Proceedings Paper:

Fang, Xinwei and Bate, Iain John orcid.org/0000-0003-2415-8219 (2017) Issues of using wireless sensor network to monitor urban air quality. In: International Workshop on the Engineering of Reliable, Robust, and Secure Embedded Wireless Sensing Systems (FAILSAFE). ACM .

Issues of using wireless sensor network to monitor urban air quality

Xinwei Fang
Department of Computer Science
University of York
xinwei.fang@york.ac.uk

Iain Bate
Department of Computer Science
University of York
iain.bate@york.ac.uk

ABSTRACT

Frequent monitoring of urban environment has now been regulated in most EU countries. Due to the design and cost of high-quality sensors, the current approach using these sensors may not provide data with an appropriate spatial and temporal resolution. As a result, using a wireless sensor network constructed by a large number of low-cost sensors is becoming increasingly popular to support the monitoring of urban environments. However, in practice, there are many issues that prevent such networks to be widely adopted. In this paper, we use data and lessons learnt from three real deployments to illustrate those issues. The issues are classified into three main categories and discussed according to the different sensing stages. In the end, we summarise a list of open challenges which we believe are significant for the future research.

ACM Reference Format:
Xinwei Fang and Iain Bate . 2017. Issues of using wireless sensor network to monitor urban air quality. In *Proceedings of ACM Conference (FAILSAFE'17).* , 8 pages.

1 INTRODUCTION

Current air pollution mitigation strategies require the air to be monitored in an appropriate spatial and temporal scale [11, 22]. High-quality sensors installed for regulatory monitoring purposes often have a high market price and demand frequent manual handling such as re-calibration. These constraints make current sensors prohibitively expensive to deploy, especially at a higher spatial resolution. Therefore, an alternative solution is urgently required. As a result, wireless sensor networks constructed by a large number of low-cost sensors are designed and implemented.

Low-cost sensors are defined as electronic sensors that cost several orders less than high-quality sensors. As a result, low-cost sensors are able to construct a higher density network with a relatively low cost, which further enables data to be obtained at the sufficient spatial and temporal resolution and addresses the issues that the current monitoring networks are having [6, 12, 15]. However, there are many challenges that make it difficult to obtain useful information from these low-cost sensors. It is noted that in this paper information is considered different from data; with data being what is sensed and information being resulting models after processing upon which decisions may be taken.

According to existing studies, low-cost sensors are prone to failures and errors, and their general accuracy can be much lower than high-quality sensors [18]. Hence, data from low-cost sensors often shows various issues. For example, the data can occur at a higher percentage of magnitude (e.g. through spikes) than the reference data [9]; or the data can be affected or biased by other substances in the air or interference from the environment (e.g. cross-sensitivity) [3]. It is aware that using these data directly would potentially bias the decision making, hence, various processes from the calibration of sensors to optimizing the deployment of networks have been proposed to address the data issue. However, we identify that the inappropriate use of some of the processes would further bias the data and needs to be avoided. Unfortunately, to the best of our knowledge, it has not yet been widely reported.

In this paper, we discuss a list of data issues that we encountered during our study and inference their possible causes. We further explain why some of the data issues are significant and cannot be compensated for by certain methods. The rest of this paper is organised as follows. We first introduce our deployments and the use of sensors in Section 2. Then, we discuss a list of data issues that occurred in different stages of sensing in an order of 1) the deployment of sensors; 2) the obtaining of sensor data; and 3) the processing of the data from the Section 2 to 5. Finally, we conclude this paper in Section 7 with some of the open challenges for the environmental monitoring using low-cost sensors.

2 SENSORS AND DEPLOYMENTS

The ELM sensor, a product from Perkin Elmer, is used as the low-cost sensor in our study [13]. It measures multiple parameters including nitrogen dioxide (NO_2), ozone (O_3), temperature, humidity, volatile organic compound (VOC), dust and noise. The parameter of dust stands for particulate matters, PM_{10} and $PM_{2.5}$; The parameter of noise presents for the level of sound in decibels. ELM sensors are powered by AC and the overall unit is about the size of a shoe box. The sensors are designed to have a life time for about 18 months as some of the sensors provide their data via chemicals that degrade. Data, by default, is uploaded to a server using GSM. However, when the GSM service is not available, data is temporally stored (within the limits of available resources) in an on-board data logger and uploaded again when the GSM communication recovers.

A deployment of sensors often has a purpose. In this work, we introduce three of our deployments in York, UK. The first deployment was in 2015 when we first had the ELM sensor. For this deployment, the aim was to understand the performance of ELM sensors in an uncontrolled environment as we only have a datasheet describing

FAILSAFE'17

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.

how it behaves in a Lab. We wanted to know if sensors can report accurate or consistent data in respect to high-quality sensors without or with a simple calibration. Hence, twenty ELM sensors were co-located with a reference sensor for more than two months at the Wolfson Atmospheric Chemistry Laboratories (WACL). The reference sensor is a high-quality sensor that is carefully maintained by WACL. This location is on the west campus of the University of York. Since it is outside of city centre and away from major roads and junctions, the environment is believed to be consistent and considered as *mild*.

