Estimating Urban Ultrafine Particle Distributions with Gaussian Process Models

Jason Jingshi Li
College of Engineering and Computer Science.
The Australian National University, Canberra, ACT 0200, Australia
jason.li@anu.edu.au

Arnaud Jutzeler
Artificial Intelligence Laboratory
EPFL, Lausanne, 1025, Switzerland
arnaud.jutzeler@epfl.ch

Boi Faltings
Artificial Intelligence Laboratory
EPFL, Lausanne, 1025, Switzerland
boi.faltings@epfl.ch

Abstract

Urban air pollution have a direct impact on public health. Ultrafine particles (UFPs) are ubiquitous in urban environments, but their distribution are highly variable. In this paper, we take data from mobile deployments in Zürich collected over one year with over 25 million measurements to build a high-resolution map estimating the UFP distribution. More specifically, we propose a new approach using a Gaussian Process (GP) to estimate the distribution of UFPs in the city of Zürich. We evaluate the prediction estimations against results derived from standard General Additive Models in Land Use Regression, and show that our method produces a good estimation for mapping the spatial distribution of UFPs in many timescales.

1 Introduction

Air pollution in urban environments have a direct impact on the health of the people. The World-Health-Organization (2011) estimated that over 1.3 million deaths per year world-wide are attributed to urban outdoor air pollution. Currently in most developed countries, a network of government-funded and operated static measurement stations continuously make highly reliable and accurate measurements on important air pollutants. However, the high cost of installation and maintenance of these stations limits the number of stations deployed in a given city. Consequently, only very limited information can be collected about the spatial distribution of air pollutants in the urban setting.

The OpenSense project, described in Aberer *et al.* (2010), is a multi-disciplinary project funded by the Swiss National Science Foundation to study mobile air quality monitoring and modelling in urban environments. It is deploying multiple mobile air quality monitoring stations on top of trams in the Swiss city of Zürich (Fig. 1), collecting measurements of ozone concentrations (O_3) and the counting of ultrafine particle (UFPs). To this date, it has publicly released over 25 million measurements over an urban area of 100 km^2. The data and their sensing methodology can be found in Li *et al.* (2012b) and Hasenfratz *et al.* (2014). These data form a sufficient basis to study the spatial variability of the pollutants in the urban environment.

In: S. Winter and C. Rizos (Eds.): Research@Locate'14, Canberra, Australia, 07-09 April 2014, published at http://ceur-ws.org

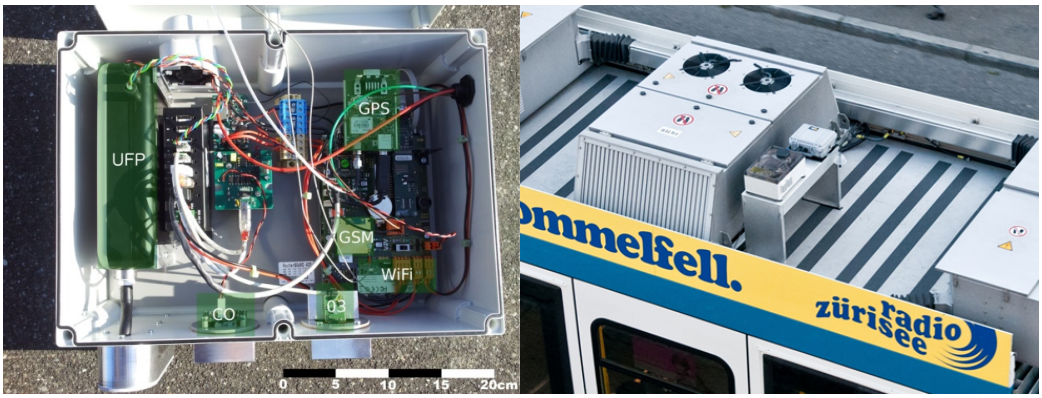

Figure 1: Left: Interior of an OpenSense sensing node; Right: deployment on a tram in Zürich

Traffic junctions, industrial installations and urban canyons all contribute to the high spatial and temporal variability of air pollution in urban areas. Small-scale spatial distribution of ambient air pollution have traditionally been studied with Land-Use-Regression (LUR) summarised in Hoek *et al.* (2008). It uses land-use and traffic characteristics of a particular grid region as explanatory variables to learn to estimate pollution concentrations under a Generalized Additive Model (GAM). The learnt model is then used to predict pollution levels for all locations with the available land-use information.

