A Practical Approach to Validation of Buildings' Sensor Data: A Commissioning Experience Report

Claudio Giovanni Mattera, Sanja Lazarova-Molnar, Hamid Reza Shaker, and Bo Nørregaard Jørgensen

Center for Energy Informatics, Mærsk Mc-Kinney Møller Institute, University of Southern Denmark
[cgim,slmo,hrsh,bnj]@mmmi.sdu.dk

***Abstract*—Often manually performed commissioning processes on building's sensors fail to systematically validate that all building's sensors operate correctly. This is so because manual processes are tedious and only inspect a limited number of sensors. As a result, sensors are often uncalibrated, biased or somehow faulty, impacting building's behaviour, comfort level and energy usage.**

We present a practical approach to automatically validate data from all building's sensors. We designed and implemented four different tests to detect out-of-range values, spikes, latency issues and non-monotonous values. Our tests are based on expert knowledge and do not need historical data.

We ran the validation tests on a newly constructed building at the campus of the University of Southern Denmark. As a result we identified two types of faulty behaviours in the building's sensors: CO_2 sensors reporting biased values and temperature sensors' readings exhibiting high latency.

We show how automatic data validation for building sensors enhances the processes of detecting issues which could severely impact building's operations, and were otherwise going unnoticed. Thus, we emphasize the importance of performing data validation as a necessity for a correct building operation.

***Index Terms*—sensor data validation, fault detection and diagnosis, smart buildings**

I. Introduction

In the past several years buildings have become more and more intelligent [1]. Complex Building Management Systems (BMSs), often, manage buildings in their entirety: ventilation, heating, lighting, and other relevant subsystems. To operate, a BMS requires access to the status of the corresponding building, both at room level and at coarser levels. For instance, when the CO_2 concentration levels in a zone are too high, the ventilation is turned on to improve the air quality. Similarly, when the temperature in a room is below its setpoint, the heating is turned on, and when a room has been empty for a certain period, the lights are turned off.

The status of a building is sampled by an increasingly large number and variety of sensors: CO_2, temperature, air flow, water flow, humidity, light intensity, occupancy, energy meters, etc. If data collected from a building is incorrect, it would impact the correct operation of the corresponding BMS. If a BMS is fed low quality data, it would produce low quality results. It is, therefore, important to perform data validation to ensure that the quality meets the requirements.

Nevertheless, in many buildings no systematic validation is performed before they are handed over to final users, even if basic tests do not require complex setups. Facility management personnel manually validate only a small subset of all sensors and might assume that the data is correct and that the building's systems are operating correctly, as well as mistake data issues with system issues. E. g., they might blame the BMS for not handling heating correctly when it is the temperature sensors that are producing lower values than real.

In this paper we propose a series of practical data validation tests that can be automatically performed on new buildings, regardless of how data is collected, and further customized. We report the application and testing of this approach on a real case study building. We, furthermore, discuss the identified building's issues and their implications on the building's operations and performance.

Most research papers on data validation make use of simulated data [2, 3, 4, 5, 6, 7], real data with simulated faults (e. g., by adding a fixed amount to model a biased sensor) [8, 9, 10, 11, 12] or real data with artificially induced faults (e. g., by manually blocking a valve) [13, 14]. In this paper we document the testing of our proposed method on real data from our case study building, through which we identified significant faults. We informed the facility management team about the discovered faults and they, subsequently, attended to them. Furthermore, we adapted tests existing in literature to threshold-based sensors.

The rest of the paper is organized as follows. The current state-of-the-art is reviewed in Section II. The data validation tests are introduced in Section III. Section IV presents the case study and discusses results and implications. Finally, conclusions are drawn in Section V.

II. Related Work

A. Data Validation: Overview and Classifications

Data validation belongs to the more general field of Fault Detection and Diagnostics (FDD), on which depends successful operation and performance of smart buildings [15]. FDD methods analyze data from buildings to detect and identify faults. The basis for any higher level FDD approach is correct data, implying that data validation must be the very first step.