Sensors may have non-unique responses in different conditions of environments [3]. Hence, the aim of the second deployment was to understand how ELM sensors would perform in a typical urban environment and to determine how the response of sensors would differ from in the mild environment. This deployment was at the Fishergate, which is in the centre of York next to a busy junction. This environment is therefore considered as *harsh*. At the Fishergate, two ELM sensors were co-located with a high-quality reference sensor for more than 8 months in early 2016. The reference sensor is managed by the City of York Council and it is a part of Automatic Rural and Urban Networks (ARUN).

The third deployment was in cooperation with ARUP to evaluate how green infrastructures in an urban environment would impact the micro-environment. Eleven ELM sensors were firstly placed at the WACL for a month co-located with a reference sensor (similar to the fisrt deployment) so that some calibration data was available and then moved to Scarcroft road, York. In this deployment, sensors were spread around the area, some sensors were on the road side while others were placed in a green park. No reference sensors were available on site. This deployment started in the middle of 2016 for more than 8 months.

For all three deployments, like most of end-users, we do not have a direct access to the hardware during and after the deployments. The data is obtained through the service providers and downloaded from their server directly via the API [13].

3 ISSUES WITH THE DEPLOYMENT

The place of the deployment is often determined in advance according to the purposes, however, real deployments are often constrained by practical limitations which can be unforeseen in a planning phase. In this section, we share a list of issues that we encountered during the deployments.

The first issue is the location of deployments in terms of spatial locations. ELM sensors are powered by mains electricity instead of battery, which is believed to address the power limitations that many low-cost sensors have. However, our deployment is then constrained by the power supplies. We determine the best practice for a consistent AC supply across a city is to utilise lamp-posts. However, locations of the sensor deployment are then constrained by the availability of lamp-posts in the city. Lamp-posts in York are managed by third parties, hence, deploying sensors on lamp-posts requires their cooperation. Furthermore, getting permissions from the local council can also be a time consuming process. Therefore, sensor deployments in urban environment can be very expensive in terms of both labour costs and arrangement.

Another issue is the location of deployments in terms of height. According to the regulatory, the air-intake of a high-quality sensor, which is used as a reference in this work, should be placed at 1.60 meters above the ground. This height is to determine pollution exposures for adults. Other studies suggest to place sensors lower as children being more vulnerable than adults are far below that height. However, our deployment failed to meet either of the requirements. Since we do not know how local community would react to the deployment and we have to prevent the sensors being physically damaged, sensors on lamp-posts are placed at three meters above the ground and sensors on top of high-quality sensors are locked in a metal cage. We are aware that our deployment indirectly ignores the variation of the height. However, considering the effect of the height is not our main interest and to the best of our knowledge no studies have shown it is significant, we did not investigate it further.

Since ideal locations of deployment can often be constrained by the practical limitations, obtaining of data in an desired location can be an issue. Furthermore, some methods of data compensation require the network or sensors in a certain topology or in a certain spatial and temporal range [4, 10, 16]. As a result, those methods may be difficult to apply in such circumstance.

4 THE ISSUES OF OBTAINING DATA

In this section, we discuss issues that we encountered during the data acquisition. The process of data acquisition collects data from the deployed sensors for the further processes. At this stage only the data pattern is checked by visual inspection. The data pattern is classified as the normal and abnormal based only on domain knowledge. An abnormal data pattern is often associated with the issues in networks or sensors, such as data gaps caused by the communication issues and constant values caused by malfunction of sensors [18]. The abnormal data pattern often suggests that a physical inspection of sensors or networks may be needed. The normal data pattern indicates that the monitoring networks is working properly in the system level, e.g. the communications. However, it would require further processes to evaluate the accuracy of the data.

4.1 Data gap

Since the sensors are designed to provide data in a consistent time interval, any gaps in the data are considered as abnormal. We have an omission failure that causes a complete loss of the data. Then, long terms and short terms data gaps can also be observed across all deployments.

Sensors deployed at the Fishergate failed to report data back to the server after the deployment. We identified that it was due to the effect of the metal cage. The cage that protects the sensors blocks all signals of the communication. For the deployment in WACL, 11 out of 20 sensors stopped getting measurements after 2 months of the deployment. It was due to the GSM service providers. Those issues can results in a permanent loss of data and a physical inspection may allow the cause to be corrected.

