

Hidden Markov Model Based Data-driven Calibration of Non-dispersive Infrared Gas Sensor

Yang You and Tobias J. Oechtering
KTH Royal Institute of Technology, Stockholm, Sweden

Abstract—Non-dispersive infrared gas sensing is one of the best gas measurement method for air quality monitoring. However, sensors drift over time due to sensor aging and environmental factors, which makes calibration necessary. In this paper, we propose a hidden Markov model approach for sensor self-calibration, which builds on the physical model of gas sensors based on the Beer-Lambert law. We focus on the statistical dependency between a calibration coefficient and the temperature change. Supervised and unsupervised learning algorithms to learn the stochastic parameters of the hidden Markov model are derived and numerically tested. The true calibration coefficient at each time instant is estimated using the Viterbi algorithm. The numerical experiments using CO_2 sensor data show excellent initial results which confirms that data-driven calibration of non-dispersive infrared gas sensors is possible. Meanwhile, the challenge in the practical design is to find an appropriate quantization scheme to keep the computation burden reasonable while achieving good performance.

Index Terms—Non-dispersive infrared gas sensor, drift, self calibration, data-driven modeling, hidden Markov model, statistical inference.

I. INTRODUCTION

Gas sensors are devices used to measure the presence or concentration of gases in an area and play an important role in many applications. A non-dispersive infrared sensor (NDIR sensor) is simple spectroscopic sensor in which a non-dispersive element is used to filter out the broad-band light into a narrow spectrum suitable to sense a specific gas [1] and [2]. However, NDIR sensors have been recognized to be sensitive to the variations of ambient temperature, atmospheric pressure, humidity and some other environmental factors [3]. Moreover, aging of the sensor components can also result in inaccuracy of the sensors. Due to this, regular calibration is needed for long-term accuracy of the sensors.

Today, the state of the art of infrared gas sensor self-calibration is the well-established ABC technology (Automatic Baseline Correction) where the sensor is calibrated to a fixed value that is assumed to be the fresh air gas concentration [4]. However, this method does not work well in mega-cities where the sensors never get exposed to fresh air. Thus, designing more robust and smart self-calibration algorithms which can be widely applied in different environments becomes more and more important. Based on the analysis of the data about a specific system, the data-driven modeling method aims to find

This research project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement No 825272 (ULISSES). We would also like to acknowledge Henrik Rödjegård and Hans Martin from SenseAir for the helpful technical discussions and for providing us the sensor data used in this work.

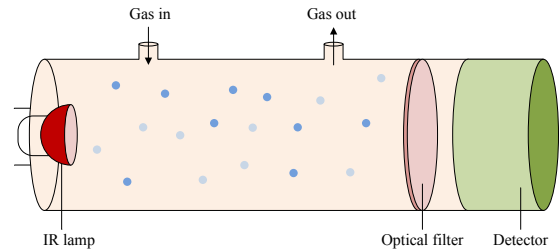


Fig. 1. Operation of NDIR sensor. Light from IR lamp is absorbed by gas particles in the tube. The intensity of remaining light is measured which provides information about the gas concentration.

relationships between the system state variables (input and output) without explicit knowledge of the physical behavior of the system [5]. Data driven modeling methods are widely applied in the area of sensor calibration since 1990s. For instance, in [6], the concept of data-driven model is proposed for calibrating the motion or position capture of the robot. By utilizing the data-driven method, the authors in [7] designed an online calibration monitoring system which can determine drifts in the sensors measurements for the oil well sensors. In [8], different data-driven models are built for laser scanner calibration. Several other research works such as [9]–[13] have applied data-driven modeling methods for calibration of different sensors or even sensor networks. Meanwhile, there is a growing interest in the machine learning approaches which become more and more promising. Machine learning approaches can provide increasing levels of automation and improved accuracy by discovering and exploiting regularities in the training data. Thus, recent research on data-driven sensor self-calibration have combined and adopted machine learning approaches to help to build a data driven model. Examples of this type of research are artificial neural networks for laser scanner calibration [8], decision tree based learning for the calibration of airborne particulate sensors [14], and support vector machine based pressure sensor calibration model [15]. In this paper, we follow this idea of combining the concept of data-driven modeling and approaches of machine learning. In more details, we utilize the hidden Markov model (HMM) [16] as a statistical learning tool to build a data-driven self-calibration system for our CO_2 sensors.

