Environmental signal processing: new trends and applications

Matthieu Puigt, Gilles Delmaire, Gilles Roussel

Environmental signal processing: new trends and applications

Matthieu Puigt, Gilles Delmaire, and Gilles Roussel

Univ. Littoral Côte d'Opale, LISIC – EA 4491, Calais, France, F-62228

Abstract. In the last years, environmental monitoring was shown to be a major application field of modern signal processing and machine learning techniques. In particular, it provides some interesting problems for which specific signal methods were proposed. In this session, we aim to review some recent advances in this topic. We propose a taxonomy of the major trends in environmental surveillance according to the characteristics of the sensing devices, i.e., (i) for a unique sensor or an array of sensors (e.g., bio-sensor, chemical sensor arrays), (ii) for remote observation, and (iii) using large-scale sensor networks.

1 Introduction

Environmental pollution is a major problem faced by the world. Indeed, studies show the effects of pollution levels—e.g., in air [1] or water [2]—on health, and more generally on biodiversity [3]. Human activity is also shown to be the main origin of climate change [4], which particularly affects billions of lifes.

As a consequence, monitoring environmental phenomena has been investigated for several decades. Environmental surveillance not only consists of monitoring air, water, or soil pollutant levels or surface temperature, but also of observing noise and light pollution in urban areas, which have consequences on health and biodiversity [5, 6, 7]. Such a surveillance has been possible using sensors, and with their increasing number emerged specific signal processing and/or machine learning techniques. In this paper, we aim to review some of these techniques—and some associated open challenges—as well as to briefly introduce the papers of the special session on environmental signal processing of the 2017 ESANN conference.

The remainder of the paper reads as follows. In Section 2, we focus on approaches for single sensors and sensor arrays met in, e.g., source apportionment and bio-chemical sensors. Section 3 shows how remote sensing can be used in environmental monitoring. Section 4 covers some approaches for large-scale sensor networks. We finally conclude in Section 5.

2 Chemometrics for single sensor and sensor arrays

Processing single environmental sensors or sensor arrays arised as a major topic in environmental monitoring, known under the name of chemometrics. Indeed, while sensors used in source apportionment provide fine particulate analyzed by chemists which must then be numerically processed to enhance their interpretability, there has also been an increasing need to process data from

chemical/bio-sensors which might have nonlinear responses. We review hereafter some of these aspects.

2.1 Source apportionment

Source apportionment aims at estimating the impacts of emissions from different sources of pollutants based on ambient data registered at monitoring sites. It has been used for many applications, to monitor indoor [8] or outdoor air [9], sea [10], and soil pollution [11], for example. The awareness of the increasing pollution and the necessity to find answers to these issues has led to some growing interest and increasing number of scientific studies, as recently reviewed in [12].

Among the classical source apportionment techniques, the receptor model assumes that a chemical mass measure is a non-negative linear combination of the masses of source profiles, weighted by their associated contribution over time. This model relies on the law of conservation of mass and implicitly assumes that air suspended particles or sedimentary matter are not reacting along their path toward the sensor. The receptor model may then be written as a non-negative matrix factorization product, i.e.,

$$X \approx G \cdot F, \tag{1}$$

where X accounts for the $n \times m$ concentration matrix (in ng/m^3), G is the $n \times p$ contribution matrix gathering the loadings of all the sources over time and F is the $p \times m$ profile matrix, where a profile is a source signature involving all the chemical species proportions (in ng/ng). In some studies, e.g., [8], the above matrices X and F are defined regarding the considered particle-size distribution.

Principal Component Analysis (PCA) has been used in numerous receptor models for a long time [13]. However, the non-negativity property was lost and, as a consequence, other techniques—i.e., Positive (PMF) / Non-negative Matrix Factorization (NMF)—emerged as an alternative [14] which became the standard receptor model. However, the unicity of their solution is not guaranteed [15] and data might be corrupted by outliers. Lastly, the profiles in F sum to one. As a consequence, some (possibly-robust) informed NMF techniques—which restrict the space of admissible solutions—were recently proposed to that end, e.g., [9, 16, 17].