The remaining nine sensors show a partial loss of the data, which have a similar pattern to Figure 1. Figure 1 shows the completeness of data received in a week time from these nine ELM sensors at

WACL. The colour is associated with a percentage of data that have been received in an hour. Light yellow indicates the data has been completely received (100%) and dark blue shows a complete data loss (0%).

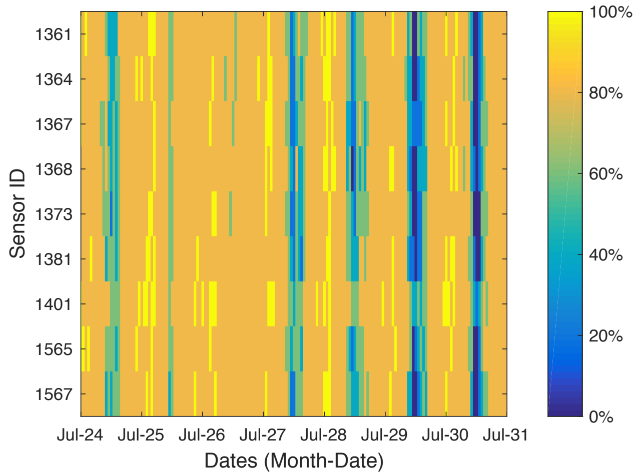

Figure 1: Data completeness

In the figure only a small percentage of data are completely received whereas the data loss can be observed frequently in different levels. Such data pattern can be widely observed across all of our deployments despite the system having a mechanism to avoid the data loss as mentioned in Section 2.

An interesting observation during the deployment is the data gaps in WACL occur more frequently during the university term time than the off-term. This relationship suggests that some of the data loss can be associated with the volume of communication traffic. Since during a term time more GSM users will share a limited bandwidth, the data collision is likely to occur. However, as we don't have the direct access to the hardware, we haven't had sufficient evidence to confirm this.

It is noted that ELM sensors were commercially available. Hence, the reliability of the system in terms of data communication should be better than many open-source systems. However, data gaps can be observed across all sensors. It suggests that data communication for WSN in urban environment may still be an issue.

4.2 Data uncertainty

Data uncertainty describes how data vary from the ground truth. Since data reflects the variations from the sensors and the environments, we believe a higher variation of a sensor will lead to a higher data uncertainty. Low-cost sensors are widely reported to have a larger variation (e.g. due to low sensitivity and selectivity) and the response of the sensor is easier to be affected by the environment (e.g. due to cross-sensitivity). Hence, data from low-cost sensors often have a larger uncertainty.

Sensor ageing and material degradation can occur in all sensors at different stages, and they are believed to influence the response of a sensor and further affect data uncertainty, like a drift of measurements.

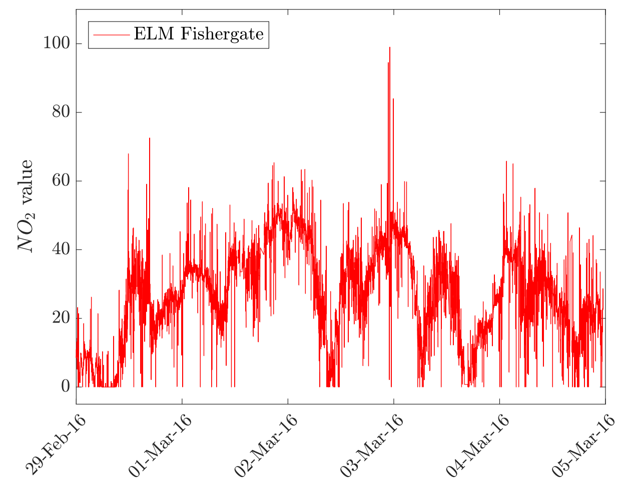

Figure 2: Data of ELM from Fishergate

Figure 2 shows a week long measurement of NO_2 from a ELM sensor at the Fishergate. In the figure, a large variation of data can be observed. Since it is impossible to separate the variations that caused by the sensor and the environment, it would not be possible to identify the drift of sensors or to know the drift rate. It shows an importance of having a ground truth during the deployment. The ground truth will indicates the variation of the environment. Hence, the variation of the sensor can be subtracted from the data.

Furthermore, environmental interference or other unknown issues can cause sensors to misbehave temporally, which results in abnormal data patterns (e.g. spikes) and affects data uncertainty. The random data patterns caused by sensor issues are referred to as anomalies in this work. For example, an unexpected spike caused by communication interference in the continuous data is considered as an anomaly. Since anomalies in the data can have random patterns, compensation of anomalies can be difficult in practice. Therefore, identifying anomalies and removing them can be important for reducing data uncertainty.

The detection of anomalies can be difficult for the environmental data. For example, we can observe a few spikes in the Figure 2. However, we would not confidently classify them as anomalies since those spikes may also be introduced by events at various levels or time-bands [23], such as spikes caused by a bus idling near a sensor (the minutes time band), York race days (the day time band), and roadworks (the week time band). As a result, knowing environmental conditions can be important for any further data processing.

