A baseline correction model for humidity and temperature compensation

WO_3 film based sensor for NO_2 detection

Sujoy Ghosh, Anujay Ghosh, Nived Kodavali, Chandra Shekhar Prajapati, Navakanta Bhat
Centre for Nanoscience and Engineering, Indian Institute of Science, Bangalore, India.
sujoyghosh@iisc.ac.in

***Abstract*— In this work, a generalized self-adaptive algorithm implementation to remove the humidity dependent gas sensor response of metal-oxide based sensors has been investigated. The sensor output analysis has been divided into steps of baseline stabilization, maximum deviation search, leak detection and recovery. The baseline search using consecutive slope comparison is accurate with an error of 0.225 to 26.88 %. Using an adaptive baseline model, the saturation value for 3 ppm NO_2 is accurately detected with a maximum error of 2.7 % and response magnitudes were found to have negligible variation. Using Ordinary Least Squares and QR Factorization, the baseline in the presence of gas was predicted using historical data of temperature and humidity relationship with current to find the point of recovery of baseline, hence considerably reducing the humidity & temperature effect which can be applied to all metal oxide-based gas sensors.**

Keywords*— *Adaptive baseline; Machine learning; Polynomial regression; WO_3 gas sensor; Baseline prediction

I. Introduction

NO_2 is formed as an equilibrium gas component in N_2O_4 storage containers which is an essential oxidizer in rocket engines [1]. In the event of leakage of NO_2 in atmosphere, quick and accurate detection of gas concentration is a major safety criterion for above stated industries. One of the major issues faced by metal oxide gas sensors is the drift and cross-sensitivity introduced by variations in humidity & temperature which affect the response of the sensor [2]. The use of an array of sensors with different sensing films for deconvolution of sensor response by taking feedback from each individual sensor and then subtracting it from the signal obtained from the target sensor has already been investigated [3]. Such methods usually rely on the accuracy of the supplementary sensors in the array and the specific weights that have been allotted to the final deconvolution algorithm, using tools like artificial intelligence, sensor fusion or genetic algorithm [4-6].

In this paper, a universally applicable, self-adapting and single sensor baseline correction model has been proposed for monitoring the leak detection. The particular case study in done with WO_3 based NO_2 sensors. Using relative variation in slope values for consecutive data points, sensor behavior has been analyzed and a baseline correction model has been proposed. Also, a self-correcting historical data based ordinary least square regression model has been developed for predicting baseline currents for leaks lasting for longer durations of time.

I. Sensor Characterization

The WO_3 film was deposited on a CMOS compatible platform which had an integrated heater as reported elsewhere [7]. A sensor was kept in open air condition for 16 days to understand the variation of sensor current with the atmospheric humidity and temperature variation. As shown in Fig. 1 (a) in green, is the region of initial sensor stabilization after which sensing is possible. The current given in yellow in Fig. 1 (a) is the region where the current variation is due to marginal variations of temperature and humidity in the environment as given in Fig. 1 (b).

II. Differential Drift Correction Algorithm

A. Baseline Search

The initial decrease in baseline can be analyzed by comparing consecutive slope values between two blocks of current. The sensor current j_i was divided into blocks of size k. As shown in Fig. 2 (a) the sensor first collects the mean value $\overline{a}$ of current in block A, where c is any arbitrary data point starting from $c-k-1^{th}$ to $c-1^{th}$ value:

$$\overline{a} = \frac{\sum_{i=c-k-1}^{c-1} j_i}{k} \tag{1}$$

Then shift to the adjacent block of current to find the mean $\overline{b}$ of block B from c^{th} to $c+k^{th}$ value:

$$\overline{b} = \frac{\sum_{i=c}^{c+k} j_i}{k} \tag{2}$$

We use Eq. 1 & 2 in further calculation of the slope m of $\overline{a}$ with respect to $\overline{b}$:

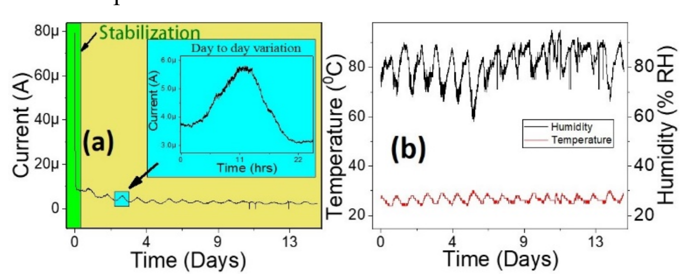

Fig. 1. (a) Sensor current variation in open air (b) Humidity and temperature experienced by the sensor during the 16-day period.

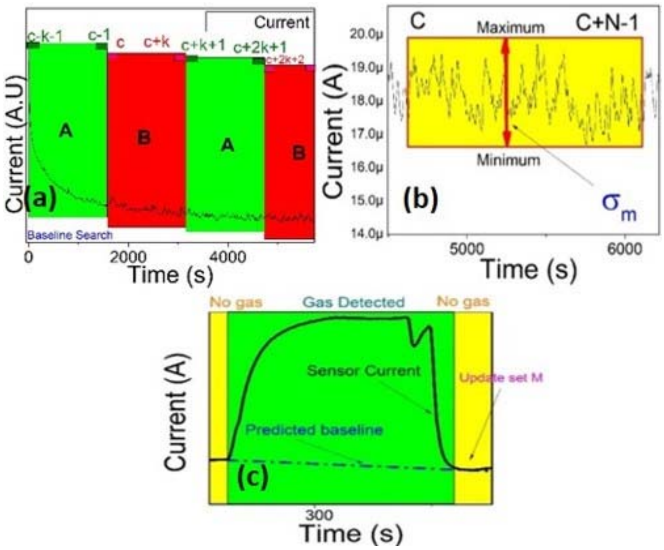

Fig. 2. (a) Baseline detection by segmenting the current into multiple blocks (b) Maximum deviation calculation between c to c+N-1 data point (c) Predicted baseline compared to actual sensor current during Gas OFF.

$$m = \frac{|\bar{a} - \bar{b}|}{k} \tag{3}$$

The slope calculated between points $c - k - 1^{th}$ to $c + k^{th}$ point is compared with the c^{th} to $c + 2k + 1^{th}$ point. The initial value of μ is set to ∞ as the previous value of slope should be higher than the present value. If the value of m is lesser than the previous value of slope μ the value of μ is updated:

$$\mu = \begin{cases} m & m < \mu \\ \mu & otherwise \end{cases} \tag{4}$$

The value of the arbitrary point c is updated to move ahead with the next iteration.

$$c_{new} = c_{old} + k + 1 \tag{5}$$

We perform this set of calculations recursively for each consecutive block until we reach a point such that μ holds a value less than or equal to zero. The sensor current value at that point is assigned to the final baseline current α.

$$\alpha = \begin{cases} \alpha & \mu > 0 \\ j_i & \mu \le 0 \end{cases} \tag{6}$$

The baseline search continually iterates till the 0 or negative value of slope has been found, marking the end of the stabilization region.

B. Maximum Deviation Search

The deviation search is based on the principle that the change in humidity and temperature in a small period of time will result in a smaller change of sensor current compared to the case of actual gas detection where the change in current will be much higher for the same interval of time [8,9]. The total time for deviation search should be a complete 24-hour cycle to account for the maximum variation in sensor current in successive N second time intervals. The total time of 24-hours is divided into multiple blocks of sets $\{S_m, S_{m+1}, S_{m+2}, S_{m+3} \ldots\}$, each of time interval size N. The sensor current j_i is collected for the given time interval N and stored in set S_m:

$$S_m = \{j_i | c \le i \le c + N - 1\} \tag{7}$$

The maximum deviation of sensor current σ_m is acquired as given in Fig. 2 (b) from the data set by finding the difference between the maximum and minimum value of current in set S_m:

$$\sigma_m = max\,\{S_m\} - min\,\{S_m\} \tag{8}$$

The recent deviation found for every N second sample is compared with the value of deviation recorded in the previous cycle. The maximum deviation in 24-hour time period is found by recursively updating the value of σ_m in every iteration:

$$\sigma_m = \begin{cases} \sigma_{m+1} & \sigma_{m+1} > \sigma_m \\ \sigma_m & \sigma_{m+1} \le \sigma_m \end{cases} \tag{9}$$