2.2 Chemical sensors

While linear chemometrics methods have been massively proposed during the last decades, an increasing number of problems require specific methods for processing chemical- or bio-sensors which may provide highly nonlinear readings. Such sensors are usually grouped into an array and aim to deliver the concentration of chemical species present in the air, rivers, or a given analyte. In some problems, the objective consists of estimating the concentrations of several chemical profiles, which may be achieved using some Ion-Selective Electrode arrays (ISE) or bio-sensors located at fixed positions. Usually, this step is achieved in a supervised way by learning neural networks—e.g., for electronic tongues

[18]—which is costly. As an alternative, some authors proposed unsupervised techniques to work in a plug and sense way. This turns out to be a nonlinear source separation problem, which is known to be a hard task [19, Ch. 14].

Nonlinear source separation is mainly solved by assuming a known nonlinear mixture model where some parameters must be estimated. In chemical sensor arrays, the Nernst equation provides a scaled logarithmic dependency between the output voltage v_i of the i-th sensor and the mixed ionic activity, i.e.,

$$v_i(n) = e_i + \frac{RT}{z_i F} \log \left(s_i(n) + \sum_{j \neq i} a_{ij} s_j(n)^{\frac{z_i}{z_j}} \right), \tag{2}$$

where $s_j(n)$ and z_j account for the ionic activity and the valence of the j-th species, respectively, R and F are the gas and Faraday constants, respectively, and T is the temperature. This model is difficult to deal with, because of the power term appearing inside the logarithm. In the case of equal valences, the authors in [20] describe this equation as a general post-nonlinear model and—assuming that the source signals are independent—propose a simplified inverting process which reads

$$y_k(n) = \sum_{i=1}^{P} w_{ki} \, \exp(f_i + h_i \, v_i(n)), \tag{3}$$

where the parameters w_{ki}, f_i, and h_i have to be tuned according to an independence criterion. To prevent the algorithm to be stuck into a local minimum, Bayesian approaches [21]—enabling to describe the complete distribution of the source signals—may also be used, to the price of a high computational load.

Despite the great number of studies in nonlinear unmixing for chemical sensing, there are many assumptions which govern the proposed solutions, which may be hard to fulfill in real case studies. Some global solutions in real conditions are still expected.

3 Environmental monitoring using hyperspectral imagery

Hyperspectral (HS) sensing technologies allowed many opportunities of new applications over the past two decades. Most HS sensors are embedded in airborne or spaceborne platforms. Spectral, spatial, and temporal resolution has reached an unprecedented level and enabled uncountable applications requiring fine identification of materials or estimation of physical parameters. HS imaging has been increasingly used for earth surveillance, e.g., for soil, sea, soil-water interfaces and also astrophysical observations [22]. Moreover, HSI can be used to monitor industrial discharges—e.g., atmospheric effluents, storage of solid residues, liquid residues and sludges—with might have an impact on environment and health [23]. Such an increasing number of applications need sophisticated signal processing algorithms because of the complexity to extract relevant information, the high dimensionality and the size of the hyperspectral data, the (linear or

nonlinear) spectral mixing, the measurement process. The framework used in these topics is rooted on signal and image processing, statistical inference, and machine learning fields.

3.1 Data and hypotheses

HS imaging (HSI) aims to acquire HS datacubes which consists of collected spectral vectors sorted as a tensor denoted $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_b}$. Each image of $n = n_1 \times n_2$ pixels represents the radiance in n_b spectral bands of the wavelength interval covered by the embedded sensor(s). Usually, the values in $\mathcal{X}$ smoothly vary along the spectral bands and the spatial pixels, and thus live in a low-dimensional manifold. Spectral information can then be denoised—using, e.g., PCA—while smooth spatial variations are well-suited for supervised classification [24]. When processing HS data, one must take into account some physical and optical considerations—e.g., the surface and ground-leaving reflectances, the effects of illumination, the angle of view, etc—and each interface between surface, atmosphere and sensor leads to a correction or a calibration process. Such considerations yield some signal processing techniques that we present below.

3.2 Hyperspectral processing techniques

As explained above, HSI gained a massive interest from the scientific community since two decades. Many data processing techniques were proposed to extract insightful information from hyperspectral data cubes, as recently reviewed in [24]. We briefly recall them below, with more emphasis on a few of them.

