ARTICLE IN PRESS

Pervasive and Mobile Computing ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Pervasive and Mobile Computing

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

Deriving high-resolution urban air pollution maps using mobile sensor nodes

David Hasenfratz [a,*], Olga Saukh [a], Christoph Walser [a], Christoph Hueglin [b], Martin Fierz [c], Tabita Arn [a], Jan Beutel [a], Lothar Thiele [a]

[a] *Computer Engineering and Networks Laboratory, ETH Zurich, Switzerland*

[b] *Laboratory for Air Pollution and Environmental Technology, EMPA, Switzerland*

[c] *Institute for Aerosol and Sensor Technology, FHNW, Switzerland*

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Mobile sensor network
Air pollution
Land-use regression
High-resolution pollution maps
Ultrafine particles

ABSTRACT

Up-to-date information on urban air pollution is of great importance for environmental protection agencies to assess air quality and provide advice to the general public in a timely manner. In particular, ultrafine particles (UFPs) are widely spread in urban environments and may have a severe impact on human health. However, the lack of knowledge about the spatio-temporal distribution of UFPs hampers profound evaluation of these effects. In this paper, we analyze one of the largest spatially resolved UFP data set publicly available today containing over 50 million measurements. We collected the measurements throughout more than two years using mobile sensor nodes installed on top of public transport vehicles in the city of Zurich, Switzerland. Based on these data, we develop land-use regression models to create pollution maps with a high spatial resolution of 100 m × 100 m. We compare the accuracy of the derived models across various time scales and observe a rapid drop in accuracy for maps with sub-weekly temporal resolution. To address this problem, we propose a novel modeling approach that incorporates past measurements annotated with metadata into the modeling process. In this way, we achieve a 26% reduction in the root-mean-square error – a standard metric to evaluate the accuracy of air quality models – of pollution maps with semi-daily temporal resolution. We believe that our findings can help epidemiologists to better understand the adverse health effects related to UFPs and serve as a stepping stone towards detailed real-time pollution assessment.

1. Introduction

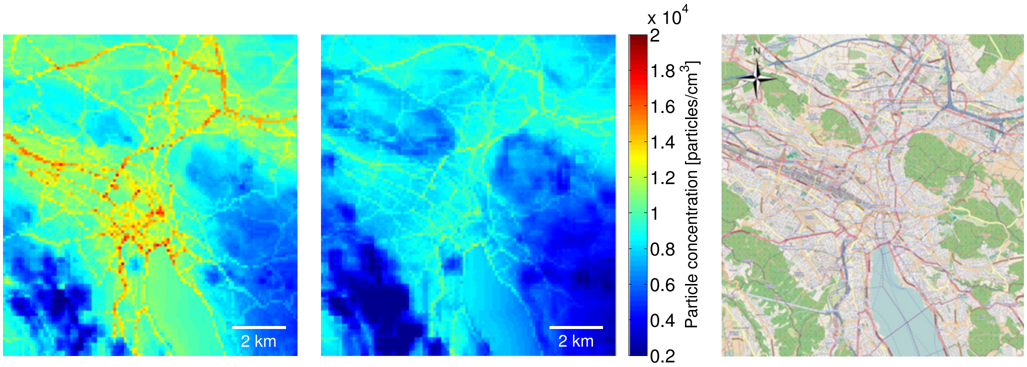
Air pollution is a major concern in many cities worldwide. Atmospheric pollutants considerably affect human health; they are responsible for a variety of respiratory and cardiovascular diseases and are known to cause cancer if humans are exposed to them for extended periods of time [1]. Additionally, air pollution is responsible for environmental problems, such as eutrophication and acidification of ecosystems.

Most countries have mass emission limits for particulate matter PM_{10} and $PM_{2.5}$ (*i.e.*, particles with a diameter of less than 10 μm and 2.5 μm, respectively), but have no restrictions on ultrafine particles (UFPs). UFPs are particles with a diameter of less than 100 nm. In ambient air, UFPs are mainly man-made as byproducts of specific high temperature processes, such

* Corresponding author. Tel.: +41 446327067.
E-mail address: hasenfratz@tik.ee.ethz.ch (D. Hasenfratz).

http://dx.doi.org/10.1016/j.pmcj.2014.11.008
1574-1192/

Please cite this article in press as: D. Hasenfratz, et al., Deriving high-resolution urban air pollution maps using mobile sensor nodes, Pervasive and Mobile Computing (2014), http://dx.doi.org/10.1016/j.pmcj.2014.11.008



(a) Monday–Saturday. (b) Sunday. (c) Map of Zurich.