In the stage of data acquisition, data is only checked visually. Gaps or certain patterns in data can indicate whether the sensors or network is working properly at some degree. However, it is difficult to compensate the data or to evaluate the data uncertainty when no ground truth is available. Moreover, in the Section 5, we will illustrate that processing methods can have different impacts on data uncertainties. Therefore, processes and techniques are needed to understand the uncertainties so that some of them can be reduced.

5 THE ISSUE OF DATA PROCESSING

The processing of data aims to improve data quality to a standard. The data quality is often defined by end-users indicating whether the accuracy of the data is sufficient for their processes. Therefore, ideally we would know the uses of the data from end-users and optimise the techniques accordingly. As we do not know the uses, our aim of the data processing becomes to improve the data quality by maximising the accuracy of the data with respect to the reference.

As explained in Section 4.2, data is sensitive to both the environment and the specific sensors. Hence, we believe that if the sensor issues are properly addressed, the accuracy of the data could be significantly improved. Therefore, our data processing is mainly focused on compensating sensor issues, such as sensors calibration and anomaly detection [3, 23]. In this section, we share a list of important findings encountered during the processing of the data.

5.1 Reference

As mentioned in Section 4, it is important to have a ground truth for data processing. However, unlike in a Lab environment, conditions in real environments do not have a control, which results in obtaining a ground truth in the environment of deployments difficult. In most practices, data from high-quality sensors is often considered as the ground truth and is referred to as the reference. Since no sensor can obtain an absolute ground truth, such practice potentially biases the result of the process. However, to the best of our knowledge, there is no better solution. Another issue of using high-quality sensors as references is the temporal resolution. Due to the design of the sensor, the reference data is often provided on an hourly basis. Consequently data obtained in a higher temporal resolution, like a minute time band, will require an aggregation process before it can be compared to the reference. As a result, the temporal scale of low-cost sensor can be affected.

For some applications, the ground truth is also suggested to be obtained using statistical estimation, like macro-calibration that will be discussed in Section 5.3. However, comparing the ground truth that obtained from high quality sensors, such approach introduces more uncertainty and has less accuracy. Hence, using an estimated reference is not suggested if a high-quality sensor is available.

5.2 Data aggregation

The main purpose of data aggregation in this work is to average data with a higher temporal resolution into the same resolution as the reference (hourly) for the evaluation. Data aggregation can be generally classified as an on-line or off-line process. An on-line process stands for the data is aggregated on the sensor before it is transmitted to a server. An advantage of this process is it significantly reduces the amount of data that needs to be transmitted and saves the costs of communication. However, such a process is not reversible, which means that if the aggregation process makes inappropriate transformations they cannot be recovered. On the contrary, an off-line process transmits everything back to a server. As a result, raw data can be securely stored and can be recovered at any time when it is necessary. Since our sensors are powered by AC, the power is not an issue. Hence, we utilise an off-line process in our study. It is noted though the off-line process can make the errors due to communications worse as discussed in section 4.1.

For the techniques of the data aggregation, arithmetic mean and median are often used. The arithmetic mean is the most often used technique to average the data. It is the sum of received values divided by the number of received values. However, it is noted that the arithmetic mean is sensitive to the sample size (the number of received values). Considering the number of samples in a window of an hour can be non-unique due to the data gaps, using arithmetic mean can result in the confidence interval of the mean value from each hour being different. Moreover, the mean is also sensitive to extreme values. For example, in Figure 2, if all observed spikes were anomalies, then the mean value could be largely influenced by the anomalies. However, it does not mean using median value is always a better option than the mean. As median value is a single value, median value will not represent for other samples. If the spikes do have an meaning, taking the median value may not be appropriate as it ignores all information from the spikes. Moreover, if a percentage of anomalies is more than 50% on an hourly samples, the median value is then biased.

Another issue is the sliding window for the hourly average. Even though a different starting point of a sliding window will not significantly affect the result, the result will still be affected by the use of different samples. Hence, it is believed that if we know how the reference is produced, a better sliding window could be determined and the uncertainty can be minimised.

Finally, depending on the type of the network, such as flat and hierarchy network, the strategy of data aggregation can be different. It is also worth considering how to archive the data and how to balance the trade-offs, such as trade-off between the communication cost and amount of data to be transmitted.

5.3 Calibration

Calibration can be important for the data process as it is believed to address a systematic variation of sensors. However, the calibration processes required by different applications may be non-unique, which could lead to various issues. We discuss the issues from calibration according to the calibration of networks and the calibration of a single sensor. We further classify the calibration of networks as macro-calibration and micro-calibration according to [3].