One main family of data processing techniques consists of enhancing the interpretability of the hyperspectral images by denoising (restoration) and fusing data cubes (spatial, spectral, spatial-spectral data fusion). In particular, several authors are interested in pansharpening [25] (or sharpening [26]), i.e., the fusion of a panchromatic (or a multispectral) image with an HS one, which provide a good spatial and spectral resolution, respectively. The pansharpened image then combines the good spatial and spectral accuracy of both images.

Classification is also an important topic met in HSI. It consists of assigning a unique label to each spatial pixel of an hyperspectral cube, corresponding to a given class. A high spatial resolution is a very desirable property for classification techniques—divided as supervised, unsupervised, and semi-supervised techniques [24]—so that most data only contain pure pixels. A pure pixel represents a predominant signature and is thus naturally linked to a class. However, a trade-off between spatial and spectral resolution must be chosen in HSI, as both cannot increase concurrently. As a consequence, hyperspectral unmixing (HU) techniques—a.k.a. blind source separation—were proposed to enhance classification performance.

In HU, the Linear Mixture Model (LMM) is widely used and most HU approaches unfold the above datacube $\mathcal{X}$ to derive a $n \times n_b$ matrix, denoted Y hereafter. Each column of Y contains the spectral information located at a given spatial pixel of $\mathcal{X}$ and the LMM methods provides the following mixture

model:

$$\min_{M,A} \|Y - MA\|_F, \quad \text{s.t. } A \geq 0,\ 1_p^T A = 1_n, \tag{4}$$

where A contains the non-negative fractional *abundance* vectors of the LMM—i.e., the mixing parameters which sum to 1—and M contains the spectral signatures of the materials—named *endmembers*—and 1_p stands for the $p \times 1$ vector of ones. Actually, HU has some similarities with the source apportionment problem introduced in Subsection 2.1, except that the rows of F sum to one in source apportionment while the columns of A sum to one in HU, thus yielding different NMF solutions. Main unmixing strategies assume the existence of pure pixels for each endmember—and are based on sparsity and/or geometric properties—or the absence of a few endmembers in some pixels, using volume-based approaches to perform HU in that case [24].

Several problems are currently investigated, i.e., the estimation of the intrinsic dimensionality [27]—i.e., the number of endmembers in Y—the endmember variability [28] for both LMM but also nonlinear mixing models which are increasingly studied [29].

Both the classification and HU can be used for HS target detection, where one aims to detect and identify man-made or natural clutters from the HSI variations [23].

4 Sensor networks for environmental monitoring

During the last decades, sensors have been massively spread around the world to provide fine grained yet accurate environmental analytics (e.g., temperature, air quality, etc) [30]. A taxonomy of such environmental sensor networks is provided in [31] and some insight about specific problems is provided. These problems have been studied in the literature since the work in [31] and we now summarize this work, together with some more recent open challenges.

4.1 Dealing with sensor inaccuracies

One of the first issues met with sensor networks is the possible lack of trust in the readings provided by the sensors. Indeed, the latter can be miscalibrated—resulting in biased scientific analyzes [32]—but their readings may also suffer from the presence of outliers, due to sensor faults for example, or to data loss. We discuss about both issues hereafter.

Sensor calibration is an important step in environmental monitoring which should be regularly performed. This is usually done in a laboratory where the sensor readings $y(t)$ are mapped to a controlled physical input $x(t)$, so that the the sensor calibration function $\mathcal{F}$ can be learned. Once learned and assuming that $\mathcal{F}$ is invertible, it is then possible to derive $x(t)$ from the sensor readings.

However, such a calibration task is not possible in many environmental problems, either because the sensors are not accessible like, e.g., in remote sensing [33], or because they are so numerous that it would be too time consuming to do so. As a consequence, some *blind calibration* or *self-calibration* techniques were

proposed. They aim to calibrate the sensors from their readings and might be possibly using other calibrated sensor readings for that purpose.

