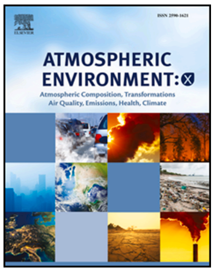
NOX and PM10 Bayesian concentration estimates using high-resolution numerical simulations and ground measurements over Paris, France

Delphy Rodriguez, Éric Parent, Laurence Eymard, Myrto Valari, Sébastien Payan

Accepted Manuscript



NO_X and PM_{10} Bayesian concentration estimates using high-resolution numerical simulations and ground measurements over Paris, France

Delphy Rodriguez, Eric Parent, Laurence Eymard, Myrto Valari, Sébastien Payan

PII: S2590-1621(19)30041-3

DOI: https://doi.org/10.1016/j.aeaoa.2019.100038

Article Number: 100038

Reference: AEAOA 100038

To appear in: *Atmospheric Environment: X*

Received Date: 24 January 2019

Revised Date: 27 May 2019

Accepted Date: 29 May 2019

Please cite this article as: Rodriguez, D., Parent, E., Eymard, L., Valari, M., Payan, Sé., NO_X and PM_{10} Bayesian concentration estimates using high-resolution numerical simulations and ground measurements over Paris, France, *Atmospheric Environment: X* (2019), doi: https://doi.org/10.1016/j.aeaoa.2019.100038.

NO_X and PM_{10} Bayesian concentration estimates using high-resolution numerical simulations and ground measurements over Paris, France

Delphy Rodriguez[1] (1), Eric Parent (2), Laurence Eymard (1), Myrto Valari (3), Sébastien Payan (1)
(1) LATMOS/IPSL, Sorbonne Université, UVSQ,CNRS, Paris, France
(2) AgroParisTech, 16 rue Claude Bernard, Paris, France
(3) LMD, Ecole Polytechnique Palaiseau, France

HIGHLIGHTS

- A Bayesian scheme combining simulated pollutant concentrations with measured ones.
- Bias correction for high-resolution air quality modeling.
- Enhanced concentration estimates of air quality monitoring network.
- Spatially resolved concentration estimates around urban monitoring sites.

Abstract

Air quality over cities is mainly monitored by in-situ surface measurements. However, these stations are too sparse to properly capture the inhomogeneity of pollutant concentrations over urban areas. The need for high-resolution concentration estimate has grown in recent years, together with the awareness of the harmful effects of air pollution. In this study, we develop a Bayesian scheme that combines the high-resolution (3×3 m^2) Particulate Micro SWIFT SPRAY numerical air quality simulations (PMSS) with operational surface measurements. The goal is to improve NO_X and PM_{10} PMSS concentrations estimates over monitoring stations and within their vicinity. For this purpose, we simulate pollutant concentrations over the city of Paris for ten days over the period of March 2016. The Bayesian model provides an enhanced estimate of pollutant concentration in space and time. At the monitoring stations location, these estimates are characterized by lower temporal dispersion compared to the simulated data. Within the vicinity of the monitor stations, enhanced concentration estimates are closer to observations. For NO_X, the improvement is stronger and occurs in a larger area for urban background stations than for traffic stations. Overall, NO_X improvement is higher than PM_{10} improvement. The initial PMSS model prediction is more biased for NO_X than for PM_{10} due to large uncertainties in NO_X emissions over the traffic network.

[1] Corresponding author.

Email address: delphy.rodriguez@latmos.ipsl.fr and delphy.rodriguez@gmail.com (D. Rodriguez)

Keywords: Urban air pollution; Surface monitoring networks; High-resolution modeling; Spatial representativeness; Bayesian modeling

1 Introduction

In large cities, air quality is often monitored by ground station networks measuring pollutant concentrations. Several types of areas are sampled: traffic, residential, etc. The spatial representativeness of these measurements is limited to a small area around the station, especially in dense urban areas: emissions from various sources, as well as air flow channeling in street canyons, result in sharp horizontal gradients of pollutant concentrations (Harrison, 2018). The need for high-resolution gridded estimates of pollutant concentrations over big cities has grown in recent years, together with the increasing awareness of the harmful health effects of atmospheric pollution.
Various approaches have been developed to address the spatial distribution of pollutant concentration around monitoring sites. These approaches use real-time measurements from local networks and expand them in space and/or time based on additional relevant information such as land-use, weather conditions, and topography. They may be divided mainly into first-principle geophysical modeling and statistical ones. Atmospheric pollutant dispersion models describe the physical and chemical processes that govern the emission, transport, mixing, chemical reaction and deposition of pollutants. High-resolution models, such as the Particulate Micro Swift Spray (PMSS) air quality model integrated into the ARIA City modeling platform (http://www.aria.fr/aria_city.php), provide realistic spatial distributions of pollutant concentrations. However, they suffer from large biases due to uncertainties on the emissions and on the parameterizations of the complex atmospheric processes governing urban air quality. Statistical approaches include land-use regression (Janssen, *et al.* 2012 and Son, *et al.* 2018), kriging interpolation schemes (Beauchamp *et al.* 2017 and Wu *et al.* 2018, Martin *et al.* 2012), methods for assimilation of measurements in models such as Kalman filtering (Hanea *et al.* 2004) or 4D variational assimilation algorithms (Elbern *et al.* 2001)) and Bayesian models (Beloconi *et al.* 2018).
The scope of this study is thus to improve NO_X and PM_{10} pollutant concentrations simulated with the PMSS high-resolution air quality model over and in the vicinity of the monitoring stations of the AIRPARIF local air quality agency in Paris, France (https://www.airparif.asso.fr). To spatially expand the measurement around monitoring stations, we use the concept of "representativeness areas" defined in our previous work (Rodriguez *et al.* 2019). These areas consist of points where concentration levels are similar and time variations correlated to those at the monitoring site's location. In practice, they are defined by applying criteria thresholds on the Normalized Root Mean Square Error (NRMSE) and on the correlation coefficient between simulated concentrations over all the pixels around the monitoring sites and simulated concentration at the monitoring site's location.

Frequentist or Bayesian statistical approaches may be used. In both approaches, the purpose is to determine as accurately as possible an unknown parameter θ. Here, the unknown parameter θ is the concentration in the so-called "representativeness

area" of an AIRPARIF monitoring site, i.e. at the vicinity of a station. In the frequentist setting, the unknown parameter θ is pre-assigned to a fixed single value. A statistical procedure is chosen to establish a confidence interval from any set of data collected in the same conditions, under the same model. It ensures the probability of this random interval to cover the fixed value θ, i.e. the reliability of the chosen statistical procedure. Then, the confidence interval is calculated from the available data as a random draw from this estimation procedure. All the information is contained in data. The frequentist procedure assumes a reproducible experience and a large data set. Frequentists claim that this is an *objective* way of reasoning.
Within the Bayesian paradigm, the unknown parameter θ is considered uncertain. What is known about θ is represented by a random variable conditioned upon the available information. The major issue is to quantify the uncertainty from all available data (Boreux *et al.*, 2010). One of the most significant differences between frequentist and Bayesian approaches is the choice of a "prior" (O'Hagan, 2008). Prior characterizes our expert knowledge by giving suggestions on an unknown parameter θ (as a distribution of θ values) before observing data and is *subjective* since it reflects a personal probabilistic judgment.

In recent years, Bayesian statistics have become a standard methodology in environmental science, it has been used to improve spatial predictions of pollutant concentration by Pirani *et al.*, 2018, Amin *et al.*, 2015, Millan *et al.*, 2009, Sahu *et al.*, 2005, among others, and deposition (Cowles *et al.*, 2003). In this study, we propose a Bayesian model that updates the distribution of current and past ground station measurements with the high-resolution output of the PMSS model ($3x3m^2$) within the city of Paris for ten days over the period of March 2016.

In Section 2, the AIRPARIF air quality network and PMSS model simulations are presented. Then, representativeness areas, as obtained in Rodriguez *et al.* (2019), are summarized, and the Bayesian framework is described. Section 3 gathers results: first, the pollutant concentrations at the monitoring site locations (Sect. 3.1) and then, concentrations within representativeness areas (Sect. 3.2). Conclusions are given in section 4.

