of nonlinear regression formulations for drift compensation and motivate the comparative investigation between linear and polynomial models presented in the following sections.

4.2 Comparison between selected linear and polynomial models

The predictive performance of the linear and polynomial regression models was first evaluated using 10-fold cross-validation to assess their ability to compensate for gas sensor drift under varying environmental conditions. The MSE was used as the primary evaluation metric to compare the prediction accuracy of the developed regression models, as it provides a direct measure of the average squared prediction error.

As shown in Figure 5, the MSE generally decreases as the linear regression models progress from M1 to M7 for all three gas sensors. For the MQ4 sensor, a substantial reduction in MSE is observed from M1 to M2, followed by gradual improvements from M3 onwards. A similar trend is observed for the MQ136 sensor, where the prediction error decreases consistently with increasing model complexity before approaching a stable value. For the TGS2442 sensor, the MSE is considerably higher for M2 than for the other models, which may be attributed to the simplified model formulation that only incorporates the interaction term ($R_S TH$) without including temperature and relative humidity as independent predictors. Consequently, the direct influence of temperature and humidity on the TGS2442 sensor response may not be sufficiently represented, resulting in a larger prediction error. From M3 onwards, however, a significant reduction in MSE is achieved, after which the prediction error remains relatively stable.

Based on the cross-validation results, Models M5, M6, and M7 exhibit comparable prediction performance across all three sensors. However, only marginal reductions in MSE are observed beyond Model M5, indicating that the additional interaction terms introduced in Models M6 and M7 provide limited improvement in prediction accuracy. Therefore, Model M5 was selected as the representative linear regression model for subsequent comparison with the proposed polynomial regression models, as it offers simpler mathematical formulation while maintaining prediction performance comparable to the more complex models.

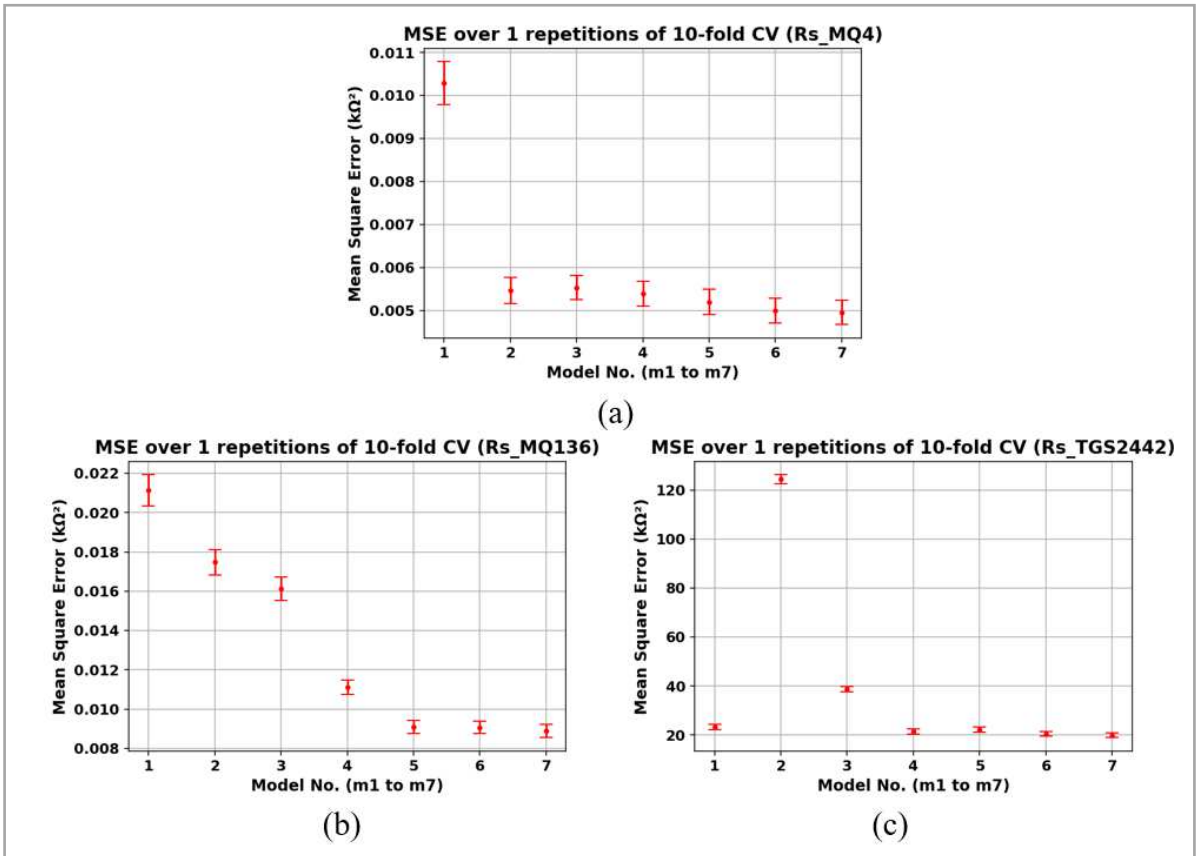


Figure 5. Averaged MSE and standard deviation for all gas sensors using 10-fold CV with 1 repetition to find the optimal model selection from m1 to m7 (linear regression), (a) MQ4, (b) MQ136, (c) TGS2442.

As shown in Figure 6, the polynomial regression models also demonstrate a progressive reduction in MSE as the model complexity increases from M1 to M6 for all three gas sensors. Compared with the linear regression models, the prediction error decreases more consistently across all sensors, indicating that the inclusion of higher-order polynomial terms provides a better representation of the nonlinear relationship between sensor resistance, temperature, and relative humidity.

Based on the cross-validation results, Models M4, M5, and M6 exhibit comparable prediction performance for all three sensors. Since only slight reductions in MSE are achieved from M4 to M6, Model M4 was selected as the representative polynomial regression model for subsequent comparison and external validation. This model provides a simpler mathematical formulation while maintaining prediction accuracy comparable to the more complex Models M5 and M6.

