

SCIENTIFIC REPORTS



OPEN

Self Calibrated Wireless Distributed Environmental Sensory Networks

Barak Fishbain¹ & Erick Moreno-Centeno²

Received: 19 October 2015

Accepted: 18 March 2016

Published: 21 April 2016

Recent advances in sensory and communication technologies have made Wireless Distributed Environmental Sensory Networks (WDESN) technically and economically feasible. WDESNs present an unprecedented tool for studying many environmental processes in a new way. However, the WDESNs' calibration process is a major obstacle in them becoming the common practice. Here, we present a new, robust and efficient method for aggregating measurements acquired by an *uncalibrated* WDESN, and producing accurate estimates of the observed environmental variable's true levels rendering the network as self-calibrated. The suggested method presents novelty both in group-decision-making and in environmental sensing as it offers a most valuable tool for distributed environmental monitoring data aggregation. Applying the method on an extensive real-life air-pollution dataset showed markedly more accurate results than the common practice and the state-of-the-art.

Problem Statement. The increasing availability of sensors and communication technologies have both facilitated^{1,2} and catalysed^{3,4} the development of Wireless Distributed *Environmental* Sensor Networks (WDESNs) that consist of low-cost Micro Sensing Units (MSUs). WDESNs present an unparalleled means for studying environmental processes such as air-pollution^{5–8}, water quality^{9,10}, smart cities^{11,12} and wildlife ecosystems^{13,14}. These networks may consist of many sensing nodes and may be deployed over large geographical areas, rendering the calibration process of the nodes as a major obstacle in them becoming the common practice.

Here we present a new, robust and efficient method for aggregating measurements acquired by an *uncalibrated* WDESN, and producing accurate estimates of the observed environmental variable's true levels. To accomplish that we introduce a new group-decision-making method – consensus aggregation of *incomplete* ratings. The suggested methodology produces accurate results without requiring the MSUs, constituting the WDESN, to be calibrated. Thus, after the aggregation process, the herein proposed methodology renders the network to be self calibrated.

Without loss of generality, let us consider now a WDESN with K sensory nodes that measure the same physical phenomenon. The same physical phenomenon can be, naturally measured when the MSUs are collocated^{5–7}. Even when the sensors are not collocated measuring the same phenomenon can be achieved when it is uniform in all measuring points⁷. With that, due to the inherent MSUs' limitations, collocating is currently the common practice^{5–8}. MSU $k \in K$ measures pollutant's levels, at a given frequency, generating a time series, \mathbf{a}^k . The goal then is to find a consensus time series, \mathbf{r} , that agrees the most with all the MSUs' acquired time series, $\{\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^K\}$. The agreement of \mathbf{r} with each acquired time series, say \mathbf{a}^k , is measured by a distance function, $d(\mathbf{a}^k, \mathbf{r})$, that fulfills a set of axioms^{15–17}. Examples for $d()$ are the L_1 and L_2 norms, and the Kemeny & Snell¹⁶ and Cook & Kress¹⁵ axiomatic distances. Thus, given all the MSUs' acquired time series, $\{\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^K\}$, the consensus time series \mathbf{r} is the one that has the minimum sum of distances to all acquired MSUs' time series:

$$\min_{\mathbf{r}} \sum_{k=1}^K d(\mathbf{a}^k, \mathbf{r}). \quad (1)$$

Problem (1) is known as the group-decision-making problem^{15–21}. The group-decision-making problem has been widely-studied and has many applications, such as: voting¹⁸, jury decisions¹⁹, consumer opinion aggregation²⁰, and project selection²¹. In general terms, the group-decision problem is defined as follows: a group of K entities or individuals (referees) collectively evaluate n objects. In our context, the evaluations are cardinal evaluations each MSU (referee) assigns to each object (location-time pair) it evaluates. The problem then is to aggregate

¹Faculty of Civil & Environmental Engineering, Technion - Israel Institute of Technology, Haifa, 32000, Israel.

²Department of Industrial & Systems Engineering, Texas A&M University, College Station, TX 77843-3131, USA. Correspondence and requests for materials should be addressed to B.F. (email: fishbain@technion.ac.il)

the referees' evaluations into a *consensus evaluation* of each and every object. Note that the referee evaluations as well as the consensus evaluation are allowed to contain ties.

For an environmental field campaign that is carried over a time window, T , an MSU's time series, \mathbf{a}^k , is considered *complete* if for all time periods $t \in T$, the MSU gives a valid measurement; otherwise we say that the MSU provides an *incomplete* time series. The latter might happen when MSUs become faulty or switch locations. The latter scenario of sensors switching locations was described by Mead *et al.*⁵, Moltchanov *et al.*⁷ and Lerner *et al.*⁸. Regardless of the incompleteness of the MSUs' time series, we require that the consensus time series is complete. All previous group-decision work has considered the specific case of complete evaluations^{15–17}. Here we introduce and solve the *incomplete* group decision making problem. To this end we (i) introduce a set of natural axioms that must be satisfied by a distance, $d()$, between *incomplete* ratings (time series); (ii) prove the uniqueness and existence of a distance, herein called the *normalized projected Cook-Kress distance* - d_{NPKK} , which satisfies these axioms; and (iii) provide an efficient and practical method for finding the optimal rating (time series) \mathbf{r}^* for problem (1) when using d_{NPKK} as the distance function. While we present the new axiomatic distance in the context of WDESNs, it can be used for data fusion in many complete or incomplete group-decision-making application in general and in distributed sensing applications in particular. In our WDESN context, and specifically when presenting the study below, we use the term *time series* when referring to a vector of measurements (provided by an MSU or the consensus time series provided by the aggregation process), while when presenting the methodology, we use the term *rating* in order to emphasize that our proposed methodology is general and applies to aggregating any set of ratings (i.e., vectors consisting of cardinal evaluations).

Air Quality Wireless Distributed Sensor Networks. Air-Pollution (AP) is known to increase risks for a wide range of diseases, such as respiratory and heart diseases. Recent data indicate that in 2010, 223,000 deaths from lung cancer worldwide resulted from air pollution²². This number is expected to grow as studies indicate that in recent years exposure levels have increased worldwide with a significant raise in rapidly industrialising countries with large populations²³. Studying AP and its impact on health, requires accurate exposure assessments. AP related exposure metrics, typically used in environmental epidemiology studies, are based either on short term sampling²⁴ or on pollutant measurements by regulatory standard Air Quality Monitoring (AQM) stations over extended time periods²⁵. AQM stations provide accurate measurements but suffer from limited deployment due to their bulkiness, high costs, and their frequent maintenance and calibration requirements. The limited deployment tampers the AQM network's ability to adequately capture air pollutant spatial concentrations because these concentrations are highly variable. In contrast, intensive sampling campaigns use a large number of AP sensors, deployed at high densities, but are limited to relatively short time periods²². Consequently, accurate exposure assessment and the study of AP-health associations are still challenging tasks²⁶.

