Meta-modeling of a simulation chain for urban air quality

Janelle K K Hammond, Ruiwei Chen, Vivien Mallet

Meta-modeling of a simulation chain for urban air quality

J.K. Hammond[a,*], R. Chen[b], V. Mallet[a]

[a] *INRIA, 2 Rue Simone IFF, 75012 Paris, France.*
[b] *CEREA, Joint Laboratory of Ecole des Ponts ParisTech - EDF R&D, Université Paris-Est, 6-8, Avenue Blaise-Pascal, Cité Descartes, Champs-sur-Marne, 77455 Marne-la-Vallée CEDEX 2, France*

Abstract

Urban air quality simulation is an important tool to understand the impacts of air pollution. However, the simulations are often computationally expensive, and require extensive data on pollutant sources. This data can be obtained through sparse measurements, or through traffic simulation. Modeling chains combine the simulations of multiple models to provide the most accurate representation possible, however the need to solve multiple models for each simulation increases computational costs even more. In this paper we construct a meta-modeling chain for urban atmospheric pollution, from dynamic traffic modeling to pollutant dispersion-reaction. Reduced basis methods (RBM) aim to compute a cheap and accurate approximation of a physical state using approximation spaces made of a suitable sample of solutions to the model. One of the keys of these techniques is the decomposition of the computational work into an expensive one-time offline stage and a low-cost parameter-dependent online stage. Traditional RBMs require modifying the assembly routines of the computational code, an intrusive procedure which may be impossible in cases of operational model codes. We propose a non-intrusive reduced order scheme, and study its application to a full chain of operational models. Reduced basis are constructed using principal component analysis (PCA), and the concentrations fields are approximated as projections onto this reduced space. We use statistical emulation to approximate projection coefficients in a non-intrusive manner. We apply a multi-level meta-modeling technique to a chain using the dynamic traffic assignment model LADTA, the emissions database COPERT IV, and the Gaussian dispersion-reaction air quality model SIRANE to a case study on the city of Clermont-Ferrand with over 45000 daily traffic observations, a 47000-link road network, a simulation domain covering 180km^2, and assess the results using hourly NO_2 concentration observations measured at stations in the agglomeration. We reduce computational times from nearly 3 hours per simulation to under 0.1 second, while maintaining accuracy comparable to the original models. The low cost of the meta-model chain and its non-intrusive character demonstrate the versatility of the method, and the utility for long-term or many-query air quality study.

Keywords: Model order reduction, Surrogate model, meta-model, Reduced basis methods, Gaussian dispersion model, Air quality simulation, Modeling chain

∗. Corresponding author
Email address: janelle.hammond@inria.fr (J.K. Hammond)

1. Introduction

Air quality simulations at urban scale are a key tool for the evaluation of population exposure to particulate matter and gaseous air pollutants. The simulations are however subject to costly computational requirements and complicated implementation. Studies in exposure estimation or uncertainty quantification, for example, require many solutions to the model. Advanced models can be rendered feasible in this context if we can reduce the computational cost without significant loss of accuracy.

Let us consider a generic stationary model over a physical domain $\Omega \subset \mathbb{R}$ and parameter domain $\mathcal{D} \subset \mathbb{R}^{N_p}$

$$\mathcal{M} : \Omega \times \mathcal{D} \to \mathbb{R}^{\mathcal{N}}$$

$$\mathbf{p} \mapsto c(\mathbf{p})$$

The model output for a given parameter vector $\mathbf{p} \in \mathcal{D}$, $c(\mathbf{p}) \in \mathbb{R}^{\mathcal{N}}$, will be a large-dimension vector representing the solution over a grid coverin Ω. $\mathcal{M}$ can represent various types of atmospheric pollution models, from highly complex formulations based on partial differential equations and fluid dynamics [1, 2] to simpler, and more commonly operational, formulations such as Gaussian dispersion models. Even in the case of the (comparatively) simpler models, the computational time necessary for the solution of $\mathcal{M}$ in practical applications over large domains with many parameters (e.g. emissions sources) can be high. This would make numerous solutions to the model too costly in practice. Methods of Model Order Reduction (MOR) can reduce computational costs without introducing significantly increased model error, and for a range of varying parameters $\mathbf{p} \in \mathcal{D}$.

Various MOR techniques have been studied in the context of air quality models (AQMs). In [3] the metamodeling technique described in section 2 was tested on pollutant concentration fields over Clermont-Ferrand approximated by the ADMS-Urban model [4] using daily profiles for traffic emissions. In [5], statistical emulation was used to evaluate the sensitivity of some input parameters on a global aerosol model. A Gaussian process emulation was used for the study of model uncertainty in [6] for accidental release scenarios. Gaussian process emulation was also used in [7] for the Sobol' sensitivity analysis of a dispersion model representing the Fukushima event.

In this paper, we will consider a modeling chain for air quality modeling over the agglomeration of Clermont-Ferrand and surrounding area in France. Air quality models are known to commit significant errors [8, 9, 1, 10], however these errors are strongly dependent on the calibration and inputs to the model. Providing more precise input data, such as data on pollutant emissions from road traffic, can greatly improve the accuracy of the modeled concentration field. The advantage of a modeling chain is the use of the best (most precise) information available on various inputs by using traffic and emissions models. In [8], the authors provide a review of modeling chain techniques for traffic pollutant emissions, atmospheric dispersion, and effects on water quality.

The modeling chain studied here consists of the dynamic traffic assignment model LADTA [11, 12], an emissions model Pollemission [13] based on COPERT-IV emissions database [14], and a Gaussian AQM, Sirane [15]. The

computation of a pollutant concentration field over the agglomeration for any given time requires the solution of each model in the chain, which proves costly for long time periods.This brings us back to MOR techniques. However in this case, we have a chain ofmultiple models to reduce, which leads us to questions on the implementation of MOR techniques : could a single reduction over the full chain be feasible, or will satisfactory results require a chain of meta-models ? How can we treat the large parameter dimension of the chain ?