Erroneous data is caused by faults in the measuring processes, when collected data points do not adequately represent the measurements. On the other hand, data may appear faulty while correctly representing a faulty system, and it is therefore important to avoid misdetecting one for the other [10]. E. g., a

	No historical data	Historical data
Single streams	[17, 18, 19, 2, 3, 20, 21, 12, 22, 13]	[17, 18, 19, 23, 13]
Multiple streams	[10]	[17, 18, 8, 9, 11, 4, 14, 5, 6, 7]

Table I
RELATED WORK IN THE DIFFERENT GROUPS

too low temperature record can be caused by a biased sensor or by a broken heating unit.

Similarly to one classification of FDD methods, data validation methods can be divided in three groups: model-based methods, data-driven methods and rule-based methods [16]. Model-based methods use physical knowledge of the system to produce estimate data points and compare them with measured ones. Data-driven methods rely on fault-free historical data which is used to learn a black-box model of the system. Rule-based methods checks whether rules obtained through expert knowledge hold for sampled sensors data.

Model-based methods need to be designed and tuned for the system under test, therefore they are often too complex or too expensive to set up. Data-driven methods are more flexible and easier to adapt but require flawless historical data, which is not always available. Methods based on expert knowledge do not require neither detailed models nor historical data: they are therefore suitable to be deployed on buildings during the commissioning, when no data has been recorded yet or it has not been validated, i. e., methods in this group are the initial step to establish ground truths.

Another way to differentiate data validation methods is whether they consider single or multiple data streams. Methods in the first group are flexible and can be adapted to many kinds of streams. Methods in the second group exploit correlations between different streams and potentially allow validating complex interactions, but they require either redundancy—sensors in similar environments should report similar values—or historical data—from which patterns are extracted and compared with real time data.

Table I shows which of the reviewed methods belong to each group. The method proposed in this paper belongs to the upper left cell, it considers single data streams and assumes no historical data is available.

B. State-of-the-Art

In the following we summarize the state-of-the-art of relevant advances in data validation. We begin by a general overview, followed by a focus on data-driven methods, and finally we review the remaining significant approaches.

1) General Overview: Siao et al. present a literature review on data validation methods. Simple tests include physical range check based on sensor's range; local realistic range, based on sensor's location and condition, possibly obtained through statistical analysis; gaps detection; flat lines detection; gradient test; tolerance band methods; and physical redundancy checks. More complex tests include statistical analysis to detect outliers, drift detection using exponential weighted moving average methods, spatial consistency methods, analytical redundancy (to check quantities correlated in a physical model), gross error detection, multivariate statistical methods (e. g., Principal Component Analysis (PCA)) and data mining techniques [24].

Pires et al. present a review and classification of data validation methods used for mobile health applications. The authors divide those methods in three groups: faulty data detection methods, data correction methods, and other assisting techniques or tools dealing with hardware errors [25].

Cugueró-Escofet et al. identify 6 levels of data validation tests—communications, physical range, trend, equipment state, spatial consistency and time series consistency—divided in low and high level tests. Low level tests concern a single sensor, while high level tests exploits correlation among different sensors. The tests, obtained from expert knowledge, are applied in sequence to incoming data which is scored accordingly. In case one or more tests fail, a data reconstruction method is used to produce valid data [19].

Branisavljević et al. consider an ordered sequence of validation methods to be applied to data: detection of zero values, detection of constant values, range check for physical limits, range check for historical limits, statistical univariate test, statistical multivariate tests, Artifical Neural Network (ANN), non-linear models, SVM and physical model. Data is augmented with contextual information (part of day and weather conditions) and validation methods perform better when tuned separately on each class [17].

2) Data-driven Methods: Castello et al. present two applications to handle data validation and correction and data provenance for buildings. Provenance includes information about the transformations through which data undergo (unit conversion, resampling and filtering). The authors present three experimental buildings as case studies [18].

Hou et al. propose a combined rough sets and ANN method for detecting biases on HVAC chillers. Several rules are defined to split data in subsets corresponding to different operating conditions, and historical fault-free data is used to train ANNs for each subset. The ANNs compute then an estimate confidence interval for the bias [8].

