Environmental Pollution xxx (2017) 1–10

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Node-to-node field calibration of wireless distributed air pollution sensor network[☆]

Fadi Kizel [a], Yael Etzion [a], Rakefet Shafran-Nathan [a], Ilan Levy [a], Barak Fishbain [a], Alena Bartonova [b], David M. Broday [a, *]

[a] *Faculty of Civil and Environmental Engineering, Technion - Israel Institute of Technology, Haifa 32000, Israel*
[b] *Norwegian Institute for Air Research (NILU), Kjeller, Norway*

ARTICLE INFO

Article history:
Received 21 March 2017
Received in revised form
24 August 2017
Accepted 3 September 2017
Available online xxx

Keywords:
Air pollution
Distributed sensor network
Chain calibration
Field deployment of sensor nodes
Exposure

ABSTRACT

Low-cost air quality sensors offer high-resolution spatiotemporal measurements that can be used for air resources management and exposure estimation. Yet, such sensors require frequent calibration to provide reliable data, since even after a laboratory calibration they might not report correct values when they are deployed in the field, due to interference with other pollutants, as a result of sensitivity to environmental conditions and due to sensor aging and drift. Field calibration has been suggested as a means for overcoming these limitations, with the common strategy involving periodical collocations of the sensors at an air quality monitoring station. However, the cost and complexity involved in relocating numerous sensor nodes back and forth, and the loss of data during the repeated calibration periods make this strategy inefficient. This work examines an alternative approach, a node-to-node (N2N) calibration, where only one sensor in each chain is directly calibrated against the reference measurements and the rest of the sensors are calibrated sequentially one against the other while they are deployed and collocated in pairs. The calibration can be performed multiple times as a routine procedure. This procedure minimizes the total number of sensor relocations, and enables calibration while simultaneously collecting data at the deployment sites. We studied N2N chain calibration and the propagation of the calibration error analytically, computationally and experimentally. The *in-situ* N2N calibration is shown to be generic and applicable for different pollutants, sensing technologies, sensor platforms, chain lengths, and sensor order within the chain. In particular, we show that chain calibration of three nodes, each calibrated for a week, propagate calibration errors that are similar to those found in direct field calibration. Hence, N2N calibration is shown to be suitable for calibration of distributed sensor networks.

1. Introduction

Air pollution is known to levy severe health effects and high risks for the public (World Health Organization, 2013; Crouse et al., 2012; Lepeule et al., 2012), hence air quality is regularly monitored in many regions worldwide. Regulatory air pollution monitoring is mainly performed by stationary and routinely calibrated reference Air Quality Monitoring (AQM) instruments, which measure the concentrations of different criteria pollutants, typically ozone (O_3), nitrogen oxides (NO_x), carbon monoxide (CO), sulfur dioxide (SO_2), and particulate matter (PM). While AQM stations provide reliable and accurate measurements, they are expensive to install and to operate, and require professional maintenance and personnel. Therefore, the spatial distribution of AQM stations is rather sparse. The use of geospatial interpolation or regression methods for estimating ambient concentrations of (and exposure to) monitored pollutants away from the AQM stations is a common procedure for bridging over the sparse spatial availability of the observations (Yuval and Broday, 2006; Eitan et al., 2010; Whitworth et al., 2011; Myers et al., 2013; Sampson et al., 2013). Yet, such a mapping is significantly affected by the spatial distribution of the stations (Yuval and Broday, 2006) and the temporal resolution of the reported data, and may involve spatially biased model errors (Yuval et al., 2017). Such model errors tend to propagate when concentration maps are used for, e.g., exposure estimation, in particular in

☆ This paper has been recommended for acceptance by Eddy Y. Zeng.

* Corresponding author. Environmental, Water and Agricultural Engineering, Faculty of Civil and Environmental Engineering, Technion – Israel Institute of Technology, Israel.
E-mail address: dbroday@tx.technion.ac.il (D.M. Broday).

https://doi.org/10.1016/j.envpol.2017.09.042
0269-7491/

areas that are characterized by considerable spatiotemporal concentration variability (Yuval et al., 2017; Zandbergen et al., 2012; O'Leary and Lemke, 2014; Reggente et al., 2015).

Recently, miniaturization of sensor technology has enabled deployment of multi-sensor Micro Sensing Units (MSUs, hereinafter nodes) as part of Wireless Distributed Sensor Networks (WDSNs) for air quality measurements (Kumar et al., 2015; Kotsev et al., 2016; Mead et al., 2013; Moltchanov et al., 2015). Dense deployment of such sensor nodes can capture the spatiotemporal variability of urban air pollution and provide more reliable exposure and risk estimates. Yet, these sensors have limited accuracy (Moltchanov et al., 2015), tendency to degrade and age relatively fast (De Vito et al., 2008; Saukh et al., 2015), and they suffer from severe interference by co-existing airborne pollutants and meteorological parameters (Fishbain et al., 2017; Castell et al., 2017). Many of these limitations are normally unaccounted for during lab testing and calibration, which are performed in controlled chambers (Mead et al., 2013; Castell et al., 2017; Williams et al., 2013). These limitations call for frequent field calibrations under real environmental conditions, to assure reliable measurements.