A. NDIR Sensor Mechanism

The general operation principle of an NDIR sensor is illustrated in Figure 1. When the sensor starts working, the sample

chamber will be filled by the gas from the external environment and the IR light will be directed through the sample chamber towards the detector. Different gas molecules will cause absorption of light with different wavelengths according to the Beer-Lambert law [17]. Accordingly, an optical filter is used at the detector side which eliminates all light except the light with the wavelength that the target gas molecules can absorb. Thus, for an NDIR CO_2 sensor, the detector will measure the attenuation which is caused by the absorption of CO_2 and further map it to the CO_2 concentration through a pre-defined deterministic mapping. This measurement process can be described by the following mappings:

$$\begin{aligned} Abs &= f(zero, IR, T) \\ CO_2 &= L(R - Abs), \end{aligned} \quad (1)$$

where Abs denotes the amount of absorbed light by CO_2 molecules, T denotes the environmental temperature, IR is the amount of IR light received by the detector, and $zero$ stands for the zero coefficient. This is the calibration parameter that is used for adjusting the sensor baseline offset and it generates a correct reading in absence of the target gas. Adjustments of the zero coefficient are done to compensate for sensor drift. After getting a measured absorption level, the value $R - Abs$ is mapped by a function $L(\cdot)$, taking care of the non-linear Lambert-Beer behavior, to a CO_2 concentration level. Here, R denotes the reference level measured by the detector when no absorbing gas is present.

B. Drift Analysis of NDIR Sensor

For the NDIR sensors, the IR signal varies according to time, temperature and/or other factors, while the same zero coefficient is being used for calculating the absorption, i.e., the calculation of the CO_2 concentration is using a wrong zero coefficient. This leads to a measurement error, which is defined as the difference between the sensor measurement and the true value. Therefore, a calibration of our CO_2 sensor is required where the zero coefficient is adjusted. To achieve a self-calibration procedure, we need to build a stochastic model of the true zero coefficient so that we later can estimate the true zero coefficient at any time instant given noisy observation of the sensor and environment using this stochastic model. In this work we only include the dependency of the measurement error with respect to the temperature. In more details, we focus on the following tasks in this paper:

- Build a joint probabilistic model between observed CO_2 measurements sequence, observed temperature sequence and true zero coefficients
- Learn the model parameters and utilize the corresponding model to predict the future true zero coefficients.

Notation: In the following, we denote a random variable by a capital letter, its realization by the corresponding lowercase letter, and its alphabet by the corresponding calligraphic letter. We further denote a random sequence (X_t, \dots, X_{t+k}) and its realization (x_t, \dots, x_{t+k}) by X_t^{t+k} and x_t^{t+k} respectively. In particular, X^t stands for X_1^t .

II. ESTIMATION FRAMEWORK FOR TRUE ZERO COEFFICIENT

In this section, we propose to use a HMM to model the probabilistic relationship between true zero coefficients and the observations. Since the HMM depends on a number of unknown parameters, we first need to learn these parameters from the data using either supervised or unsupervised learning methods. Following this, the most likely sequence of hidden states is found by using Viterbi algorithm, which provides the maximum likelihood estimates of the true zero coefficient at each specific time step. Since the HMM only allows the hidden states to be traversed in a sequential manner, the sampling intervals should always remains the same¹.

A. Probabilistic State Space Model

We consider an HMM where $\{X_t\}_{t=1}^T$ denotes the stochastic process that describes the hidden states, and $\{Y_t\}_{t=1}^T$ denotes the stochastic process of the observations. At each sampling instance, the random variable (or vector) X_t takes a value from the hidden state space \mathcal{X} , and the random variable Y_t takes a value from the observation space \mathcal{Y} . With the above definitions, our probabilistic State Space Model (SSM) is then fully characterized by the following distributions:

- Transition probability: $X_t \sim P_{X_{t+1}|X_t}$,
- Emission probability: $Y_t \sim P_{Y_t|X_t}$,
- Prior distribution: $X_0 \sim P_{X_0}$.