Since AP-MSUs cost significantly less than AQM stations, MSUs can be spread more densely and thus provide data with higher spatial resolution. However, MSUs are error-prone, may become faulty, have limited coherence over time and are inaccurate when compared to AQM stations^{5–8}. Early studies that evaluated MSUs' capabilities in a controlled lab environment^{27,28} stressed the need for a calibration process in order to sustain reliable measurements. Field deployments of such MSUs, measuring ambient O_3 levels by metal-oxide sensors⁶, and measuring CO , NO and NO_2 by electrochemical⁵ or metal-oxide²⁹ probes, have shown that calibration processes applicable for controlled lab environments do not work in the field, when the calibrated data is compared to data collected at a collocated standard AQM station^{6,7} (even after an initial field calibration has been applied⁶). Thus, the field calibration process is a critical hurdle that one must overcome, in order to make WDESN a viable tool for AP exposure assessment. Having said that, the suggested method is applicable to many WDESN applications, even though the examples here focus on AP-WDESNs.

Methods

In this section the set of axioms that a distance metric between *incomplete* ratings must fulfil so that, when using this distance within problem (1), the obtained consensus rating appropriately minimizes the disagreement of the judges (MSUs' measurements in our context) is presented. In doing so, our aim is to have a distance that is appropriate to aggregate the measurements obtained by uncalibrated MSUs.

Each sensor presents two types of errors - normal measurement error and calibration error. The former is typically considered to be additive, normally distributed with zero mean and constant standard deviation over time^{30–32}. The latter is assumed to be independent from other sensors' errors; and roughly stable throughout the measurement collection process/timeframe. The mean calibration error is assumed to be zero, though no assumption is made on the shape of the distribution. Finally the calibration error is considered to be additive. In case of multiplicative error, the algorithm can deal with this in two ways: (1) The algorithm can be applied as is and still obtain meaningful results to the extent that the multiplicative error is significantly smaller w.r.t. the readings themselves. (2) Otherwise, one can take the logarithm of each measurement and apply the (unchanged) algorithm to this re-scaled data because this data re-scaling effectively transforms a multiplicative error to an additive error (since $\log(ab) = \log(a) + \log(b)$).

The essence of the proposed method is that, due to the calibration error, given any one of the MSUs the difference between any pair of its measurements is significantly more reliable than the absolute value of the measurements themselves. As such, these differences among the same MSU measurements will be the focus of the following definitions, the axioms proposed, and the resulting distance. Specifically, the main aim of our method is to extract as much information as possible from the reliable measurement differences, and then, with that information in hand, solve the bias problem as a second step (this will be most evident in the solution procedure described at the end of this section).

Notation and Definitions. Let us consider two arbitrary incomplete ratings, \mathbf{a} and \mathbf{b} , in a universe V of n objects; each rating evaluating the objects in $\mathcal{A} \subseteq V$ and $\mathcal{B} \subseteq V$, respectively. Hereafter, we represent a rating as a vector of the form $\mathbf{a} = (a_1, a_2, \dots, a_n)$, where a_i is the score (a cardinal evaluation) of object i if object i is evaluated in \mathbf{a} and a_i is undefined otherwise. We also assume without loss of generality that the possible scores are contained in some pre-specified interval $[\ell, u]$; this assumption is without loss of generality, as the MSU's have a limited measurement range $R \equiv u - \ell$. Given two arbitrary incomplete ratings \mathbf{a} and \mathbf{b} , the following concepts are defined.

Definition 1. Given a rating \mathbf{a} and a subset S of the object universe V , the projection of \mathbf{a} on S , denoted as $\mathbf{a}|_S$, is the rating of the objects in S that preserves the scores specified by \mathbf{a} to the objects in S (similarly, the objects in S that were not evaluated by \mathbf{a} , will remain un-evaluated in $\mathbf{a}|_S$).

The following three definitions are natural extensions for incomplete ratings of the corresponding definitions given for complete ratings by Cook & Kress¹⁵:

Definition 2. Rating \mathbf{a} is said to be adjacent to rating \mathbf{b} if $|(a_i - a_j) - (b_i - b_j)| \leq 1$ for every pair of objects i and j in the set $\mathcal{A} \cap \mathcal{B}$. That is, if for every pair of objects their score difference in rating \mathbf{a} is either the same as in rating \mathbf{b} or differs by exactly one unit.

Definition 3. Rating \mathbf{a} is said to be adjacent of degree k to rating \mathbf{b} if \mathbf{a} is adjacent to \mathbf{b} and

$$|(i, j) \text{ such that } |(a_i - a_j) - (b_i - b_j)| = 1, i < j| = k.$$

That is, if the number of object pairs in the set $\mathcal{A} \cap \mathcal{B}$ for which their score difference differs by one unit is k .

Definition 4. Rating \mathbf{b} is between ratings \mathbf{a} and \mathbf{c} if, for every pair of objects i and j in the set $\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}$, either $a_i - a_j \leq b_i - b_j \leq c_i - c_j$ or $a_i - a_j \geq b_i - b_j \geq c_i - c_j$.

Definition 5. Ratings \mathbf{a} and \mathbf{b} are opposite ratings on $\mathcal{A} \cap \mathcal{B}$, if (1) \mathbf{a} rates $|\mathcal{A} \cap \mathcal{B}|/2$ objects ($\lfloor |\mathcal{A} \cap \mathcal{B}|/2 \rfloor$ if $|\mathcal{A} \cap \mathcal{B}|$ is odd) with the highest possible score, while \mathbf{b} rates those objects with the lowest possible score, and (2) \mathbf{a} rates the remaining objects with the lowest possible score, while \mathbf{b} rates those remaining objects with the highest possible score. (Intuitively, two opposite ratings are ratings in total disagreement when considering only the objects evaluated by both ratings).

Axioms. The objective is to design a distance such that, when used within problem (1), the obtained consensus rating minimizes the disagreement of the judges (uncalibrated MSUs in our context). Following is a set of axioms that a distance metric between *incomplete* ratings must satisfy so that our objective is achieved. Remark: when designing these axioms, we have in mind that, (i) given any MSU, the difference between any pair of its measurements is significantly more reliable than the absolute value of its measurements themselves; and (ii) an MSU providing a large amount of measurements for a particular location is not necessarily more reliable/accurate than an MSU providing a comparatively smaller amount of measurements.