In this paper, we propose a novel approach of estimating urban ultrafine particle levels across different temporal aggregates from measurements collected from the trams. Similar to standard models in land use regression, it estimates the pollution levels within different grid-cells in the urban environment from a set of land-use features. Our model is based on constructing a Gaussian Process described in Rasmussen and Williams (2006), with additional consideration to spatial features in the covariance matrix. Following the practice in previous work of Hasenfratz *et al.* (2014), we evaluate the models (GAM, pure land use and mixed spatial-land-use) using standard random 10-fold cross validation.

The outline of this paper is as follows: we begin with a summary of the background to the paper: the data, the traditional models used in land-use regression, and introducing Gaussian Process Regression. We then introduce a new approach for estimating UFP levels, and evaluate it against the previous approach over a benchmark dataset.

2 Background

2.1 The Aggregate Datasets

The data were selected from UFP measurements collected on Zürich trams between April 2012 and March 2013 as part of the OpenSense project and the sensing methodology is described in Li *et al.* (2012b) and Hasenfratz *et al.* (2014). The data were partitioned into 13,200 grid cells of size 100m $\times$ 100m. The profile of a typical grid cell, such as the one containing Centralplatz in Zürich, is shown in Fig 2. We can see that instead of being fitted to a normal distribution (solid line shows the best-fit), the measurements fits much better as log-normal distribution (dotted line). This is consistent with literature on particle count concentrations in urban environments described in Mølgaard *et al.* (2012). The data were captured and transmitted in real time to a back-end server running Global Sensor Network (GSN) by Aberer *et al.* (2006), and removed to a local database to be preprocessed and aggregated before entered into the model.

Several preprocessing steps were used before the data were prepared for the model, including removal of measurements within the indoor tram depot, measurements with bad GPS data, and measurements with extraordinary high levels >100'000 particles per cm^3. These steps were described in detail in Hasenfratz *et al.* (2014), with the purpose of avoiding bias due to erroneous measurements.

We then aggregate the data within the different grid cells according to the different time windows, such as yearly, seasonal, monthly, biweekly, weekly, daily and half-daily. This is done to understand the trade-off between long and short term aggregate data. In order to evaluate and compare our results to previous work, we followed the convention of selecting only the 200 grid cells with the highest measurements count for the purpose of modelling and validation, as Hasenfratz *et al.* (2014) showed that the state-of-the-art models produced the

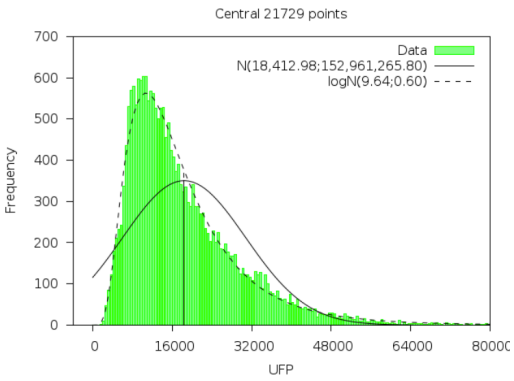

Figure 2: Distribution of UFP measurements collected in the grid cell near Centralplatz in Zürich during winter 2013. The black line shows the maximum likelihood estimated normal distribution whereas the dashed line shows the maximum likelihood estimated log-normal distribution.

most reliable predictions when only the top 200 grid cells with the highest measurement count are considered.

2.2 Land Use Regression

In literature, land-use regression models are used to assess intra-urban air pollution distributions, and a comprehensive review of these techniques can be found in Hoek *et al.* (2008). They typically combine monitoring of air pollution at 20-100 locations spread over the study area, and develop a model using predictor variables obtained through geographic information systems (GIS). The predictor variable generally include some traffic information, population density, designated land use and features of the landscape such as attitude and slope. Due to the cost of deployments, studies usually last 1-2 weeks in duration.

For particulate matter such as $PM_{2.5}$ and PM_{10} and UFPs, Generalized Additive Models (GAMs) have been used in land use regression to study their spatial and temporal variability. It typically use the following equation to model the relationship between the pollution level p and a set of explanatory variables $A_1, \dots A_n$.

$$ln(p) = a + s_1(A_1) + s_2(A_2) + \cdots + s_n(A_n) + \epsilon \tag{1}$$

where a is known as the intercept, ϵ the error term, and $s_1 \dots s_n$ are typically smooth regression splines with an upper limit of 3 on the degree of freedom. In this paper, we use the GAM data from Hasenfratz *et al.* (2014) as a benchmark to compare our model predictions.