At the end of every iteration, the point c is updated as follows:

$$c_{new} = c_{old} + N \tag{10}$$

C. Leak Detection and Recovery

On exposure to NO_2 there is a sudden rise in sensor current j_i as the temperature used in our experiments in 270^0C, where there is a breakdown of NO_2 molecule into reducing gas NO, which due to its reducing nature increases the current of the metal oxide [10,11]. As shown in Algorithm 1, once the deviation search has been completed, the algorithm continually compares the present sensor current value j_i to previous current value j_{i-1} separated by time interval N. If the difference D between present and previous value of currents is greater than σ_m, then gas leak is detected and the baseline current value is stored in a variable α. Otherwise the value of j_{i-1} gets updated with j_i and the cycle continues. After entering the leak detection loop, the sensor current in every N second block is averaged to find h which is further used to calculate the response $R\%$ [12]. In metal oxide gas sensors, the absence of gas leak is indicated by the recovery of the sensor current j_i to the baseline current recoded at the beginning of purging of gas. Such a model works well in ideal conditions where the baseline stays the same before and after gas leak but in open environment, where there exists a lack of control over temperature and humidity, the baseline keeps drifting with time [13]. Nonlinear polynomial regression

Algorithm 1 Leak Detection and Recovery

```
D = j_i − j_{i−1}
    if D > σ_m then
α = j_{i−1}
    while j_i > (α + σ_m) do
    Input humidity & temperature in α = f_r (H, T)
    Calculate averaged current h = (Σ_i^{i+N} j_i) / N
    Calculate R% = (h − α)/α X 100
    end while
    else
    Update data set M
    Generate equation with data set M
    end if
```

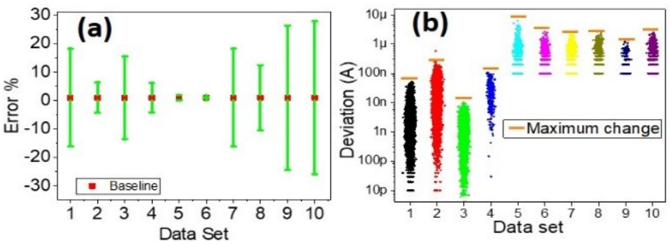

Fig. 3. Maximum percentage error of subsequent data points relative to predicted baseline after baseline detection (b) For repeated 10 second samples the deviation in current is recorded and the maximum deviation is found with the highest value for each data set.

function f_r was used to predict the baseline values using past humidity and temperature data as reference [14,15].

The algorithm stores the sensor current with temperature and humidity data in a row stacked dimensional data matrix, where each column stores a particular set of values of various factors affecting the sensor current. Using QR factorization process and Ordinary Least Squares Regression, an equation denoting the relationship of baseline current with humidity H and temperature T and coefficients K_1 to K_5 is found:

$$f_r = K_1 H + K_2 T + K_3 HT + K_4 T^2 + K_5 H^2 \qquad (12)$$

As shown in Fig. 2 (c), in the absence of NO_2 leak, the data set M containing current, humidity and temperature data points and equation for their relationship obtained using nonlinear polynomial regression function f_r is continuously updated. In the event of gas exposure, the latest equation generated before gas leak is used for predicting the future α values, which is then used to verify that whether the predicted baseline value converges with the sensor current j_i. When the sensor current j_i is less than or equal to $\alpha + \sigma_m$, we know that there is no NO_2.