Axiom 1 (Relevance) $d(\mathbf{a}, \mathbf{b}) = d(\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B})}, \mathbf{b}|_{(\mathcal{A} \cap \mathcal{B})})$

Axiom 2 (Nonnegativity) $d(\mathbf{a}, \mathbf{b}) \geq 0$

Axiom 3 (Commutativity) $d(\mathbf{a}, \mathbf{b}) = d(\mathbf{b}, \mathbf{a})$

Axiom 4 (Incomplete Ratings Triangular Inequality) $d(\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}, \mathbf{b}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}) + d(\mathbf{b}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}, \mathbf{c}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}) \geq d(\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}, \mathbf{c}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})})$, and equality holds if and only if $\mathbf{b}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}$ is between $\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}$ and $\mathbf{c}|_{(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})}$

Axiom 5 (Proportionality) The distance between any two adjacent ratings is proportional to the degree of adjacency

Axiom 6 (Normalization) $d(\mathbf{a}, \mathbf{b}) \leq 1$; and $d(\mathbf{a}, \mathbf{b}) = 1$ if and only if $\mathbf{b}|_{(\mathcal{A} \cap \mathcal{B})}$ and $\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B})}$ are opposite ratings

It is important to note that Axioms 2 to 4 for incomplete ratings are natural extensions of Cook & Kress' non-negativity, commutativity and triangular inequality axioms for complete ratings. Indeed, these two sets of axioms are identical when restricted to complete ratings. Similarly, Axioms 5 and 6 are a natural extension of Cook & Kress' proportionality axiom; the only minor difference is that Cook & Kress' axiom fixes the proportionality constant to '1'; while our normalization Axiom 6 (as shown later), sets the proportionality constant to the reciprocal of $2 \cdot R \cdot \left\lfloor \frac{|\mathcal{A} \cap \mathcal{B}|}{2} \right\rfloor \cdot \left\lfloor \frac{|\mathcal{A} \cap \mathcal{B}|}{2} \right\rfloor$. This minor difference is critical in the context of aggregating incomplete ratings. Specifically, normalization guarantees that when solving problem (1) all of the incomplete ratings are given the same importance regardless of the number of objects that each evaluated—this is critical since larger amounts of data/measurements does not necessarily mean higher accuracy.

Normalized Projected Cook-Kress Distance. Having this set of axioms, we now define the Normalized Projected Cook-Kress (d_{NPK}) distance. To do this we use the Cook & Kress distance (d_{CK}) for complete ratings. From our above discussion, it follows that d_{CK} satisfies the nonnegativity, commutativity, triangular inequality, and proportionality axioms (i.e. Axioms 2 to 5) when focusing on complete ratings¹⁵. Cook & Kress' distance is

$$d_{CK}(\mathbf{a}, \mathbf{b}) = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n |(a_i - a_j) - (b_i - b_j)|. \tag{2}$$

The Normalized Projected Cook-Kress (NPCK) distance is given by:

$$d_{NPCK}(\mathbf{a}, \mathbf{b}) = \begin{cases} \frac{d_{CK}(\mathbf{a}|_{\mathcal{A} \cap \mathcal{B}}, \mathbf{b}|_{\mathcal{A} \cap \mathcal{B}})}{2 \cdot R \cdot \left\lfloor \frac{|\mathcal{A} \cap \mathcal{B}|}{2} \right\rfloor \cdot \left\lfloor \frac{|\mathcal{A} \cap \mathcal{B}|}{2} \right\rfloor} & \text{if } |\mathcal{A} \cap \mathcal{B}| \geq 1, \\ 0 & \text{if } |\mathcal{A} \cap \mathcal{B}| = 0. \end{cases} \tag{3}$$

The following sequence of results will allow us to prove that the d_{NPCK} distance is the unique distance satisfying Axioms 1 to 6 simultaneously.

Lemma 6. Given a set V of n objects and a rating interval $[\ell, u]$ (evaluation range is $R = u - \ell$), the maximum Cook & Kress distance, $d_{CK}(\bullet, \bullet)$, between any two (complete) ratings is $2 \cdot R \cdot \left\lfloor \frac{n}{2} \right\rfloor \cdot \left\lfloor \frac{n}{2} \right\rfloor$. Moreover, this maximum distance is attained by any two opposite ratings.

Proof Sketch. The lemma can be restated as follows: “Any pair of opposite (complete) ratings is a global maximizer of problem (4) with an optimal objective value of $2 \cdot R \cdot \left\lfloor \frac{n}{2} \right\rfloor \cdot \left\lfloor \frac{n}{2} \right\rfloor$ ”

$$\max_{\mathbf{a}, \mathbf{b}} d_{CK}(\mathbf{a}, \mathbf{b}) \tag{4}$$

Using Eq. (2) and since \mathbf{a} and \mathbf{b} must be complete ratings within the interval $[\ell, u]$, problem (4) can be re-written as

$$\begin{aligned} & \text{Maximize}_{\mathbf{a}, \mathbf{b}} \quad \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n |(a_i - a_j) - (b_i - b_j)| \\ & \text{s.t.} \quad \ell \leq a_i \leq u \quad \text{for } i = 1..n \\ & \quad \quad \ell \leq b_i \leq u \quad \text{for } i = 1..n. \end{aligned}$$

It can be shown that, when \mathbf{a} and \mathbf{b} are assigned values so that they are opposite ratings, (i) one obtains a local maximum of the problem (all feasible directions are non-increasing), and (ii) the objective value of such assignment is equal to $2 \cdot R \cdot \left\lfloor \frac{n}{2} \right\rfloor \cdot \left\lfloor \frac{n}{2} \right\rfloor$. Since the above optimization problem is convex, every local maximum is a global maximum and thus the result follows. \square

Lemma 7. The NPCK distance satisfies Axioms 1 to 6.

Proof. The NPCK distance, d_{NPCK} , satisfies Axiom 1 directly from its definition, Eq. (3). The fact that d_{NPCK} satisfies Axioms 2, 3 and 5 follows from Eq. (3) and d_{CK} satisfying Cook & Kress’ non-negativity, commutativity and proportionality axioms.

d_{NPCK} satisfying Axiom 4 follows from Eq. (3); the fact that $\mathbf{a}|_{\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}}$, $\mathbf{b}|_{\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}}$, and $\mathbf{c}|_{\mathcal{A} \cap \mathcal{B} \cap \mathcal{C}}$ are complete ratings on the set $(\mathcal{A} \cap \mathcal{B} \cap \mathcal{C})$; and d_{CK} satisfying Cook & Kress’ triangular inequality axiom.

Finally, from Lemma 6 and Eq. (3), d_{NPCK} satisfies Axiom 6. \square

Corollary 8. Axioms 1 to 6 are consistent.

Theorem 9. The d_{NPCK} distance is the unique distance satisfying Axioms 1 to 6 simultaneously.

Proof. The fact that d_{NPCK} satisfies axioms 1 to 6 was established in Lemma 7. Thus, we only need to show that no other distance satisfies axioms 1 to 6 simultaneously. Let d be a generic distance satisfying axioms 1 to 6. We prove the theorem by showing that, for any two ratings \mathbf{a} and \mathbf{b} , $d(\mathbf{a}, \mathbf{b}) = d_{NPCK}(\mathbf{a}, \mathbf{b})$. We divide our analysis in the following two cases:

Case 1: Both \mathbf{a} and \mathbf{b} are complete ratings (i.e., $\mathcal{A} \cap \mathcal{B} = V$).