2.3 Gaussian Process Regression

Also known as Kriging, Gaussian process regression (GPR) has been extensively used for decades in Geostatistics to model various spatial phenomena such as soil concentrations, weather-related events, etc., and in-depth overviews can be found in Cressie and Cassie (1993) and Rasmussen and Williams (2006). Similar to other non-parametric approaches, GPR does not require prior structural knowledge about the phenomenon. Indeed, the idea is precisely that structure is directly inferred from the data. Furthermore, GPR outputs statistical predictions and thus represents an adequate candidate to model phenomena that are inherently noisy and which one can only observe through noisy instruments. Recently it has been successfully applied in many machine learning tasks such as bioinformatics in Chu *et al.* (2005), sensor calibration in Monroy *et al.* (2012) and crowdsourcing Venanzi *et al.* (2013) It still represents a very active ongoing research area as seen in e.g. Bonilla *et al.* (2010), Cao *et al.* (2013) and Nguyen and Bonilla (2014). To allow the reader to have a better understanding of our models, in the following we will provide a very brief technical overview of Gaussian Process Regression.

A Gaussian Process (GP) is used to model a phenomenon that takes place in a certain input space $\mathcal{X} \subseteq \mathbb{R}^d$. We formally write $f(\mathbf{x})$ where $\mathbf{x} \in \mathcal{X}$ the function that models the phenomenon. The general idea is to assume that the function $f(\mathbf{x})$ is a specific realization of a prior Gaussian Process $\mathcal{GP}$, which is the generalization of a multivariate normal distribution to an infinity of random variables, that is to say a distribution over whole functions. A GP is fully defined by its mean function $m(\mathbf{x})$ and its covariance function $k(\mathbf{x}, \mathbf{x}')$ (also called

kernel) that are the generalization of the mean vector, respectively the covariance matrix of a multivariate normal distribution.

Regression with a GP is typically performed as follows. In general, we can only make from the phenomenon noisy observations $y_i = f(\mathbf{x_i}) + \epsilon_i$ where the additive noise ϵ is also assumed to be Gaussian $\epsilon \sim \mathcal{N}(0, \sigma_n^2)$. By using the marginalization property of GPs and the additive nature of the noise ϵ we know the joint distribution of the observations $\mathbf{y}$ at locations X and the values $\mathbf{f}_*$ at test points X_* to be:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{f}_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(X) \\ m(X_*) \end{bmatrix}, \begin{bmatrix} k(X, X + \sigma_n^2 I) & k(X, X_*) \\ k(X_*, X) & k(X_*, X_*) \end{bmatrix} \right) \tag{2}$$

Then for those test points X_* the regression consists in computing the predictive distribution $p(\mathbf{f}_*|\mathbf{y})$. Fortunately by the conditioning property of a joint multivariate Gaussian distribution this expression is tractable and even admit a closed formula. It results in another multivariate Gaussian distribution. For any single test points $\mathbf{x}_* \in X_*$ the predictive mean and variance are given by:

$$\bar{f}(\mathbf{x}_*) = m(\mathbf{x}_*) + k(\mathbf{x}_*, X)(k(X, X) + \sigma_n^2 I)^{-1}(\mathbf{y} - m(\mathbf{x})) \tag{3a}$$

$$\mathbb{V}[f(\mathbf{x}_*)] = k(\mathbf{x}_*, \mathbf{x}_*) - k(\mathbf{x}_*, X)(k(X, X) + \sigma_n^2 I)^{-1} k(X, \mathbf{x}_*) \tag{3b}$$

The main challenge is to create and choose prior mean and covariance functions that carry adequate assumptions about the phenomenon. We describe in detail how we derived such functions in the following section.

3 Our Model

3.1 The Land-Use Model

Our first GP model uses only land-use variables as features to generate predictions on the mean UFP concentration measured by the sensors within the respective grid cells in the timeframe of the specified dataset. They follow from the features used in Hasenfratz *et al.* (2014). The model takes a vector $\mathbf{x_{LU}}$ containing the land-use variables values of a certain 100m × 100m grid cell as input. These land-use features were taken from the following sources:

- Swiss Federal Statistical Office
  - Population density, industry density, building heights, heating type, terrain elevation, terrain slope
- Canton of Zürich government
  - Average daily traffic volume
- *OpenStreetMaps.org*
  - Main road type, distance to next major road, distance to major traffic signal