Sharifi et al. propose a MPPCA model for non-linear sensor faults detection. The input space is divided in few locally linear regions and on each of them a PPCA model is trained. When a new measurement is available, it is first mapped to the correct region and then validated. A drawback of the proposed method is that in case of large error it is difficult to obtain the correct region and, therefore, to correctly isolate the fault [14].

Tsang et al. propose a method to validate sensor data using polynomial predictive filter and fuzzy logic. Three sets of fuzzy rules are considered: data is in range, data frequency is in range, and data variance satisfies the F-ratio test. Polynomial predictive filter is applied to historical data to obtain estimates for the rules' lower and upper bounds [13].

3) Model-based and Rule-based Methods: Tsang proposes a gray model method for sensor data validation. The authors consider three fault indicators: limit indicator, where a signal or

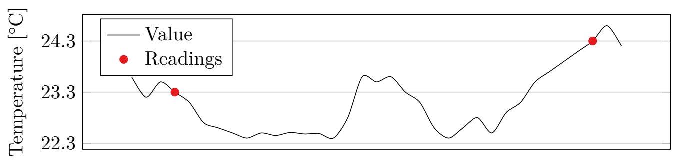

Figure 1. Threshold-based sensors produce a new reading when the difference from the previous value is larger than a threshold (1 °C).

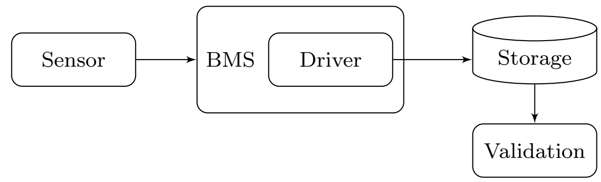

Figure 2. Data validation flowchart. Sensors push records to the BMS. A driver collects records from the BMS and forwards it to a centralized data storage, on which validation is performed.

its rate of change are out of prescribed bounds, jump indicator, when there is a sudden change in the signal corresponding to a spike, and noise indicator, when there is a change in the predicted signal's error [22].

Näsi et al. propose fuzzy limits centered on the signal's current mean to identify outliers within the range limits. The authors compare distribution-based and density-based limits, showing that the latter are less sensitive to non-evenly divided values. Adaptive fuzzy limits contain a damping term based on distance from the average, in order to prevent the outliers to affect the limits [20, 21].

III. Validation of Sensors and Meters Data

The methodology presented in this paper is mostly inspired by the M1–M3 tests from [17], threshold-based tests from [18] and level 0–2 tests from [19]. Implementations from these papers were not available to deploy on our building. Moreover, the mentioned approaches assumed sensors can be sampled with constant frequency, which was not the case in our setup. Therefore, we designed and implemented our own methodology.

Sensors perform measurements and make them available to the BMS. We assume sensors to be threshold-based, i.e., to continuously sense values and produce a new reading only when the difference from the previous one is larger than a threshold (see Figure 1). Rate-of-change tests, while popular in literature [2, 3, 11, 20, 21, 12, 22, 13], are ineffective for threshold-based sensors, since their rate-of-change is constant ($\pm$ the threshold).

We define the following terms:

- *Value*: the physical value sensed by a sensor;
- *Reading*: the act of receiving a value from a sensor;
- *Record*: a pair (t, v), denoting a reading at instant t with reported value v;
- *Sensor-threshold*: difference between the current value and the last record needed for a sensor to make a new reading.

Figure 2 shows an overview of the data path in our system. When a reading happens on a sensor, a record is stored on the BMS. Over time a sensor produces a time series of records $[(t_0, v_0), (t_1, v_1), (t_2, v_2), \ldots]$. A *driver* is a program running on the BMS which forwards readings to a centralized data storage. Validation tests are performed on all incoming records going to the data storage.

A. Validation Tests

We implemented four different tests to validate records from the building sensors: range, latency, spikes and monotonicity. Each test can detect different issues (Table II).