For complete ratings, Axiom 1 is a tautology and, as argued above, axioms 2 to 5 are identical to all of Cook & Kress’ axioms except for the proportionality constant. Therefore, for complete ratings, Axioms 2 to 5 uniquely determine d_{CK} except for a proportionality constant. Consequently, since $d(\mathbf{a}, \mathbf{b})$ satisfies Axioms 1 to 5 we conclude that, for complete ratings,

$$d(\mathbf{a}, \mathbf{b}) = \alpha \cdot d_{CK}(\mathbf{a}, \mathbf{b}) \tag{5}$$

for some constant α that may depend only on $|V|$ and R .

Also, from Eq. (3) and since $\mathcal{A} \cap \mathcal{B} = V$, we have that

$$d_{NPCK}(\mathbf{a}, \mathbf{b}) = \frac{1}{2 \cdot R \cdot \left\lfloor \frac{|V|}{2} \right\rfloor \cdot \left\lfloor \frac{|V|}{2} \right\rfloor} d_{CK}(\mathbf{a}, \mathbf{b}). \tag{6}$$

In view of eqs (5) and (6), in order to conclude that $d(\mathbf{a}, \mathbf{b}) = d_{NPCK}(\mathbf{a}, \mathbf{b})$ for complete ratings, we only need to prove that

$$\alpha = \frac{1}{4 \cdot R \cdot \left\lfloor \frac{|V|}{2} \right\rfloor \cdot \left\lfloor \frac{|V|}{2} \right\rfloor}.$$

This result follows since both $d(\mathbf{a}, \mathbf{b})$ and $d_{NPCK}(\mathbf{a}, \mathbf{b})$ attain their extreme values (zero and one) at lots of rating pairs. Specifically, given *any* two opposite ratings, say \mathbf{a}' and \mathbf{b}' , axiom 6 stipulates that $d(\mathbf{a}', \mathbf{b}') = d_{NPCK}(\mathbf{a}', \mathbf{b}') = 1$. Similarly, given *any* rating, say \mathbf{a}' , eqs (5) and (6) and the definition of d_{CK} (Eq. (2)) imply that $d(\mathbf{a}', \mathbf{a}') = d_{NPCK}(\mathbf{a}', \mathbf{a}') = 0$.

Case 2: At least one of \mathbf{a} or \mathbf{b} is an incomplete rating.

The following equalities show that $d(\mathbf{a}, \mathbf{b}) = d_{NPCK}(\mathbf{a}, \mathbf{b})$ for any two ratings \mathbf{a} and \mathbf{b} under the assumptions of this case:

$$\begin{aligned} d(\mathbf{a}, \mathbf{b}) &= d(\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B})}, \mathbf{b}|_{(\mathcal{A} \cap \mathcal{B})}) \\ &= d_{NPCK}(\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B})}, \mathbf{b}|_{(\mathcal{A} \cap \mathcal{B})}) \\ &= d_{NPCK}(\mathbf{a}, \mathbf{b}) \end{aligned}$$

The first and last equalities follow from Axiom 1, while the second equality follows from our analysis of case 1 and the fact that $\mathbf{a}|_{(\mathcal{A} \cap \mathcal{B})}$ and $\mathbf{b}|_{(\mathcal{A} \cap \mathcal{B})}$ are complete ratings over the set $\mathcal{A} \cap \mathcal{B}$. □

Finding the Consensus Rating. The NPCK distance generalizes the distance between complete ratings proposed by Cook & Kress¹⁵. Hochbaum and Levin³³ showed that this complete-rating aggregation problem is a special case of their own separation-deviation model, and thus efficiently solvable. Similarly, given all the MSUs' acquired time series, $\{\mathbf{a}^1, \mathbf{a}^2, \dots, \mathbf{a}^K\}$, the incomplete-rating aggregation problem (Eq. (1)) using the NPCK distance is a special case of the separation-deviation problem and can be reformulated as:

$$\min_{\mathbf{r}, \mathbf{z}} \sum_{k=1}^K \sum_{i \in \mathcal{A}^k} \sum_{j \in \mathcal{A}^k} \frac{|z_{ij} - (a_i^k - a_j^k)|}{4 \cdot R \cdot \left\lfloor \frac{|\mathcal{A}^k|}{2} \right\rfloor \cdot \left\lfloor \frac{|\mathcal{A}^k|}{2} \right\rfloor} \tag{7}$$

$$\text{subject to } z_{ij} = r_i - r_j \text{ for } i = 1, \dots, n; j = 1, \dots, n \tag{8}$$

$$\ell \leq r_i \leq u \text{ for } i = 1, \dots, n \tag{9}$$

Problem (7) is a special case of the convex dual of the minimum cost network flow problem, and thus it can be solved in $O\left(n^3 \log\left(\frac{u-\ell}{\varepsilon}\right)\right)$ time³⁴, where n is the number of objects (in our context, number of time points when the measurements were taken), and ε is the desired accuracy.

Finally, recall that our aim when designing the distance function was to extract as much information as possible from the reliable measurement differences (in contrast to the unreliable absolute measurements). Indeed in Problem (7) the bias of each MSU is completely ignored; specifically, Problem (7) can be interpreted as finding the vector \mathbf{r} , whose pairwise differences, \mathbf{z} , are as close as possible to the given MSU's pairwise measurement differences $(a_i^k - a_j^k)$. This is precisely what we aimed for because the MSU's are uncalibrated and thus the MSU's pairwise measurement differences are significantly more reliable than the absolute values of the measurements. Now, note that given any optimal solution to Problem (7), say \mathbf{r}^* , the vector $\mathbf{r}' = \mathbf{r}^* + c$, for any given scalar constant c , has exactly the same pairwise differences, \mathbf{z}^* , and thus is also an optimal solution to Problem (7). As such, the last step of our MSU aggregation method, is to calibrate our aggregated/consensus "measurements", \mathbf{r}^* . In particular, we need to find the best calibration constant, c , to calibrate our consensus measurement vector \mathbf{r}^* (keeping fixed all of its pairwise differences, \mathbf{z}^*). This is achieved by solving the problem

$$\min_c \sum_{k=1}^K \sum_{i \in \mathcal{A}^k} |(r_i^* + c) - a_i^k| \tag{10}$$

$$\ell \leq r_i^* \leq u \text{ for } i = 1, \dots, n \tag{11}$$

We note that Problem (10) is efficiently solvable by a simple binary search procedure over c . Indeed it can be shown that the objective functions of problems (7) and (10) can be combined in a single objective function by adding them and multiplying the objective function of Problem (7) by a large constant so that it is lexicographically more important than that of Problem (10). Moreover, the resulting combined optimization problem would still be a special case of the separation deviation problem and thus efficiently solvable.