5.3.1 Macro-calibration. Macro-calibration utilises the consistency of nearby environment and maximises the similarity of measurements from the neighbouring sensors. This calibration requires the ground truth to be estimated from corresponding sensors. As mentioned in Section 5.1, the estimated ground truth may have large uncertainties, which can further affect the result of calibration.

Existing studies that utilise macro-calibration often require a ultra dense network [2]. However, considering the issues discussed in Section 2, it can be difficult to deploy a network to fit the requirement of calibration in cities.

The authors in [1] propose a technique that can estimate ground truth without requiring a dense network by assuming a network can oversample the underlying signals. However, our understanding of the paper suggests that such a method requires targeted signals being consistent over the space, e.g. temperature. If the signal is not or less spatially consistent, like NO_2, a dense network will still be required. Otherwise, the method will obtain a biased ground truth and affect the result of calibration.

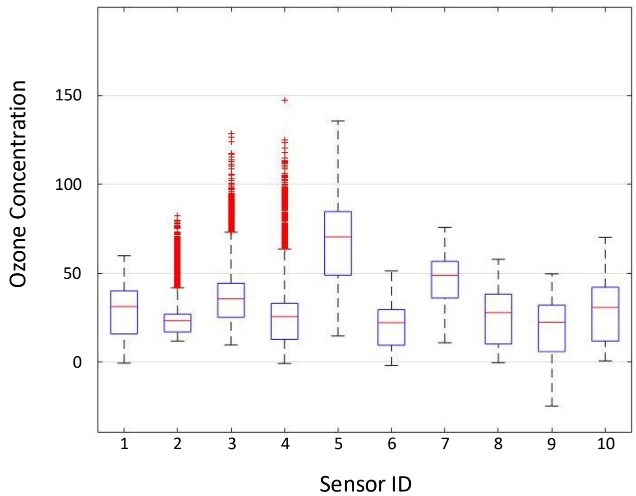

Figure 3: Inconsistency

Moreover, we also need to aware if the calibration process is appropriate for the end-users. Figures 3 shows boxplot of 2 months worth of O_3 from 10 ELM sensors at WACL. The figure shows O_3 measurement is inconsistent. Ideally, we would like to use a reference sensor co-located with those sensors to identify whether the inconsistency of the measurements was due to the environment or sensors. However, since the reference would not also be available, such evaluation may not be performed in many real deployment. As a result, the actual variation of the environment could be considered as the variation of sensors and be leveraged in a macro-calibration process. If the calibrated data, later on, is used to determine the environmental difference between locations, the determined result is likely to be biased.

Above examples show the importance of having references in the calibration process.

5.3.2 Micro-calibration. Micro-calibration, on the contrary, relies on reference sensors. However, this would require a reference sensor to be co-located with every low-cost sensor. Since the requirement can be practically difficult, solutions that utilise a fresh calibrated sensors to propagate the calibration has been widely used [5, 16].

The biggest issue in using calibrated sensors to propagate the calibration is the calibration errors as calibration errors of individual sensor will propagate through the calibration path. As a result, the error of micro-calibration will be closely associated with the calibration of errors of individual sensor and the size of the network.

Furthermore, the propagation of calibration will also require similarity or consistency of the environment. As a result, it may face to a similar problem that discussed in Section 5.3.1. Some suggest using the calibration function obtained during the night time as the air pollutant is more spatially consistent than during the day time. However, the pattern of pollutant in a day and night time can be different as shown in Figure 4. The Figure 4 shows the measurements of O_3 in the day and the night respectively. The measurements were averaged into hourly basis using median value. It suggests that the measurements are generally lower during the night than the day. As a result, the calibration function determined using the night time data may not be optimal for the data from day time [3].

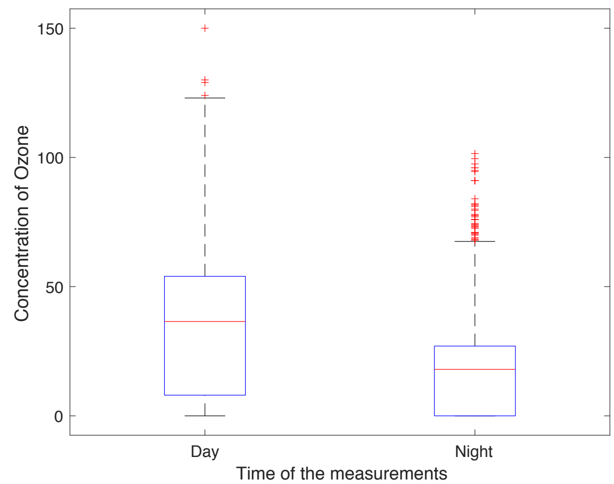

Figure 4: Day and night pattern of Ozone

5.3.3 Calibration of a single sensor. As errors from a single sensor calibration can be propagated in the calibration of networks, reducing the errors can be significantly important for the calibration of the network.