2 Data and methods

As in Rodriguez *et al.* (2019), a simulation set of ten non-consecutives days in March 2016 are used. During this period (March 2 – 21), a high-pressure episode occurred, characterized by strong atmospheric stability and the accumulation of PM_{10} within the boundary layer. It resulted in a sharp rise in PM_{10} concentration on March 11. The effect of this synoptic-scale episode is less pronounced for NO_X than for PM_{10}. Indeed, in highly urbanized areas, the spatial and temporal variability of pollutant concentration is more driven by traffic patterns for NO_X than for PM_{10}. We note that the contribution of the traffic sector to the total emission concentrations is two times higher for NO_X than for PM_{10} in the Paris area (56% vs. 28% AIRPARIF, 2012). In the close proximity to roads, the NO_X emissions decrease much higher for NO_X (Pasquier *et al.*, 2017). Data from eight local monitoring stations from the AIRPARIF network have been collected for the same period.

2.1 Surface measurements

The location of AIRPARIF monitoring sites have been chosen to sample various urban landscapes in Paris. They are classified as “traffic” or “urban background” depending on their proximity (or not) to the street network. A thorough description of the AIRPARIF network can be found on the agency’s website (https://www.airparif.asso.fr/en/stations/index). Data from three traffic-oriented and five urban background monitoring stations are used in this study as listed in **Table1**. Traffic monitoring stations are located on sidewalks along roads and on major crossroads, whereas urban background monitors are located on low emission areas such as parks, pedestrian squares, schoolyards, and leisure parks. All traffic monitors measure both NO_X and PM_{10}, whereas PM_{10} is measured only at two out of five urban background sites.
These measurements are considered as reference, or “ground truth” of pollutant concentration, assuming that the measurement error is negligible.

Type	ID station	Characteristic features	Pollutants
Traffic-oriented stations	HAUS	Bld Haussmann*, side roads	NO_X, PM_{10}
	OPERA	Crossroad*, road traffic	NO_X, PM_{10}
	ELYS	Champs-Élysées Avenue*: large road bordered with trees, side roads	NO_X, PM_{10}
Urban background stations	PA15L	Stadium* and leisure park, “Boulevard Périphérique”, traffic road	NO_X, PM_{10}
	PA12	Schoolyard*, railways, traffic road	NO_X
	PA13	Park*, traffic road, crossroad	NOX
	PA04C	Place*, park, the Seine River, traffic road	NO_X, PM_{10}
	PA07	Square close to the Eiffel Tower*, Champ-de-Mars Garden, Seine banks, the Seine River, Bridge	NO_X

Table1: Characteristic features of the study areas. Locations marked with an asterisk indicate the location of the monitoring site.

2.2 Numerical air quality simulations

NO_X and PM_{10} concentrations were simulated using the high-resolution model PMSS, in a 12×10 km^2 grid covering the city of Paris. The model set-up is the same as in Rodriguez *et al.* (2019).
The model has thirty-five vertical layers from the ground up to an altitude of 800 m. The thickness of the surface layer is two meters and increases with the altitude (Moussafir *et al.*, 2015). The model simulates the turbulent flow around the buildings and pollutant dispersion through a three-meters horizontal resolution grid. Emission fluxes from urban sources are taken from the AIRPARIF inventory. Boundary conditions are taken from the CHIMERE chemistry-transport model (Mailler *et al.*, 2017). Meteorological conditions are simulated with the MM5 weather mesoscale model (http://www2.mmm.ucar.edu/mm5/). Building contours and heights are also accounted for as obstacles for the air flow. They are taken from the BDTOPO database developed by the French National Geographic Institute. Atmospheric dispersion is simulated with the PSPRAY Lagrangian particles model. Traffic emissions are calculated with the HEAVEN chain (Healthier Environment through the Abatement of Vehicle Emissions and Noise, http://www.airparif.asso.fr/etat-air/air-et-climat-emissions-heaven) HEAVEN calculates in near real-time the traffic circulation and allows obtaining pollutant emissions on main traffic roads of the Paris area based on a “bottom-up” approach. Hourly traffic emissions are derived from the combination of traffic model (car flow, mean speed, cold engines percentage) and emission factors (car speed, road type, temperature, car type, etc.) applied by loop-based counting systems.

At each monitor, simulated concentrations are compared to measurements for both pollutants. PMSS simulations are in a very good agreement with PM_{10} measurements, both for traffic and urban background sites, with a correlation coefficient (R^2) equal to 0.83 and 0.92 respectively, and a root mean squared error (RMSE) equal to 9.57 μg/m^3 and 6.21 μg/m^3 respectively (Rodriguez *et al.*, 2019). Model performance is not as good for NO_X with an R^2 value of 0.65 and 0.45, and RMSE of 81.57 μg/m^3 and 21.12 μg/m^3 for traffic and urban background sites, respectively.

Fig.1 shows scatter plots of concentrations at the ELYS monitoring station. The PMSS model fails to represent NO_X concentration maxima. As discussed in Rodriguez et al., (2019), this may be due to the simple chemical mechanism of the PMSS model that fails to represent the concentration of the highly reactive pollutants such as NO and/or to uncertainties in the NO_X emission derived by the HEAVEN traffic emission modeling chain. Note that the error for NO_2 is smaller than for NO for traffic stations (RMSE(NO_2)=21.9 μg/m^3 vs RMSE(NO)=72.3 μg/m^3). A similar result is obtained for urban background stations.

The model is not always able to represent the high concentration levels during pollution episodes at the studied monitoring sites in Paris due to omissions in emission sources, errors in the meteorological predictions, or inaccuracy of the parameterization. However, it still provides realistic dispersion patterns of pollutant according to other studies during specific measurements campaigns in wind tunnel and read field experiment (Trini Castelli *et al.*, 2018) and for academic project (Hanna *et al.*, 2011). To perform this study, we assume that the model captures the actual spatial variability of pollutant in Paris (Rodriguez *et al.*, 2019).

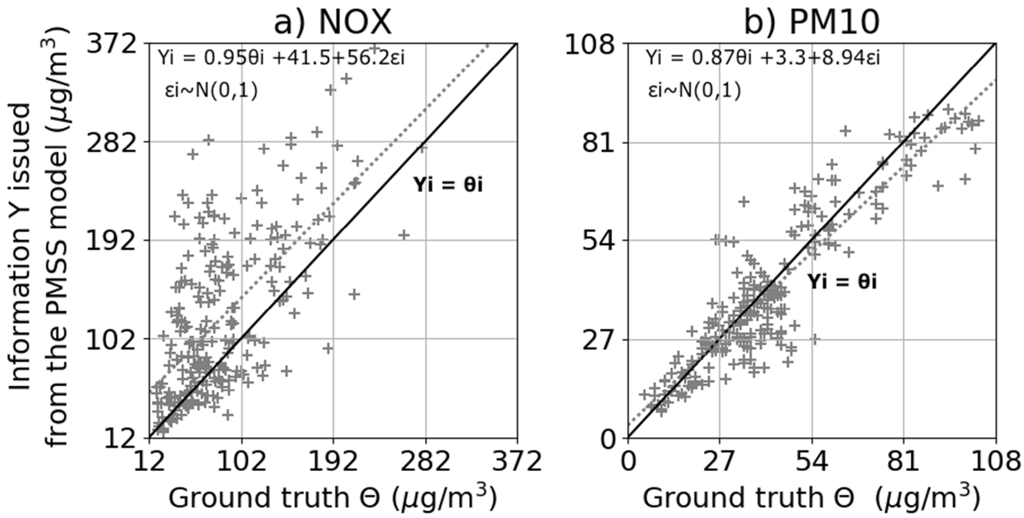


Fig.1: Scatter plots of measured vs simulated hourly concentrations for NO_X (a) and PM_{10} (b) of Champs-Élysées avenue monitoring station during the ten days of March 2016. The black line represents Y=θ. The grey dotted line represents the linear regression between AIRPARIF measurements and PMSS simulations.

2.3 Spatial representativeness areas

PMSS simulations were used in Rodriguez et al. (2019) to determine representativeness areas around the AIRPARIF monitoring sites. The daily spatial representativeness area around a monitoring site was defined as the ensemble of adjacent pixels where the pollutant concentration is (1) strongly correlated with the concentration at the pixel of the station; (2) has a low statistical error (Normalized Root Mean Squared Error) with respect to the concentration at the pixel of the monitoring station. Correlation coefficient and NRMSE thresholds selection are based on an analysis over the 10-day with a step of 0.05 for both criteria. For each threshold, a score is estimated by counting the number of days on which the size of the selected area including the monitoring site is comprised between 300 m^2 and $400 \times 10^3 m^2$.

The maximum score for each site is 10. Finally, a total score is calculated from all stations for the 10 days and reported as the percentage of success. In the case where two different combinations of criteria thresholds give the same score (percentage of success), the retained combination is the more restrictive (i.e. the higher □ correlation coefficient and/or the lowest NRMSE value).