Campaign	Measured Pollutants	Sensor Platform	Sensors #	AQM	Dates
1	O ₃	MO	407, 414, 415, 416, 418, 420, 422, 423, 424	Igud	27/12/2012–04/04/2013
2	O ₃	MO	414, 420, 422, 619, 624, 625, 626	Tel-Hai	16/12/2013–19/02/2014
3	O ₃ , NO ₂	MO	414, 422, 624, 626	Tel-Hai	29/04/2014–28/05/2014
4	O ₃ , NO ₂	MO	418, 620, 621	Tel-Hai	09/06/2014–10/07/2014
5	NO, CO	EC	135, 136, 468	Atzmaut	03/02/2015–26/02/2015
6	NO, CO	EC	220, 465, 471	Atzmaut	03/02/2015–26/02/2015

Table 1. Field Campaigns.

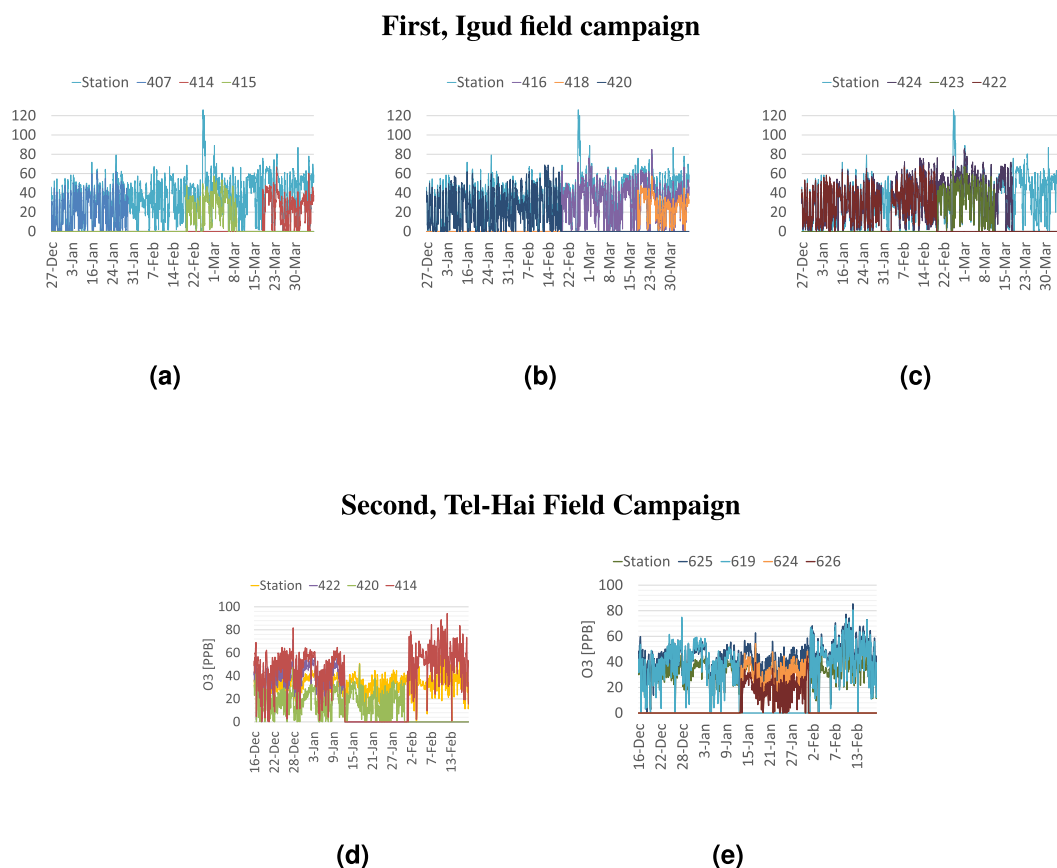


Figure 1. O₃ time series acquired by an AQM and collocated MSUs. First, Igud field campaign, Figures (a–c), (a) 407, 414, 415 and AQM; (b) 416, 418, 420 and AQM; (c) 422, 423, 424 and AQM. Second, Tel-Hai Field Campaign, Figures (d,e). (d) 414, 420, 422 and AQM; (e) 619, 624, 625, 626 and AQM.

The end result after the integration process (i.e., solving problems (7) and (10)) is a set of aggregated and calibrated measurements $r_i^* + c^*$ for $i = 1, \dots, n$ from the WDESN.

Results

Experimental Setup. The suggested methodology was applied on six longitudinal air-quality field campaigns. All campaigns were held in the city of Haifa, located at the eastern Mediterranean sea at the north of Israel (~595,000 residents in Haifa and its satellite cities). The city is built on and around the Carmel Ridge, from the shore at the foot of the ridge to its top at ~400 [m] above sea level (a.s.l.). Ambient levels of nitrogen dioxide (NO₂) and ozone (O₃) were acquired by metal-oxide (MO) sensors (Elm, by Perkin-Elmer, USA) and nitrogen oxide (NO) and carbon monoxide (CO) by electrochemical (EC) based MSUs (model AQMesh, produced by GeoTech, UK). EC and MO are currently the only available technologies for measuring gaseous pollutants with MSUs^{5–9}. In all campaigns an array of MSUs was placed next to a standard AQM station. Three different AQM stations were involved in the research reporting pollutants ambient levels. “Igud” AQM station, which is located at the Haifa District Municipalities Association for the Environment (HDMAE) headquarters, at the heart of the Haifa bay heavy industrial area (~30 [m] a.s.l., Lat:32.789379, Lon:35.040452); “Tel-Hai” HDMAE AQM station, located

Campaign	Pollutant	Count	R^2			Std CI		
			L_1	L_2	$NPCK$	L_1	L_2	$NPCK$
1	O_3	3,439	0.7039	0.7209	0.8556	1.2850	1.2883	0.8410
2	O_3	1,038	0.7071	0.7922	0.7922	0.9323	0.9628	0.9357
3	NO_2	1,298	$2 \cdot 10^{-5}$	$2 \cdot 10^{-5}$	$2 \cdot 10^{-5}$	2.6715	2.6896	2.5783
	O_3	1,299	0.8730	0.86693	0.88768	3.192	3.0301	3.1271
4	NO_2	1,420	0.0025	0.0002	0.0032	2.999	13.014	2.9756
	O_3	1,431	0.7931	0.8428	0.8402	1.4507	1.6261	1.6478
5	CO	2,001	0.1256	0.0851	0.1205	3.7435	3.0432	3.7111
	NO	1,904	0.98088	0.9814	0.9824	5.0644	5.3579	5.0349
6	CO	2,001	0.1446	0.1279	0.2113	3.8027	3.9326	3.7678
	CO	1,766	0.9462	0.9428	0.9421	7.9352	7.9078	7.8121

Table 2. R^2 and Std-CI of the consensus time series with respect to the AQM measurements.