As we wanted to start with no particular *a priori* structural knowledge, only very simple mean functions were tried such as the trivial fixed 0 function and a constant c. Deriving a suitable covariance function was, however, a bit more complex. Indeed, to be valid a covariance function must be positive definite. It is common practice to start from well-known parametrized families of positive definite functions and fit the parameters (that in the scope of GPR are called hyperparameters) using the data. All the covariance functions that were tried are stationary that is to say every points of the space shows the exact same covariance structure with its own surroundings or more formally we have $k(\mathbf{x}, \mathbf{x}') = k(\mathbf{x} - \mathbf{x}')$. Stationary covariance functions such as squared exponential, and various flavours of the Matérn class were tried, each one carrying different assumption about the smoothness of the process. Finally from preliminary tests results, a constant mean function and a squared exponential covariance function were selected. We note that the chosen covariance function was the one that carried the strongest smoothness assumptions. The prior GP is thus defined by:

$$m(\mathbf{x_{LU}}) = c \tag{4a}$$

$$k(\mathbf{x_{LU}}, \mathbf{x'_{LU}}) = \sigma_f^2 \exp\Big(-\frac{1}{2}(\mathbf{x_{LU}} - \mathbf{x'_{LU}})^\top M(\mathbf{x_{LU}} - \mathbf{x'_{LU}})\Big) \text{ where } M = \operatorname{diag}(\boldsymbol{\ell_{LU}})^{-2} \tag{4b}$$

The σ_f^2 is the magnitude hyperparameter, and the $\boldsymbol{\ell_{LU}}$ are the length-scale hyperparameters that determine the relevance of some or other land-use variables. To learn the values of all the hyperparameters $\theta = (c, \sigma_f^2, \boldsymbol{\ell_{LU}}, \sigma_n^2)$ one can either use optimization or sampling techniques. In our case, we used the standard approach that consists in optimizing the log marginal likelihood:

$$\log p(\mathbf{y}|X, \theta) = -\frac{1}{2}\mathbf{y}^\top (K + \sigma_n^2 I)^{-1}\mathbf{y} - \frac{1}{2}\log|K + \sigma_n^2 I| - \frac{n}{2}\log 2\pi$$

Every evaluation of this expression takes $O(n^3)$ with n being the number of training points X. From then the evaluation of its derivatives with regards to hyperparameters takes $O(n^2)$ per hyperparameter.

3.2 The Mixed Spatial Land-Use Model

Even though the explanation of the phenomenon given by the land-use variables may already be quite good, it is very likely that part of it still elude us because of some contributions to the phenomenon that are badly or not at all reflected in the variables. To address this matter, we tried to incorporate geographical informations into the model with the hope that such missed contribution will at least be partly explained locally.

The problem with parametric models such as GAM is that we cannot easily add geographic informations into the model in a sensible way. For example if we naively add the longitude and latitude as covariates, we would be making very strong assumptions rather unrealistic.

However, with GPR (and this is why it has been extensively used in Geostatistics) it is natural to include such informations in the reasoning. This is done by including a consideration for geographical distance in the covariance function. We call our second model a mixed spatial-land-use model, which is a variant of the first one in which we added a term in the covariance structure. We also tried different isotropic kernels to be this additional term. From the preliminary experiments the following covariance function was selected:

$$k(\begin{bmatrix}\mathbf{x_{LU}}\\ \mathbf{x_S}\end{bmatrix}, \begin{bmatrix}\mathbf{x'_{LU}}\\ \mathbf{x'_S}\end{bmatrix}) = \sigma_{f_{LU}}^2 \exp\Big(-\frac{1}{2}(\mathbf{x_{LU}} - \mathbf{x'_{LU}})^\top M(\mathbf{x_{LU}} - \mathbf{x'_{LU}})\Big) + \sigma_{f_S}^2 \exp\Big(-\frac{\|\mathbf{x_S} - \mathbf{x'_S}\|}{\ell_S}\Big) \tag{5}$$

It is worth noting that it is the exponential function, the less smooth of the considered covariance functions, that was chosen to be the additional term in function of the geographical distance. The values of the hyperparameters $\theta = (c, \sigma_{f_{LU}}^2, \sigma_{f_S}^2, \ell_S, \boldsymbol{\ell_{LU}}, \sigma_n^2)$ were once again fixed using marginal likelihood maximization.