1) Range test: Sensors from the building measure different physical quantities, e. g., CO_2 concentration level in the air, light intensity in rooms, temperature of air in rooms or in ventilation units or temperature of water in the heating system and humidity in rooms. Most of these quantities should have measure within a given range. E. g., in normal rooms CO_2 concentration level cannot be lower than in atmosphere and air temperature should be within comfort level range. Sensible ranges can be obtained for many of the measured quantities from expert knowledge, sensors data sheets and validated historical data. Given upper and lower bounds $v_{\min}$ and $v_{\max}$ the record (t, v) is labeled as erroneous if $v \lnapprox v_{\min} \vee v \gnapprox v_{\max}$.

2) Latency test: For threshold-based sensors at any time the uncertainty of a measure is twice the sensor-threshold. Long periods without a reading can be caused by a decrease in sensor accuracy, a sensor hard failure, communication problems or BMS failure.

Given a maximal latency $\Delta t_{\max}$ and two consecutive records (t_0, v_0) and (t_1, v_1) *the former record* is labeled as erroneous if $t_1 - t_0 > \Delta t_{\max}$. For this test it is not the records themselves which are erroneous, but the interval between them. More readings were expected between t_0 and t_1, therefore the record (t_0, v_0) was labeled as erroneous, since it is not known until when its value can be trusted.

3) Spikes test: A spike is a large variation in a very short time window. Occasionally a driver may fail to parse a record from the BMS, or the sensor itself can generate an erroneous value. Typical examples are zero, negative numbers, or random numbers. Sometimes spikes are invalid values (e. g., a negative CO_2 concentration level) which are easy to filter out, but they can also be valid (e. g., temperature can be negative).

Issue \ Test	Range	Latency	Spikes	Monotonicity
Sensor bias	✓			
Misplaced sensor	✓	✓		
Driver fault		✓	✓	✓
Accuracy degradation		✓		
Communication problem		✓		✓

Table II
Issues detected by each test

A naïve way to check for spikes is to check when the difference between two consecutive records is above a given threshold. This is however susceptible to false positives: for instance, if the building's communication network was down for a short period of time the first record after it is back on line might be significantly different from the previous one, but this should not be considered a spike. Given the parameters δ_v and δ_t and two consecutive records (t_0, v_0) and (t_1, v_1), the latter record is labeled as a spike if

$$\frac{v_1 - v_0}{t_1 - t_0} > \frac{\delta_v}{\delta_t}. \tag{1}$$

4) Monotonicity test: Some meters, in particular energy meters that record the total energy consumption, record an incremental quantity. Such values are monotonically increasing, as it is not possible to recover previously consumed energy, and therefore every record value must be greater or equal to the previous one. Given two consecutive records (t_0, v_0) and (t_1, v_1) the latter record is labeled as erroneous if $v_1 \ngeq v_0$.

IV. Case Study

In this paper we present Building OU44 as a case study. The building is located on the campus of the University of Southern Denmark. It contains classrooms, offices and study rooms, it has been operating and collecting data since October 2015.

Every room has the following sensors:

- Temperature $[°C]$, relevant for heating and ventilation;
- CO_2 $[\text{ppm}]$, relevant for ventilation;
- PIR $[\text{boolean}]$, room occupancy;
- Light $[\text{lx}]$, relevant for automatic lighting control.

Some rooms have additional sensors or meters. For instance some have separate meters for plug load or sensors for humidity. Four rooms are equipped with occupancy counting cameras that provide an estimate of people in the room. In addition to that, the building has a weather station that records external temperature, wind speed, rain and solar radiation. There are also several energy meters: for heating, ventilation, hot water, lighting, plug load, usually aggregate by floor or area.

All sensors are accessible through a KNX bus [26] and broadcast records to the BMS according to their configuration. Custom drivers fetch data from the BMS and publish it to a centralized data base using Simple Measurement and Actuation Profile (sMAP) protocol, so that it is available to other applications, like occupancy prediction [27] and model development and calibration [28].

Validation tests were executed on all the available rooms for a number of selected sensors and meters. Table III shows a summary of the tests and the sensors along with the corresponding parameters. Details about number of streams and average frequency are shown in Table IV.