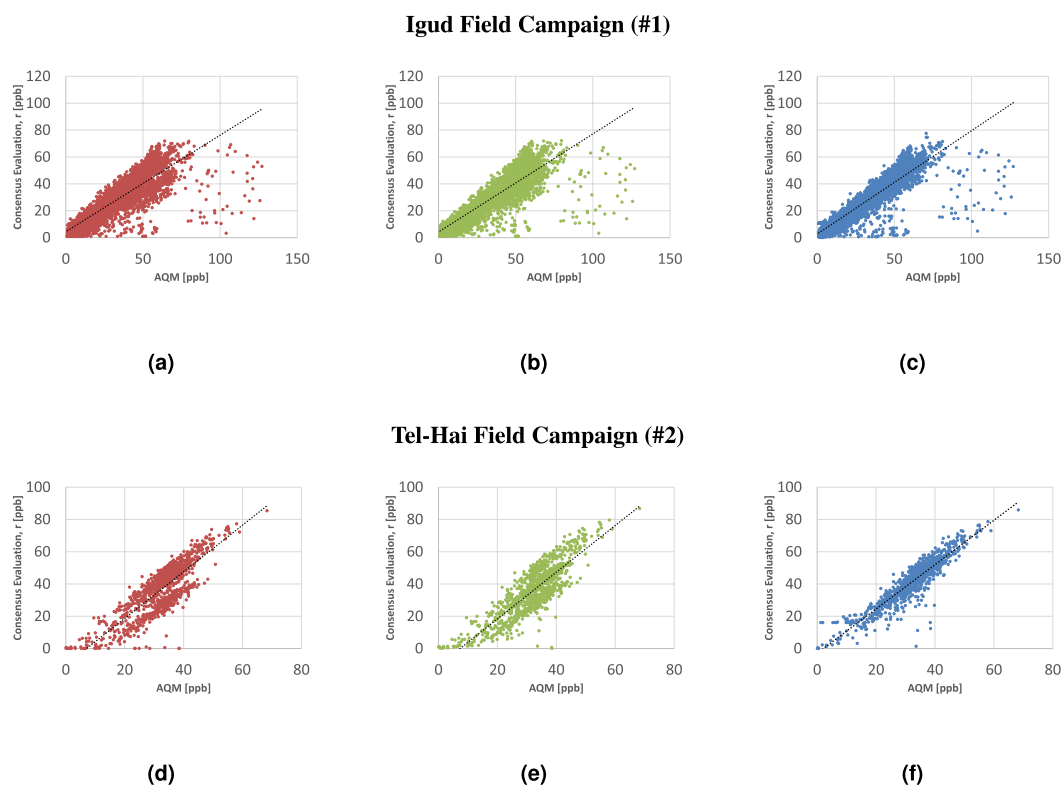


Figure 2. Consensus time series Vs. AQM measurements (Metal-Oxide MSUs). Igud Field Campaign, Figures (a–c), Tel-Hai Field Campaign, Figures (d–f). L_1 (a,c); L_2 (b,d); and $NPCK$ (c,f).

at a Haifa residential neighbourhood on the Carmel Ridge (~200 [m] a.s.l., Lat:32.787293, Lon:35.021072); and “Atzmaut” AQM station (~8 [m] a.s.l., Lat:32.81644, Lon:35.00167), which is operated by the Israeli ministry of environmental protection and is dedicated to measure transportation related pollutants, i.e., NO and CO . The campaigns details are reported in Table 1.

To illustrate the acquired data, Fig. 1 depicts ozone time series acquired in Igud and Tel-Hai in the first and second campaigns. The AQMs’ complete data is plotted alongside the incomplete time series acquired by 9 (ID numbers 407, 414, 415, 416, 418, 420, 422, 423 & 424) and 7 (numbers 414, 420, 422, 619, 624, 625 & 626) collocated MSUs at the Igud and Tel-Hai stations respectively. Note that, at different dates, MSUs 414, 420, and 422 switched locations between Igud and Tel-Hai—respectively of the common MSUs, the data for each campaign was aggregated independently of that of the other campaign.

Consensus time series Evaluation. For each campaign and each pollutant three consensus time series were obtained by solving problem (1) with the three different distance metrics— L_1 , L_2 and d_{NPCK} . Table 2 presents the coefficients of determination, R^2 , between each of the three consensus and the AQM measurements. This table also details the confidence intervals (CI) of the standard deviations of consensus measurements grouped by the AQM measurements. Thus, how disperse are the consensus measurements for each time the AQM reported

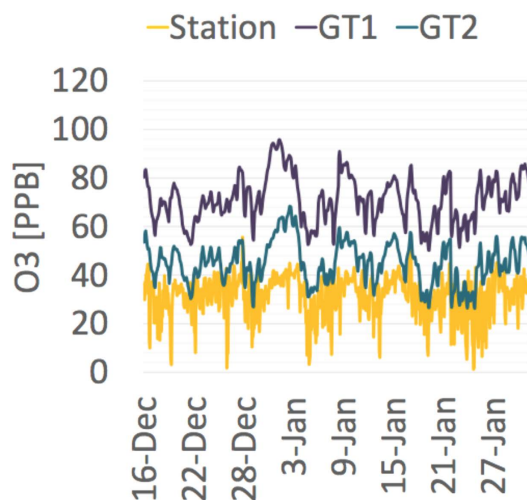


Figure 3. GT135, GT136 and the AQM O_3 measurements throughout campaign 2 (Dec 16th, 2013 and Feb. 19th, 2014).

	414	420	422	619	624	625	626	135	136
Correlation	0.8710	0.8869	0.9257	0.9147	0.9175	0.9290	0.8903	0.4779	0.5496
MSE	0.6827	0.5434	0.4180	0.3656	0.3926	0.3243	0.7325	1.5021	0.5862

Table 3. Second Field Campaign MSUs' characteristics.