Ideally, every sensor will undertake a series of lab tests. In the lab, the environment is encapsulated and sensors will not be influenced by other gases or conditions. However, since sensors in real practice will be exposed in an environment that contains multiple gases and some of them can react each other under certain environmental conditions, calibration of sensors in a real environment can be different from the lab.

A conventional way to calibrate a sensor in real environments is to determine a predictive model between an uncalibrated sensor and its reference. However, due to cross-sensitivity and other unknown reasons, the correlation between the data from low-cost sensor and the reference is often weak [9]. As a result, the calibration can be difficult and the calibration errors are often large. Figure 5 shows an example of a low correlation between an ELM sensor and a reference. It suggests that a more comprehensive calibration method may be required.

The authors in [9, 19–21] suggest utilising the cross-sensitive parameters, temperature and humidity in the calibration as the response of sensor can be closely associated with those variables. Their result shows the calibration errors are significantly reduced. However, it is noted that no two sensors are identical, hence, it is necessary to calibrate every sensor in the network. Furthermore, locations would affect the calibration as the environmental conditions can vary in different places. As a result, having the ground truth for the calibration would be an issue.

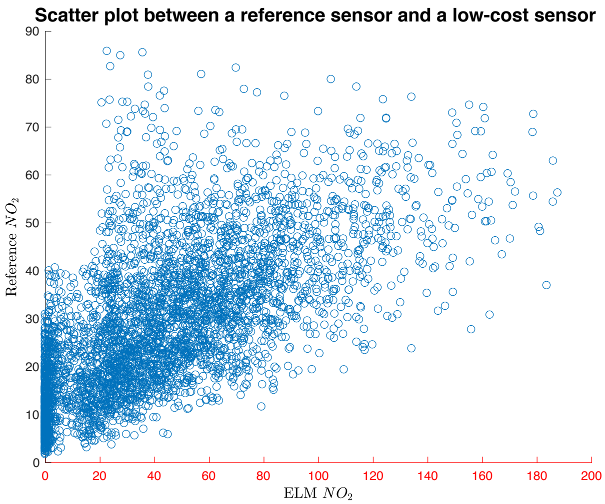

Figure 5: Correlation

6 DETECTION OF ANOMALIES

As explained in Section 4.2, the anomalies can potentially increase data uncertainty and hard to be compensated. As a result, detection and removal anomalies is another important process to ensure data quality.

For the detection of anomalies, understanding the pattern, magnitude and percentage of anomalies can be important. As our sensors are deployed in real environments which means that the ground truth of anomalies may be unavailable, it can be difficult to accurately label or deeply understand anomalies. As a result, most of quantitative analysis for anomaly detections are evaluated on an artificial data. However, since the understanding of anomalies from environmental data is not comprehensive, the artificial data may not fully simulate the reality, which leads to an issue of detecting anomalies in real applications.

Detection of anomalies has been an active research in many domains. Since anomalies in our dataset may not have an unique pattern as explained in 4.2, many well-known methods or techniques that rely on data pattern to distinguish anomalies are not applicable. For example, supervised learning methods may face issues when the anomalies are not correctly labelled in the training phase; or threshold based methods may be difficult to accurately classify the anomaly as anomalies do not have an unique data pattern.

The authors in [23] suggests that data that is difficult to be separated in one feature space may be easier to be separated in another feature space. Since the fact that the correct measurements are likely to be related in a certain context, whilst the anomalies may be stochastically unrelated to the correct measurements, utilising contextual information is believed to help to detect anomalies. For example, in the Figure 2, if the spikes were also observed in the data of its co-located sensors, the spikes are unlikely to be anomalies. Otherwise, we would have a higher confidence to classify the spikes as anomalies. Even though this observation is intuitive simple, understanding the stochastic relationships and how context affects the relationships is not for many systems including wireless ones [7, 8].

Temporal dependency and spatial dependency are mostly used contextual information in the literature [14, 17]. However, they may not be ideal for our application. Figure 6 shows an auto-correlation of NO_2 between time t and time $t-1$. It suggests that in our application the temporal dependency can be weak, which suggests that utilising temporal dependency can be insufficient for anomaly detection in our application. Figure 7 shows the data of NO_2 from WACL and Fishergate respectively, at each location an ELM sensor is co-located with a reference. Comparing the sensors at WACL with those in Fishergate, it is believed that utilising spatial dependency for the detection of anomalies is also inappropriate as the spatial variation can be significant for certain parameters.

Real data suggests that existing contextual information may not be applicable for our application. Since there is a lack of studies in understanding anomalies in environmental data, it is still not clear what contextual information is relevant or should be utilised for the detection of anomalies of a particular parameter.