In order to perform blind calibration, it is first necessary to know the sensor model $\mathcal{F}$ in order to estimate its intrinsic parameters. Most authors—e.g., [34, 35, 36]—assume it is affine, i.e., $\mathcal{F}$ is defined by an offset and a gain. A few ones consider an extended nonlinear model which might be piecewise linear [37] or polynomial [38, 39]. As an alternative to both above models which are function of a single input $x(t)$, many sensors—e.g., miniaturized gas sensor—are driven by multiple—say M—inputs, e.g., the sensed gas concentration, the ambient temperature and the humidity [40]. In that case, the simplest model of $\mathcal{F}(.)$ is multilinear. Moreover, as the sensors must be regularly calibrated, this means that their response drifts with time. Some authors propose to incorporate this drift inside the sensor response model [40]. Once the sensor model is known, calibration can be performed, either simultaneously for the whole sensor network (macro-calibration [41]), or sequentially for each individual sensor (micro-calibration [42]). Many macro-calibration techniques are based on the statistical moments of the sensed phenomenon [33, 35], or assume the sensed phenomenon to be low-rank and project the readings on a previously learned subspace [34]. These approaches can be applied to mobile and fixed sensors. On the contrary, most micro-calibration methods were proposed for mobile sensors. While a few of them are based on moments [43], many assume the sensors to be in *rendezvous* [44]. Two sensors in rendezvous are in the same spatio-temporal vicinity and should thus acquire the same phenomenon. As a consequence, assuming that some sensors are calibrated, it is possible to perform calibration by regression of both the calibrated and the uncalibrated sensor readings in rendezvous [42]. As an alternative to this strategy, the authors in [36, 39] revisited macro-calibration by discretizing along space and time an observed area, using the rendezvous definition. Within this framework, self-calibration can be tackled as a structured matrix factorization problem with missing entries.

As explained above, sensor readings might also suffer from outliers [45] and missing data [46]. Interestingly, both problems can be tackled by some approaches—denoted robust principle component analysis or matrix/tensor completion, respectively—which use the same intrinsic assumption. If the sensor network is dense enough, it should oversample the observed phenomenon, i.e., if the observed data are arranged into a matrix or a tensor, the latter should be low-rank. Low-rank modeling attracted a lot of interest in the last decade and recent reviews on the topic can be found in, e.g., [47] (for sparse outlier removal) and [48] (for low-rank completion). Specific techniques for environmental applications were proposed, to deal with outliers in ozone measurements [49] and with missing information in seismic data [50] or in wireless sensor networks [46].

4.2 Physics-driven inverse problems

At one point, environmental monitoring aims to derive some accurate information about a sparsely sensed phenomenon. Several authors combine models with some signal processing tools. For example, the authors in [51, 52, 53] proposed

an approach to optimize the location of the sensors with respect to the sensed field. While one may assume the sensors to move in order to reduce the uncertainties of a physical model [51], the authors in [52, 53] propose to only use a subset of available sensors which minimize such uncertainties.

Similarly, some work was performed to estimate (i) the positions of the punctual sources of emission [54], (ii) the shape of non-punctual sources [55], and (iii) the source emission wavefronts [56] as an inverse problem.

Lastly, some authors propose to decompose a sensed phenomenon as linear combinations of interpretable modes, e.g., using non-negative decomposition of latent variables [57].

4.3 Privacy-preserving issues in partipatory science

Nowadays, environmental sensor networks are heterogeneous as some sensors can be fixed or mobile [31]. In the latter case, they can be carried by robots [58] or volunteers in participatory science [59]. This last situation has been particularly studied in the literature. Indeed, with the increase use of smartphones and their abilities to connect to sensing devices, it is quite easy to propose a large-scale experiment but also to reverse enginneer the geolocated data to find who are the volunteers. As a consequence, privacy-preserving techniques were proposed [60] and mainly consists of modifying (i) the sensed values and (ii) the locations of these readings, which both might affect the performance of the tasks discussed in this section. A few methods, e.g., [61, 62], take into consideration such an issue by processing data at the sensor scale while sharing as few information as possible with neighbour sensing devices but such a hot topic is still open.

5 Conclusion and discussion

To conclude, environmental monitoring is a major application field for signal processing and machine learning. Due to the impact of climate change and pollution levels, together with the intrinsic interest from the politics and the population, environmental data processing techniques must be carefully driven to keep a high level of trust. We reviewed some current problems regarding the sensor scale (single sensor/sensor array, remote sensing, large-scale sensor networks) but could not cover other very interesting problems, such as the multimodal fusion of data issued from all these sensor technologies.