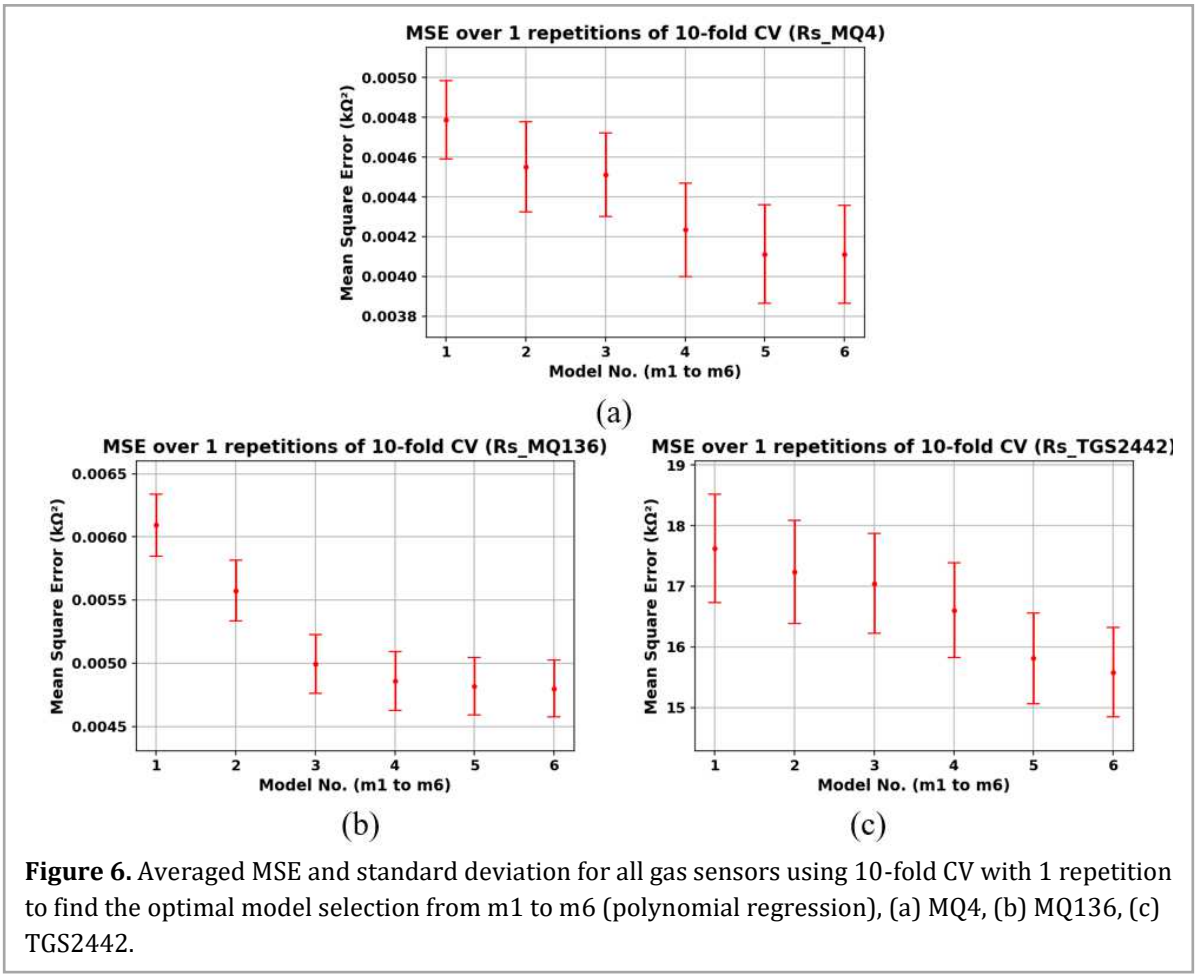


Figure 6. Averaged MSE and standard deviation for all gas sensors using 10-fold CV with 1 repetition to find the optimal model selection from m1 to m6 (polynomial regression), (a) MQ4, (b) MQ136, (c) TGS2442.

Based on the cross-validation results, Model M5 and Model M4 were selected as the representative linear and polynomial regression models, respectively. These models were subsequently compared to evaluate the effectiveness of polynomial regression in improving gas sensor drift compensation. The comparison was performed using the mean MSE and standard deviation obtained from the 10-fold cross-validation results, as summarized in Table 2.

As shown in Table 2, the polynomial regression model consistently achieves lower MSE values than the corresponding linear regression model for all three gas sensors. For the MQ4 sensor, the MSE is reduced from 5.196×10^{-3} kΩ² to 4.235×10^{-3} kΩ², representing an 18.5% reduction. Similarly, the MQ136 sensor exhibits a reduction from 9.076×10^{-3} kΩ² to 4.860×10^{-3} kΩ², corresponding to a 46.5% improvement. For the TGS2442 sensor, the MSE decreases from 22.136 kΩ² to 16.606 kΩ², yielding a 25.0% reduction. These results demonstrate that the polynomial regression model consistently provides higher prediction accuracy than the corresponding linear regression model across

all three sensors by more effectively capturing the nonlinear influence of temperature and humidity on sensor resistance.

In addition to the lower MSE values, the polynomial regression model also exhibits lower standard deviations for all three sensors, indicating more consistent prediction performance throughout the cross-validation process. As summarized in Table 2, the standard deviation is reduced from 0.295×10^{-3} kΩ² to 0.235×10^{-3} kΩ² for the MQ4 sensor (20.3% reduction), from 0.327×10^{-3} kΩ² to 0.233×10^{-3} kΩ² for the MQ136 sensor (28.7% reduction), and from 1.063 kΩ² to 0.7773 kΩ² for the TGS2442 sensor (26.9% reduction). The lower standard deviations indicate that the polynomial regression model produces more consistent predictions across different cross-validation folds, demonstrating improved robustness, reduced variability, and greater generalization capability than the corresponding linear regression model.

The improved performance of the polynomial regression model can be attributed to the inclusion of second order and third-order polynomial terms, which enable the model to better capture the nonlinear relationship between sensor resistance, ambient temperature, and relative humidity. Unlike the linear regression model, the polynomial formulation can represent the complex environmental effects that contribute to gas sensor drift, thereby producing more accurate and stable drift compensation.

Table 2: Cross-validation MSE and standard deviation for linear regression models (M1-M7) and polynomial regression models (M1-M6).