Table2 shows the applied criteria thresholds to select representativeness areas.

Pollutants	ρ	NRMSE
NO_X	0.7	0.45
PM_{10}	0.75	0.3

Table2: Thresholds of representativeness criteria: correlation coefficient (ρ) and Normalized Root Mean Square Error (NRMSE) for NO_X and PM_{10}

The retained thresholds for PM_{10} are more restrictive than for NO_X. This is due to the higher spatial variability of NO_X concentrations, which requires less restrictive thresholds in order to retain areas large enough to get beyond the pseudo-station.

This method is applied independently each day of the simulation to select the representative pixels. A representativeness probability is assigned to each pixel, depending on the frequency (i.e. number of days out of ten) at which each pixel is selected. An example of representativeness areas around the monitoring site of “Avenue des Champs-Elysées” is given in Fig.2. This monitoring station is located on the sidewalk of a wide urban boulevard. NO_X most representative area includes the sidewalk where the instrument is located along a distance of 500 m (yellow to red color, indicating that this area is selected at least 7 days), whereas the dark blue area shows street portions which have been selected only one day. The much larger PM_{10} area among the most probable representative ones (yellow to red, 7 days at least) includes both sidewalks along a distance of 700 m (boulevard portion including the station) as well as a 400-meters segment of sidewalk in northwestern direction from the station.

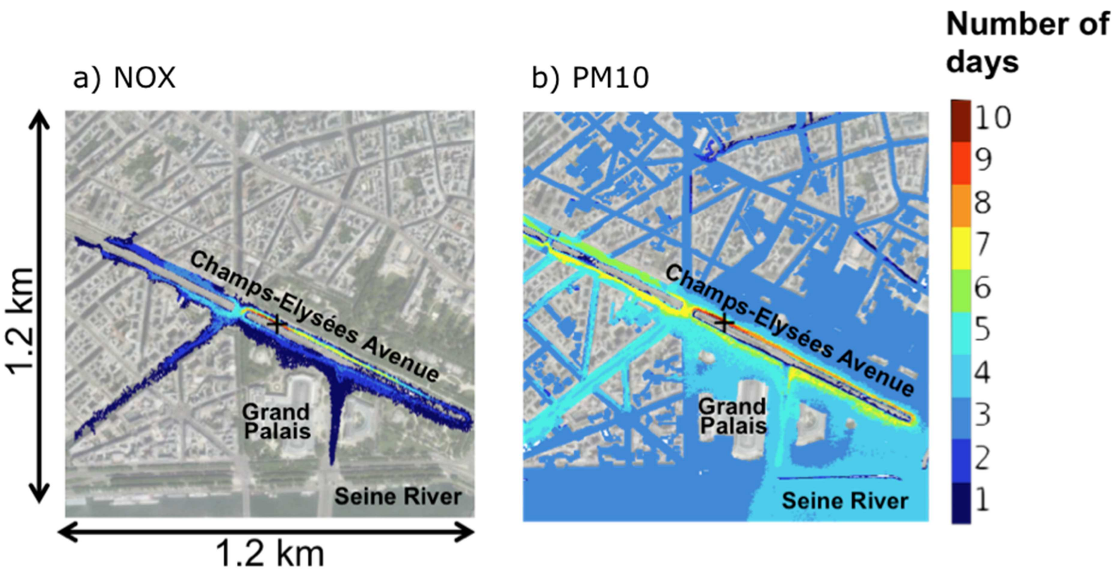


Fig.2: NO_X (a) and PM_{10} (b) representativeness areas cumulated over the 10 days of the study for ELYS monitoring site. The black cross indicates the location of the monitoring site.

2.4 The Bayesian framework

2.4.1 Model formulation

The goal of this study is to improve the PMSS pollutant simulations at, and around, each AIRPARIF station by using past and current station measurements. The scheme of the Bayesian model applied at the monitoring station location is represented with the Directed Acyclic Graph (DAG) of **Fig.3**.

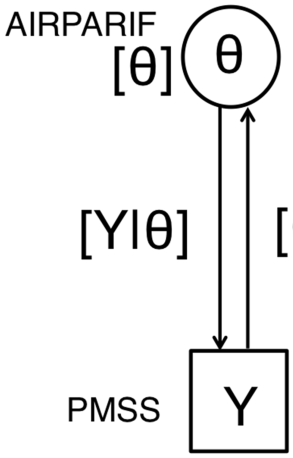


Fig.3: Directed Acyclic Graph (DAG) describing our scientific question with [θ]: prior distribution, [Y|θ]: likelihood, [θ|Y]: posterior distribution.

Past AIRPARIF measurements represent the information on which the prior expert knowledge of the unknown parameter θ (i.e. pollutant concentration at a given point in space and time) is based (before getting the predicted data from the PMSS simulated concentration).

After integrating data information from the PMSS simulation, the prior information (past measurements from the station) is updated through the likelihood function to obtain the posterior knowledge. The likelihood function is the probability distribution of the data (PMSS simulations) conditional on the parameter θ (AIRPARIF measurements to the monitoring station), commonly used in other statistical approaches. Simple linear regression is used as a probabilistic model (PMSS simulated concentrations vs AIRPARIF measurements, **Eq.1**).

$$Yi = a\theta i + b + \sigma\,\varepsilon i \qquad \textbf{(Eq.1)}$$

with error term $\varepsilon i \sim N(0,1)$

where Yi is the initial PMSS simulations (µg/m^3), θi is the AIRPARIF monitoring station measurements (µg/m^3) during the ten days in March 2016 and a, b and σ respectively the slope, intercept, and the standard deviation of the residuals of the linear regression. Parameter values (i.e. a high slope, a low intercept, and standard deviation) for which the likelihood is high, are those that have a high probability of producing the observed data.

Bayes'rule (**Eq.2**) updates the prior knowledge by learning from data (Y:PMSS simulated concentration) and provides a posterior distribution from which the *a posteriori* most probable concentration is estimated, associated with its uncertainty (the posterior standard deviation) (O'Hagan, 2008).

All three, prior, likelihood and posterior are distributions describing the probability that an event occurs. According to the Bayes'rule, the posterior distribution [θ|Y] is the product of the prior distribution [θ] and the likelihood function [YIθ], except for a constant scaling factor (**Eq.2**).

$$[\theta|Y] = \frac{[Y|\theta][\theta]}{[Y]} \quad \textbf{(Eq.2)}$$

Where $[\theta|Y]$ is the posterior distribution, $[Y|\theta]$ is the likelihood, $[\theta]$ prior distribution, $[Y]$ a constant of normalization that we can omit to formulate Bayes'rule by saying that the posterior is proportional to the prior times the likelihood.

The distribution of NO_X and PM_{10} concentrations are close to log-normal (Fig.4a). Neperian logarithm concentrations (Fig.4b) are used instead of concentrations in order to use normal distributions, which are convenient to get an explicit posterior distribution, following the *a posteriori* formula of the conjugate normal model (O'Hagan, 2008) (**Eq.3).**

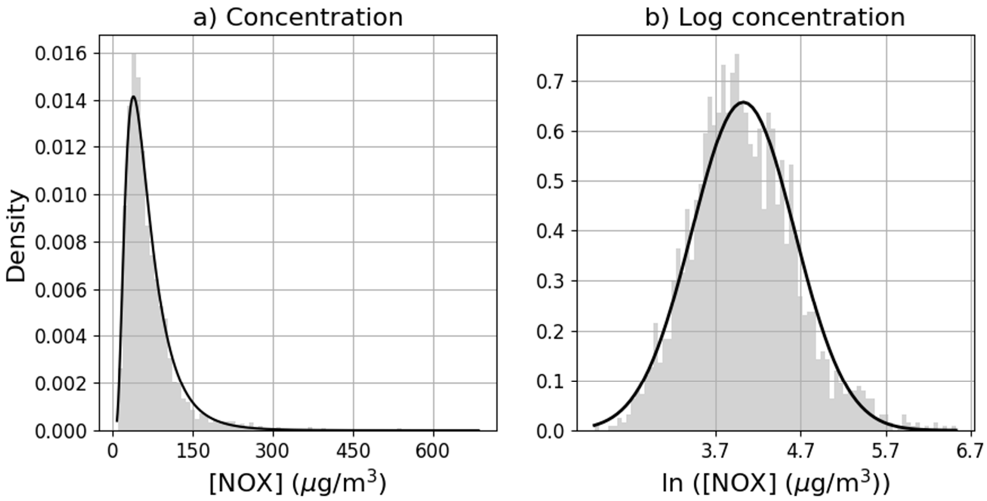


Fig.4: NO_X prior mean concentration of Champs-Élysées avenue monitoring site over ten years (2006 to 2015) at 1 AM during Monday and Friday: (a) concentration and (b) Neperian logarithm concentration.