4 Evaluations

We implemented our own Java framework to perform GPR. However, the conjugate gradient optimizer, used to maximize the log marginal likelihood, was taken from the Matlab toolbox GPML v.2 (see Rasmussen and Nickisch (2010)) and translated in Java. Most of linear algebra operations were carried out using EJML[1] library. The experiments were conducted on a server with 64 AMD Opteron processing cores and 96 GB of RAM. In the experiments, we compared the following three different type of models on the UFP datasets described earlier.

1. **GAM** A General Additive Model from Hasenfratz *et al.* (2014);
2. **GP_LU** Our land-use only GP model;
3. **GP_LUXY** Our mixed spatial-land-use GP model.

From the benchmarking data supplied by Hasenfratz *et al.* (2014), we get 989 datasets comprise of 597 half-daily, 309 daily, 44 weekly, 23 biweekly, 11 monthly, 4 seasonally and a single yearly aggregated dataset from measurements taken from Zürich trams between April 2012 and March 2013. For each aforementioned type of model, we trained yearly to half-daily models to predict mean pollution level within grid cells (in particle count per cm^3). We evaluated the quality of the models predictions using standard randomised 10-fold cross validation. That is, for each dataset, we randomly partitioned the data into 10 equal parts, and iteratively we used 9 parts as training set of the model to generate predictions to be compared against the 1 remaining part.

Fig. 3 shows the satellite image of the urban area covered in the deployment, the output of the pollution map for the season of summer in 2012, and the comparison of the prediction against ground truth of the same season under random 10-fold cross validation. Fig.4 shows the scatter plots of model predictions against ground truth

[1] http://code.google.com/efficient-java-matric-library/

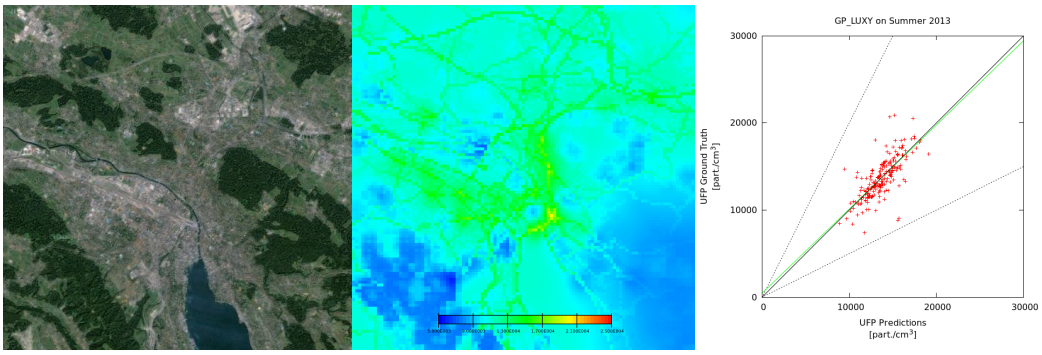

Figure 3: left: a satellite image of Zurich; centre: the predicted summer mean UFP level from mixed spatial-land-use GP model; and right: the scatter plot of the predictions from the same model against ground truth under random 10-fold cross validation.

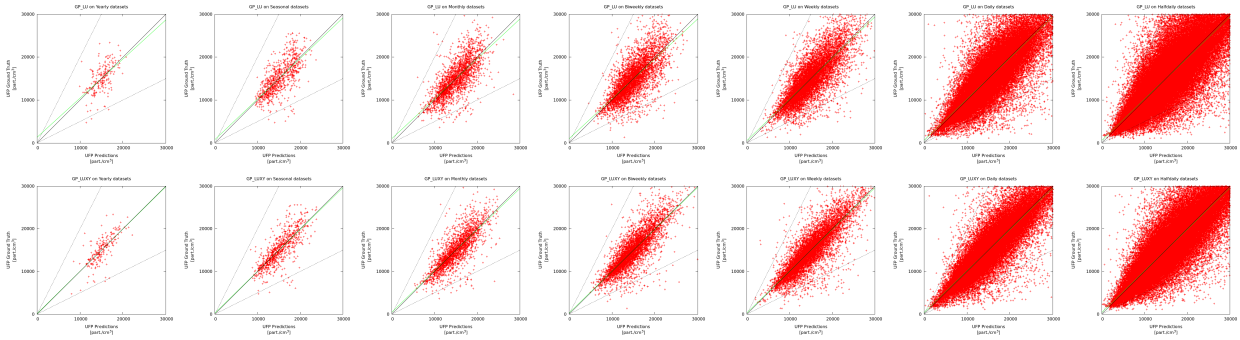

Figure 4: Scatter plot of prediction vs. ground truth after random 10-fold cross validation on yearly, seasonal, monthly, biweekly, weekly, daily and half-daily data. Top row: GP_LU model, bottom row: GP_LUXY model.

across all time scales, where all predictions of the same time scale are located on the same plot. It is worthy to note that similar to the previous model presented in Hasenfratz *et al.* (2014), our models also show little to no bias, as evident from the fact that across all cases the linear regression lines (in green) are very close to the optimal 1-to-1 lines. It indicates the absence of systematic model errors.