Sensor	Model (Linear)	Mean Square Error ($k\Omega^2$)	Standard Deviation Error ($k\Omega^2$)	Model (Polynomial)	Mean Square Error ($k\Omega^2$)	Standard Deviation Error ($k\Omega^2$)
MQ4	M1	10.28×10^{-3}	0.5×10^{-3}	M1	4.788×10^{-3}	0.196×10^{-3}
	M2	5.466×10^{-3}	0.303×10^{-3}	M2	4.552×10^{-3}	0.227×10^{-3}
	M3	5.531×10^{-3}	0.282×10^{-3}	M3	4.512×10^{-3}	0.209×10^{-3}
	M4	5.385×10^{-3}	0.287×10^{-3}	M4	4.235×10^{-3}	0.235×10^{-3}
	M5	5.196×10^{-3}	0.295×10^{-3}	M5	4.113×10^{-3}	0.248×10^{-3}
	M6	4.992×10^{-3}	0.29×10^{-3}	M6	4.112×10^{-3}	0.248×10^{-3}
	M7	4.952×10^{-3}	0.283×10^{-3}			
MQ136	M1	21.12×10^{-3}	0.799×10^{-3}	M1	6.092×10^{-3}	0.244×10^{-3}
	M2	17.47×10^{-3}	0.636×10^{-3}	M2	5.575×10^{-3}	0.238×10^{-3}
	M3	16.13×10^{-3}	0.589×10^{-3}	M3	4.996×10^{-3}	0.231×10^{-3}
	M4	11.11×10^{-3}	0.35×10^{-3}	M4	4.86×10^{-3}	0.233×10^{-3}
	M5	9.076×10^{-3}	0.327×10^{-3}	M5	4.817×10^{-3}	0.228×10^{-3}
	M6	9.076×10^{-3}	0.324×10^{-3}	M6	4.799×10^{-3}	0.226×10^{-3}
	M7	8.89×10^{-3}	0.334×10^{-3}			
TGS2442	M1	23.222	1.064	M1	17.622	0.8943
	M2	124.379	1.834	M2	17.237	0.8504
	M3	38.706	1.028	M3	17.046	0.8248

M4	21.449	1.018	M4	16.606	0.7773
M5	22.136	1.063	M5	15.813	0.7454
M6	20.513	0.969	M6	15.583	0.7394
M7	19.932	0.983			

4.3 Drift compensation performance using selected models

Based on the cross-validation results presented in Section 4.2, Linear Model M5 and Polynomial Model M4 were selected as representative regression models for subsequent drift compensation analysis. After model selection, both models were reconstructed using the complete dataset to maximize the utilization of the available information, eliminating the need for data partitioning used during the cross-validation stage. The resulting regression equations for the selected linear and polynomial models are presented in Equations (10) to (12) and Equations (13) to (15), respectively.

.

Selected Linear Regression Model (M5):

$$R_{S_MQ4} * = -0.5788 - 0.0913R_S - 0.0045T + 0.0140H + 0.0235\,R_S T + 0.0071R_S H \quad (10)$$

$$R_{S_MQ136} * = 4.3688 - 1.5023R_S - 0.0837T - 0.0302H + 0.0448R_S T + 0.0199R_S H \quad (11)$$

$$R_{S_TGS2442} * = -158.4698 + 0.5129R_S + 5.2487T - 0.0310\,H + 0.0069R_S T - 0.0000R_S H \quad (12)$$

Selected Polynomial Regression Model (M4):

$$R_{S_MQ4} * = 2.9012 + 0.1526R_S - 0.5958T + 0.0920H + 0.1493{R_S}^2 + 0.0229T^2 - 0.0018H^2 - 0.0257{R_S}^3 - 0.0003T^3 + 0.0000H^3 + 0.0127R_S T + 0.0041R_S H \quad (13)$$

$$R_{S_MQ136} * = -15.4160 + 2.7395R_S + 0.0026T + 0.8646H - 0.6994{R_S}^2 - 0.0006T^2 - 0.0185H^2 + 0.0527{R_S}^3 + 0.0000T^3 + 0.0001H^3 + 0.0113R_S T + 0.0039R_S H \quad (14)$$

$$R_{S_TGS2442} * = -249.9664 + 3.5469R_S - 9.3579T - 8.2503H - 0.0046{R_S}^2 + 0.9588T^2 + 0.1596H^2 + 0.0000{R_S}^3 - 0.0123T^3 - 0.0010H^3 - 0.0428R_S T + 0.0004R_S H \quad (15)$$

The derived regression equations were subsequently applied to compensate for environmentally induced drift in the MQ4, MQ136, and TGS2442 sensors. The effectiveness of the selected linear and polynomial regression models was evaluated by analyzing the corrected sensor resistance under varying temperature and relative humidity conditions. Figures 7 to 12 present the three-dimensional response surfaces together with the corresponding temperature and humidity response curves, providing a visual comparison of the drift compensation performance achieved by the two regression models.

A comparison between the measured sensor responses (Figure 2 to 4) and the drift-compensated responses obtained using the selected Linear Model M5 (Figure 7 to 9) shows a noticeable reduction in

the influence of temperature and relative humidity on the sensor resistance. Before compensation, the measured response surfaces exhibit clear slopes and variations across the investigated environmental conditions, while the corresponding temperature and humidity response curves display significant changes in sensor resistance as the environmental parameters vary. After applying the linear regression model, the response surfaces become noticeably flatter and the corrected response curves exhibit smaller variations with respect to temperature and humidity. These observations indicate that the selected linear regression model effectively reduces the environmental drift and improves the stability of the sensor responses under varying operating conditions.