Field calibration of WDSN sensors has been studied using the so-called collocation procedure, where the nodes are placed next to a standard AQM station and the time series recorded by the sensors are regressed against the co-measured AQM data (Mead et al., 2013; Moltchanov et al., 2015; Fishbain et al., 2017; Castell et al., 2017; Williams et al., 2013; Holstius et al., 2014; Deary et al., 2016; Spinelle et al., 2015; Jiao et al., 2016). Specifically, this approach relies on placing the sensor next to a reference device for a certain time-period, averaging the rich sensor data to fit the lower sampling frequency of the reference device, and performing a pairwise linear-regression between the sensor and the AQM datasets. The regression coefficients are then used to correct the sensor measurements and make them follow the reference data.

Let $\mathbf{y}$ and $\mathbf{x}$ be the registered measurements by the reference device and by the WDSN sensor, respectively. Assuming a linear relationship between $\mathbf{y}$ and $\mathbf{x}$ (Moltchanov et al., 2015; Spinelle et al., 2015),

$$\mathbf{y} = \alpha \cdot \mathbf{x} + \beta + \mathbf{e}, \tag{1}$$

where α and β are the slope and intercept of the linear model, respectively, and $\mathbf{e}$ is a vector of the model errors, which are assumed to have a zero mean. Let $\alpha^\wedge$ and $\beta^\wedge$ be the estimated coefficients that are obtained using the collocation data. The calibrated measurements, $\mathbf{x}^\wedge$, are given by:

$$\mathbf{x}^\wedge = \alpha^\wedge \cdot \mathbf{x} + \beta^\wedge. \tag{2}$$

It is noteworthy that the length of the collocation period in which the sensors are adjacent to the AQM station until a reliable calibration is obtained may vary, depending on the environmental conditions (Moltchanov et al., 2015; Saukh et al., 2015; Balzano and Nowak, 2007; Hasenfratz et al., 2012) and the sensor technology (Williams et al., 2013; Holstius et al., 2014). Moreover, relocating the sensor nodes to the AQM station for calibration is labor intensive, and for a WDSN with a large number of nodes can become cumbersome. Frequent relocations of nodes to the AQM station for calibration involve also loss of measurements until the sensors are returned to their prescribed deployment sites. As such, this strategy counteracts the main advantage of the WDSN concept – richness and continuous data.

A field calibration procedure that does not require collocation at an AQM station has been suggested (Fishbain and Moreno-Centeno, 2016) for cases where the measurement errors comply with certain limitations. Yet, since the sensors are calibrated against the mean reading of all the reporting WDSN nodes, they may still provide values that do not conform with those measured by a reference device. For example, if all the sensors have a systematic measurement error this method will come short of reporting accurate concentrations (Moltchanov et al., 2015).

We propose here an alternative strategy, designated node-to-node (N2N) calibration. The idea is to employ chain calibration of the sensors in the field, with minimal interruption to the continuous measurement and fewer hops of the nodes between their deployment sites and the reference (AQM) site. Whereas N2N calibration is not limited to stationary nodes, for simplicity we assume in the following WDSNs with stationary nodes. WDSN sensors require proactive frequent calibrations, therefore a calibration procedure that involves a smaller number of collocations at AQM stations is advantageous as it enables versatile calibration logistics. Moreover, continuous measurement at the deployment sites guarantees little missing data and better spatial and temporal analyses. Reducing the number of collocations is also cost effective and environmental friendly, since WDSNs may be deployed quite far from AQM stations, i.e. the nodes may be closer to each other than to a distant AQM station.

Let $\text{AQM} \leftarrow u_1 \leftarrow u_2 \leftarrow u_3 \leftarrow \cdots \leftarrow u_{n-1} \leftarrow u_n$ represent a sequence of collocated nodes, such that sensor u_1 is collocated next to an AQM instrument for a period T. Then it is relocated and collocated with sensor u_2 (during a non-overlapping period T). Next, sensor u_2 is relocated and collocated with sensor u_3 (during a non-overlapping period T), etc. Finally, the last sensor u_n is collocated next to sensor u_{n-1}. At this stage, sensor u_n can be N2N calibrated against the AQM data. Yet, the process can end also by relocating sensor u_n to the AQM station, such that the N2N calibration process can be evaluated. Namely, the N2N calibration procedure proposes that all the sensors $\{u_1, \ u_2, \ \cdots, \ u_n\}$ are calibrated one against the other in a sequential manner, with all of them (but u_1) not collocated at the AQM station. In fact, N2N calibration has been suggested before but its mathematical model for stationary nodes was developed only for two sequential sensor pairings (Hasenfratz et al., 2012; Bychkovskiy et al., 2003). Similarly, N2N calibration of mobile sensors was also suggested by pairing events, inherent for roaming sensors mounted on vehicles (Saukh et al., 2015), using Geometric Mean Regression (GMR) to reduce the propagation of the calibration error relative to Ordinary Least Squares (OLS) regression. However, the study accounted only for the slope and disregarded the effect of the intercept on the accumulated calibration error.

Here, we study N2N calibration of stationary sensors both analytically, computationally, and experimentally, demonstrating the effect of the number and order of the nodes on the propagation of calibration coefficient errors (slope and intercept) and the overall calibration uncertainty. We present a detailed derivation of chain calibration equations and of the respective error propagation, followed by computational results that confirm the analytical derivation and reveal certain limitations of the process. Next, experimental results of WDSN nodes that were first collocated at an AQM station and then deployed in the field are presented, and the N2N calibration process and the propagation of calibration errors throughout the network are demonstrated. We conclude by discussing the suitability of the method for field calibration of air quality WDSNs.

2. Methods

2.1. Theoretical aspects of node-to-node calibration

Let sensor u_1 be collocated next to an AQM reference device for a time-period T_1 and let sensor u_2 be collocated next to sensor u_1 for