A. Results

1) Results for CO_2CO2 Sensors: Figure 3 shows spikes test violations for CO_2 concentration level for selected room and period. Some spikes have value zero, which is impossible for CO_2, while some other spikes have unusually large, but in principle plausible, values.

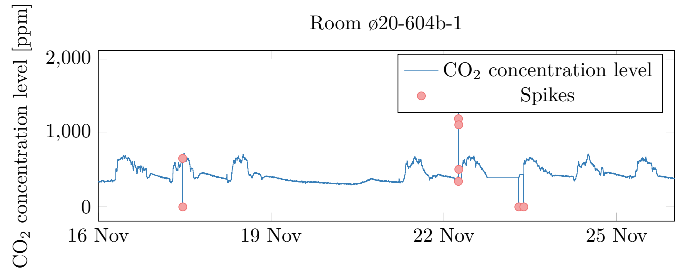

Figure 3. Spikes test violations for CO_2 concentration level

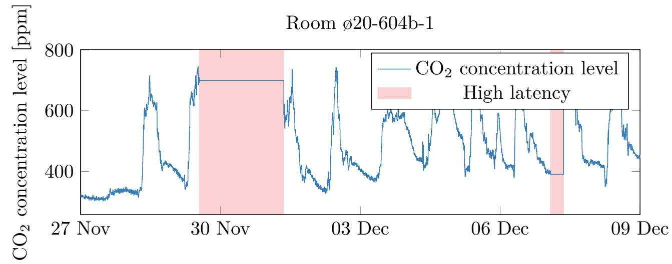

Figure 4. Latency test violations for CO_2 concentration level

Figure 4 shows latency test violations for CO_2 concentration level for a selected room and a time period. There are two intervals with missing data due to a driver crash. The data was, however, still available on the BMS, although not yet forwarded to the storage.

Figure 5 shows range test violations for CO_2 concentration level for a selected room and a time period. Results for most of the rooms in the building are similar to these. For the given room, the CO_2 sensor was reporting CO_2 concentration levels lower than the current atmospheric level. The CO_2 sensors have an accuracy of $\pm 125\,\text{ppm}$, therefore, some violations were expected when the CO_2 concentration level was close to the minimum. However, records were consistently below range nearly all the time, sometimes for several days, which suggests these sensors are faulty.

2) Results for Temperature Sensors: Figure 6 shows latency test violations for temperature measurements for a selected

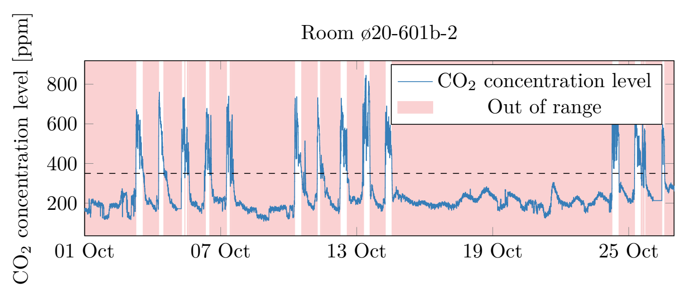

Figure 5. Range test violations for CO_2 concentration level

Test \ Sensor	Temperature	Humidity	CO_2	Energy Meter
Range	10 °C to 40 °C	0 % to 100 %	350 ppm to 1200 ppm	
Latency	180 min	20 min	10 min	
Spikes	20 °C in 5 min	5 % in 10 min	300 ppm in 10 min	
Monotonicity				Yes

Table III
IMPLEMENTED TESTS PARAMETERS

	Temperature	Humidity	CO_2	Energy Meter
Number of streams	140	10	231	88
Average frequency	30 min	5 s	5 min	1 min

Table IV
NUMBER OF STREAMS AND AVERAGE FREQUENCY

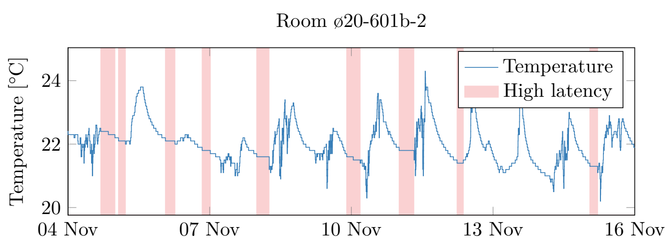

Figure 6. Latency test violations for temperature

room and time period. Even if the maximal latency was rather large (180 min), there were many erroneous intervals. The sensor-threshold in the sensor was set to 0.1 °C, so more frequent readings were expected. However, most of the faults occurred during the night, when the room was empty and temperature should stay close to the setpoint. It is also possible that boxes enclosing sensors isolate them too much from the environment and they, therefore, shield them from high-frequency variations.