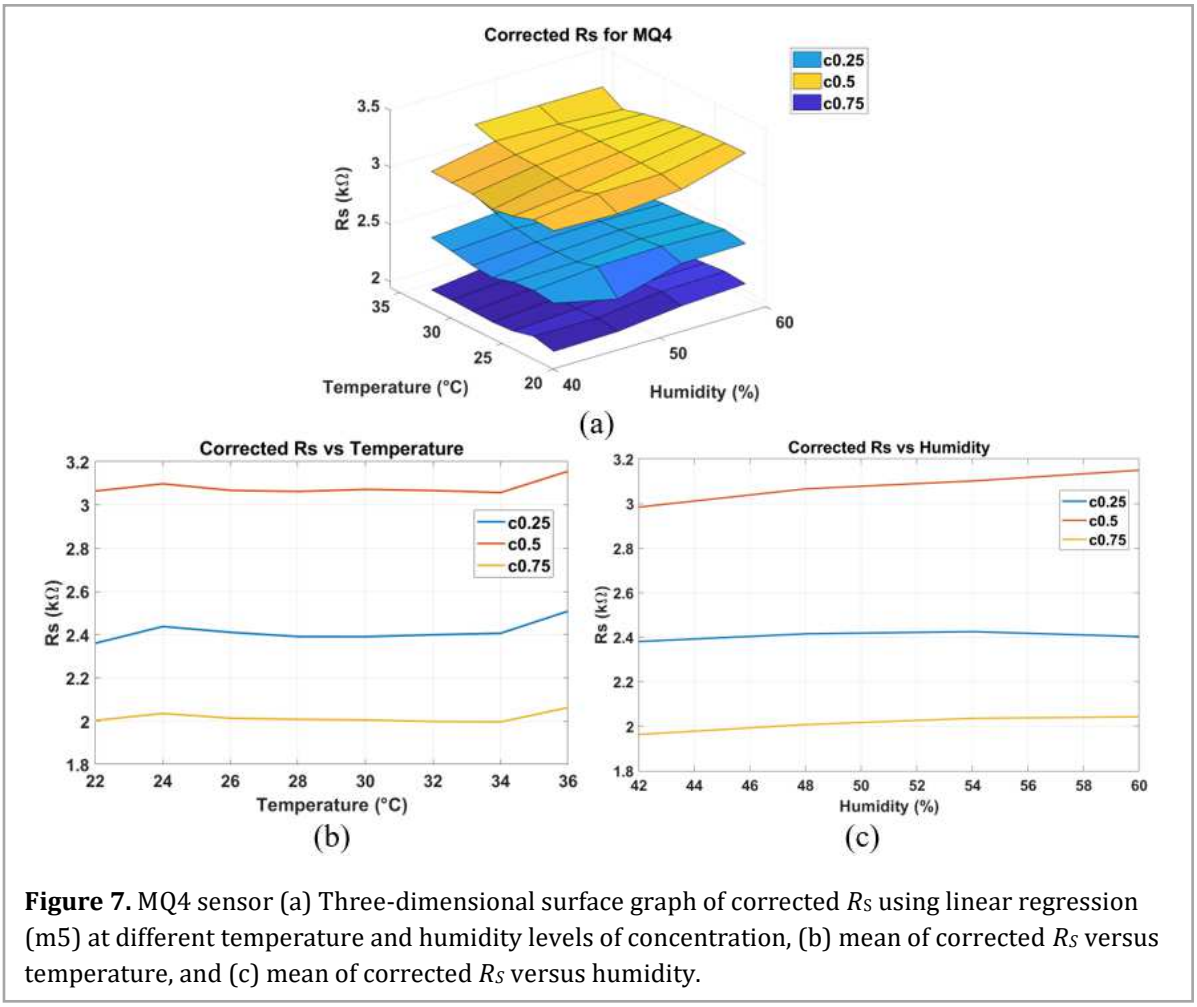


Figure 7. MQ4 sensor (a) Three-dimensional surface graph of corrected R_S using linear regression (m5) at different temperature and humidity levels of concentration, (b) mean of corrected R_S versus temperature, and (c) mean of corrected R_S versus humidity.

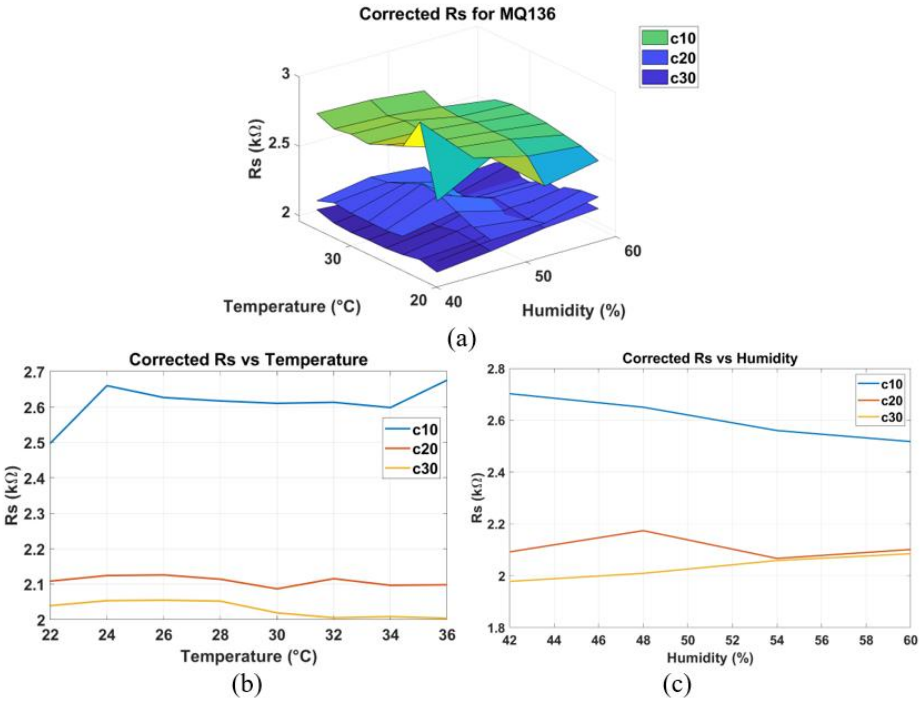


Figure 8. MQ136 sensor (a) Three-dimensional surface graph of corrected R_S using linear regression (m5) at different temperature and humidity levels of concentration, (b) mean of corrected R_S versus temperature, and (c) mean of corrected R_S versus humidity.