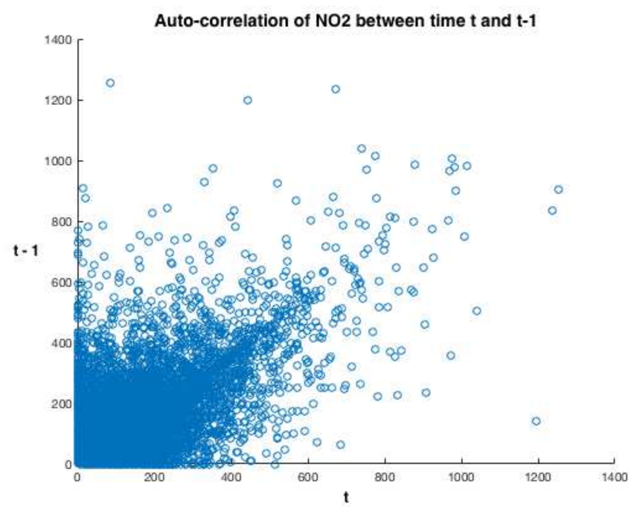

Figure 6: Auto-correlation

7 CONCLUSION AND OPEN ISSUES

In this paper, we share a list of issues that we encountered during our deployments.

In the first deployment, temperature and humidity show a consistent measurement over the number of sensors. But the measurements of NO_2 and O_3 have a large variation among the different sensors. Considering the distance between sensors are within 5 meters range and the environment condition is relatively consistent, we suspect the inconsistency of data was caused by sensors.

In the second deployment, the magnitude of measurement is constantly and significantly higher than the reference, especially at the a high temporal resolution. Apart from low-cost sensors having a lower accuracy, we believe the difference in measurement is also associated with the dynamic of the environment, such as emissions

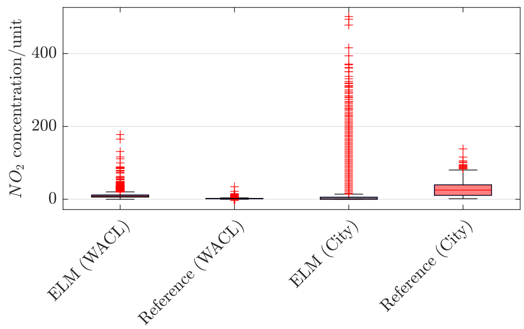

Figure 7: boxplots for NO_2 in two locations

from vehicles. Since low-cost sensors sample the environment more frequently than high-quality sensors, every 20s, low-cost sensors are more affected by the dynamics of the urban environment.

In the third deployment, the calibration functions determined at WACL was soon found to be invalid in the deployment. It is because some of parameters change with time and space. For example, the temperature is different between summer and winter; the NO_2 concentration is inconsistent between mild and harsh environments.

From those deployments, we observe various data issues and some of them can easily be compensated. However, others are more difficult to address. For example, the calibration of the networks can be one of them. As explained in Section 5.3, the propagation of calibration function in a network relies on the consistency of the environment. However, an urban environment is dynamic, which would result in large calibration errors in the result. On the other hand, it is also practically impossible to have reference sensors co-located with every low-cost sensor in the network. Therefore, the calibration of networks is still an open challenge.

We believe that having a transferable calibration model can be significantly useful for the calibration of networks. Ideally, the transferable model is able to obtain an calibration function in one location, e.g. in a place where the reference is available. Then, the calibration function is still working when sensors are moved to a different location. As the sensor response is closely associated with different environmental conditions (e.g. different concentration and combination of gases) and various environmental variables (e.g. vehicles, wind, sunlight), if we can identify how the sensor response is affected by those variables and conditions, the transferable model can be obtained by automatically adjusting the differences.

Furthermore, the detection of anomalies in the data can be another issue. Currently, we often do not know why and how an anomaly is caused. Hence, distinguishing anomalies from real events faces a high false positive rate, which can have a significant impact on data analysis. For example, if data caused by an event is wrongly classified as anomalies, a decision may be taken differently. Therefore, a better understanding of anomalies and their causes is also important.

Apart from the examples above, there are still many open issues that prevent low-cost sensors to be widely adapted for the urban environmental monitoring. We identify some of the core issues as:

- What environmental variables affect the response of the sensor most and how to determine them?
- What are the properties of anomalies in environmental data and what are their main causes?
- How can we have more reliable and diverse ground truth for the evaluation of low-cost sensors or networks in different time-bands?
- How context affects the approaches to calibration and anomalies?
- How to balance the available resources against the usefulness of data, e.g. when deciding between off-line and on-line methods?