We resort to projection-based MOR techniques based on Reduced Basis (RB) [16] to construct cheap and accurate meta-models. A projection-based meta-model for the dynamic traffic model was built in [17]. Here we will complete the model chain with the conversion between traffic assignment and emissions model outputs and pollutant dispersion model inputs, then construct a meta-model for the AQM to build a low-cost meta-model chain for the entire system. The motivation for this choice will be explained in section 3.

In section 2, we will describe the meta-modeling technique based on RB methods. In section 3, we will describe the case study over Clermont-Ferrand : input and measurement data, computational domain, and selected models. In section 4, we will summarize the results of the meta-model on the AQM chain, studying accuracy, precision, and computational savings.

2. meta-modeling Methods

Computation times for large problems are commonly on the order of hours, making many-query contexts, such as sensitivity analysis and optimization, hardly feasible. Model reduction methods are of great interest to applications of parametrized problems involving many-query or real-time study. We will rely on a projection-based method of model order reduction in which the output solution space to the model $\mathcal{X} = \{c(\mathbf{p})|\mathbf{p} \in \mathcal{D}\} \subset \mathbb{R}^{\mathcal{N}}$, where the parameter dimension is N_p, is represented by a reduced basis of small dimension. While the model output is of high dimension $\mathcal{N}$, the reduced order solution will be of dimension $N \ll \mathcal{N}$. We will begin here by detailing the MOR method as applied to the AQM part of the chain, and we will discuss the details of the full meta-model chain in section 3.

2.1. Reduced Basis Method

Let us consider a model, or model chain, $\mathcal{M}$ which takes input parameter vector $\mathbf{p} \in \mathcal{D} \subset \mathbb{R}^{N_p}$ and computes an output vector $c(\mathbf{p})$ over a grid of $\mathcal{N}$ points. Our objective is to construct a reduced basis $\{\Psi_n^{AQ}\}_{1 \leq n \leq N}$ of N basis functions approximating the concentration solution space $\mathcal{X}$ such that the projection of any simulated state, $\Pi_N c(\mathbf{p})$, onto the reduced basis is sufficiently precise. The basis representing atmospheric concentration fields will be denoted by AQ (air quality). To construct a RB we first need to sample a large number of solutions in $\mathcal{X}$. This so-called *training* set should represent the variability in the solution states. We will sample the solution space by Latin Hypercube Sampling (LHS). Then we will construct the RB by Principal Component Analysis (PCA).

We use LHS to select N_{train} sample points $(\mathbf{p}_1, \ldots, \mathbf{p}_{N_{train}})$ in the parameter domain $\mathcal{D}$, and compute model simulations from each point to build the training ensemble $\mathbf{Y}^{AQ} = [c(\mathbf{p}_1), \ldots, c(\mathbf{p}_{N_{train}})]$ to train the model reduction. As is common practice in PCA applications, we will first compute the ensemble mean $\bar{c} = \frac{1}{N_{train}} \sum_{i=1}^{N_{train}} c(\mathbf{p}_i)$ of the training ensemble. PCA is computed on the centered ensemble $\bar{\mathbf{Y}}^{AQ} = [c(\mathbf{p}_1) - \bar{c}, \ldots, c(\mathbf{p}_{N_{train}}) - \bar{c}]$. The eigenvalues $\{\lambda_k\}_{1 \leq k \leq \mathcal{N}}$ and eigenvectors $\{\Psi_k^{AQ}\}_{1 \leq k \leq \mathcal{N}}$ of the covariance matrix $\bar{\mathbf{C}}^{AQ} = \left(\bar{\mathbf{Y}}^{AQ}\right)^T \left(\bar{\mathbf{Y}}^{AQ}\right)$ of the training ensemble are such that

$$\sum_{i=1}^{N_{train}} \left\| c(\mathbf{p}_i) - \bar{c} - \sum_{n=1}^{N} \Psi_n^{AQ} \Psi_n^{AQ^T} (c(\mathbf{p}_i) - \bar{c}) \right\|_2^2 = \sum_{k=N+1}^{\mathcal{N}} \lambda_k, \tag{1}$$

for eigenvalues λ arranged in decreasing order. N principle component basis functions ψ_n^{AQ} are selected to represent $I_N = 98\%$ of the variability in the concentration state. This means that the error of projecting any member of the training ensemble onto the basis, Err_N, will be bounded by the tolerance $Err_N \leq \epsilon_N = \sqrt{1 - I_N}$. For any new parameter, we can thus represent the solution as

$$c(\mathbf{p}) \simeq \Pi_N c(\mathbf{p}) = \bar{c} + \sum_{n=1}^{N} \alpha_n^{AQ} \Psi_n^{AQ} \tag{2}$$

with projection coefficients $\alpha_n^{AQ} = \Psi_n^{AQ^T} (c(\mathbf{p}) - \bar{c})$.

2.2. Statistical Emulation

Once we have constructed the reduced basis by PCA, we need a reduced order modeling scheme to approximated new solutions. Classical reduced basis methods which replace the approximation space with the reduced basis space are intrusive and require the modification of the computational code. We would like to use a non-intrusive method which can be applied to a black-box model or model chain, which is particularly pertinent in the context of operational models. We consider meta-modeling by the emulation of projection coefficients α_n^{AQ}, $1 \leq n \leq N$. First we select a linear *trend*, which will be a least squares regression $\mathcal{R}_n(\mathbf{p}) = \sum_{k=1}^{N_p} \beta_{n,k} p_k$, calculated from the training simulations $\{c(\mathbf{p}_i)\}_{1 \leq i \leq N_{train}}$. To this we add an interpolation term on the residuals $\alpha_n(\mathbf{p}_i) - \mathcal{R}_n(\mathbf{p}_i)$, $\mathcal{I}_N(\mathbf{p}) = \sum_{i=1}^{N_{train}} \omega_{n,i} \phi\big(d_\theta(\mathbf{p}, \mathbf{p}_i)\big)$. We chose to compute this interpolation using radial basis functions (RBF). We chose cubic RBFs ϕ and a weighted Euclidean distance $d_\theta(\cdot, \cdot)$ to represent the varying ranges of each input parameter.