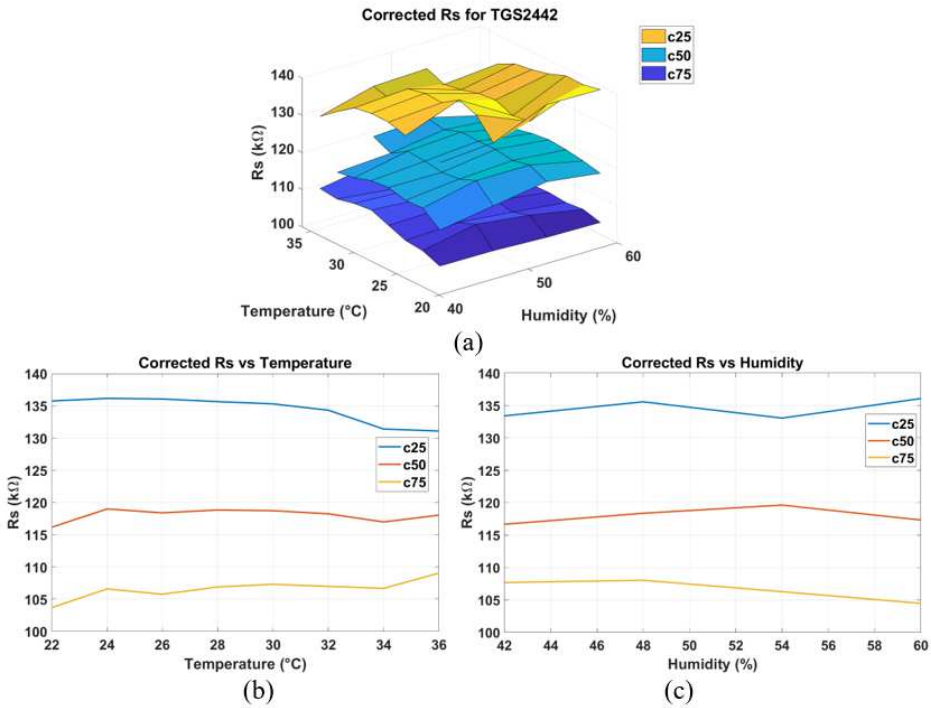


Figure 9. TGS2442 sensor (a) Three-dimensional surface graph of corrected R_S using linear regression (m5) at different temperature and humidity levels of concentration, (b) mean of corrected R_S versus temperature, and (c) mean of corrected R_S versus humidity.

Although the selected Linear Model M5 effectively reduces the influence of temperature and relative humidity on the sensor responses, further improvements can be observed when the selected Polynomial Model M4 is applied, as illustrated in Figure 10 to 12. Compared with the linear regression model, the polynomial regression model produces flatter three-dimensional response surfaces together with more uniform temperature and humidity response curves for all three sensors. The corrected sensor resistance exhibits smaller variations across the investigated environmental conditions, indicating that the higher-order polynomial formulation provides a more accurate representation of the nonlinear relationship between sensor resistance, temperature, and relative humidity. Consequently, the polynomial regression model achieves more effective environmental drift compensation while maintaining stable sensor responses over the entire operating range.

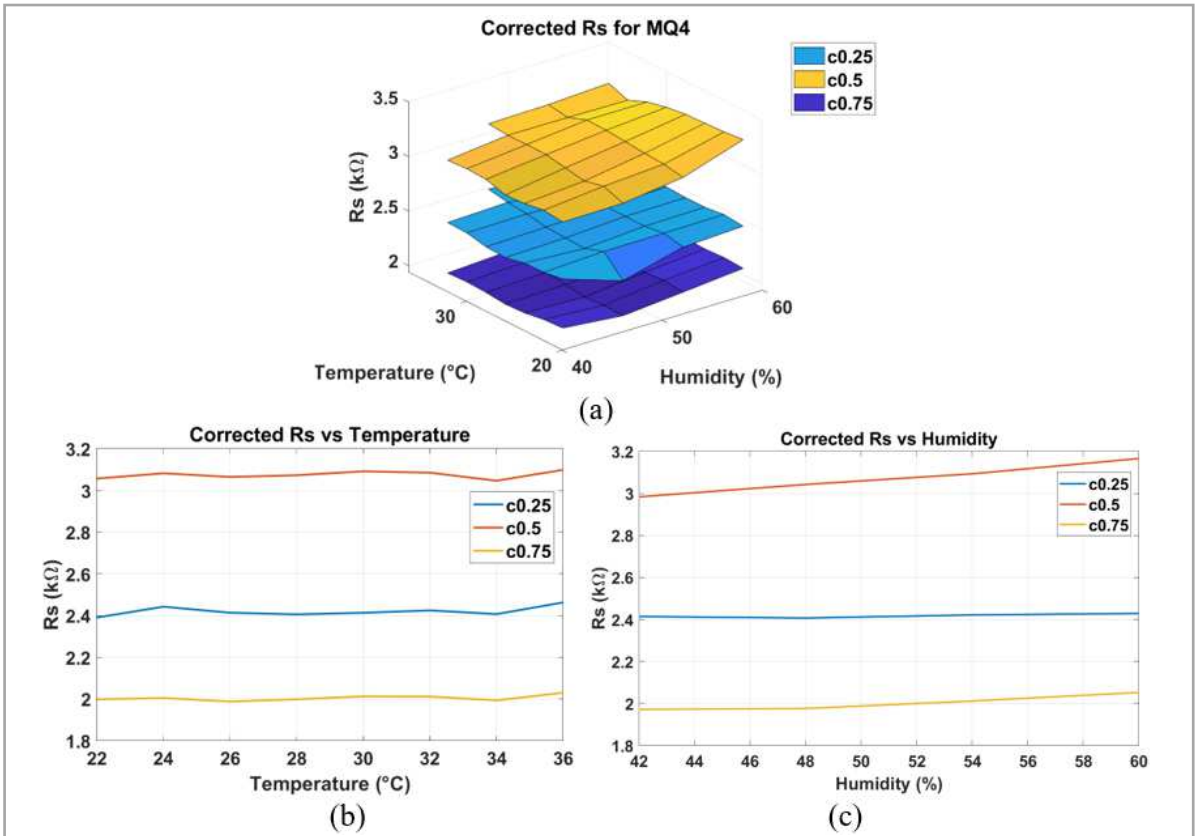


Figure 10. MQ4 sensor (a) Three-dimensional surface graph of corrected R_S using polynomial regression (m4) at different temperature and humidity levels of concentration, (b) mean of corrected R_S versus temperature, and (c) mean of corrected R_S versus humidity.