The posterior distribution is a normal distribution with a mean (m_{y-log}) and a standard deviation (s_{y-log}) obtained from the Bayesian updating formulae **Eq.3**. Model precision is calculated as the inverse of the variance of the *posterior* distribution $\left(\frac{1}{s_{y-log}^2}\right)$. It reflects the degree of uncertainty related to the obtained mean of the estimated parameter.

$$\frac{m_{y-log}}{s_{y-log}^2} = \left(\frac{m_{o-log}}{s_{o-log}^2} + \frac{(Y-\mathrm{b})}{\mathrm{a}}\frac{a^2}{\sigma^2}\right)$$

$$\frac{1}{\mathrm{s}_{\mathrm{y-log}}^2} = \frac{1}{\mathrm{s}_{\mathrm{o-log}}^2} + \frac{a^2}{\sigma^2} \quad \textbf{(Eq.3)}$$

Where m_{o-log} and s^2_{o-log} are respectively the mean and variance of the prior log concentration at the monitoring site's location, Y is the Neperian logarithm of the initial PMSS simulated concentration at a given pixel, a , b and σ^2 are respectively the slope, the intercept, and the variance of the residuals of the linear regression model used as likelihood function (see Section 2.4.2).

At the end of the processing, the mean and standard deviation of the posterior distribution of NO_X and PM_{10} concentration are obtained by applying a log-normal transformation to the model output.

Fig.5 shows an example of the application of Bayes'rule for NOx log-concentration with concentration expressed in µg/m^3 at the OPERA monitoring station at 3 PM on March 2th. It is shown that the initial belief (prior) lies within a value ranging between 4.8 and 6 (Neperian logarithm concentration) with a maximum weight of 5.4. The likelihood values range between 4.4 and 5.3 with a maximal weight of 4.9. The posterior distribution is the combination between prior and likelihood. It is a peaky distribution centered around 5, which appears as a compromise between likelihood and prior information. In this case, data (likelihood) are more informative than the prior because the variance of the prior distribution is high. The posterior mean properly fits the AIRPARIF measurements whereas the PMSS mean initial simulated concentration was biased, at 5.2.

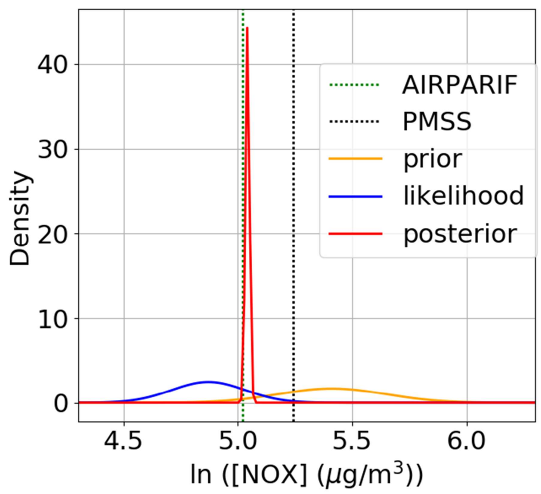


Fig.5: Example of Bayesian update of the NO_X log-concentration at the Opera crossroad monitoring station at 3 PM on March 2, 2016.

2.4.2 Finding the optimal probabilistic model at the monitor location

A major challenge when using Bayesian statistics is to properly choose the prior distribution [θ] and the likelihood function [YIθ].

Prior distribution [θ]
The prior is a distribution of possible values of the unknown parameter θ. It describes the expert knowledge before the observation. In the present case, to build a prior, we must answer the question: which data classification is the best source of information, to predict hourly NO_X and PM_{10} concentration during the ten days of March 2016 at any monitoring station location?

NO_X and PM_{10} concentrations in the center of Paris are strongly correlated to traffic emissions. They also depend on meteorological conditions responsible for their transport and mixing. Consequently, two priors are tested, both based on data collected at the AIRPARIF station in past years: (1) the first one based on traffic variations, by classifying concentration data along three classes: working days, Saturdays, Sundays and public holidays; (2) the second one based on weather conditions, by distinguishing low and high-pressure days, as well as fast-flow situations. Because of the high spatial and temporal variability within the studied domain, different priors for each station and time of the day must be considered.

AIRPARIF monitoring stations have been operating starting from different years: 2006 for PA13, PA12, PA07 and ELYS stations, 2011 for OPERA and PA04C, 2010 for HAUS and 2014 for PA15L.

Fig.6 shows daily mean profiles of NO_X concentration based on ten years of data (2006-2015) at the ELYS monitoring station. The diurnal cycle of mean concentrations is marked with a morning and afternoon peak corresponding to rush hours during working days. During Saturday, Sunday and public holidays, the night peak is more marked than the morning peak which is also shifted later on. The prior with weather classification (not shown) gives similar results for $PM_{10,}$ but does not work properly for NO_X. The prior classification which best predicts NO_X and PM_{10} hourly concentrations for the ten days in March 2016 is, therefore, a classification per day-of-week.

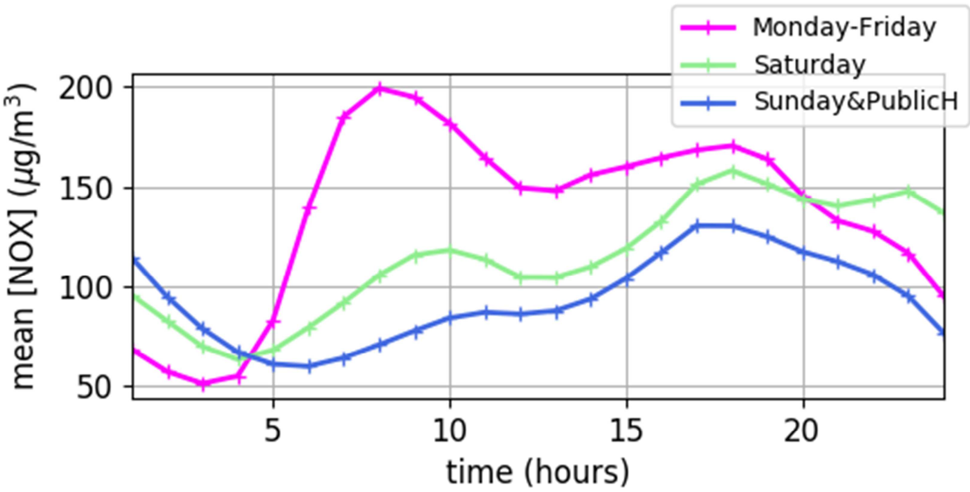


Fig.6: Daily profiles of mean concentration for NO_X prior at the Champs-Élysées Avenue monitoring site. The magenta line represents working days data (Monday to Friday), the green line represents Saturday data and blue line Sunday and Public holidays data.

Likelihood function [YIθ]

The second challenge is the choice of the appropriate likelihood function. This is the probability distribution of the data (PMSS simulation) conditional to the unknown parameter θ (AIRPARIF measurements from the monitoring station), commonly used in most statistical approaches. Through this function, the knowledge about the pollutant concentrations (AIRPARIF measurements from the monitoring station during the studied period) is updated, based on new information (PMSS simulation).

A simple linear regression of PMSS simulated concentrations vs. AIRPARIF measurements is used here as a probabilistic model. A strong constraint is imposed by the small size of our dataset (hourly concentrations during ten days) to define the appropriate classification method. We couple simulated and observed concentration data over time-slots so that at least fifteen points are used for the linear regression **(Eq.1)**. For each station, different time frames are tested to group the hourly data. Time windows of 2, 3, 4, 6, 8 and 12 consecutive hours, and whole day are tested as shown in **Table 3**.

A sensitivity analysis is performed for each type of likelihood and both proposed priors (traffic and weather conditions). NO_X and PM_{10} PMSS simulated concentration at the monitoring station's location are compared with AIRPARIF monitoring station measurements before and after the Bayesian updating. To evaluate the best configuration, we calculate a percentage of Root Mean Square Error (RMSE) difference in each case (**Eq.4**). For the most relevant choice, to complete this first evaluation approach, we also calculate a percentage of standard deviation difference.

$$\%\,RMSE\ difference = \frac{(RMSE_{after\ bayesian\ updating} - RMSE_{initial})}{RMSE_{initial}} * 100 \quad (\textbf{Eq.4})$$

where $RMSE(ref, val) = \sqrt{\frac{1}{n}} (\sum_{t=0}^{t=240}(ref_{(t)} - val_{(t)})^2)$ is the root mean square error, $ref_{(t)}$ is the PM_{10} or NO_X hourly concentration at the AIRPARIF monitoring site. For $RMSE_{initial}$, $val_{(t)}$ is the PM_{10} or NO_X hourly PMSS simulated concentration at the monitoring station location for the ten studied days. For $RMSE_{after\ bayesian\ update}$, $val_{(t)}$ is the PM_{10} or NO_X hourly posterior mean concentration resulting from the Bayesian updating.