The remainder of the *Environmental Signal Processing* session follows the same structure as this paper: chemical sensor processing [8, 18], remote sensing HU [28], and signal processing for/using sensor networks [55, 57, 46].

References

[1] J. Lelieveld, J.S. Evans, M. Fnais, D. Giannadaki, and A. Pozzer. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature*, 525(7569):367–371, 2015.

[2] S. L. Wright, R. C. Thompson, and T. S. Galloway. The physical impacts of microplastics on marine organisms: A review. *Environmental Pollution*, 178:483–492, 2013.

[3] R. Bobbink, M. Hornung, and J. G. M. Roelofs. The effects of air-borne nitrogen pollutants on species diversity in natural and semi-natural european vegetation. *Journal of Ecology*, 86(5):717–738, 1998.

[4] IPCC. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. 2014. [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds)].

[5] Stephen A Stansfeld and Mark P Matheson. Noise pollution: non-auditory effects on health. *British medical bulletin*, 68(1):243–257, 2003.

[6] C. D. Francis, C. P. Ortega, and A. Cruz. Noise pollution changes avian communities and species interactions. *Current Biology*, 19(16):1415–1419, 2009.

[7] T. Longcore and C. Rich. Ecological light pollution. *Frontiers in Ecology and the Environment*, 2(4):191–198, 2004.

[8] R. Ouaret, A. Ionescu, O. Ramalho, and Y. Candau. Indoor air pollutant sources using blind source separation methods. In *Proc. of ESANN*, 2017.

[9] A. Limem, G. Delmaire, M. Puigt, G. Roussel, and D. Courcot. Non-negative matrix factorization under equality constraints—a study of industrial source identification. *Applied Numerical Mathematics*, 85:1–15, November 2014.

[10] X.-J. Luo, S.-J. Chen, B.-X. Mai, G.-Y. Sheng, J.-M. Fu, and E. Y. Zeng. Distribution, source apportionment, and transport of PAHs in sediments from the Pearl River Delta and the northern South China Sea. *Archives of Environmental Contamination and Toxicology*, 55(1):11–20, 2008.

[11] Q. Zuo, Y.H. Duan, Y. Yang, X.J. Wang, and S. Tao. Source apportionment of polycyclic aromatic hydrocarbons in surface soil in tianjin, china. *Environmental Pollution*, 147(2):303–310, 2007.

[12] P. K Hopke. Review of receptor modeling methods for source apportionment. *Journal of the Air & Waste Management Association*, 66(3):237–259, 2016.

[13] P. K Hopke. It's time to drop principal components analysis as a 'receptor model'. *Journal of Atmospheric Chemistry*, 72(2):127–128, 2015.

[14] P. Paatero and U. Tapper. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5:111–126, 1994.

[15] M. Viana, T. A. J. Kuhlbusch, X. Querol, and A. Alastuey. Source apportionment of particulate matter in europe: a review of methods and results. *Journal of Aerosol Science*, 39(10):827–849, October 2008.

[16] M. Plouvin, A. Limem, M. Puigt, G. Delmaire, G. Roussel, and D. Courcot. Enhanced NMF initialization using a physical model for pollution source apportionment. In *Proc. of ESANN*, pages 261–266, 2014.

[17] R. Chreiky, G. Delmaire, C. Dorffer, M. Puigt, G. Roussel, and A. Abche. Robust informed split gradient NMF using Alpha Beta-divergence for source apportionment. In *Proc. of MLSP*, 2016.

[18] A. González-Calabuig, G. Faura, and M. del Valle. High dimensionality voltammetric biosensor data processed with artificial neural networks. In *Proc. of ESANN*, 2017.

[19] P. Comon and C. Jutten. *Handbook of blind source separation : independent component analysis and applications*. Academic Press, 2010.

[20] L. T. Duarte, S. Moussaoui, and C. Jutten. Source separation in chemical analysis : Recent achievements and perspectives. *IEEE Sig. Proc. Magazine*, 31(3):135–146, 2014.