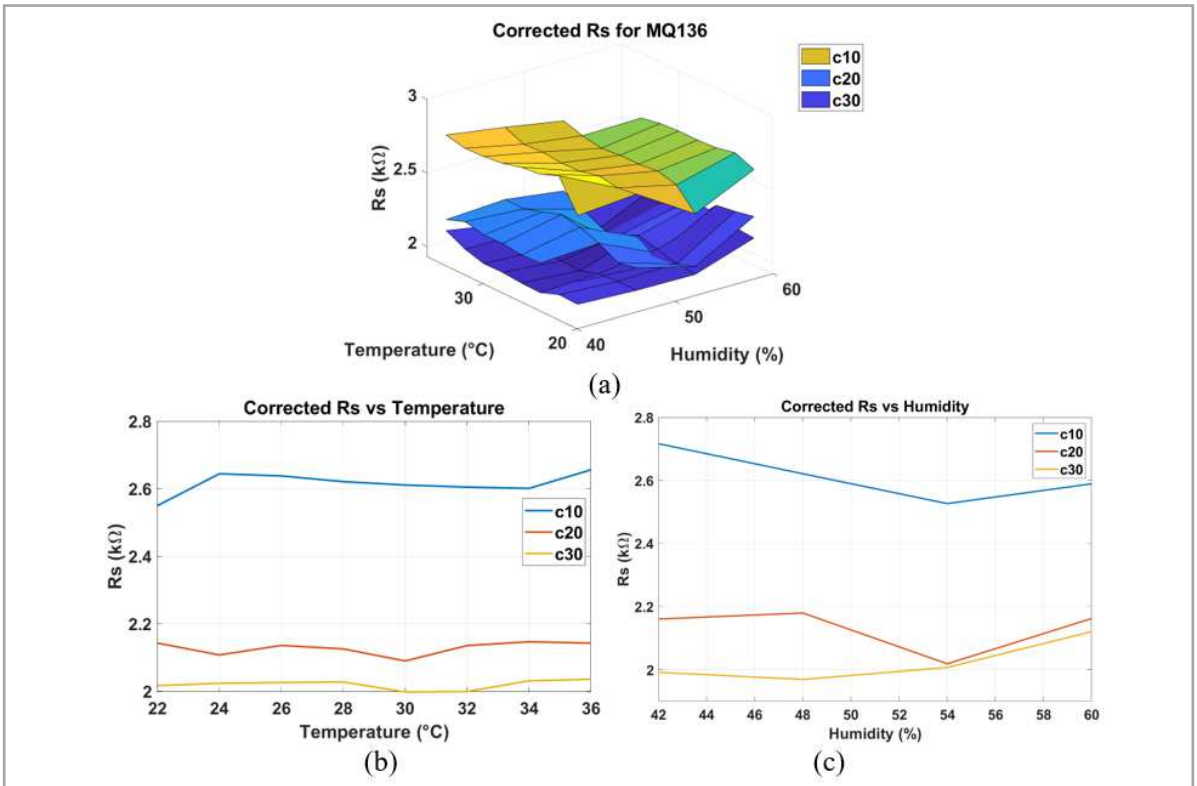


Figure 11. MQ136 sensor (a) Three-dimensional surface graph of corrected R_S using polynomial regression (m4) at different temperature and humidity levels of concentration, (b) mean of corrected R_S versus temperature, and (c) mean of corrected R_S versus humidity.

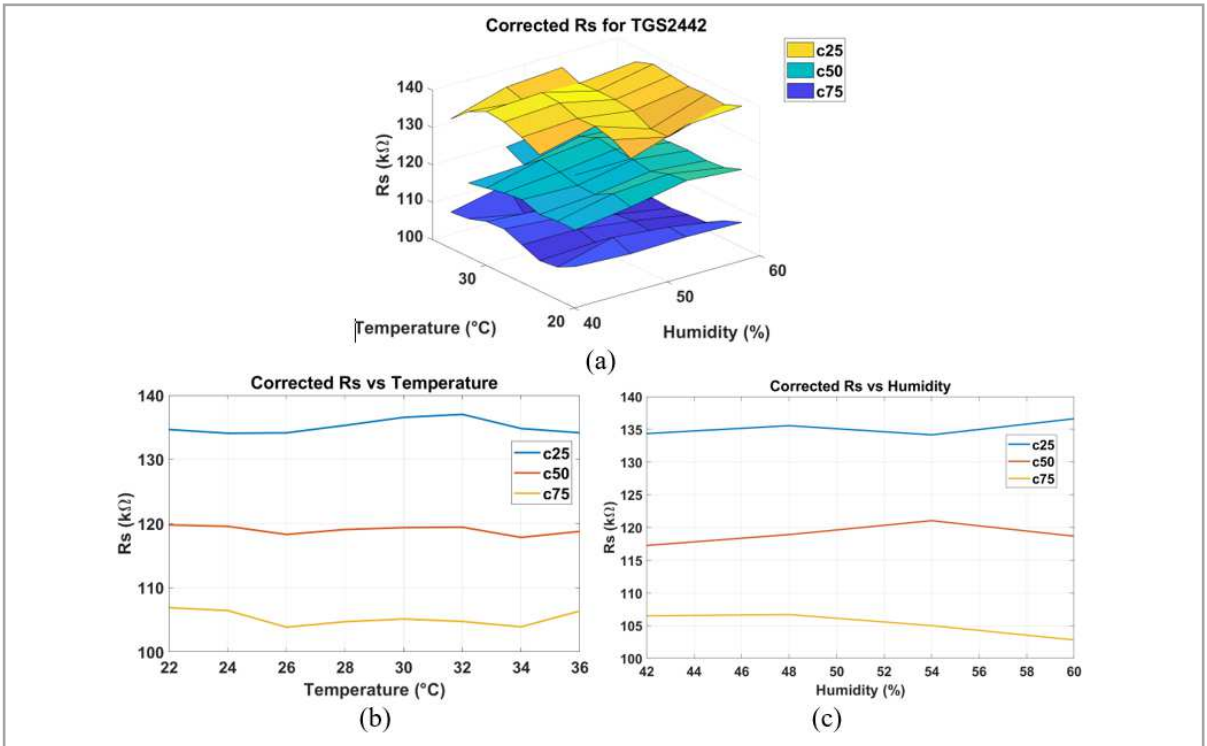


Figure 12. TGS2442 sensor (a) Three-dimensional surface graph of corrected R_S using polynomial regression (m4) at different temperature and humidity levels of concentration, (b) mean of corrected R_S versus temperature, and (c) mean of corrected R_S versus humidity.