8 ACKNOWLEDGEMENT

This work is funded by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 608014 (CAPACITIE)

REFERENCES

[1] L. Balzano and R. Nowak. Blind calibration of sensor networks. In *Proceedings of the 6th International Conference on Information Processing in Sensor Networks*, pages 79–88. ACM, 2007.
[2] V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak. A collaborative approach to in-place sensor calibration. In *Proceedings of the 2nd International Conference on Information Processing in Sensor Networks*, pages 301–316. Springer, 2003.
[3] X. Fang and I. Bate. Using multi-parameters for calibration of low-cost sensors in urban environment. In *International Conference on Embedded Wireless Systems and Networks (EWSN)*, 2017.
[4] K. Fu, W. Ren, and W. Dong. Multihop calibration for mobile sensing: k-hop calibratability and reference sensor deployment. In *Proceedings of IEEE International Conference on Computer Communications (INFOCOM)*, 2017.
[5] D. Hasenfratz, O. Saukh, and L. Thiele. On-the-fly calibration of low-cost gas sensors. In *International Conference on Embedded Wireless Systems and Networks (EWSN)*, pages 228–244. Springer, 2012.
[6] P. Kumar, L. Morawska, C. Martani, G. Biskos, M. Neophytou, S. Di, M. Bell, L. Norford, and R. Britter. The rise of low-cost sensing for managing air pollution in cities. *Environment International*, 75:199–205, 2015.
[7] H. Lau, I. Bate, P. Cairns, and J. Timmis. Adaptive data-driven error detection in swarm robotics with statistical classifiers. *Robotics and Autonomous Systems*, 59:1021–1035, 2011.
[8] T. Lim, I. Bate, and J. Timmis. A self-adaptive fault-tolerant systems for a dependable Wireless Sensor Network. *Design Automation for Embedded Systems*, 18(3-4):223–250, 2014.
[9] B. Maag, O. Saukh, D. Hasenfratz, and L. Thiele. Pre-deployment testing, augmentation and calibration of cross-sensitive sensors. In *International Conference on Embedded Wireless Systems and Networks (EWSN)*, pages 169–180. ACM, 2016.
[10] B. Maag, Z. Zhou, O. Saukh, and L. Thiele. Scan: Multi-hop calibration for mobile sensor arrays. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. ACM, 2017.
[11] A. Makri and N. Stilianakis. Vulnerability to air pollution health effects. *International journal of hygiene and environmental health*, 211(3):326–336, 2008.
[12] H. Messer, A. Zinevich, and P. Alpert. Environmental monitoring by wireless communication networks. *Science*, 312(5774):713–713, 2006.
[13] Perkin Elmer. Elm sensor. https://elm.perkinelmer.com/, February 2015. [Online; accessed on 22-September-2017].
[14] M. Rassam, A. Zainal, and M. Maarof. Advancements of data anomaly detection research in wireless sensor networks: A survey and open issues. *Sensors*, 13(8):10087–10122, 2013.
[15] B. Resch, M. Mittlboeck, F. Girardin, R. Britter, and C. Ratti. Live geography: Embedded sensing for standarised urban environmental monitoring. *International Journal on Advances in Systems and Measurements*, 2009.
[16] O. Saukh, D. Hasenfratz, and L. Thiele. Reducing multi-hop calibration errors in large-scale mobile sensor networks. In *Proceedings of the 14th International*

Conference on Information Processing in Sensor Networks, pages 274–285. ACM, 2015.

[17] N. Shahid, I. Naqvi, and S. Qaisar. Characteristics and classification of outlier detection techniques for wireless sensor networks in harsh environments: a survey. *Artificial Intelligence Review*, 43(2):193–228, 2015.

[18] A. Sharma, L. Golubchik, and R. Govindan. Sensor faults: Detection methods and prevalence in real-world datasets. *ACM Transactions on Sensor Networks (TOSN)*, 6(3):23, 2010.

[19] L. Spinelle, M. Gerboles, and M. Aleixandre. Report of laboratory and in-situ validation of micro-sensor for monitoring ambient air-ozone micro-sensors, αSense, model:$B4$ O_3 sensors. *Publications Office of the European Union*, 26681, 2013.

[20] L. Spinelle, M. Gerboles, M. Villani, M. Aleixandre, and F. Bonavitacola. Calibration of a cluster of low-cost sensors for the measurement of air pollution in ambient air. In *SENSORS*, pages 21–24. IEEE, 2014.

[21] L. Spinelle, M. Gerboles, M. Villani, M. Aleixandre, and F. Bonavitacola. Field calibration of a cluster of low-cost available sensors for air quality monitoring Part A: Ozone and nitrogen dioxide. *Sensors and Actuators B: Chemical*, 215:249–257, 2015.

[22] L. Trasande and G. Thurston. The role of air pollution in asthma and other pediatric morbidities. *Journal of allergy and clinical immunology*, 115(4):689–699, 2005.

[23] Y. Zhang, N. Meratnia, and P. Havinga. Outlier detection techniques for wireless sensor networks: A survey. *IEEE Communications Surveys & Tutorials*, 12(2):159–170, 2010.