$$d_\theta(\mathbf{p}_1, \mathbf{p}_2) = \sqrt{\sum_{i=1}^{N_p} \theta_i (\mathbf{p}_1^i - \mathbf{p}_2^i)^2}, \tag{3}$$

where $\theta_i = \dfrac{1}{\left(\max\limits_{\mathbf{p}\in\mathcal{D}} \mathbf{p}^i - \min\limits_{\mathbf{p}\in\mathcal{D}} \mathbf{p}^i\right)^2}$. We then define the emulated projection coefficients as follows.

$$\hat{\alpha}_n^{AQ}(\mathbf{p}) = \underbrace{\sum_{k=1}^{N_p} \beta_{n,k} p_k}_{\text{Least squares regression}} + \underbrace{\sum_{i=1}^{N_{train}} \omega_{n,i} \phi\big(d_\theta(\mathbf{p}, \mathbf{p}_i)\big)}_{\text{Residual interpolation}}. \tag{4}$$

The weights $\{\omega_{n,i}\}_{1\le n\le N; 1\le i\le N_{train}}$ are chose such that the interpolation is exact for the sample points $\{\mathbf{p}_i\}_{1\le i\le N_{train}}$,

$$\hat{\alpha}_n^{AQ}(\mathbf{p}_j) = \alpha_n^{AQ}(\mathbf{p}_j) = \sum_{k=1}^{N_p} \beta_{n,k} p_{j,k} + \sum_{i=1}^{N_{train}} \omega_{n,i} \phi\big(d_\theta(\mathbf{p}_j, \mathbf{p}_i)\big). \tag{5}$$

The emulated solution is finally

$$\hat{c}_N(\mathbf{p}) = \bar{c} + \sum_{n=1}^{N} \hat{\alpha}_n^{AQ} \Psi_n^{AQ}. \tag{6}$$

The regression represents the relation between the model parameters and the RB projection coefficients, and computed from the training set $(\mathbf{p}_i, \alpha(\mathbf{p}_i))_{1\le i\le N_{train}}$. This provides an initial trend to be corrected by the interpolation. In practice, the interpolation of the residual is the most important part of the emulation. In [3] this method of approximating projection coefficients is compared to approximation by Kriging. The two meta-models showed similar results, and we chose RBF emulation for its simpler (and thus more accessible in operational applications) implementation and lower computational cost.

3. Case Study on Clermont-Ferrand

In this work we will apply the meta-modeling method described in section 2 to a modeling chain over the city of Clermont-Ferrand in France. We will build a meta-model chain representing road traffic emissions and the dispersion and reaction of pollutants over the urban agglomeration and surrounding area using data over a two-year period form 2013-2015. The model chain is represented in figure 1.

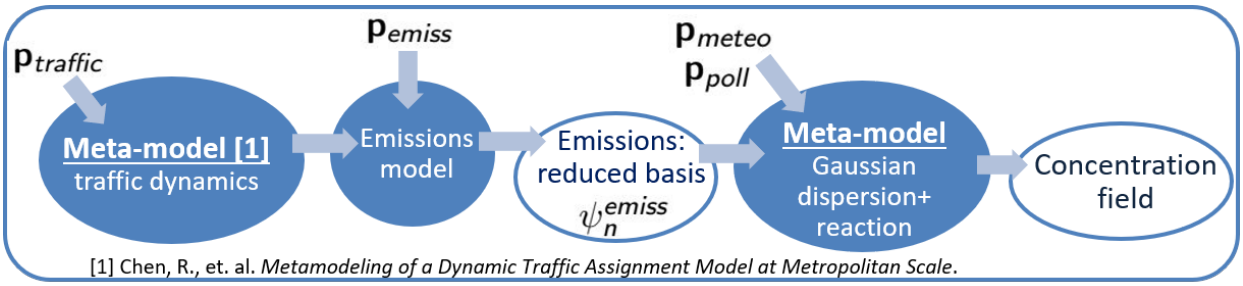

FIGURE 1: meta-modeling chain over Clermont-Ferrand.

3.1. Traffic Emissions Modeling

Traffic emissions modeling is done using the dynamic traffic assignment model LADTA. A meta-model was constructed [17] to represent the traffic flow and speed simulations over a road network of $19,628$ oriented links, where nearly $45,000$ traffic flow observations are available each day. Emissions of NOx and PM are computed using Pollemission code [18] based on the COPERT-IV emissions database [19, 20]. A detailed description of this section of the modeling chain and its input parameters can be found in [17]. The varying input parameters consist of 23 traffic parameters and 6 emissions parameters. These parameters are time-dependent or considered sources of uncertainty. They include temporal traffic demand, computed using traffic observations, the capacity and speed limits of traffic network links, multiplicative coefficients on origin-destination matrices representing spatial trends of traffic, traffic direction (morning versus evening), engine size, bype, and emission standards of the vehicle fleet, and ratio of heavy-duty vehicles to personal cars.

The emissions model provides traffic emissions estimations for NOx and PM_{10}. However the atmospheric pollution model incorporates chemical reaction parametrizations which treat NO_2, NO, $PM_{2.5}$, and PM_{10}. In order to approximate emissions of NO_2, NO, $PM_{2.5}$, and PM_{10}, we would like to estimate what proportion of NOx consists of NO, and what proportion of PM_{10} is $PM_{2.5}$. In the deterministic case, we set the ratio $\frac{NO_2}{NOx} = 0.15$ [21, 22, 23], and the ratio $\frac{PM_{2.5}}{PM_{10}} = 0.75$ [24, 25]. In order to construct a meta-model which can account for varied or uncertain speciation ratios, we will draw LHS parameters for the training ensemble in the intervals $(p_{NO_2}, p_{PM_{2.5}}) \in [0.1, 0.25] \times [0.65, 0.8]$. The output of the traffic-emissions coupling is the emissions on each link of the traffic network in $g/15\text{min}$.