Table 3 summarizes results of the sensitivity analysis, associating the previously chosen prior and the time frames tested to construct the likelihood function. The best data classification corresponds to the largest decrease of the %RMSE (the highest negative value), meaning that the Bayesian updating provides less biased concentration estimates compared to the initial PMSS simulation. Note that we only focus on the evaluation of the most probable concentration (posterior distribution mean). The overall best model performance is obtained when gathering the hourly data in 2-hours time frames. Data from all ten days corresponding to each 2-hours time frame are gathered to obtain twenty points for calculating the linear regression.

a) NO_X

STATION		**Initial RMSE [Initial standard deviation (%)]**	**RMSE difference (%) [Standard deviation difference (%)]**						
TYPE	**ID**		Day	**2h**	3h	4h	6h	8h	12h
Traffic	**HAUS**	60.0 [42%]	-49%	**-45% [-31%]**	-43%	-43%	-43%	-41%	-40%
	OPERA	108.6 [50%]	-68%	**-72% [-67%]**	-71%	-70%	-69%	-69%	-67%
	ELYS	67.8 [75%]	-27%	**-29% [-21%]**	-27%	-27%	-27%	-25%	-22%
Urban backg round	**PA12**	20.1 [24%]	-19%	**-31% [-35%]**	-27%	-27%	-26%	-21%	-20%
	PA13	16.6 [33%]	-13%	**-26% [-22%]**	-21%	-22%	-18%	-15%	-14%
	PA04C	20.8 [35%]	-22%	**-30% [-31%]**	-25%	-25%	-28%	-21%	-21%
	PA07	25.5 [36%]	-27%	**-17% [-12%]**	-14%	-15%	-11%	-13%	-12%
	PA15L	23.3 [45%]	-17%	**-26% [-24%]**	-22%	-25%	-12%	-19%	-13%

b) PM_{10}

STATION		**Initial RMSE [Initial standard deviation**	**RMSE difference (%) [Standard deviation difference (%)]**						
TYPE	**ID**		Day	**2h**	3h	4h	6h	8h	12h

		(%)]							
Traffic	HAUS*	8.6 [23%]	12%	**-12% [-13%]**	-14%	-12%	-7%	-10%	-5%
	OPERA	10.5 [30%]	-30%	**-38% [-49%]**	-36%	-36%	-37%	-34%	-32%
	ELYS	9.6 [19%]	12%	**-10% [-8%]**	-8%	-8%	-5%	-7%	-4%
Urban backg round	PA04C	5.1 [13%]	4%	**-1% [-8%]**	3%	2%	5%	3%	5%
	PA15L	7.1 [21%]	19%	**-8% [-19%]**	-3%	-4%	-4%	1%	3%

Table3: % RMSE and standard deviation (in brackets) difference between simulated and observed data before and after the Bayesian updating for several time frames of classification of the hourly data (per day, by 2,3,4,6,8 and 12-hours period for the ten days). [HAUS* - PM_{10} results are averaged on 8/10 days, due to the lack of AIRPARIF measurements for two days].

2.4.3 Expanding the measured value within the representative area

We applied the previous Bayesian model, for each monitoring station and each hour of the day, to each pixel of the representativeness areas (see Section 2.3), by accounting for the probability of each pixel to belong to this area.

The Bayesian model update is performed on every pixel within the representativeness area, using the initial PMSS simulated concentration at the given pixel and with the selected prior and likelihood function at the station (Section 2.4.2). The fraction of the number of days, for which the pixel is selected to belong to the representativeness areas on the total number of days of the study, is used as weighting coefficient (**Eq. 5**). Thus, for pixels within a representativeness area selected ten days out of ten, we apply a weight equal to one. In this case, the concentration at the pixel is the posterior mean concentration at the given pixel (i.e. concentration after the Bayesian updating). For pixels selected only one day out of ten the final concentration will be calculated as the sum of 1/10th of the posterior mean concentration at the pixel and 9/10th of the initial PMSS simulated concentration at the pixel.

$$C_{final} = a * C_{after\ bayesian\ updating} + (1-a) * C_{initial} \quad \textbf{(Eq5)}$$

where C_{final} is the new concentration estimate at the pixel, a the weighting coefficient, $C_{after\ bayesian\ updating}$ the enhanced concentration estimate at the pixel and $C_{initial}$ the PMSS simulated concentration at the pixel.

We note here that the weighting coefficient used to expand the measured value in the representativeness areas depends on representativeness criteria statistics (daily correlation coefficient and NRMSE) calculated from the ten-day period of the study.

Given the shortness of the time period and the random nature of meteorology, we expect to underestimate and/or misplace the daily variability of the concentration field. However, by taking the most probable representativeness area (representativeness area shared seven days on ten) instead of daily representativeness area, the uncertainty on daily representative area and thresholds effects are minimized.

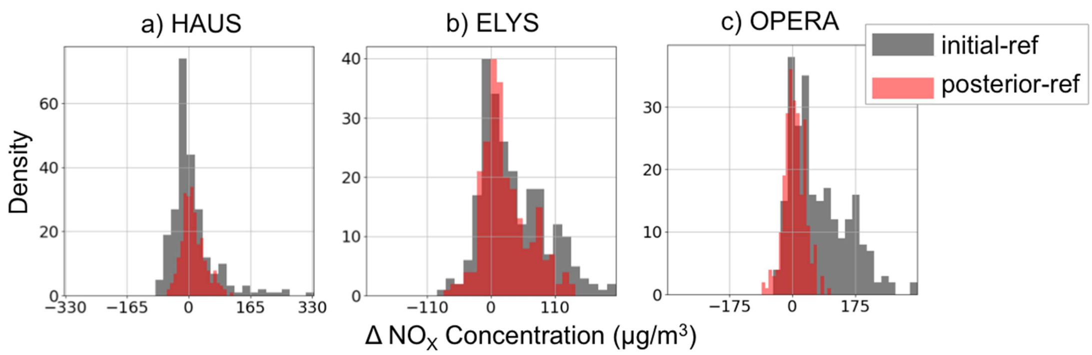
The methodology should be applied on longer study periods covering a larger variety of atmospheric conditions, to reduce uncertainties. Representativeness areas could then be defined based on dispersion pattern regimes. Alternative approaches include geostatistical kriging such as in Beauchamp *et al.* (2018c).

3 Results

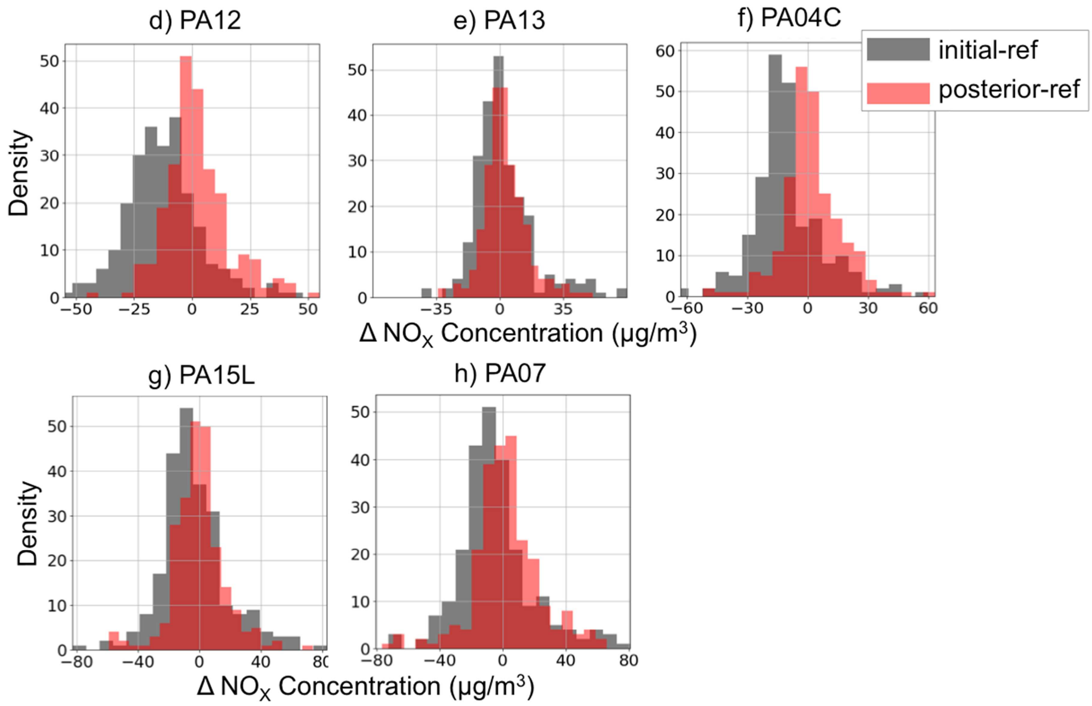
3.1 Pollutant concentration estimates at the monitoring site location