Figure 7 shows range test violations for temperature for selected room and period. For a short time during the night the temperature dropped below the lower bound. Since this was an isolated instance, it occurred during the night and the temperature went back to the setpoint in the morning, it might suggest that a window was left open during the night.

B. Implications and Discussion

After the CO_2 sensors' faults were detected, the supplier replaced the sensors. The range test violations for CO_2 after

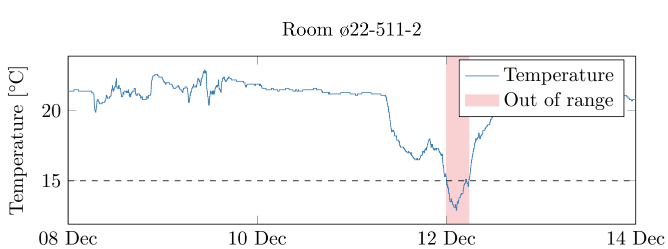

Figure 7. Range test violations for temperature

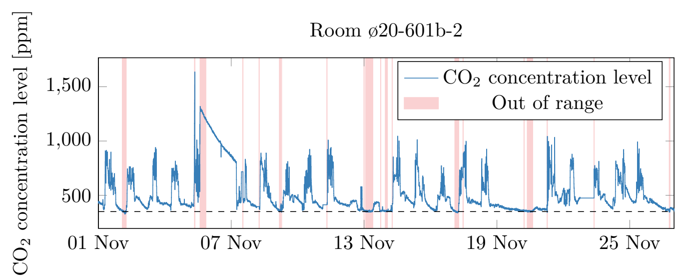

Figure 8. Range test violations for CO_2 concentration level after sensors were replaced

the replacement are shown in Figure 8 (in the same room of Figure 5). There were still few intervals below the minimum range, but they happened when the CO_2 concentration level was close to the minimum, and they therefore expected due to the sensor accuracy (±125 ppm).

To summarize, running the tests on the building led us to detecting two particular faults:

- many CO_2 sensors were often below their minimal range,
- for some temperature sensors new readings occurred with a very high latency.

Further investigation showed that CO_2 sensors were biased, mostly to a lower value. The BMS uses these records to determine when to turn on ventilation, therefore, this fault resulted in bad air quality in the building and lower ventilation energy consumption.

Additional exploration helped in discovering that for some temperature sensors (especially the ones in the weather station) the sensor-threshold is too large, up to 1 °C, which means that the accuracy of temperature is 2 °C and smaller variations are not recorded. This configuration is probably due to constraints on the KNX bus bandwidth, which supports a relatively small number of simultaneously transmitting sensors. The effects of decreasing the sensor-threshold need to be investigated.

V. CONCLUSIONS

We presented a sensor data validation approach that was designed and developed driven by the real need of our research project, aimed at improving smart buildings' energy efficiency. The data validation tests performed on the project's case study building exposed at least two faults: CO_2 sensors "out of range" and temperature sensors showing high latency. CO_2 sensors bias impacted the ventilation system and other ongoing research projects regarding model calibration and parameter estimation on the building [29], and the issue was detected only after few months. Temperature sensors' high latency might have impacted model calibration as well. If those tests had being running earlier during the commissioning phase, a long period of ventilation issues could have been avoided and bus limitations could have been addressed.

This experience showed that validating sensors values in new buildings is indeed an essential necessity for a correct building operation. Expert knowledge based practical tests for single streams are sufficient to expose issues with sensors and meters that greatly affect building's performance.