4.1 RMSE

First we compare the Root Mean Square Error (RMSE) of the predictions derived from the models under random 10 fold cross validation (Fig. 5). It is a standard metric of predictive power for measuring the accuracy of prediction models. It is obtained by:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(p_i - g_i)^2}{N}} \tag{6}$$

where p_i denotes the i^{th} prediction, g_i the ground truth of the i^{th} prediction, and N the total number of predictions. In Fig. 5, the plot on the left displays the overall mean of the RMSE, while the box-plot on the right displays the minimum, lower quartile, median, upper quartile and maximum of the average RMSE of the whole 10-fold validation tests on all the datasets of the same time scale. The yearly data came from a single dataset, thus it is presented as a single value. It shows that as expected, the higher temporal resolution leads to higher uncertainty in the prediction, the GP models outperforms GAM across all temporal resolutions, and the mixed spatial-land-use model produced less error than the land-use only model.

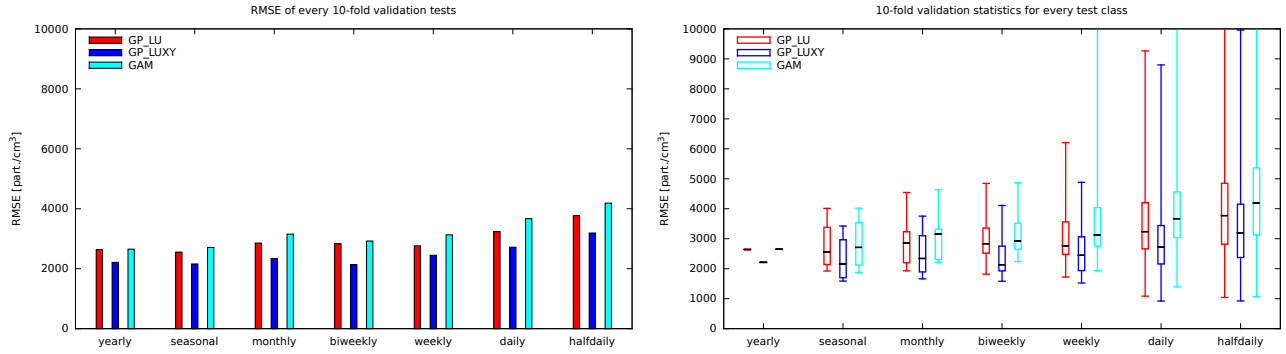

Figure 5: Mean (left) and distribution (right) RMSE of model predictions across all datasets in random 10 fold cross validation (the lower the better)

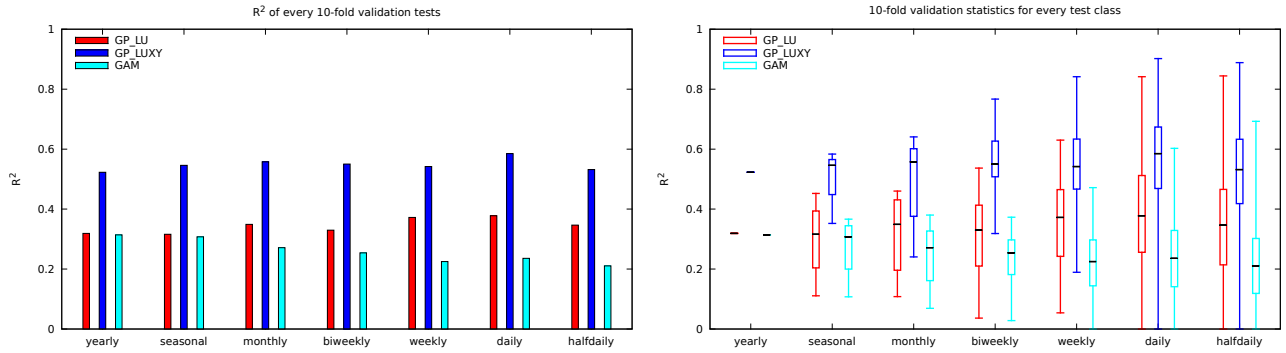

Figure 6: Mean (left) and distribution (right) R^2 score of model predictions across all datasets in random 10 fold cross validation (the higher the better)