As shown in **Table 3**, concentration estimates by Bayesian updating present in almost all cases a lower RMSE than the initial PMSS simulated hourly concentrations. For both NO_X and PM_{10} pollutants, this improvement is higher for traffic stations (up to 72% decrease in RMSE at the OPERA monitor for NO_X) than for urban stations (31% at PA12 for NO_X). As shown in the histograms of **Fig.7**, the distribution of the PMSS simulation bias with respect to the monitoring station is not normal (grey histograms). This suggests the presence of systematic errors in the PMSS simulation. Especially for traffic stations, and for both pollutants, we find extreme values, suggesting high model overestimation of the measured value. After the Bayesian updating, the bias distribution is closer to a normal distribution with errors centered around zero. This shows that the applied methodology is an efficient bias-correction method for the PMSS model.

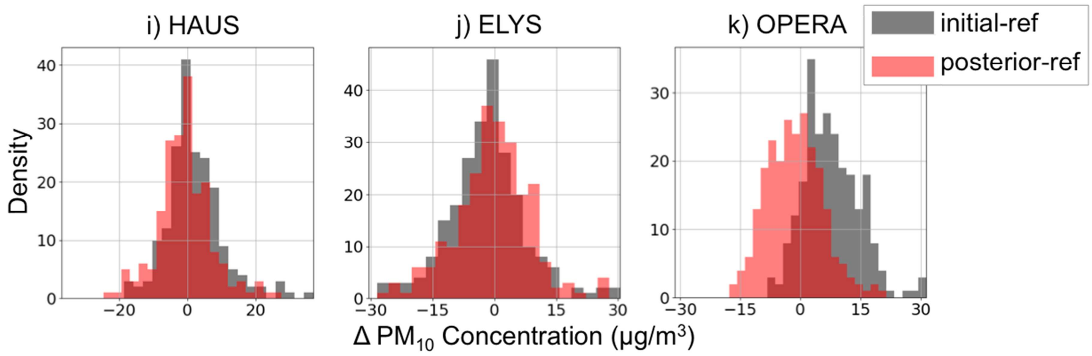
NO_X traffic stations



NO_X urban background stations



PM_{10} traffic stations



PM_{10} urban background stations

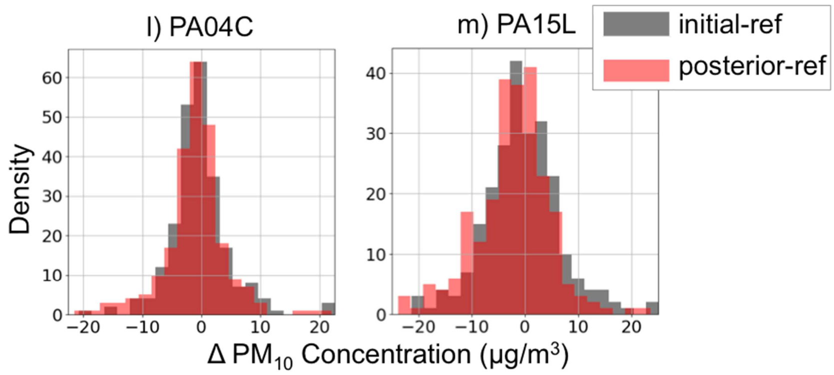


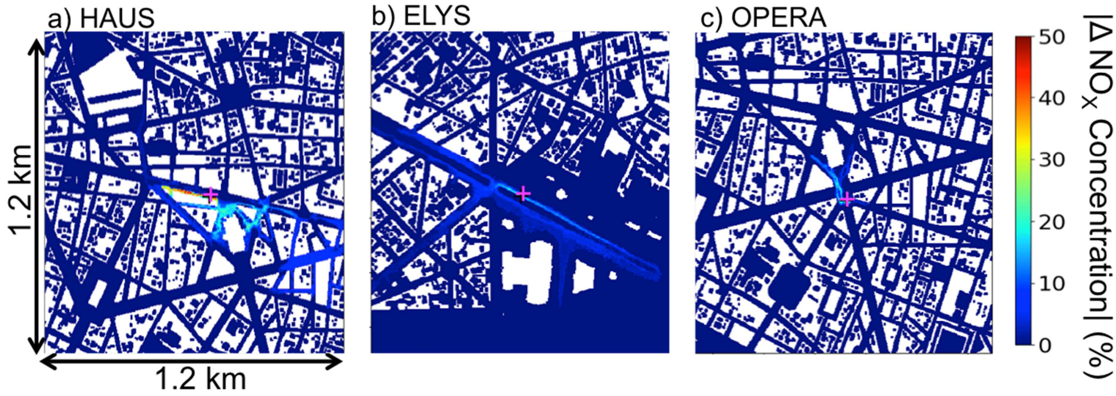
Fig.7: Distribution of the bias of initial PMSS simulated concentrations (grey) and of the posterior mean concentration (red), with respect to the monitoring station measurements.

3.2 Pollutant concentration estimates in the representativeness areas

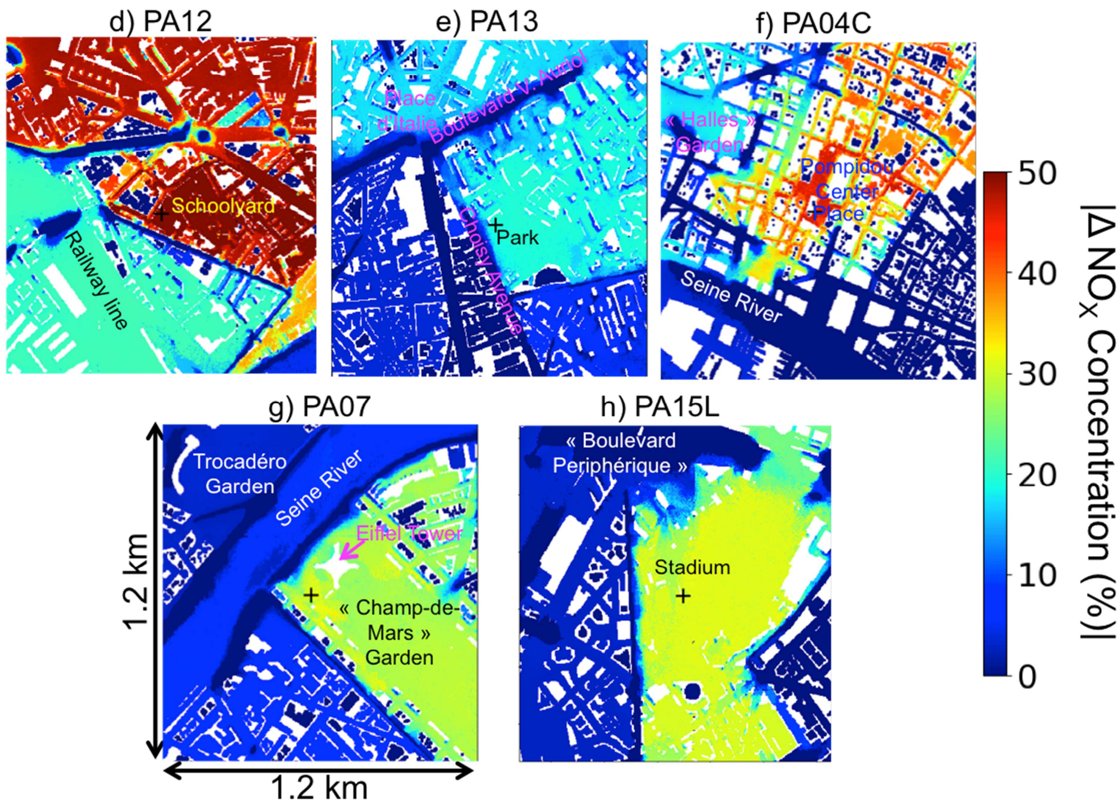
The same Bayesian model (same prior and likelihood) is applied to the pixels of the representativeness areas in the vicinity of the monitoring station to update NO_X and PM_{10} PMSS simulated concentrations. To account for the probability of each pixel to belong in the representativeness area, the weighting function is applied to the posterior distributions to provide the final concentration estimates (see Section 2.4.3).