[21] L. T. Duarte, C. Jutten, and S. Moussaoui. A bayesian nonlinear source separation method for smart ion-selective electrode arrays. *IEEE Sensors Journal*, 9(12):1763–1771, 2009.

[22] H. Kaufmann *et al.* EnMAP a hyperspectral sensor for environmental mapping and analysis. In *Proc. of IGARSS*, pages 1617–1619, 2006.

[23] D. Manolakis, S. Golowich, and R.S. DiPietro. Long-wave infrared hyperspectral remote sensing of chemical clouds: A focus on signal processing approaches. *IEEE Signal Processing Magazine*, 31(4):120–141, 2014.

[24] J. M Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geoscience and remote sensing mag.*, 1(2):6–36, 2013.

[25] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, J.-Y. Tourneret, M. A. Veganzones, G. Vivone, Q. Wei, and N. Yokoya. Hyperspectral pansharpening: A review. *IEEE Geoscience and remote sensing magazine*, 3(3):27–46, 2015.

[26] Z. Chen, H. Pu, B. Wang, and G.-M. Jiang. Fusion of hyperspectral and multispectral images: A novel framework based on generalization of pan-sharpening methods. *IEEE Geoscience and Remote Sensing Letters*, 11(8):1418–1422, 2014.

[27] C.-I. Chang and Q. Du. Estimation of number of spectrally distinct signal sources in hyperspectral imagery. *IEEE Trans. on geoscience and remote sensing*, 42(3):608–619, 2004.

[28] C. Revel, Y. Deville, V. Achard, and X. Briottet. Impact of the initialisation of a blind unmixing method dealing with intra-class variability. In *Proc. of ESANN*, 2017.

[29] R. Heylen, M. Parente, and P. Gader. A review of nonlinear hyperspectral unmixing methods. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6):1844–1868, 2014.

[30] P. W. Rundel, E. A. Graham, M. F. Allen, J. C. Fisher, and T. C. Harmon. Environmental sensor networks in ecological research. *New Phytologist*, 182(3):589–607, 5 2009.

[31] J. K. Hart and K. Martinez. Environmental sensor networks: A revolution in the earth system science? *Earth-Science Reviews*, 78(3):177–191, 2006.

[32] T. R. Karl, A. Arguez, B. Huang, J. H. Lawrimore, J. R. McMahon, M. J. Menne, T. C. Peterson, R. S. Vose, and H.-M. Zhang. Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, 2015.

[33] H. Carfantan and J. Idier. Statistical linear destriping of satellite-based pushbroom-type images. *IEEE Trans. on Geoscience and Remote Sensing*, 48(4):1860–1871, April 2010.

[34] L. Balzano and R. Nowak. Blind calibration of sensor networks. In *Proc. of IPSN*, pages 79–88, 2007.

[35] C. Wang, P. Ramanathan, and K.K. Saluja. Moments based blind calibration in mobile sensor networks. In *Proc. of ICC '08*, pages 896–900, May 2008.

[36] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel. Blind calibration of mobile sensors using informed nonnegative matrix factorization. In *Proc. of LVA-ICA*, volume 9237 of *LNCS*, pages 497–505. 2015.

[37] C. Wang, P. Ramanathan, and K. K. Saluja. Blindly calibrating mobile sensors using piecewise linear functions. In *Proc. of SECON*, pages 1–9, Jun. 2009.

[38] C. Wang, P. Ramanathan, and K. K. Saluja. Calibrating nonlinear mobile sensors. In *Proc. of SECON*, pages 533–541, Jun. 2008.

[39] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel. Nonlinear mobile sensor calibration using informed semi-nonnegative matrix factorization with a Vandermonde factor. In *Proc. of SAM*, 2016.

[40] R. Piedrahita, Y. Xiang, N. Masson, J. Ortega, A. Collier, Y. Jiang, K. Li, R.P. Dick, Q. Lv, M. Hannigan, and L. Shang. The next generation of low-cost personal air quality sensors for quantitative exposure monitoring. *Atmospheric Measurement Techniques*, 7(10):3325, 2014.

[41] K. Whitehouse and D. Culler. Macro-calibration in sensor/actuator networks. *Mobile Networks and Applications*, 8(4):463–472, Aug. 2003.