Automated testing of single streams is the first step in data validation. Such tests, however, can only detect a subset of faults and issues. Moreover, expert knowledge is necessary to set the appropriate test parameters. Our future work in this problem domain anticipates utilizing peer validation and exploiting interactions between different data streams with more advanced data validation methods. Once quality of historical data is ensured, our focus will be on data-driven methods.

ACKNOWLEDGEMENT

This work is supported by the Innovation Fund Denmark for the project COORDICY.

REFERENCES

[1] Bo Nørregaard Jørgensen, Mikkel Baun Kjærgaard, Sanja Lazarova-molnar, Hamid Reza Shaker, and Christian T Veje. "Advancing Energy Informatics to Enable Assessable Improvements of Energy Performance in Buildings". In: *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems*. Bangalore, India, 2015, pp. 77–82.

[2] R Jeyanthi and K Anwamsha. "Fuzzy-based sensor validation for a nonlinear bench- mark boiler under MPC". English. In: *10th International Conference on Intelligent Systems and Control (ISCO)*. Coimbatore: IEEE, 2016, pp. 1–6.

[3] Nithya Kancherla and R. Jeyanthi. "Design of a generic Fuzzy-Based Sensor Data Validation algorithm for a chemical process". English. In: *International conference on Circuits, Controls and Communications (CCUBE)*. Bengaluru: IEEE, 2013, pp. 1–6.

[4] Ihab Samy, Ian Postlethwaite, and Dawei Gu. "Neural network based sensor validation scheme demonstrated on an unmanned air vehicle (UAV) model". In: *2008 47th IEEE Conference on Decision and Control* (2008), pp. 1237–1242.

[5] S W Wang and Y M Chen. "Sensor validation and reconstruction for building central chilling systems based on principal component analysis". English. In: *Energy Conversion and Management* 45.5 (2004), pp. 673–695.

[6] Shengwei Wang and Fu Xiao. "AHU sensor fault diagnosis using principal component analysis method". In: *Energy and Buildings* 36.2 (2004), pp. 147–160.

[7] S W Wang and J Y Qin. "Sensor fault detection and validation of VAV terminals in air conditioning systems". English. In: *Energy Conversion and Management* 46.15-16 (2005), pp. 2482–2500.

[8] Zhijian Hou, Zhiwei Lian, Ye Yao, and Xinjian Yuan. "Data mining based sensor fault diagnosis and validation for building air conditioning system". English. In: *Energy Conversion and Management* 47.15-16 (2006), pp. 2479–2490.

[9] MinJeong Kim, Hongbin Liu, Jeong Tai Kim, and ChangKyoo Yoo. "Sensor fault identification and reconstruction of indoor air quality (IAQ) data using a multivariate non-Gaussian model in underground building space". English. In: *Energy and Buildings* 66 (2013), pp. 384–394.

[10] Alexey Kozionov, Mikhail Kalinkin, Alexey Natekin, and Alexander Loginov. "Wavelet-based sensor validation: Differentiating abrupt sensor faults from system dynamics". In: *2011 IEEE 7th International Symposium on Intelligent Signal Processing* (2011), pp. 1–5.

[11] J Kullaa. "Sensor validation using minimum mean square error estimation". English. In: *Mechanical Systems and Signal Processing* 24.5 (2010), pp. 1444–1457.

[12] Z G Shen and Q Wang. "Data Validation and Confidence of Self-validating Multifunctional Sensor". English. In: *2012 Ieee Sensors Proceedings* (2012), pp. 1045–1048.

[13] K M Tsang and W L Chan. "Data validation of intelligent sensor using predictive filters and fuzzy logic". English. In: *Sensors and Actuators a-Physical* 159.2 (2010), pp. 149–156.

[14] R Sharifi and R Langari. "Nonlinear sensor fault diagnosis using mixture of probabilistic PCA models". English. In: *Mechanical Systems and Signal Processing* 85 (2017), pp. 638–650.

[15] Sanja Lazarova-Molnar, Hamid Reza Shaker, Nader Mohamed, and Bo Norregaard Jorgensen. "Fault detection and diagnosis for smart buildings: State of the art, trends and challenges". In: *IEEE 3rd MEC International Conference on Big Data and Smart City (ICBDSC 2016)*. Muscat, Oman, 2016, pp. 1–7.