Overall, the graphical analysis demonstrates that both regression models can reduce the influence of temperature and relative humidity on the sensor responses, with the polynomial regression model providing more effective environmental drift compensation than the corresponding linear regression model. The flatter response surfaces and more uniform temperature and humidity response curves indicate that the proposed polynomial model is better able to capture the nonlinear environmental characteristics affecting the sensor responses. However, these results were obtained using the dataset employed for model development. Therefore, external validation using an independent dataset was subsequently performed to evaluate the generalization capability and practical applicability of the proposed polynomial regression model.

4.4 External verification

The results presented in Sections 4.2 and 4.3 demonstrate that the proposed Polynomial Model M4 consistently outperforms the representative Linear Model M5 in terms of prediction accuracy and drift compensation performance. Based on these findings, Polynomial Model M4 was selected as the final compensation model because it provides an effective balance between prediction accuracy and model complexity. To further assess its generalization capability and practical applicability, an external verification test was conducted using gas concentration conditions that were not included during the model development process.

The purpose of this verification was to determine whether the selected polynomial model could reduce environmental drift and improve the stability of the sensor responses under unseen gas conditions. The verification experiment was conducted using the same hardware configuration described in Abdullah [5]. Single-gas cylinders were used to introduce CO, CH_4, and H_2S at selected concentrations. These concentrations were chosen within the overall experimental range but were different from the concentration levels used during model development. The gas types and concentrations used for the verification test are listed in Table 3. This allowed the selected model to be evaluated using unseen gas conditions and provided an assessment of its generalization capability.

Table 3 : Gas types and concentrations used for model verification

Gas types	Concentrations
Carbon monoxide (CO)	100ppm
Methane (CH4)	3%
Hydrogen sulphide (H2S)	50ppm

Table 4 presents the external verification results for the polynomial regression model M4 applied to the three gas sensors, namely MQ4, MQ136, and TGS2442. The table compares the mean and standard deviation of the measured resistance values with the corrected resistance values after applying the selected model. This comparison is important to evaluate how effectively the model reduces signal variation and improves stability under real operating conditions.

Table 4 : Comparison of mean and standard deviation for measured and corrected data using polynomial regression model M4.

Sensor	Mean Measured (kΩ)	Standard Deviation Measured (kΩ)	Mean Corrected (kΩ)	Standard Deviation Corrected (kΩ)
MQ4	3.106	0.415	2.935	0.120
MQ136	2.925	0.201	2.705	0.086
TGS2442	147.996	36.136	74.188	30.542

The external verification results indicate that the proposed Polynomial Model M4 successfully reduces the variability of the sensor responses under unseen operating conditions. As summarized in Table 3, the standard deviation of the MQ4 sensor decreases from 0.415 kΩ to 0.120 kΩ, corresponding to a reduction of 71.1%. Similarly, the MQ136 sensor exhibits a reduction from 0.201 kΩ to 0.086 kΩ, representing a 57.2% decrease. Although the reduction in standard deviation for the TGS2442 sensor is smaller, decreasing from 36.136 kΩ to 30.542 kΩ (15.5%), a noticeable improvement is still achieved. This indicates that the proposed polynomial regression model remains effective in reducing environmentally induced drift, although the compensation performance varies depending on the sensor characteristics. Overall, these results demonstrate that the proposed model can improve the stability of sensor responses when applied to an independent dataset.

Overall, the external verification results demonstrate that the proposed Polynomial Model M4 can reduce environmentally induced drift across all three gas sensors when applied to an independent dataset. Although the degree of improvement varies among the sensors, the observed reduction in response variability confirms that the developed model generalizes beyond the data used during model development and is effective under previously unseen operating conditions.

5. Conclusion

This study developed a low-complexity polynomial regression framework for compensating temperature- and humidity-induced drift in MOX gas sensing systems. The framework systematically evaluated model complexity across three commercial MOX sensors and selected a polynomial formulation that balanced prediction accuracy and mathematical simplicity. Compared with the selected linear model, the polynomial model reduced cross-validation error across all sensors and improved response stability during external verification. These results demonstrate the potential of polynomial regression as a transparent and computationally efficient compensation method for embedded gas sensing applications.

Acknowledgements

The authors would like to acknowledge the financial support from the Fundamental Research Grant Scheme under grant number FRGS/1/2022/TK08/UNIMAP/02/65 from the Ministry of Higher Education Malaysia.

References

[1] P. Saxena and P. Shukla, “A review on recent developments and advances in environmental gas sensors to monitor toxic gas pollutants,” Environ. Prog. Sustain. Energy, vol. 42, no. 5, Art. no. e14126, 2023.

[2] E. Poeta, E. Núñez-Carmona, and V. Sberveglieri, “A review: Applications of MOX sensors from air quality monitoring to biomedical diagnosis and agro-food quality control,” J. Sens. Actuator Netw., vol. 14, no. 3, Art. no. 50, 2025.

[3] H. Chai, Z. Zheng, K. Liu, J. Xu, K. Wu, and Y. Luo, “Stability of metal oxide semiconductor gas sensors: A review,” IEEE Sensors J., vol. 22, no. 6, pp. 5470–5481, 2022.

[4] A. Khorramifar, H. Karami, L. Lvova, A. Kolouri, E. Łazuka, and M. Piłat-Rożek, “Environmental engineering applications of electronic nose systems based on MOX gas sensors,” Sensors, vol. 23, no. 12, Art. no. 5716, 2023.

[5] S. Deepak, P. Anilkumar, J. Jasmin, and S. Preetha, “Room-temperature detection of harmful gases: A comprehensive review of metal oxide nanostructured chemiresistive sensors,” J. Nanopart. Res., vol. 27, no. 9, 2025.

[6] S. Giri, J. W. Gardner, P. K. Guha, A. Ghosh, and S. Santra, “Nanoscale metal oxide heterojunction chemiresistive gas sensor: A review,” Nano Futures, vol. 9, no. 3, Art. no. 032001, 2025.

[7] N. Goel, K. Kunal, A. Kushwaha, and M. Kumar, “Metal oxide semiconductors for gas sensing,” Eng. Rep., vol. 5, no. 6, Art. no. e12604, 2023.

[8] F. Liu, Z. Liu, X. Zhang, Z. Liu, and Y. Liu, “Machine learning-enhanced metal oxide gas sensor arrays: A review of selectivity and performance improvements,” Sensor Rev., vol. 45, pp. 1–19, 2025.