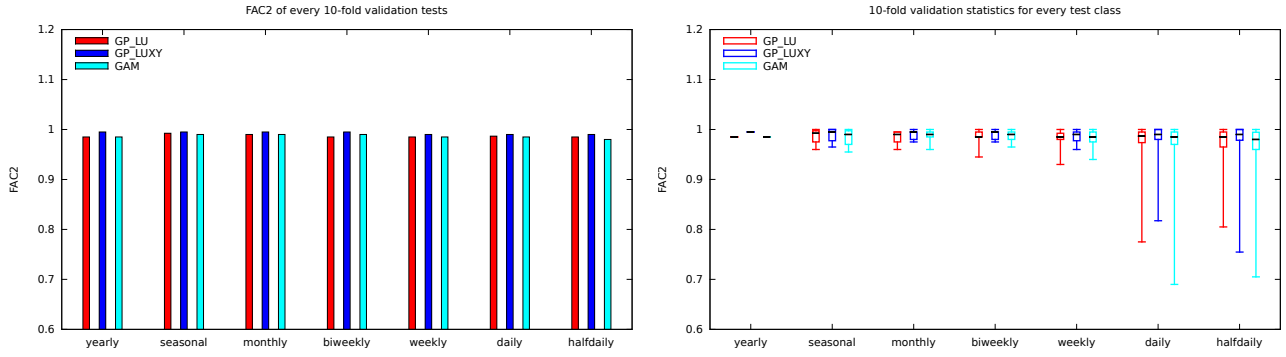

Figure 7: Mean (left) and distribution (right) of FAC2 score of model predictions across all datasets in random 10 fold cross validation (the higher the better)

4.2 R^2 Score

We then compare the R^2 coefficient, also known as the coefficient of determination of the model predictions (Fig. 6). It indicates how well the observed outcomes are replicated by the model predictions as the proportional variation of outcomes explained by the model. Its formula is given by:

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(p_i - g_i)^2}{\sum_{i=1}^{N}(g_i - \bar{g})} \tag{7}$$

where p_i denotes the i^{th} prediction, g_i the ground truth of the i^{th} prediction, and $\bar{g}$ is the mean of the ground truth. In Fig. 6 observe that the variance of the R^2 score also increases as the the time scale shrinks across all models. We see that the results of GP models in general have a higher R^2 than GAMs, and introducing the spatial covariance in the GP model also improves the R^2 score across all time scales.

4.3 FAC2 Score

Finally, we compare the $FAC2$ score of the model predictions (Fig. 7). It measures the fraction of data points that lie inside the factor of two area. It is a robust measure of prediction as it is not overly influenced by high and low outliers. It is derived by:

$$FAC2 : 0.5 \leq \frac{p_i}{g_i} \leq 2 \tag{8}$$

The box plots in Fig. 7 show the FAC2 distributions for all models across all temporal scales. We can see that they all have very high FAC2 values for yearly, seasonal, monthly, biweekly and weekly data. Daily and half daily predictions have lower FAC2 values, with GP models perform slightly better than GAM.

5 Conclusion and Future Work

We implemented two schemes based on Gaussian Process for estimating mean UFP concentrations in urban areas of Zürich, Switzerland. We show that they provide an alternative to GAM approaches in land-use regression, and there is a general trade off between the length of the time scale and the quality of the model predictions. We also show that across the timescales the proposed GP models presents an improvement on the current state of the art. The resulting maps may be useful for application such as assessing population exposure to air pollutants similar to that of Carroll *et al.* (1997), uncover areas of high air pollution for persons with allergies, or evaluate the trustworthiness of measurements contributed by a community of sensors as described in Li *et al.* (2012a) and Faltings *et al.* (2014).

Possible future work includes moving away from a grid-based model to make use of urban spatial features described in Li *et al.* (2012b), developing models that handles different aspects of sensor reliability and measurement bias; detecting and filtering spurious measurements, and combining meteorological information and real time data to produce the best real-time estimations for individual exposure analysis and route planning. Our approach based on Gaussian Process Regression is very general, and it is interesting to see if it can be generalised to particulate dispersion outside urban environments to applications such as bush-fire detection; and whether it can be applied to estimating other air-borne or water-borne pollutant dispersions.