[16] Srinivas Katipamula and Michael Brambley. "Review Article: Methods for Fault Detection, Diagnostics, and Prognostics for Building Systems-A Review, Part I". In: *HVAC&R Research* 11.1 (2005), pp. 3–25.

[17] Nemanja Branisavljević, Zoran Kapelan, and Dušan Prodanović. "Improved real-time data anomaly detection using context classification". In: *Journal of Hydroinformatics* 13.3 (2011), p. 307.

[18] Charles C. Castello, Jibonananda Sanyal, Jeffrey Rossiter, Zachary Hensley, and Joshua R. New. "Sensor data management, validation, correction, and provenance for building technologies". In: *ASHRAE Transactions* 120 (2014), pp. 370–382.

[19] Miquel À. Cugueró-Escofet, Diego García, Joseba Quevedo, Vicenç Puig, Santiago Espin, and Jaume Roquet. "A methodology and a software tool for sensor data validation/reconstruction: Application to the Catalonia regional water network". English. In: *Control Engineering Practice* 49 (2016), pp. 159–172.

[20] Jari Näsi and Aki Sorsa. *On-line measurement validation through confidence level based optimal estimation of a process variable*. Tech. rep. Oulu: University of Oulu, 2004.

[21] Jari Näsi, Aki Sorsa, and Kauko Leiviskä. "Sensor validation and outlier detection using fuzzy limits". English. In: *44th IEEE Conference on Decision and Control, and the European Control Conference, CDC-ECC '05*. 2005, pp. 7828–7833.

[22] K M Tsang. "Sensor data validation using gray models". English. In: *ISA Transactions* 42.1 (2003), pp. 9–17.

[23] Hector Rodriguez, Vicenc Puig, Juan J Flores, and Rodrigo Lopez. "Combined holt-winters and GA trained ANN approach for sensor validation and reconstruction: Application to water demand flowmeters". In: *2016 3rd Conference on Control and Fault-Tolerant Systems (SysTol)*. IEEE, Sept. 2016, pp. 202–207.

[24] Sun Siao, Jean-LucBertrand Krajwewski, Anders Lynggaard-Jensen, Joep Van den Broeke, Florian Edthofer, Maria do Ceu Almeida, Ribeiro Alvaro Silva, and Jose Menaia. *Literature review for data validation methods*. 2011.

[25] Ivan Miguel Pires, Nuno M. Garcia, Nuno Pombo, Francisco Flórez-revuelta, and Natalia Díaz Rodríguez. "Validation Techniques for Sensor Data in Mobile Health Applications". English. In: *Journal of Sensors* 2016 (2016), pp. 1–9.

[26] KNX Association. *What is KNX?* 2017. URL: https://www.knx.org/knx-en/knx/association/what-is-knx/ (visited on 01/25/2017).

[27] Sanja Lazarova-Molnar, Mikkel Baun Kjærgaard, Hamid Reza Shaker, and Bo Nørregaard Jørgensen. "Commercial Buildings Energy Performance within Context - Occupants in Spotlight". In: *Proceedings of the 4th International Conference on Smart Cities and Green ICT Systems*. Lisbon, Portugal: SCITEPRESS - Science, 2015, pp. 306–312.

[28] Muhyiddine Jradi, Christian Veje, and Bo Nørregaard Jørgensen. "Towards Energy Efficient Office Buildings in Denmark : The Maersk Building Case Study". In: *29th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems (ECOS2016)*. Portorož, Slovenia, 2016.

[29] Krzysztof Arendt, Ana Ionesi, Muhyiddine Jradi, Ashok Kumar Singh, Mikkel Baun Kjærgaard, Christian T Veje, and Bo Nørregaard Jørgensen. "A Building Model Framework for a Genetic Algorithm Multi-objective Model Predictive Control". In: *CLIMA 2016 - proceedings of the 12th REHVA World Congress*. Aalborg, Denmark, 2016.