III. Results and Discussion

A. Baseline Search and Gas Detection

Baseline search was performed on ten different sensor stabilization data sets. The k was set to 100 and baseline was found to stabilize within a range of 159 to 1284 seconds. The α value was compared to subsequent data points to find the maximum fluctuation due to ambient drift as given in Fig. 3 (a). The range of error % with respect to α was found to be 0.225 to 26.88 % which can be attributed to baseline drift due to various environmental factors. Maximum deviation search followed with gas detection was implemented to continuously keep updating the baseline set point and then find the response relative to that value. As shown in Fig. 3 (b) the sensor current was recorded for multiple 10 second blocks and the maximum deviation in current at every 10 seconds was plotted for different sensors. The maximum deviation σ_m in sensor current in an interval of 10 seconds over the complete data set which is well above the usual current fluctuations observed in the data sets. 3 ppm NO_2 gas was repeatedly purged on the sensor as shown in Fig. 4 (a) and the gas was accurately detected using leak detection algorithm. The sensor response as analyzed by the algorithm is as shown in Fig. 4 (b), which was in the range of 405.5 to 520.31 %. The maximum error in value of stabilized peak current h was 62 nA.

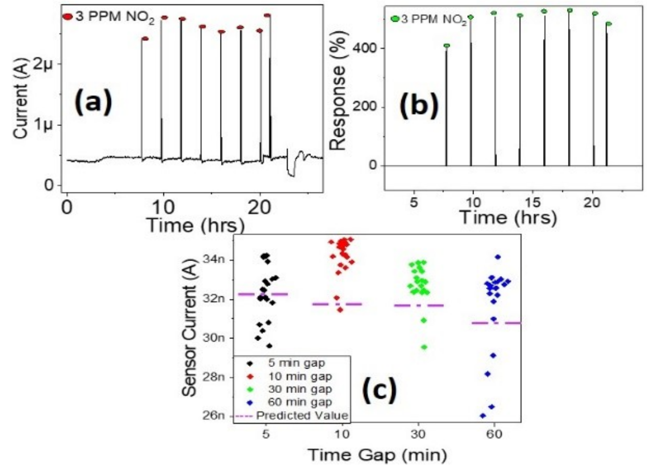

Fig. 4. (a) Sensor current for 3 ppm gas purge (b) Sensor response after baseline correction (c) Comparison of predicted and actual current values in time interval from 5 to 60 minutes.

B. Recovery Phase

The sensor was kept in open air conditions in the absence of gas leak. Using previously recorded values of humidity and temperature with current, a polynomial equation was derived for different time intervals to check the maximum time for gas leak up to which the baseline current can accurately be predicted. As shown in Fig. 4 (c), the maximum error in prediction was found to be 2.1, 3.4, 4.3 & 4.7 nA for 5,10,30 & 60 mins time gap respectively. The point at which the predicted current is less than or equal to the sensor current the gas leak can be understood to be absent. The predictor equations can further be improved by taking larger sets of historical data to produce much accurate polynomial equations. A gas purge up to 60 mins can provide reasonable accuracy of 4.7 nA for baseline prediction which is approximately 12 % error with reference to the baseline.

IV. Conclusions

The algorithm was able to find the stabilization time on sensor switching ON between 159 to 1284 seconds which had an accuracy of 0.225 to 26.88 % with respect to subsequent data points. The sensor response as calculated by leak detection approach was able to find response of 405.5 to 520.31 %. The peak saturation current as calculated had an error of 62 nA. Using polynomial regression of historical data, the error in baseline prediction was found to be 2.1,3.4,4.3 & 4.7 nA for 5, 10, 30 & 60 minutes of time interval between gas ON and OFF. The error between predicted and actual value of base current at the end of gas leak was found to be only 12 % , which is a very low error value as compared to the response for 3 ppm NO_2 gas leak hence, providing accurate baseline stabilization.