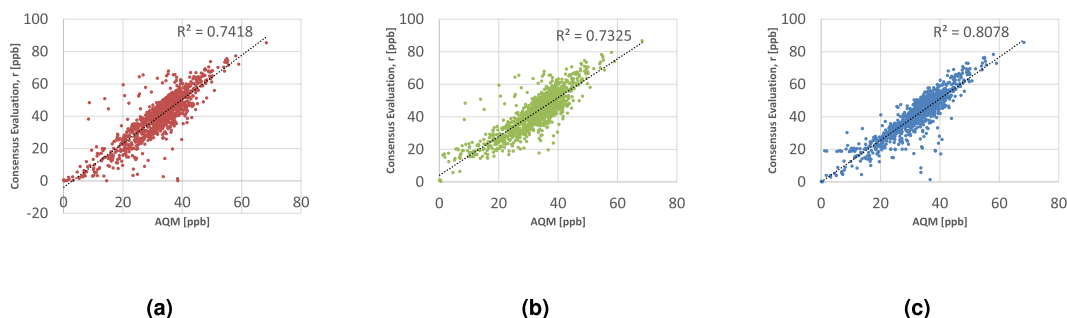


Figure 4. Consensus time series Vs. AQM measurements (Electro-Chemical and Metal-Oxide). (a) L_1 , (b) L_2 and (c) NPCK.

the same value. The CI is calculated as the standard error multiplied by the critical two-tailed value of z for $\alpha = 0.05$ ³⁵. Note that the consensus time series obtained when using the NPCK distance when solving problem (1) present, most of the times, higher R^2 and lower CI values as compared to those obtained when using either the L_1 and L_2 distances. Specifically, the NPCK has shown higher R^2 and lower CI for both NO_2 campaigns and for three, out of the four O_3 campaigns. In addition, when the NPCK does not present the best results, it is not far behind presenting almost the same score. Therefore, we conclude that the consensus measurements/time series obtained when using d_{NPCK} , is the best fit for estimating the real AQM measurements/time series based on the consensus of all MSUs.

To illustrate the notions above visually, Fig. 2 plots three consensus time series against the AQM time series obtained for first two campaigns. Each point in the graphs corresponds to a specific time, its x-coordinate is the “measurement” of the consensus time series at that time and its y coordinate is the measurement taken by the AQM at that time. For the Igud campaign, comparing Fig. 2a,b with Fig. 2c, it is evident that the linear relation between the AQM measurements and the consensus time series is stronger for the NPCK as the measurements spread around the linear line is smaller; this exact same result holds for the Tel-Hai campaign (as evident when comparing Fig. 2d,e with Fig. 2f). Supporting the quantitative analysis above.

Robustness Analysis. The robustness of the suggested scheme is presented next. For this purpose, two time series (#135 and #136), acquired in the second campaign in conjunction with the data of Fig. 1d,e were added into the aggregation process. These two time series were acquired using EC ozone MSUs (AQMesh of GeoTech, UK). While EC MSUs have been previously used for ozone measurements, this technology suffers heavily from

interferences^{5,6} and thus, produces measurements that are less accurate than those obtained by using metal-oxide ozone MSUs (which was the type of MSUs used to obtain the data in Fig. 1d,e). Figure 3 presents the Tel-Hai AQM station's complete time series alongside the incomplete time series measurements acquired from the GT135 and GT136 MSUs. Table 3 depicts the correlation coefficient and the Mean Squared Error (MSE) between the AQM measurements and all sensors that took part in this campaign (see Table 1) and the two added time series (Fig. 3). Note that the last two MSUs added to the process, GT135 and GT136, have a significantly lower correlation and higher MSE than the rest of the MSUs.

Figure 4 is analogous to Fig. 2 and plots, against Tel-Hai's AQM, the consensus "measurements" aggregating both metal-oxide and electro-chemical MSUs when using the L_1 , L_2 and d_{NPCK} metrics within problem (1). Figure 4 also presents the coefficients of determination, R^2 , between the three consensus and the AQM measurements. Similarly to the results when using only metal-oxide MSUs, the correlation coefficient of consensus measurements obtained with d_{NPCK} is by far the largest one. Therefore, we again conclude that the consensus measurements/time series obtained when using d_{NPCK} is the best fit for estimating the real AQM measurements/time series based on the consensus of all MSUs.

Discussion

This paper introduces a scheme for the aggregation of incomplete ratings into a group consensus decision making. The core of the method is the herein-developed axiomatic Normalized Projected Cook-Kress (NPCK) distance. The NPCK distance is derived from a set of axioms any distance between *incomplete* ratings should fulfil so the consensus rating aggregates the given ratings. The consensus rating is the rating that minimizes the sum of all distances from the different ratings. The NPCK approach is an extension of Cook and Kress complete rating aggregation problem, making it suitable to many new applications. An efficient algorithm for finding the consensus rating is also provided.

Wireless Distributed Environmental Sensory Networks (WDESN) have become technically and economically feasible. However, WDESNs may consist of many sensors and thus, the calibration process is a major obstacle. The suggested NPCK distance presents a new, robust and efficient method for aggregating measurements acquired by an uncalibrated, inexpensive and error-prone WDESN, and producing accurate estimates of the observed environmental variable's true levels. Given a set of collocated Micro Sensing Units (MSUs), the NPCK incomplete ratings scheme is applied, where each measurement (defined by time and location) is considered as a referee evaluation. These time series can be incomplete as sensors might become faulty or shift locations. Based on a set of collocated measurements (in time and in space) a consensus measurement is derived using the NPCK scheme.

The methods have been applied to a wide set of pollutants measurements (i.e., ozone, nitrogen oxide, nitrogen dioxide and carbon monoxide) acquired by all available MSU technologies (metal oxide and electrochemical). When compared to a standard regulatory Air Quality Monitoring (AQM) station, the suggested methodology has shown markedly more accurate results than the common practice and the state-of-the-art, without requiring the Micro Sensing Units (MSUs), constituting the WDESN, to be calibrated, rendering the network to be self-calibrated. To achieve this, some assumptions on the error behaviour are made (i.e., additive, zero mean error). While these assumptions are commonly accepted, we have also presented a simple logarithmic data re-scaling technique which enables the method to handle multiplicative errors. Therefore, generalising the suggested scheme even further.

This research has addressed the challenging problem of data aggregation, where only measurements of a single pollutant are aggregated. The interplay between gases in the atmosphere^{36–38}, for some gases (e.g., NO_2 and O_3) is known and may allow for the aggregation of data acquired from an heterogeneous set of sensors. Finally, the availability of the code, with the accurate results, present a great potential for making the NPCK the tool of choice for aggregating measurements acquired by uncalibrated WDESNs.