[9] J. Li, E. Na, X. Liang, Q. Liang, M. Fan, and H. Chen, “Surface oxygen chemistry of metal oxide semiconductors for gas-sensing applications,” Inorg. Chem. Front., vol. 11, no. 24, pp. 8602–8626, 2024.

[10] P. Wang, S. Xu, X. Shi, J. Zhu, H. Xiong, and H. Wen, “Recent advances in resistive gas sensors: Fundamentals, material and device design, and intelligent applications,” Chemosensors, vol. 13, no. 7, Art. no. 224, 2025.

[11] A. Rossi, E. Spagnoli, A. Visonà, D. Ahmed, M. Marzocchi, V. Guidi, and B. Fabbri, “SO2 detection over a wide range of concentrations: An exploration on MOX-based gas sensors,” Chemosensors, vol. 12, no. 6, Art. no. 111, 2024.

[12] I. S. Chang, S. W. Byun, T. B. Lim, and G. M. Park, “A study of drift effect in a popular metal oxide sensor and gas recognition using public gas datasets,” IEEE Access, vol. 11, pp. 26383–26392, 2023.

[13] G. Müller and G. Sberveglieri, “Origin of baseline drift in metal oxide gas sensors: Effects of bulk equilibration,” Chemosensors, vol. 10, no. 5, Art. no. 171, 2022.

[14] H. Yu, C. Guo, X. Zhang, Y. Xu, X. Cheng, S. Gao, and L. Huo, “Recent development of hierarchical metal oxides based gas sensors: From gas sensing performance to applications,” Adv. Sustain. Syst., vol. 6, no. 4, Art. no. 2100370, 2022.

[15] S. Kim, H. Sung, S. Kim, M. Je, and J. H. Kim, “ML-based humidity and temperature calibration system for heterogeneous mox sensor array in ppm-level BTEX monitoring,” in Proceedings - IEEE International Symposium on Circuits and Systems, Institute of Electrical and Electronics Engineers Inc., 2021.

[16] T. Dutta, T. Noushin, S. Tabassum, and S. K. Mishra, “Road map of semiconductor metal-oxide-based sensors: A review,” Sensors, vol. 23, no. 15, Art. no. 6849, 2023.

[17] R. Ariyaratne, M. A. Elangasinghe, M. L. Zamora, D. G. G. P. Karunaratne, A. Manipura, K. B. S. N. Jinadasa, and K. H. N. Abayalath, “Understanding the effect of temperature and relative humidity on sensor sensitivities in field environments and improving the calibration models of

multiple electrochemical carbon monoxide (CO) sensors in a tropical environment," Sens. Actuators B Chem., vol. 390, Art. no. 133935, 2023.
[18] R. Leturcq, R. Bhusari, and E. Barborini, "Physical mechanisms underpinning conductometric gas sensing properties of metal oxide nanostructures," Adv. Phys. X, vol. 7, no. 1, Art. no. 2044904, 2022.
[19] S. Zhang, Z. Yuan, Y. Liu, and F. Meng, "Humidity-counteraction identification of VOCs driven by semiconductor sensor temperature modulation response data," IEEE Trans. Instrum. Meas., vol. 74, pp. 1–9, 2024.
[20] A. N. Abdullah, K. Kamarudin, L. M. Kamarudin, A. H. Adom, S. M. Mamduh, Z. H. Mohd Juffry, and V. H. Bennetts, "Correction model for metal oxide sensor drift caused by ambient temperature and humidity," Sensors, vol. 22, no. 9, Art. no. 3301, 2022.
[21] A. Khorramifar, H. Karami, L. Lvova, A. Kolouri, E. Łazuka, and M. Piłat-Rożek, "Environmental engineering applications of electronic nose systems based on MOX gas sensors," Sensors, vol. 23, no. 12, Art. no. 5716, 2023.
[22] P. Chesler and C. Hornoiu, "MOX-based resistive gas sensors with different types of sensitive materials (powders, pellets, films), used in environmental chemistry," Chemosensors, vol. 11, no. 2, Art. no. 95, 2023.
[23] H. Mahdavi, S. Rahbarpour, S. M. Hosseini-Golgoo, and H. Jamaati, "Reducing the destructive effect of ambient humidity variations on gas detection capability of a temperature modulated gas sensor by calcium chloride," Sens. Actuators B Chem., vol. 331, Art. no. 129091, 2021, doi: 10.1016/j.snb.2020.129091.
[24] D. Maulud and A. M. Abdulazeez, "A review on linear regression comprehensive in machine learning," J. Appl. Sci. Technol. Trends, vol. 1, no. 4, pp. 140–147, 2020, doi: 10.38094/jastt1457.
[25] D. Sedera and M. Atapattu, "Polynomial regression and response surface methodology: Theoretical non-linearity, tutorial and applications for information systems research," Australas. J. Inf. Syst., vol. 23, 2019.
[26] A. Aurora, "Algorithmic Correction of MOS Gas Sensor for Ambient Temperature and Relative Humidity Fluctuations," IEEE Sens. J., vol. 22, no. 15, pp. 15054–15061, Aug. 2022, doi: 10.1109/JSEN.2022.3185176.
[27] S. Kanan, K. Obeideen, M. Moyet, H. Abed, D. Khan, and A. Shabnam, "Recent advances on metal oxide based sensors for environmental gas pollutants detection," Crit. Rev. Anal. Chem., vol. 55, no. 5, pp. 911–944, 2025.
[28] M. Iqbal, R. K. Dinata, and R. Suwanda, "Comparison of linear regression and polynomial regression for predicting rice prices in Lhokseumawe City," J. Sisfokom (Sistem Inf. dan Komput.), vol. 14, no. 3, pp. 387–394, 2025.
[29] X. Huang, H. Wang, B. Song, Y. Mao, Z. J. Zhao, B. Yu, and L. Qian, "Advances in high-temperature sensor technologies: Material system, sensing principles, drift compensation, and encapsulation reliability," ACS Omega, vol. 11, no. 4, pp. 4843–4867, 2026.
[30] A. N. Abdullah, K. Kamarudin, S. Sudin, S. M. Mamduh, M. A. Islam, and Z. H. M. Juffry, "Dataset for characterization of gas sensor under varying ambient temperature and humidity conditions," Eng. Res. Express, vol. 8, no. 8, Art. no. 085308, 2026.