[42] O. Saukh, D. Hasenfratz, and L. Thiele. Reducing multi-hop calibration errors in large-scale mobile sensor networks. In *Proc. of IPSN*, 2015.

[43] B.-T. Lee, S.-C. Son, and K. Kang. A blind calibration scheme exploiting mutual calibration relationships for a dense mobile sensor network. *IEEE Sensors Journal*, 14(5):1518–1526, May 2014.

[44] O. Saukh, D. Hasenfratz, C. Walser, and L. Thiele. On rendezvous in mobile sensing networks. In *Proc. of REALWSN*, volume 281 of *LNCS*, pages 29–42, 2014.

[45] 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.

[46] M. Giannopoulos, S. Savvaki, G. Tsagkatakis, and P. Tsakalides. Application of tensor and matrix completion on water treatment data. In *Proc. of ESANN*, 2017.

[47] L. Zhouchen. A review on low-rank models in data analysis. *Big Data and Information Analytics*, 1(2/3):139–161, 2016.

[48] M. Davenport and J. Romberg. An overview of low-rank matrix recovery from incomplete observations. *Journal on Special Topics in Signal Processing*, 10(4):608–622, 2016.

[49] F. Harrou, F. Kadri, S. Khadraoui, and Y. Sun. Ozone measurements monitoring using data-based approach. *Process Safety and Environmental Protection*, 100:220–231, 2016.

[50] R. Kumar, C. Da Silva, O. Akalin, A. Y. Aravkin, H. Mansour, B. Recht, and F. J. Herrmann. Efficient matrix completion for seismic data reconstruction. *Geophysics*, 80(5):V97–V114, 2015.

[51] G. Roussel, L. Bourgois, M. Benjelloun, and G. Delmaire. Estimation of a semi-physical GLBE model using dual EnKF learning algorithm coupled with a sensor network design strategy: Application to air field monitoring. *Information Fusion*, 14(4):335–348, 2013.

[52] J. Ranieri, A. Chebira, and M. Vetterli. Near-optimal sensor placement for linear inverse problems. *IEEE Trans. on Signal Processing*, 62(5):1135–1146, 2014.

[53] V. Roy, A. Simonetto, and G. Leus. Spatio-temporal sensor management for environmental field estimation. *Signal Processing*, 128:369–381, 2016.

[54] J. Ranieri, I. Dokmanić, A. Chebira, and M. Vetterli. Sampling and reconstruction of time-varying atmospheric emissions. In *Proc. of ICASSP*, pages 3673–3676, 2012.

[55] J. Murray-Bruce and P. L. Dragotti. Solving inverse source problems for sources with arbitrary shapes using sensor networks. In *Proc. of ESANN*, 2017.

[56] M. Martinez-Camara, I. Dokmanić, J. Ranieri, R. Scheibler, M. Vetterli, and A. Stohl. The fukushima inverse problem. In *Proc. of ICASSP*, pages 4330–4334, 2013.

[57] M. Lopez-Radcenco, A. Aïssa-El-Bey, P. Ailliot, and R. Fablet. Non-negative decomposition of geophysical dynamics. In *Proc. of ESANN*, 2017.

[58] M. Dunbabin and L. Marques. Robots for environmental monitoring: Significant advancements and applications. *IEEE Robotics & Automation Magazine*, 19(1):24–39, 2012.

[59] H.E. Roy, M.J.O. Pocock, C.D. Preston, D.B. Roy, J. Savage, J.C. Tweddle, and L.D. Robinson. Understanding citizen science & environmental monitoring. *NERC–NHM*, 2012. Final report on behalf of UK-EOF.

[60] D. Christin, A. Reinhardt, S.S. Kanhere, and M. Hollick. A survey on privacy in mobile participatory sensing applications. *Journal of systems and software*, 84(11):1928–1946, 2011.

[61] J.-F. Markert, M. Budde, G. Schindler, M. Klug, and M. Beigl. Private rendezvous-based calibration of low-cost sensors for participatory environmental sensing. In *Proc. of the Urb-IoT*, pages 82–85, 2016.

[62] I.E.K. Harrane, R. Flamary, and C. Richard. Toward privacy-preserving diffusion strategies for adaptation and learning over networks. In *Proc. of EUSIPCO*, 2016.