Maps of the differences between the updated concentration estimates and the initial PMSS simulation within the representativeness areas are shown in **Fig.8** for NO_X and PM_{10} pollutants respectively. For NO_X, around traffic stations, a small number of pixels are modified, due to the small size of the representativeness area.
Changes can be seen in larger areas around the urban stations. The change in PM_{10} concentrations is smaller than in NO_X concentrations with a maximal correction of about 15% for PM_{10} vs. 50% for NO_X. As mentioned in Section 2.2, this is most probably due to the better PMSS model performance for PM_{10} than for NO_X.
Moreover, for NO_X, around traffic stations, we only find small changes in the initial PMSS concentration (by 20-30% in average) whereas around urban background stations, up to 50% changes are observed on large areas for PA12 and PA04C. For PM_{10}, the highest modification is obtained for PA15L with 15% change and around the three traffic stations within a small-sized area surrounding the station.

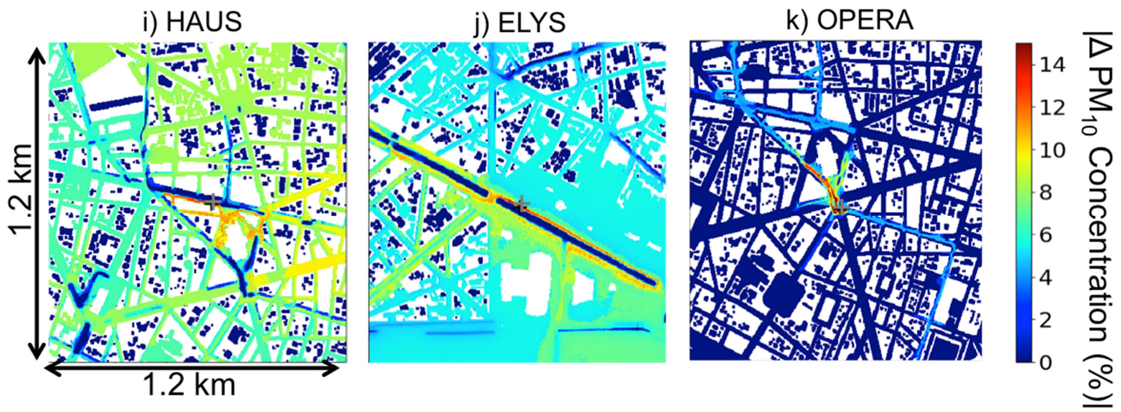
NO_X traffic stations



NO_X urban background stations



PM_{10} traffic stations



PM_{10} urban background stations

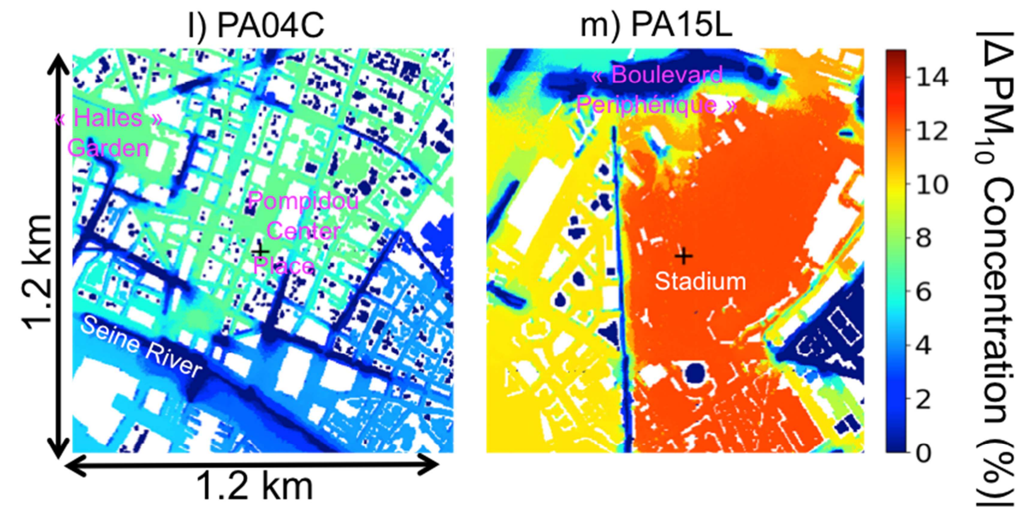


Fig.8: Percentage difference in PM_{10} concentrations between the updated estimate and the initial PMSS simulation averaged over the ten days of the study. Crosses indicate the location of the monitoring sites.

Table4 shows results obtained within the most probable representativeness area around each monitoring station, i.e. the area selected at least seven days out of ten (see Rodriguez *et al.*, 2019). The results are given, as in Section 2.4.2, in terms of percentage decrease of the RMSE and standard deviation values between the updated concentration estimate and the initial PMSS simulated concentration with respect to the AIRPARIF measurement (**Eq.4**). RMSE and standard deviation shown in **Table4** are spatially averaged across the pixels of the most probable representativeness areas.

NO_X concentrations after Bayesian updating are improved in all cases, with a maximal correction for the OPERA station (-46% for RMSE difference and - 39% for the standard deviation difference).

PM_{10} concentrations are also closer to the measured values after Bayesian updating within the most probable representativeness area by considering RMSE and standard deviation (**Table4.b**).

a) NO_X

TYPE STATION	ID STATION	Initial RMSE [Initial standard deviation (%)]	RMSE difference (%) [Standard deviation difference (%)]
Traffic	HAUS	72.0 [48%]	-37% [-33%]
	OPERA	105.5 [47%]	-46% [-39%]
	ELYS	64.0 [71%]	-20% [-14%]
Urban background	PA12	20.4 [34%]	-21% [-25%]
	PA13	18.5 [36%]	-21% [-19%]
	PA04C	25.3 [37%]	-17% [-22%]
	PA07	26.6 [36%]	-19% [-22%]
	PA15L	24.6 [48%]	-22% [-19%]

b) PM_{10}

TYPE STATION	ID STATION	Initial RMSE [Initial standard deviation (%)]	RMSE difference (%) [Standard deviation difference (%)]
Traffic	HAUS	25.3 [52%]	-8% [-12%]
	OPERA	10.0 [26%]	-18% [-24%]
	ELYS	10.5 [21%]	-12% [-13%]
Urban background	PA04C	5.5 [13%]	-5% [-8%]
	PA15L	7.2 [21%]	-10% [-18%]

Table4: Average RMSE, standard deviation and % RMSE, and standard deviation differences between simulated and observed NO_X (a), and PM_{10} (b) concentrations, before and after Bayesian updating, spatially averaged within the most probable representativeness area, for the ten days of the study.

4 Conclusion

In this study, our main goals are (1) to improve NO_X and PM_{10} simulated concentrations over AIRPARIF monitoring stations and (2) to extend this enhanced concentration estimate within representative areas at the vicinity of the station.
We show that the bias in the ten-day PMSS simulation with respect to the AIRPARIF measurements does not follow a normal distribution. Uncertainties due to the emission inventory, meteorological conditions and model parameterizations lead to systematic errors. Bayesian statistics is especially appropriate to handle model uncertainties and provide bias correction in such cases.
The proposed Bayesian model combines PMSS model simulations with current and past surface pollutant concentration measurements of the AIRPARIF stations (3 traffic-stations, 2 and 5 urban background stations respectively for PM_{10} and NO_X). Combination of these two sources of information results in PMSS model error reduction at the station location and provides a spatially resolved concentration estimate in the vicinity of the monitoring site.

The most probable NO_X and PM_{10} concentrations at the monitoring station location are given by the Bayesian posterior distribution. A sensitivity analysis is performed to find the optimal probabilistic model at each station, and determine the parameters of the prior and likelihood distributions. Past hourly AIRPARIF pollutant concentration measurements are used to establish the prior distribution for each monitoring station. Two different data classifications, based either on the intensity of traffic circulation or meteorological conditions, are tested. The classification per weekday, accounting for traffic circulation appears more appropriate. The hourly data were grouped in several time frames for the linear regression in order to test different likelihood functions. The two-hour period likelihood is found to give the best results by reducing RMSE and standard deviation between simulated concentrations and AIRPARIF measurements. For example, at the OPERA crossroad site, a -72% difference in the RMSE is obtained for NO_X concentrations.