3.2. Air Quality Modeling

Air quality modeling is done using the Gaussian dispersion and reaction model Sirane [15, 26] over a simulation domain of 180 km^2. Sirane is used as a static model which approximates the solution at a given time of the transport-reaction equations satisfied by the pollutant concentrations. The traffic emissions over a relatively coarse road network are converted to $g/s/\text{link}$ on a finer network representing over $47,000$ line sources. For the calculation of NO_2 concentrations, we provide the so-called background concentrations of pollutant species involved directly or indirectly in the formation of NO_2. The background concentrations represent the imported concentrations of pollutants, that is, concentrations transported from distant locations to the city, and possibly dispersed from previous emissions in the case of stationary solution. We will provide inputs on NO_2, NO, O_3, $PM_{2.5}$, and PM_{10}. Input data on meteorological conditions and surface emissions sources are also provided. The AQM output is the NO_2 concentration over a grid at ground level, at $20m$ resolution. Hourly concentration observations are available over two years at 5 stations, or around $90,000$ NO_2 observations for analysis of model simulation outputs.

3.3. Modeling Chain

The modeling chain consists of these three steps - traffic modeling, emissions calculation, and dispersion-reaction modeling - and the conversions between outputs and inputs. In figure 2 we can see the traffic flow (veh/h/link) and associated emissions (g/km/s), and NO_2 concentration ($\frac{\mu g}{m^3}$) simulations at 8a.m. on a Tuesday in November 2014, provided by the traffic meta-model and full air quality model. The task remains to reduce the computational time required to obtain concentration fields by constructing a meta-model for the entire chain.

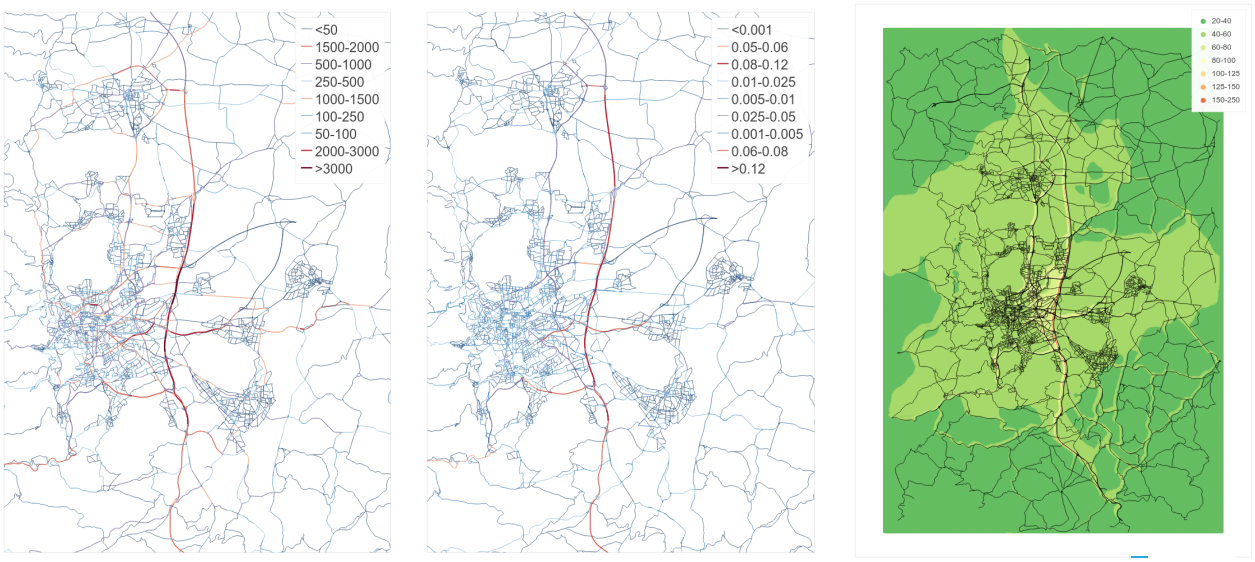

FIGURE 2: Simulations over Clermont-Ferrand on 18/11/2014 at 8am. Traffic flow (left), NO_2 emissions (center), NO_2 concentration (right).

3.3.1. Surrogate Modeling Chain Construction

As noted above, the traffic emissions on a geographically finer road network provided as input to the air quality model represent over $47,000$ line sources. In the context of model order reduction, this represents as many parameters, which in the practice of projection-based reduction methods makes the identification of the projection coefficients α_n^{AQ} dependent on $47,000$ parameters unfeasible (or impossible). We thus need to reduce the complexity of the problem bu reducing the dimension of the input parameters. To do so we will construct a reduced basis of the traffic emissions, again using PCA.

Reduction of line emissions.

We currently have the full chain parameter vector $\mathbf{p}_{full}^T = (\mathbf{p}_{traffic}^T, \mathbf{p}_e^T, \mathbf{p}_{AQ}^T)$, where the outputs of the emissions model consist in a coefficient for each of the links in the road network. These coefficients are then treated as the (very large) input parameter vector for the air quality model. To reduce the dimension of this vector, we will use the same method as in section 2.1. We first select a set of training parameters $(\mathbf{p}_{traffic}^T, \mathbf{p}_e^T)$ by LHS

to represent the variations of these parameters in the admissible parameter space $\mathcal{D}$. We compute the emissions solutions $E(\mathbf{p}_{traffic}, \mathbf{p}_e)$ to construct a reduced basis $\{\Psi_n^E\}_{1\leq n\leq N_{lin}}$ by PCA, representing the variations of the emissions fields centered around $\bar{E} = \frac{1}{N_{train}} \sum_{i=1}^{N_{train}} E(\mathbf{p}_{traffic}^i, \mathbf{p}_e^i)$. We can compute the orthogonal projection of any emissions field onto the traffic emissions RB as follows.

$$E(\mathbf{p}_{traffic}, \mathbf{p}_e) \simeq \Pi_{N_{lin}} E(\mathbf{p}_{traffic}, \mathbf{p}_e) = \bar{E} + \sum_{n=1}^{N_{lin}} \big((E(\mathbf{p}) - \bar{E})^T \Psi_n^E\big) \Psi_n^E = \bar{E} + \sum_{n=1}^{N_{lin}} \alpha_n^{lin} \Psi_n^E. \tag{7}$$