Acknowledgment

The authors would like to acknowledge support provided by National Nano Fabrication Facility (NNFC), Micro and Nano Characterization Facility (MNCF) at the Centre for Nano Science and Engineering (CeNSE) Bangalore, India. CSP would like to acknowledge DST for inspire faculty award. NB would also like to thank INAE and SERB for Abdul Kalam Technology Innovation National Fellowship.

REFERENCES

[1] H. A. Bent, "Dimers of Nitrogen Dioxide. II. Structure and Bonding," Inorg. Chem., vol. 2, no. 4, pp. 747–752, Aug. 1963.

[2] K. Kamarudin *et al.*, "Cross-sensitivity of metal oxide gas sensor to ambient temperature and humidity: Effects on gas distribution mapping," *AIP Conf. Proc.*, vol. 1808, no. 1, p. 020025, Mar. 2017.

[3] S. M. Scott, D. James, and Z. Ali, "Data analysis for electronic nose systems," *Microchim. Acta*, vol. 156, no. 3–4, pp. 183–207, Dec. 2006.

[4] J. W. Gardner, P. Boilot, and E. L. Hines, "Enhancing electronic nose performance by sensor selection using a new integer-based genetic algorithm approach," *Sensors Actuators, B Chem.*, vol. 106, no. 1 SPEC. ISS., pp. 114–121, Apr. 2005.

[5] R. Dutta, E. L. Hines, J. W. Gardner, K. R. Kashwan, and M. Bhuyan, "Tea quality prediction using a tin oxide-based electronic nose: An artificial intelligence approach," *Sensors Actuators, B Chem.*, vol. 94, no. 2, pp. 228–237, Sep. 2003.

[6] S. De Vito, E. Massera, M. Piga, L. Martinotto, and G. Di Francia, "On field calibration of an electronic nose for benzene estimation in an urban pollution monitoring scenario," *Sensors Actuators, B Chem.*, vol. 129, no. 2, pp. 750–757, Feb. 2008.

[7] C. S. Prajapati and N. Bhat, "Ppb level detection of NO2 using a WO3 thin film-based sensor: Material optimization, device fabrication and packaging," *RSC Adv.*, vol. 8, no. 12, pp. 6590–6599, 2018.

[8] W. J. Parton and J. A. Logan, "A model for diurnal variation in soil and air temperature," *Agric. Meteorol.*, vol. 23, no. C, pp. 205–216, Jan. 1981.

[9] K. D. Hage, "Urban-Rural Humidity Differences," *J. Appl. Meteorol.*, vol. 14, no. 7, pp. 1277–1283, 1975.

[10] R. P. Pohanish, *Sittig's Handbook of Toxic and Hazardous Chemicals and Carcinogens*. Elsevier, 2012.

[11] T. Akamatsu, T. Itoh, N. Izu, and W. Shin, "NO and NO2 sensing properties of WO3 and Co3O4 based gas sensors," *Sensors (Switzerland)*, vol. 13, no. 9, pp. 12467–12481, 2013.

[12] C. S. Prajapati, R. Soman, S. B. Rudraswamy, M. Nayak, and N. Bhat, "Single Chip Gas Sensor Array for Air Quality Monitoring," *J. Microelectromechanical Syst.*, vol. 26, no. 2, pp. 433–439, Apr. 2017.

[13] G. Müller, A. Friedberger, P. Kreisl, S. Ahlers, O. Schulz, and T. Becker, "A MEMS toolkit for metal-oxide-based gas sensing systems," in *Thin Solid Films*, 2003, vol. 436, no. 1, pp. 34–45.

[14] M. Gu and S. C. Eisenstat, "Efficient Algorithms for Computing a Strong Rank-Revealing QR Factorization," *SIAM J. Sci. Comput.*, vol. 17, no. 4, pp. 848–869, 1996.

[15] H. A. L. Kiers, "Weighted least squares fitting using ordinary least squares algorithms," *Psychometrika*, vol. 62, no. 2, pp. 251–266, 199.