In the following, we identify hidden states and observations for our NDIR sensor calibration problem. Let $\Delta T_t = T_t - T_{t-1}$ denote the change of temperature between two consecutive sampling instances. We then define the two-dimensional hidden state as the pair of quantized true zero coefficient and temperature change. We assume there are N such quantized pairs in total, $X_t = (ZERO_t, \Delta T_t)$ is thus one of the two-dimensional hidden state in the set $\mathcal{X} = \{(zero, \Delta T)_i : \forall i \in [1 : N]\}$. In fact, the sequence of temperature change is the available observation to us and thus not hidden. However, in order to maintain a single and stationary Markov chain for our HMM, we take the change of temperature as one dimension of the hidden state. Further, let the CO_2 measurements be the corresponding observation sequence of the HMM. For our sensor model, the CO_2 measurements can only take integer values within a certain range. In this case, assuming there are M possible CO_2 measurements in total, the random variable $Y_t = CO_{2t}$ is one of CO_2 measurement from the set $\mathcal{Y} = \{CO_{2j} : \forall j \in [1 : M]\}$.

Further, the state transition probabilities are denoted by $A_{ik} = P_{X_{t+1}|X_t}(x_k|x_i)$, $A = \{A_{ik}\}_{i,k}$, which models the dynamic of the $(ZERO_t, \Delta T_t)$ pairs. The stationary emission probabilities are denoted by $B_i(j) = P_{Y_t|X_t}(y_j|x_i)$, $B = \{B_i(j)\}_{i,j}$, which models the likelihood of CO_{2t} given different $(ZERO_t, \Delta T_t)$ pairs. Lastly, define $\pi_i = P_{X_0}(x_i)$, $\pi = \{\pi_i\}_i$, as the initial prior distribution of the hidden

¹To keep this property, in our numerical experiments, we need to split our whole dataset into several subsets so that the data points in each subset are strictly synchronized.

states. In this case, the parameter $\lambda = \{\pi, A, B\}$ can fully characterize the statistics of our HMM. Thus, the learning of HMM becomes equivalent to learning the above parameters.

B. Learning of HMM

1) *Expectation maximization on HMM*: The expectation-maximization algorithm is an iterative method for estimating the parameters of a statistical model that depends on some hidden states. When applied to HMMs, the algorithm is known as the Baum-Welch algorithm [16], which is used to find an approximation to the maximum likelihood (ML) estimate:

$$\hat{\lambda}^* = \arg \max_{\lambda} P(y_1, y_2, \dots, y_L | \lambda). \quad (2)$$

Here, for computational reasons, the Baum-Welch algorithm is only applied to a subsequence y^L of the total set of available measurements y^T .

Assuming that the sequential parameter estimates are $\{\hat{\lambda}_1, \hat{\lambda}_2, \dots\}$, the parameter estimates are updated until:

$$\frac{\log P(y^L | \hat{\lambda}_{i+1}) - \log P(y^L | \hat{\lambda}_i)}{\log P(y^L | \hat{\lambda}_i)} < \gamma, \quad (3)$$

where the threshold γ is a small positive number. Given the current estimate of parameters $\hat{\lambda} = [\hat{\pi}, \hat{A}, \hat{B}]$, define several quantities as following.