For our case study, we chose $N_{lin} = 11$ to represent 95% of the variability of the emissions solutions. This corresponds to a relative projection error tolerance over the training samples of $\epsilon_{lin}^2 = 0.05$. In the model chain, the over $47,000$ line source parameters will henceforth be replaced by the $N_{lin} = 11$ projection coefficients $\{\alpha_n^{lin}\}_{n\leq N_{lin}}$, and the traffic emissions field for a given parameter approximated by its projection $\Pi_{N_{lin}} E(\mathbf{p}_{traffic}, \mathbf{p}_e)$ onto the traffic emissions RB. We perform the same reduction over the hourly surface emissions with $N_{surf} = 1$ and projection coefficient α_{surf}. In figure 3, we can see the largest singular values of the PCA step, and the relative mean projection errors of the training traffic emissions simulations onto the RB $\{\Psi_n^E\}_{1\leq n\leq N_{lin}}$, as defined by

$$Err_N = \frac{1}{N_t} \sum_{i=1}^{N_{train}} \frac{\|\Pi_N E(\mathbf{p}_i) - E(\mathbf{p}_i)\|_2}{\|E(\mathbf{p}_i)\|_2}. \tag{8}$$

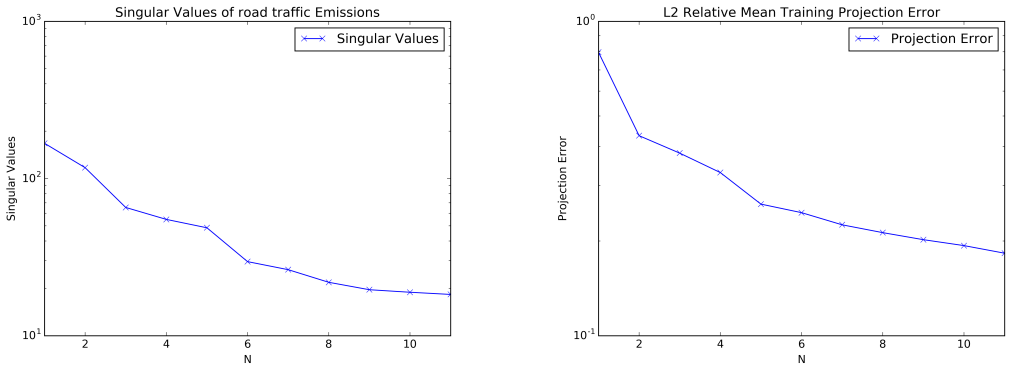

FIGURE 3: Left : Singular values of the emissions mass matrix. Right : L^2 relative mean projection errors of the LHS training ensemble of road traffic emissions fields onto the RB.

In figure 4 we can see the first 4 principal components of the traffic emissions RB.

Construction of the air quality meta-model.

We now can write the reduced concentration model parameters $\mathbf{p}_c^T = (\alpha_{lin}^T, \alpha_{surf}^T, \mathbf{p}_{AQ}^T)$. We will construct a meta-model of the air quality model to complete the meta-modeling chain, with reduced full parameters as described in table 1. The choice to build a separate air quality meta-model to complete the chain of meta-models (as opposed to a meta-model of the chain) was for multiple reasons. First, to allow multi-level assessment using traffic flow and air quality measurement data (a possibility particularly pertinent in a study of uncertainty quantification).