Fig. 1. Novel ultrafine particle concentration maps for Zurich (Switzerland). The particle concentrations are higher during the week (Monday–Saturday) than on weekends (Sunday) due to higher traffic volumes.

as combustion reactions in car engines. The adverse health effects of UFPs are most probably underestimated when they are traditionally monitored by mass as part of PM_{10} and $PM_{2.5}$ [2]. This is because UFPs make a dominant contribution to the total number of urban particle concentrations, but their contribution to the total particle mass is small [3]. Therefore, UFPs were not considered particularly hazardous in the past. There are strong indications, however, that adverse health effects are more related to particle number concentration rather than to particle mass [2]. To better understand the adverse health effects of UFPs, it is essential to have spatially resolved UFP concentration measurements at hand [4].

Nowadays, air pollution is monitored by networks of static measurement stations operated by official authorities. These stations are highly reliable and able to accurately measure a wide range of air pollutants. However, their high acquisition and maintenance costs severely limit the number of installations. As a result, very little is known about the spatial distribution of air pollutants in urban environments and there is a lack of accurate intraurban air pollution maps. However, for air pollutants with high spatial variability, such as UFPs, the public availability of reliable pollution maps is essential. They raise the citizens' awareness about air pollution and empower environmental scientists to craft and evaluate new policies.

Contributions and road-map. To tackle the challenges above, we propose to use a mobile measurement system [5–8]. Node mobility trades off temporal resolution against spatial resolution, enabling a high spatial resolution across large areas without the need for a huge number of fixed sensors. However, due to the reduced temporal resolution of any covered location, it is a formidable challenge to derive pollution maps with a high temporal resolution at daily or hourly time scales. In this paper, we demonstrate that a mobile measurement system can effectively be used to derive accurate UFP pollution maps with high spatio-temporal resolution.

Our mobile measurement system consists of ten sensor nodes installed on top of public transport vehicles, which cover a large urban area on a regular schedule. The sensor nodes are equipped with a novel measurement device (MiniDiSCs [9]) to monitor UFP concentrations. Throughout more than two years, we collected over 50 million UFP measurements. Based on these data, we develop land-use regression (LUR) models to produce accurate pollution maps with high spatio-temporal resolution, such as those depicted in Fig. 1. LUR models use a set of explanatory variables (land-use and traffic data) to model pollution concentrations at locations not covered by the mobile sensor nodes. In a first step, we evaluate the dependencies between the explanatory variables and the measurements. Then, we exploit these relationships to predict the pollution levels for all locations without measurements but with available land-use and traffic information. Using this method and our mobile measurement system, we derive accurate and fine-grained pollution maps, which are valuable to environmental scientists, epidemiologists, and the general public. Finally, we use the developed pollution maps to analyze how much urban dwellers from Zurich (Switzerland) can reduce their exposure to UFPs by not taking the shortest path between origin and destination but a healthier and slightly longer alternative route.

In summary, this paper makes the following contributions:

- We introduce in Section 2 our mobile measurement system, which is deployed in the city of Zurich (Switzerland) collecting a highly spatially resolved data set of UFP measurements. From April 2012 to April 2014, we collected more than 50 million UFP measurements.
- Assessing the quality of the measurements is difficult due to very sparse ground truth data. We post-process the measurements (calibration and filtering) and propose in Section 3 a three-fold validation approach to evaluate the quality of the processed data. Our analysis indicates a high data quality.
- We use the validated measurements in Section 4 to derive LUR models for UFP pollution maps with a high spatial resolution of 100 m $\times$ 100 m. In Section 5 we apply standard metrics to analyze the quality of the models from yearly up to semi-daily temporal resolutions. We find a good quality of pollution maps with yearly to weekly time scales, while models with sub-weekly temporal resolutions perform less well.
- To tackle this problem, we propose in Section 6 a novel modeling approach that incorporates past measurements (annotated with metadata, such as environmental and meteorological conditions) into the modeling process. In this

Please cite this article in press as: D. Hasenfratz, et al., Deriving high-resolution urban air pollution maps using mobile sensor nodes, Pervasive and Mobile Computing (2014), http://dx.doi.org/10.1016/j.pmcj.2014.11.008