- The probability of seeing the partial observation sequence (y_1, y_2, \dots, y_t) and ending up in state i at time t :

$$\alpha_i(t) = P(Y_1 = y_1, \dots, Y_t = y_t, X_t = x_i | \hat{\lambda}). \quad (4)$$

- The probability of seeing partial observation sequence $(y_{t+1}, y_{t+2}, \dots, y_L)$ given in state x_i at time t :

$$\beta_i(t) = P(Y_{t+1} = y_{t+1}, \dots, Y_L = y_L | X_t = x_i, \hat{\lambda}). \quad (5)$$

- The probability of being the hidden state is equal to x_i at time t given the full observation sequence (y_1, y_2, \dots, y_L) :

$$\gamma_i(t) = P(X_t = x_i | Y_1 = y_1, \dots, Y_L = y_L, \hat{\lambda}). \quad (6)$$

And it can be further calculated by using $\alpha_i(t)$ and $\beta_i(t)$:

$$\gamma_i(t) = \frac{\alpha_i(t)\beta_i(t)}{\sum_{j=1}^N \alpha_j(t)\beta_j(t)}. \quad (7)$$

- The probability of being in state x_i at time t also being in state x_k at time $t+1$ given the full observation sequence (y_1, y_2, \dots, y_L) :

$$\xi_{ik}(t) = P(X_{t+1} = x_k, X_t = x_i | Y_1 = y_1, \dots, Y_L = y_L, \hat{\lambda}), \quad (8)$$

which can be further calculated by using:

$$\xi_{ik}(t) = \frac{\gamma_i(t)A_{ik}B_{ij}\beta_k(t+1)}{\beta_i(t)}. \quad (9)$$

With the above definitions, by applying the Baum-Welch algorithm, the estimates of parameters are updated by calculating the following approximated relative frequencies:

- i) Estimate of the initial prior distribution of hidden states:

$$\hat{\pi}_i(t) = \gamma_i(1). \quad (10)$$

TABLE I
QUANTIZATION OF TRUE ZERO COEFFICIENT.

Interval	Level	Interval	Level
[12434, 12485)	zero ₁	[12555, 12565)	zero ₈
[12485, 12495)	zero ₂	[12565, 12575)	zero ₉
[12495, 12505)	zero ₃	[12575, 12637]	zero ₁₀
...	...		

TABLE II
QUANTIZATION OF ΔT FOR UNSUPERVISED LEARNING.

Interval	Level	Interval	Level
$(-\infty, -0.2)$	$\Delta T(1)$	$(0.2, +\infty)$	$\Delta T(2)$
$[-0.2, 0.2]$	$\Delta T(3)$		

- ii) Estimate of the transition probability:

$$\hat{A}_{ik}(t) = \frac{\sum_{t=1}^{L-1} \xi_{ik}(t)}{\sum_{t=1}^{L-1} \gamma_i(t)}. \quad (11)$$

- iii) Estimate of the emission probability:

$$\hat{B}_{ij}(t) = \frac{\sum_{t=1}^{L-1} \delta_{Y_t, y_j} \gamma_i(t)}{\sum_{t=1}^{L-1} \gamma_i(t)}, \quad (12)$$

where δ denotes the indicator such that $\delta_{Y_t, y_j} = 1$ if $Y_t = y_j, \forall t \in [1 : L], j \in [1 : M]$.

Then, for each iteration, the Baum-Welch algorithm updating rule is given as follows:

$$\hat{\lambda}_n \xrightarrow{(4-9)} \{\xi_{ik}(t), \gamma_i(t)\} \xrightarrow{(10-12)} \hat{\lambda}_{n+1} \quad (13)$$

However, the estimates are not guaranteed to converge to the optimal one, the convergence depends highly on the initialization of the model. An appropriate initialization that fits the true hidden model would result in a fast convergence of the EM algorithm. Thus, in the following section, we emphasize on discussing how to find a reasonable initialization for our algorithm.

C. Algorithm Initialization

The data we use is acquired from the 10 sensors which are put in a station at highway E18 in Sweden. Within this dataset, raw CO₂ measurements, current temperature, and the current time are recorded. The sampling interval is 15 minutes during most of the time. However, during some periods some data records are missing, the samples are thus not strictly synchronized. Thus, the whole sequence does not satisfy the requirement for the HMM modeling. To guarantee the performance of our algorithm, we need to have a minimum number of data points. Thus, we use the data from the sensor which is most synchronized and pick the subset with longest synchronized time series.