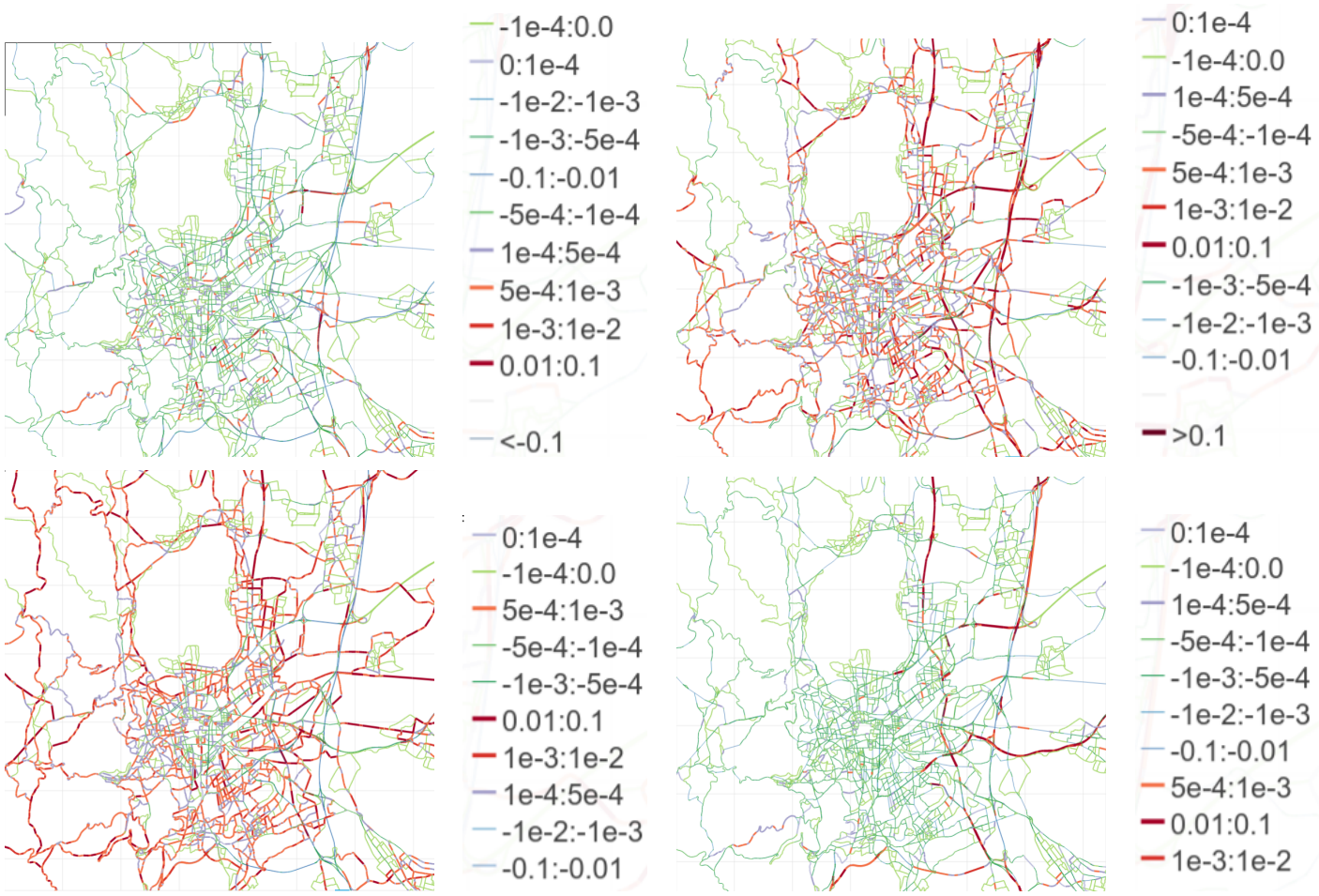

FIGURE 4: Principal components of the emissions mass matrix, first four.

Secondly, to maintain the level of precision of the reduced-order simulations by using $N_{traffic} + N_{AQ} = 11 = 5$ basis functions for traffic and concentration respectively.

When constructing the training ensemble for the air quality meta-model, we chose to draw LHS parameters for the full modeling chain $\mathbf{p}_{full}$. This choice lead to reduced variations in the emissions projection coefficients $\{\alpha_n^{lin}\}_{1 \leq n \leq N_{lin}}$ versus LHS selection over uniform distributions of the emissions projection coefficients $\alpha^{lin} \in [\alpha_{min}^{lin}, \alpha_{max}^{lin}]^{N_{lin}}$. The projection coefficients are in practice not independent; a strong first coefficient is often associated to a weaker second or third coefficient, as these principal components tend to represent different spatial distributions of the emissions. This means that the entire space $[\alpha_{min}^{lin}, \alpha_{max}^{lin}]^{N_{lin}}$ represents significantly more variation in the state $E(\mathbf{p}_{traffic}, \mathbf{p}_e)$ than the traffic-emissions model produces. By performing LHS over the full chain parameters $\mathbf{p}_{full} = (\mathbf{p}_{traffic}, \mathbf{p}_e, \mathbf{p}_{AQ}) \in \mathbb{R}^{41}$, the emissions projection coefficients are computed during the conversion of traffic meta-model outputs to concentration meta-model inputs.

In figure 5 we compare the parameters α_n^{lin} selected by these two methods by plotting the parameter spaces $(\alpha_1^{lin}, \alpha_2^{lin})$ and $(\alpha_1^{lin}, \alpha_4^{lin})$. We can see that the parameter spaces in red, which correspond to performing LHS on

Dynamic Traffic Model LADTA	Emissions Database COPERT IV	Air Quality Model Sirane
$\mathbf{p}_{traffic} \in \mathbb{R}^{23}$	$\mathbf{p}_{emiss} \in \mathbb{R}^{6}$	$\mathbf{p}_{AQ} \in \mathbb{R}^{22}$
Temporal traffic demand ; link capacity and speed limit ; distance traveled data ; direction coefficient	Vehicle fleet data	Meteorological data ; traffic, surface and railway emissions ; background pollution

TABLE 1: Summary of input parameters to the full meta-model chain.

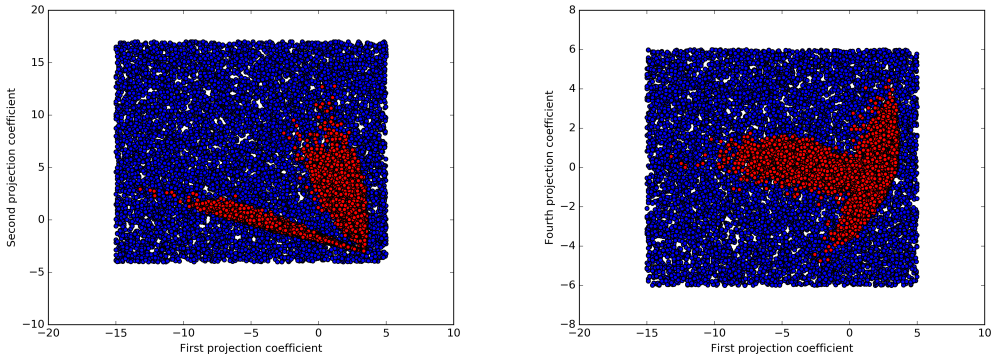

FIGURE 5: Projection coefficients on the traffic emissions basis from LHS performed directly on $[\alpha_{min}^{lin}, \alpha_{max}^{lin}]^{N_{lin}}$ (blue) compared to the projection coefficients of traffic emissions model outputs $E(\mathbf{p}_{traffic}, \mathbf{p}_e)$ (red) over a training ensemble of parameters $(\mathbf{p}_{traffic}^i, \mathbf{p}_e^i)$ selected by LHS. Left : the parameter space of $(\alpha_1^{lin}, \alpha_2^{lin})$. Right : the parameter space of $(\alpha_1^{lin}, \alpha_4^{lin})$.

$\mathbf{p}_{full}$ and computing the projection coefficients α_n^{lin} of the traffic emissions model output $E(\mathbf{p}_{traffic}, \mathbf{p}_e)$ represents significantly less variation than LHS selection directly on the parameters α_n^{lin}. This tactic avoids building a meta-model unnecessarily representing additional variation of the state by only considering realistic traffic emissions. In table 2, we set the ranges of each input parameter which defines the parameter space $\mathcal{D}$.

From a training set of $N_{train} = 9347$, we compute the NO_2 concentration fields $c(\mathbf{p}_{full})$ to construct a reduced basis $\{\Psi_n^{AQ}\}_{1 \leq n \leq N}$ by PCA, representing the variations of the concentration fields centered around the sample concentration mean $\bar{c}$. We set the RB dimension $N = 5$ to represent 98% of this variability. In figure 6, we can see singular values of the matrix of centered solutions $\bar{\mathbf{Y}}_{AQ}$ defined in section 2, and the projection errors of the training ensemble of the atmospheric pollution model, as defined by equation (8).