We spatially extend the updated concentration estimates within the representativeness areas of each monitoring site, by applying a weighting function that considers the probability of each pixel to belong to the area. We propose to estimate the updated concentration at each pixel, by taking a weighted average between the posterior mean concentration and the initial PMSS simulation at the given pixel. The fraction of the number of days for which the pixel is selected to belong to the representativeness areas on the total number of days of the study is used as the weighting coefficient.
Results show that final NO_X and PM_{10} concentration estimates within the most probable representative area (pixels selected seven days among ten) are closer to AIRPARIF measurements than the initial PMSS simulation with a reduced error (decrease of % RMSE and of % standard deviation).
Bias correction is larger for NO_X concentrations than for PM_{10} because the initial PMSS model error is smaller for PM_{10}. Modifications are observed over larger areas around the urban stations than around traffic ones due to the size of the representativeness area.
The Bayesian model developed in this study is an innovative and low computational cost method to spatially extend pollutant concentration measurements in the vicinity of the station. By providing low-bias high-resolution pollutant concentration estimates over urban areas, the method could contribute to a better assessment of human

exposure to atmospheric pollution. This method should be further validated, by performing local measurements inside representativeness areas.

Finally, longer PMSS simulations would increase the available dataset for the linear regression, leading to more robust likelihood functions. In particular, the impact of the specific air pollution episode with a sharp PM_{10} increase in the whole region would be attenuated in favor of more general statistics reflecting the baseline conditions.

Acknowledgments:
Acknowledgments to ARIA Technologies for providing the PMSS simulations for ten days.

References / bibliography

Amin, N.A.M., Adam, M.B., Aris, A.Z., 2015. Bayesian Extreme for Modeling High PM10 Concentration in Johor. Procedia Environmental Sciences, Environmental Forensics 2015 30, 309–314. https://doi.org/10.1016/j.proenv.2015.10.055

Beauchamp, M., de Fouquet, C., Malherbe, L., 2017. Dealing with non-stationarity through explanatory variables in kriging-based air quality maps. Spatial Statistics 22, 18–46. https://doi.org/10.1016/j.spasta.2017.08.003

Beauchamp, M., Malherbe, L., de Fouquet, C.,and Létinois, L, 2018. A necessary distinction between spatial representativeness of an air quality monitoring station and the delimitation of exceedance areas. Environmental Monitoring and Assessment, 190(7):441, Jun 2018c. ISSN 1573-2959. https://doi.org/10.1007/s10661-018-6788-y.

Beloconi, A., Chrysoulakis, N., Lyapustin, A., Utzinger, J., Vounatsou, P., 2018. Bayesian geostatistical modelling of PM10 and PM2.5 surface level concentrations in Europe using high-resolution satellite-derived products. Environment International 121, 57–70. https://doi.org/10.1016/j.envint.2018.08.041

Boreux, J.J., Parent, E., Bernier, J., 2010. Pratique du calcul bayésien, Statistique et probabilités appliquées. Springer-Verlag.

Cowles, M.K., Zimmerman, D.L., 2003. A Bayesian space□time analysis of acid deposition data combined from two monitoring networks. Journal of Geophysical Research: Atmospheres 108. https://doi.org/10.1029/2003JD004001

Elbern, H.,and Schmidt, H., 2001. Ozone episode analysis by four-dimensional variational chemistry data assimilation, J. Geophys. Res., 106, No. D4, 3569-3590. https://doi.org/10.1029/2000JD900448

Fuentes, M., Raftery, A.E., 2005. Model Evaluation and Spatial Interpolation by Bayesian Combination of Observations with Outputs from Numerical Models. Biometrics 61, 36–45. https://doi.org/10.1111/j.0006-341X.2005.030821.x

Hanea, R.G., Velders G.J. and Heemink, A., 2004. Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model J. Geophys. Res 109, D10302. doi:10.1029/2003JD004283

Hanna, S., White, J., Trolier, J., Vernot, R., Brown, M., Gowardhan, A., Kaplan, H., Alexander, Y., Moussafir, J., Wang, Y., Williamson, C., Hannan, J., Hendrick, E., 2011. Comparisons of JU2003 observations with four diagnostic urban wind flow and Lagrangian particle dispersion models. Atmospheric Environment 45, 4073–4081. https://doi.org/10.1016/j.atmosenv.2011.03.058

Harrison, R.M., 2018. Urban atmospheric chemistry: a very special case for study. npj Climate and Atmospheric Science 1, 5. https://doi.org/10.1038/s41612-017-0010-8

Janssen, S., Dumont, G., Fierens, F., Deutsch, F., Maiheu, B., Celis, D., Trimpeneers, E., Mensink, C., 2012. Land use to characterize spatial representativeness of air quality monitoring stations and its relevance for model validation. Atmospheric Environment 59, 492–500. https://doi.org/10.1016/j.atmosenv.2012.05.028

Mailler, S., Menut, L., Khvorostyanov, D., Valari, M., Couvidat, F., Siour, G., Turquety, S., Briant, R., Tuccella, P., Bessagnet, B., Colette, A., Létinois, L., Markakis, K., Meleux, F., 2017. CHIMERE-2017: From urban to hemispheric chemistry-transport modeling. Geoscientific Model Development. 10. 2397-2423. 10.5194/gmd-10-2397-2017.

Martín, F., Palomino, I. and Vivanco M.G. ,2012. Combination of measured and modelling data in air quality assessment in Spain. Int. J. Environment and Pollution, Vol. 49, Nos. 1/2, 36-44. https://doi.org/10.1504/IJEP.2012.049773

McMillan, N.J., Holland, D.M., Morara, M., Feng, J., 2010. Combining numerical model output and particulate data using Bayesian space–time modeling. Environmetrics 21, 48–65. https://doi.org/10.1002/env.984

Moussafir, J., Olry, C., Nibart, M. Albergel, A., Armand, P., Duchenne, C., Mahé, F.,Thobois, L., Loaëc, S., Oldrini, O., 2014. AIRCITY: A Very High Resolution Atmospheric Dispersion Modeling System for Paris. American Society of Mechanical Engineers, Fluids Engineering Division (Publication) FEDSM. 1. 10.1115/FEDSM2014-21820.

O'Hagan, A., 2008. The Bayesian approach to statistics. In Rudas, T. *Handbook of probability: Theory and applications* (pp. 85-100). Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781452226620

Pasquier, A.,and Ré, M., 2017. Considering criteria related to spatial variabilities for the assessment of air pollution from traffic. Transportation Research Procedia, World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016 25, 3354–3369. https://doi.org/10.1016/j.trpro.2017.05.210

Pirani, M., Gulliver, J., Fuller, G.W., Blangiardo, M., 2014. Bayesian spatiotemporal modelling for the assessment of short-term exposure to particle pollution in urban areas. J Expo Sci Environ Epidemiol 24, 319–327. https://doi.org/10.1038/jes.2013.85

Rodriguez, D., Valari, M., Payan, S., Eymard, L., 2019. On the spatial representativeness of NO_X and PM_{10} monitoring-sites in Paris, France. *Atmospheric Environment: X.* https://doi.org/10.1016/j.aeaoa.2019.100010

Sahu, Sujit K., and Kanti V. Mardia. A Bayesian Kriged Kalman Model for Short-Term Forecasting of Air Pollution Levels. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 54, no. 1, 2005: 223-44. http://www.jstor.org/stable/3592609

Son, Y., Osornio-Vargas, Á.R., O'Neill, M.S., Hystad, P., Texcalac-Sangrador, J.L., Ohman-Strickland, P., Meng, Q., Schwander, S., 2018. Land use regression models to assess air pollution exposure in Mexico City using finer spatial and temporal input parameters. Science of The Total Environment 639, 40–48. https://doi.org/10.1016/j.scitotenv.2018.05.144

Trini Castelli, S., Armand, P., Tinarelli, G., Duchenne, C., Nibart, M., 2018. Validation of a Lagrangian particle dispersion model with wind tunnel and field experiments in urban environment. Atmospheric Environment 193, 273–289. https://doi.org/10.1016/j.atmosenv.2018.08.045

Wu, C.-D., Zeng, Y.-T., Lung, S.-C.C., 2018. A hybrid kriging/land-use regression model to assess PM2.5 spatial-temporal variability. Science of The Total Environment 645, 1456–1464. https://doi.org/10.1016/j.scitotenv.2018.07.073