Acknowledgements

We thank our collaborators at ETHZ David Hasenfratz and Olga Saukh for supplying the benchmarking data and model from their previous work. This work is supported by OpenSense project funded by NanoTera.ch, and the ARC Discovery Project (DP120103758) "Artificial Intelligence Meets Sensor Networks".

References

K. Aberer, M. Hauswirth, and A. Salehi. A middleware for fast and flexible sensor network deployment. In *VLDB*, 2006.

K. Aberer, S. Sathe, D. Chakraborty, A. Martinoli, G. Barrenetxea, B. Faltings, and L. Thiele. OpenSense: Open community driven sensing of environment. In *ACM IWGS*, 2010.

Edwin V Bonilla, Shengbo Guo, and Scott Sanner. Gaussian process preference elicitation. In *NIPS*, pages 262–270, 2010.

Yanshuai Cao, Marcus A Brubaker, David Fleet, and Aaron Hertzmann. Efficient optimization for sparse gaussian process regression. In *NIPS*, pages 1097–1105, 2013. URL http://papers.nips.cc/paper/5087-efficient-optimization-for-sparse-gaussian-process-regression.pdf.

RJ Carroll, R Chen, EI George, TH Li, HJ Newton, H Schmiediche, and N Wang. Ozone exposure and population density in harris county, texas. *Journal of the American Statistical Association*, 92(438):392–404, 1997.

Wei Chu, Zoubin Ghahramani, Francesco Falciani, and David L Wild. Biomarker discovery in microarray gene expression data with gaussian processes. *Bioinformatics*, 21(16):3385–3393, 2005.

Noel AC Cressie and Noel A Cassie. *Statistics for spatial data*, volume 900. Wiley New York, 1993.

Boi Faltings, Jason Jingshi Li, and Radu Jurca. Incentive mechanisms for community sensing. *IEEE Transactions on Computers*, 63(1):115–128, 2014.

D. Hasenfratz, O. Saukh, C. Walser, C. Hueglin, M. Fierz, and L. Thiele. Pushing the spatio-temporal resolution limit of urban air pollution maps. In *Proceedings of the 12th International Conference on Pervasive Computing and Communications (PerCom'14)*, 2014.

Gerard Hoek, Rob Beelen, Kees de Hoogh, Danielle Vienneau, John Gulliver, Paul Fischer, and David Briggs. A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment*, 42(33):7561 – 7578, 2008. ISSN 1352-2310. doi: http://dx.doi.org/10.1016/j.atmosenv.2008.05.057. URL http://www.sciencedirect.com/science/article/pii/S1352231008005748.

Jason Jingshi Li, Boi Faltings, and Radu Jurca. Incentive schemes for community sensing. In *The 3rd International Conference in Computational Sustainability*, 2012a.

Jason Jingshi Li, Boi Faltings, Olga Saukh, David Hasenfratz, and Jan Beutel. Sensing the air we breathe - the opensense zurich dataset. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence (AAAI12), Toronto, Canada*, July 2012b.

Bjarke Mølgaard, Tareq Hussein, Jukka Corander, and Kaarle Hmeri. Forecasting size-fractionated particle number concentrations in the urban atmosphere. *Atmospheric Environment*, 46(0):155 – 163, 2012. ISSN 1352-2310. doi: http://dx.doi.org/10.1016/j.atmosenv.2011.10.004. URL http://www.sciencedirect.com/science/article/pii/S1352231011010491.

J Monroy, Achim Lilienthal, J Blanco, Javier González-Jimenez, and Marco Trincavelli. Calibration of mox gas sensors in open sampling systems based on gaussian processes. In *IEEE Sensors'12*, pages 1743–1746, 2012.

Trung V Nguyen and Edwin V Bonilla. Fast allocation of gaussian process experts. In *International Conference on Machine Learning*, 2014.

Carl Edward Rasmussen and Hannes Nickisch. Gaussian processes for machine learning (gpml) toolbox. *J. Mach. Learn. Res.*, 11:3011–3015, December 2010. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=1756006.1953029.

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006. ISBN 0-262-18253-X.

Matteo Venanzi, Alex Rogers, and Nicholas R Jennings. Crowdsourcing spatial phenomena using trust-based heteroskedastic gaussian processes. In *First AAAI Conference on Human Computation and Crowdsourcing*, 2013.

World-Health-Organization. Air quality and health. In *Fact Sheet No. 313*, 2011.