In figure 7 we see the first 4 principal components of the concentration RB. We can see that the first basis function represents urban background concentration in the denser urban areas. The second seems to represent additional pollution from traffic. The third appears to represent strong wind from the East, while the fourth displays wind from the North.

For any new parameter value, the concentration field can be approximated by the orthogonal projection onto

Input	Dimension	Range
Traffic meta-model parameters		
Temporal traffic profile	13	$[0.0, 2.0]$
Traffic capacity	2	$[0.5, 1.5]$
Speed limit coefficient	2	$[0.5, 1.5]$
Traffic trip distance coefficient	5	$[0, 1.8]$
Traffic direction	1	$[0, 1]$
Traffic emissions parameters		
Fleet engine type	1	$[10, 100]\%$ gasoline
Fleet engine capacity	2	$[1., 100]\%$ gasoline/diesel small-midsize
Emission standard	2	$[0.1, 1]\%$ gasoline/diesel after Euro3
Heavy duty vehicles	1	$[0.0, 0.3]\%$ heave duty vehicles
Conversion $\frac{NO_2}{NOx}$, $\frac{PM_{2.5}}{PM_{10}}$	2	$[0.1, 0.25] \times [0.65, 0.8]$
Air quality model parameters		
Wind velocity	1	$[0, 17] \frac{m}{s}$
Wind direction	1	$[0, 360]^\circ$
Temperature	1	$[-10, 40]^\circ C$
Precipitation coefficient	1	$[0, 26]$
Cloud coefficient	1	$[0, 8]$
Background concentration NO_2, O_3 PM_{10}	3	$[0, 200] \times [0, 180] \times [0, 120] \frac{\mu g}{m^3}$

TABLE 2: Model chain input parameter ranges.

the RB, for projection coefficients $\{\alpha_n^{AQ}\}_{1 \leq n \leq N}$,

$$c(\mathbf{p}_{full}) \simeq \Pi_N c(\mathbf{p}_{full}) = \bar{c} + \sum_{n=1}^{N} \alpha_n^{AQ} \Psi_n^{AQ}. \tag{9}$$

Finally we use the statistical emulation method described in section 2.2 to construct an emulator of the concentration projection coefficients α_n^{AQ}. The full chain can be computed with a single code which applies the traffic-emissions meta-model, the calculation of emissions RB projection coefficients, and the atmospheric pollutant meta-model. This meta-model chain provides outputs on traffic flow, speed, and traffic emissions over

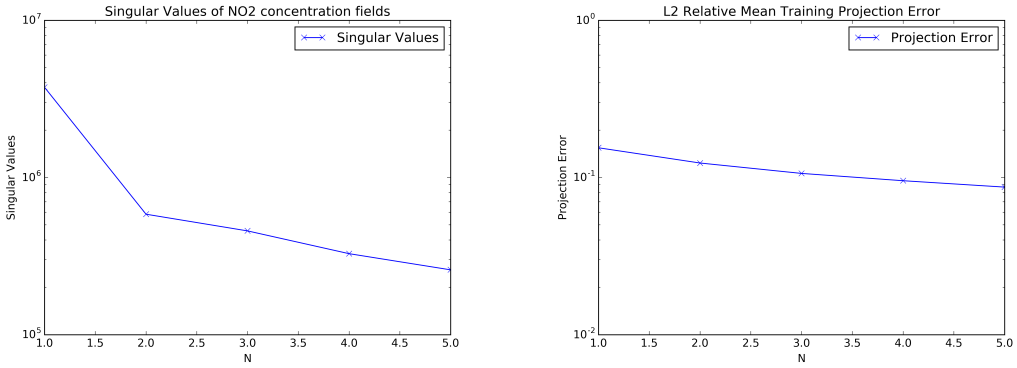

FIGURE 6: Left : Singular values of the NO_2 concentration field mass matrix. Right : L^2 relative mean projection errors of the LHS training ensemble of NO_2 concentration fields onto the RB.

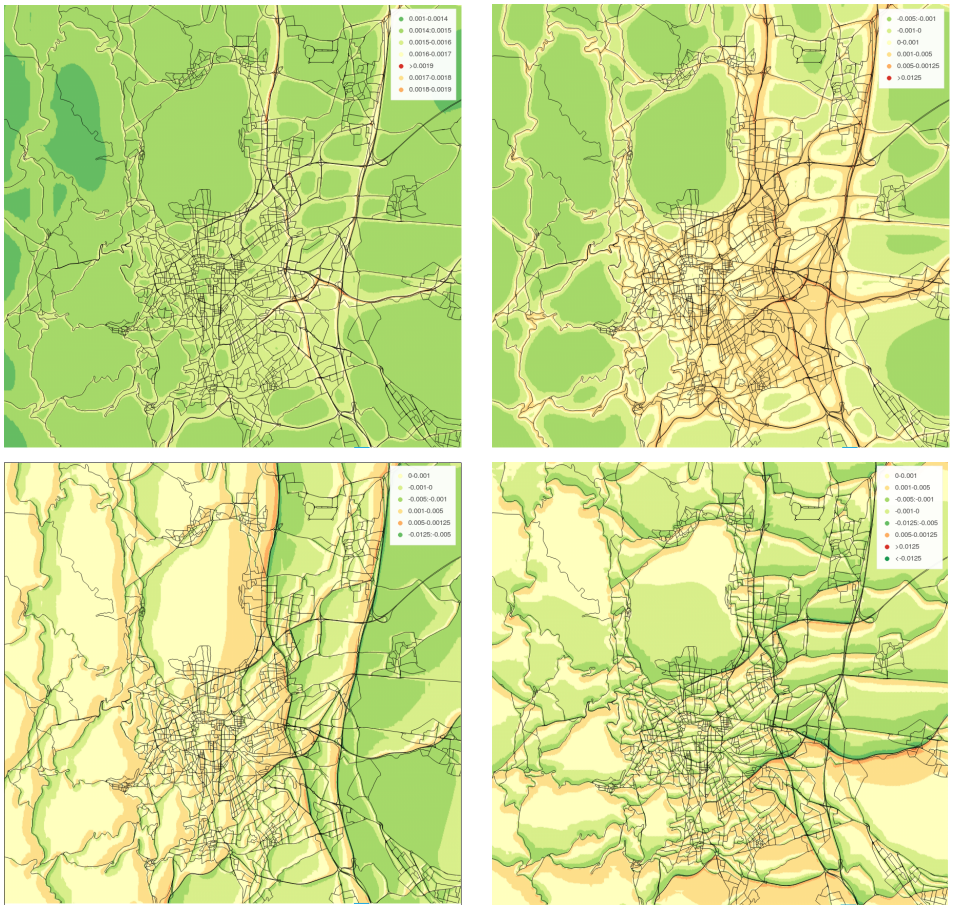

FIGURE 7: Principal components of the NO_2 concentration field mass matrix.