1) *Initialization for Baum-Welch algorithm*: As the prior knowledge, we know the range of the true zero coefficient is [12434, 12637] and most of the true zero coefficients lie in the range [12485, 12575]. But we do not know the exact value of true zero coefficient at each sampling instance. Following the equal-width quantization principle, we wish to quantize

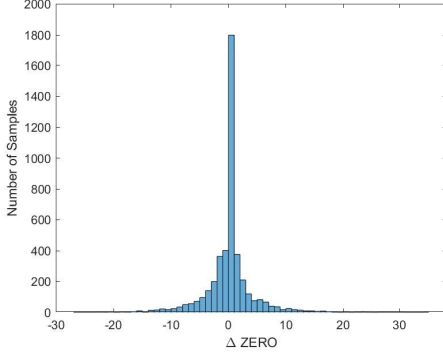


Fig. 2. Empirical distribution of $\Delta zero$.

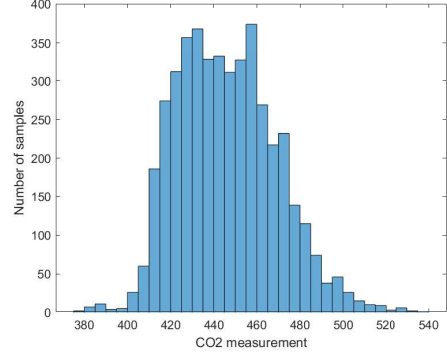


Fig. 3. Empirical distribution of CO_2 .

the whole range of true zero coefficient [12434, 12637] into N intervals each of width 10. To ensure we have enough data points for each interval, we propose the quantization scheme as shown in Table. I. Thus, we quantize the true zero coefficient into ten different values $\{zero_i\}_{i=1}^{10}$, where $zero_1 = 12470$, $zero_{10} = 12590$, and $zero_i$ equals to the value of the midpoint of the interval for $i \in [2, 9]$. We further quantize ΔT according to Table. II. These three different levels represent the cases when the temperature decreases, increases or stays the same at a given sampling instance. Thus, with ten quantized true zero coefficients and three quantized ΔT , we have 30 hidden state pairs in total. The prior distribution of these 30 hidden state pairs is initialized as uniform distribution.

Given the prior knowledge on the effect of the change of temperature on the transition of true zero coefficient, and the experience that the true zero coefficient will most probably stay the same at two adjacent time steps, it can also transfer to its most adjacent values with certain probabilities (as illustrated in Fig. 2). We propose the following rules for initialization of the state transition probabilities:

- When ΔT_t is 0, the zero coefficient has equal probabilities of increasing, decreasing or staying the same, and the ΔT_{t+1} at the next time step can also be 'positive', '0', and 'negative' with equal probability.
- When ΔT_t is positive, the zero coefficient will increase with higher probability, and the ΔT_{t+1} at the next time will be 'positive' with higher probability.
- When ΔT_t is negative, the zero coefficient will decrease higher probability, and the ΔT_{t+1} at the next time will be 'negative' with higher probability.

Regarding the initialization of the emission probability, we first identify the range of the CO_2 measurements in our dataset which is [379, 536]. We further make assumptions that each possible zero coefficient corresponds to the CO_2 measurements within a certain range, and the value of zero coefficient is in proportion to the value of CO_2 measurement. Next, we should divide the whole range [379, 536] into 10 subsets $\Omega_i, i = 1, \dots, 10$ and each Ω_i represent the set of possible observations for $zero_i$. As it can be observed from Fig. 3, most of the CO_2 measurements lie in the range

TABLE III
DIFFERENT SUBSETS OF THE CO_2 MEASUREMENT.

Interval	Level	Interval	Level
[379, 410)	Ω_1	[470, 480)	Ω_8
[410, 420)	Ω_2	[480, 490)	Ω_9
[420, 430)	Ω_3	[490, 536]	Ω_{10}
...	...		

[410, 490]. In order to have enough data points for each subset, we divide the whole range [379, 536] into the 10 subsets as shown in Table. III.