References

- Liu, Y. & Zhou, G. Key technologies and applications of internet of things. In *Intelligent Computation Technology and Automation (ICICTA), 2012 Fifth International Conference on 197–200* (2012) doi: 10.1109/ICICTA.2012.56.
- Chi, Q., Yan, H., Zhang, C., Pang, Z. & Xu, L. D. A reconfigurable smart sensor interface for industrial wsn in iot environment. *IEEE T Industrial Informatics* **10**, 1417–1425 (2014).
- Kelly, S., Suryadevara, N. & Mukhopadhyay, S. Towards the implementation of iot for environmental condition monitoring in homes. *IEEE Sens J* **13**, 3846–3853 (2013).
- Lazarescu, M. Design of a wsn platform for long-term environmental monitoring for iot applications. *IEEE J Emerging and Selected Topics in Circuits and Systems* **3**, 45–54 (2013).
- Mead, M. *et al.* The use of electrochemical sensors for monitoring urban air quality in low-cost, high-density networks. *Atmos. Environ.* **70**, 186–203 (2013).
- Williams, D. E. *et al.* Validation of low-cost ozone measurement instruments suitable for use in an air-quality monitoring network. *Meas Sci Technol* **24**, 065803 (2013).
- Moltchanov, S. *et al.* On the feasibility of measuring urban air pollution by wireless distributed sensor networks. *Sci Total Environ* **502**, 537–547 (2014).
- Lerner, U. *et al.* The effect of ego-motion on environmental monitoring. *Sci Total Environ* **533**, 8–16 (2015).
- Zhuykov, S. Solid-state sensors monitoring parameters of water quality for the next generation of wireless sensor networks. *Sensor Actuator B* **161**, 1–20 (2012).
- Dinh, T. L. *et al.* Design and deployment of a remote robust sensor network: Experiences from an outdoor water quality monitoring network. In *Local Computer Networks, 2007. LCN 2007. 32nd IEEE Conference on* vol. 1, 799–806 (2007) doi: 10.1109/LCN.2007.39.
- Chinrungrueng, J., Sunantachaikul, U. & Triamlumlert, S. Smart parking: An application of optical wireless sensor network. In *Applications and the Internet Workshops, 2007. SAINT Workshops 2007. International Symposium on 66–66* (2007) doi: 10.1109/SAINT-W.2007.98.
- Mitton, N., Papavassiliou, S., Puliafito, A. & Trivedi, K. Combining cloud and sensors in a smart city environment. *EURASIP J Wirel Comm* **2012**, 1–10 (2012) doi: 10.1186/1687-1499-2012-247.

13. Luoto, M., Kuussaari, M. & Toivonen, T. Modelling butterfly distribution based on remote sensing data. *J Biogeography* **29**, 1027–1037 (2002).
14. Osborne, P., Alonso, J. & Bryant, R. Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *J Appl Ecol* **38**, 458–471 (2001).
15. Cook, W. D. & Kress, M. Ordinal ranking with intensity of preference. *Manage Sci* **31**, 26–32 (1985).
16. Kemeny, J. G. & Snell, L. J. Preference ranking: An axiomatic approach. In *Mathematical Models in Social Science* 9–23 (Ginn, Boston, 1962).
17. Cook, W. D., Kress, M. & Seiford, L. M. Information and preference in partial orders: a bimatrix representation. *Psychometrika* **51**, 197–207 (1986).
18. Wallis, W. Voting systems. In *Mathematics in the Real World* 169–185 (Springer New York, 2013).
19. Saaty, T. L. & Peniwati, K. *Group Decision Making: Drawing Out and Reconciling Differences* (RWS Publications, Pittsburg, PA, 15213, USA, 2013).
20. Hochbaum, D. & Moreno-Centeno, E. Rating customers according to their promptness to adopt new products. *Oper Res* **59**, 1171–1183 (2011).
21. Ebrahimnejad, S., Mousavi, S., Tavakkoli-Moghaddam, R., Hashemi, H. & Vahdani, B. A novel two-phase group decision making approach for construction project selection in a fuzzy environment. *Appl Math Model* **36**, 4197–4217 (2012).
22. Hoek, G. *et al.* Long-term air pollution exposure and cardio-respiratory mortality: a review. *Environ Health* **12**, 43 (2013).
23. Kurt Straif, J. S. & Cohen, A. Air Pollution and cancer. Tech. Rep., International Agency for Research on Cancer (IARC) (2013).
24. Crouse, D., Goldberg, M. & Ross, N. A prediction-based approach to modelling temporal and spatial variability of traffic-related air pollution in montreal, canada. *Atmos. Environ.* **43**, 5075–5084 (2009).
25. Pope, C. A. *et al.* Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *J AM Med. Assoc* **287**, 1132–1141 (2002).
26. Rao, S. *et al.* Environmental modeling and methods for estimation of the global health impacts of air pollution. *Environ Model Assess* **17**, 613–622 (2012).
27. Lee, D. Environmental gas sensors. *IEEE Sens J* **1**, 214–224 (2001).
28. Becker, T. *et al.* Air pollution monitoring using tin-oxide-based micro-reactor system. *Sensor Actuator B* **69**, 108–119 (2000).
29. Piedrahita, R. *et al.* The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurements Techniques Discussions* **7**, 2425–2457 (2014).
30. Rao, S. & Zurbenko, I. Detecting and tracking changes in ozone air quality. *Air Waste* **44**, 1089–1092 (1994).
31. Schwartz, J. & Marcus, A. Mortality and air pollution in london: a time series analysis. *AM J Epidemiol* **131**, 185–194 (1990).
32. Varotsos, C. & Ondov, J. M. E. Scaling properties of air pollution in athens, greece and baltimore, maryland. *Atmos. Environ.* **39**, 4041–4047 (2005).
33. Hochbaum, D. S. & Levin, A. Methodologies and algorithms for group-rankings decision. *Manage Sci* **52**, 1394–1408 (2006).
34. Ahuja, R. K., Hochbaum, D. S. & Orlin, J. B. Solving the convex cost integer dual network flow problem. *Manage Sci* **49**, 950–964 (2003).
35. Howell, D. C. *Statistical Methods for Psychology* (Wadsworth, Cengage Learning, 2013) 8 edn.
36. Eeftens, M. *et al.* Development of land use regression models for pm 2.5, pm 2.5 absorbance, pm 10 and pmcoarse in 20 european study areas; results of the escape project. *Environ Sci Technol* **46**, 11195–11205 (2012).
37. Eeftens, M. *et al.* Spatial variation of pm 2.5, pm 10, pm 2.5 absorbance and pm coarse concentrations between and within 20 european study areas and the relationship with no 2–results of the escape project. *Atmos. Environ.* **62**, 303–317 (2012).
38. Crutzen, P. J. The influence of nitrogen oxides on the atmospheric ozone content. *Q J Roy Meteor Soc* **96**, 320–325 (1970).

Acknowledgements

This work was partially supported by the 7th European Framework Program (FP7) ENV.2012.6.5-1, grant agreement no. 308524 (CITI-SENSE), the Technion Center of Excellence in Exposure Science and Environmental Health (TCEEH) budget #2018567, the Environmental Health Foundation (EHF) and the New York Metropolitan Research Fund budget #2020306.

Author Contributions

B.F. came up and designed the study. E.M.-C. formulated the theoretical framework. Both B.F. and E.M.-C. wrote the software for analyzing the data and did the analysis, wrote the paper and reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Fishbain, B. and Moreno-Centeno, E. Self Calibrated Wireless Distributed Environmental Sensory Networks. *Sci. Rep.* **6**, 24382; doi: 10.1038/srep24382 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>