Based on the above, we propose the following initialization. For $i = 1$:

$$P(y|zero_i, \Delta T) = \begin{cases} \frac{0.9}{|\Omega_i|}, y \in \Omega_i, \\ \frac{0.1}{|\Omega_{i+1}|}, y \in \Omega_{i+1}. \end{cases} \quad (14)$$

For $i = 2, \dots, 9$:

$$P(y|zero_i, \Delta T) = \begin{cases} \frac{0.8}{|\Omega_i|}, y \in \Omega_i, \\ \frac{0.1}{|\Omega_{i+1}|}, y \in \Omega_{i+1}, \\ \frac{0.1}{|\Omega_{i-1}|}, y \in \Omega_{i-1}. \end{cases} \quad (15)$$

For $i = 9$:

$$P(y|zero_i, \Delta T) = \begin{cases} \frac{0.9}{|\Omega_i|}, y \in \Omega_i, \\ \frac{0.1}{|\Omega_{i-1}|}, y \in \Omega_{i-1}. \end{cases} \quad (16)$$

Even if the specified HMM would have been the true underlying model of the studied data, the Baum-Welch algorithm would not have been guaranteed to converge to the globally optimal solution, i.e., the ML estimate. However, our specified initialization is shown to always result in parameter estimates that agreed with basic intuition. Now, with the learned HMM parameters, our task is to decode the observed CO_2 measurement sequence into the true zero coefficient sequence.

D. Viterbi Algorithm

Assume the optimal parameter estimate given by EM algorithm is $\hat{\lambda}$, in order to find out the most possible corresponding true zero coefficient sequence, we need perform the following maximum a posterior (MAP) estimation [16]:

$$\hat{x}^T = \arg \max_{x^T} P(x^T | y^T, \hat{\lambda}). \quad (17)$$

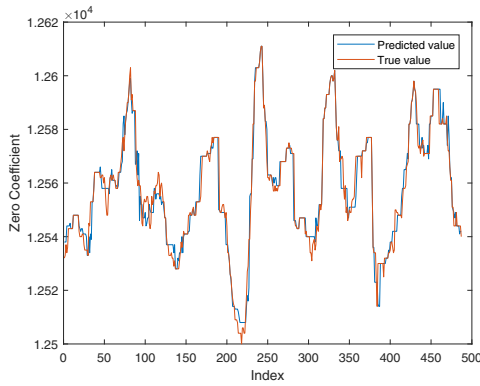


Fig. 4. Supervised learning result with high-resolution quantization.

Note that our hidden state here is a two-dimensional vector $(zero_t, \Delta T_t)$. However, since the temperature change ΔT_t is already known to us, we can directly plug the sequence ΔT^T into the above function and only find the best estimate for the sequence $zero^T$.

III. NUMERICAL RESULTS

In this section, we provide our experimental results on both supervised and unsupervised learning approaches, the predicted true zero coefficients are presented and are compared to the actual true zero coefficients. The supervised learning approach is implemented under a high-resolution quantization scheme, where the true zero coefficient is quantized to each integer within the range [12485, 12575]. Fig. 4 shows the decoded sequence of supervised learning approach compared to the true zero coefficients. While Fig. 5 shows the decoded sequence of our proposed unsupervised learning (Baum-Welch) approach compared to the quantized true zero coefficients. As we can see from the figures, even though the supervised learning is obviously with a better performance, our proposed unsupervised algorithm is also guaranteed to a certain level accuracy, where most of the decoded values are same with true values. And the inaccuracy for the small or large values are caused due to our rough quantization². But this inaccuracy for zero coefficient will lead to the small difference between CO_2 measurements, which can practically be neglected.

IV. CONCLUSION

In this paper, we studied the drift of the NDIR CO_2 sensors and developed two HMM-based approaches for data driven self-calibration mechanism for temperature dependent NDIR CO_2 sensors. It turns out that the drift of the sensor can be fully described by the drift of the zero coefficient which serves as a calibration parameter. We formulated the self-calibration problem as a statistical inference problem of the true calibration parameter, i.e., true zero coefficient. We only studied the dependency of the zero coefficient drift with

²In order to guarantee a fast convergence of the EM algorithm, a rough quantization is made here to have a small state space dimension.

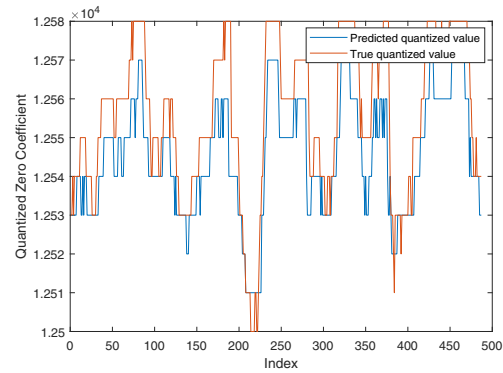


Fig. 5. EM result with rough quantization.

respect to the temperature. The experimental results show that the supervised learning outperforms the unsupervised learning approach in terms of accuracy. However, the unsupervised learning is still preferable since it is much more applicable and efficient to deal with more complex systems and is also guaranteed with a certain level of accuracy.

REFERENCES

- [1] J. Wong, "Ndir gas sensor," Aug. 22 1995, US Patent 5,444,249.
- [2] L. Miller, *Principles of Infrared Technology: A Practical Guide to the State of the Art.*, Springer, 2001.
- [3] J. Park, H. Cho, and S. Yi, "NDIR CO_2 gas sensor with improved temperature compensation," *Procedia Engineering*, pp. 303–306, 2010.
- [4] "TN-011," Tech. Rep., SenseAir, 2000.
- [5] D. Solomatine, L. M. See, and R. J. Abraham, "Data-driven modelling: concepts, approaches and experiences," in *Practical hydroinformatics*, pp. 17–30. Springer, 2009.
- [6] N. Roy and S. Thrun, "Online self-calibration for mobile robots," in *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C)*, May 1999, pp. 2292–2297 vol.3.
- [7] A. A. Boechat, U. F. Moreno, and D. Haramura Jr, "On-line calibration monitoring system based on data-driven model for oil well sensors," *IFAC Proceedings Volumes*, vol. 45, no. 8, pp. 269–274, 2012.
- [8] T. Wissel, B. Wagner, P. Stber, A. Schweikard, and F. Ernst, "Data-driven learning for calibrating galvanometric laser scanners," *IEEE Sensors Journal*, vol. 15, no. 10, pp. 5709–5717, Oct 2015.
- [9] Y. Wang and M. Neff, "Data-driven glove calibration for hand motion capture," in *Proceedings of the 12th ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2013, pp. 15–24.
- [10] J. W. Hines, A. Gribok, I. Attieh, and R. Uhrig, "Improved methods for on-line sensor calibration verification," in *Proceedings of the 8th International Conference on Nuclear Engineering*, 2000.
- [11] L. Chen et al., "Data-driven calibration of soil moisture sensor considering impacts of temperature: A case study on fld sensors," *Sensors*, vol. 19, no. 20, pp. 4381, 2019.
- [12] D. Wang, J. Liu, and R. Srinivasan, "Data-driven soft sensor approach for quality prediction in a refining process," *IEEE Transactions on Industrial Informatics*, vol. 6, no. 1, pp. 11–17, Feb 2010.
- [13] K. Shafique, A. Hakeem, O. Javed, and N. Haering, "Self calibrating visual sensor networks," in *2008 IEEE Workshop on Applications of Computer Vision*. IEEE, 2008, pp. 1–6.
- [14] L. Wijeratne, D. Kiv, A. Aker, S. Talebi, and D. Lary, "Using machine learning for the calibration of airborne particulate sensors," *Sensors*, vol. 20, no. 1, pp. 99, 2020.
- [15] W. Xie and P. Bai, "A pressure sensor calibration model based on support vector machine," in *2012 24th Chinese Control and Decision Conference (CCDC)*. IEEE, 2012, pp. 3239–3242.
- [16] C. M. Bishop, *Pattern recognition and machine learning*, springer, 2006.
- [17] D. F. Swinehart, "The beer-lambert law," *Journal of Chemical Education*, vol. 39, no. 7, pp. 333, 1962.