231 the road network, and NO_2 concentrations over a $20m$-resolution grid.

4. Results

In this section, we will summarize the results of the method described in section 2 to the case study in section 3 using data over the month of November 2014. Traffic flow measurement data serves as inputs to the model chain for deterministic simulation, and data on pollutant concentration serves to study model and meta-model performance. We will compare the meta-model output to simulations from the full model Sirane, as well as to concentration observation data, and we will assess computational savings.

4.1. Meta-model performance

We introduce the following statistical scores commonly used for evaluation of models [3] : the normalized mean square error (NMSE), the normalized root mean square error (NRMSE), and the correlation. We define here the output functionals $\ell_m : \Omega \to \mathbb{R}$ associated to each of the concentration sensors m, such that the observation data $y_m^{obs}(\mathbf{p}) = \ell_m(c^{true}(t))$. We denote by $c^{true}(t)$ the unknown true concentration field at time t.

$c_m = \ell_m(c(\mathbf{p}))$ is the value of the output functional associated to sensor m applied to the simulated state estimate, for the full model output $c(\mathbf{p})$ or the meta-model output $\hat{c}(\mathbf{p})$. M is the total number of data available. $\bar{c}$ and $\bar{y}^{obs}$ are respectively the mean of $(c_m)_{1\leq m\leq M}$ and $(y_m^{obs})_{1\leq m\leq M}$.

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{m=1}^{M}(c_m - y_m^{obs})^2}. \tag{10}$$

$$\sqrt{\text{NMSE}} = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\frac{(c_m - y_m^{obs})^2}{\bar{c}\bar{y}^{obs}}}. \tag{11}$$

$$\text{Correlation} = \frac{\sum_{m=1}^{M}(c_m - \bar{c})(y_m^{obs} - \bar{y}^{obs})}{\sqrt{\sum_{m=1}^{M}(c_m - \bar{c})^2}\sqrt{\sum_{m=1}^{M}(y_m^{obs} - \bar{y}^{obs})^2}}. \tag{12}$$

$$\text{Bias} = \frac{1}{M}\sum_{m=1}^{M}(y_m^{obs} - c_m) \tag{13}$$

Finally we define the NRMSE as $\frac{\text{RMSE}}{\bar{y}^{obs}}$, and the MNRMSE as the mean over all sensors (or grid points) of the NRMSE calculated over the concentration c_i at each sensor (or grid point) over the month.

$$\text{MNRMSE} = \frac{1}{N_{grid}}\sum_{i=1}^{N_{grid}} NRMSE(c_i) \tag{14}$$

4.1.1. Comparison with the full model chain

We first analyze the precision of the meta-modeled concentration fields as compared to the full model Sirane. This will help us understand the ability of the meta-model to reproduce the concentration state and quantify the loss of precision caused the the dimensional reduction.

In figure 8, we see statistical scores spatially mapped over the meta-model domain. The NRMSE shows that the emulated solutions perform well in approximating the urban background concentration levels, but don't capture the highest concentrations along the large highways as well. The correlation map also shows low correlation between the meta-model and full model only along the roadways, where the dimensional reduction has failed to capture the extent of the increased concentrations due to traffic emissions. Finally the bias map shows that the meta-model generally predicts higher concentrations in the denser urban areas when compared to the full model, again matching the trend of the dimensional reduction reducing the sensitivity of the meta-model to sharp spatial variations in concentrations. However, the areas with poor scores remain limited, and we must also consider the significant error that will inevitably be committed by the full model in the next section.

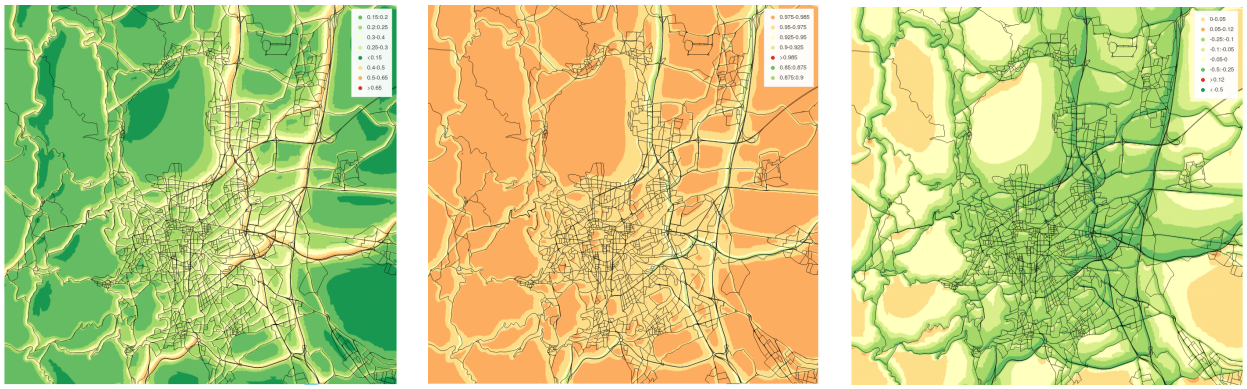

FIGURE 8: Left : NRMSE (10) of the emulated NO_2 concentration field compared to the projected solution, for parameters over the month of November 2014. Center : correlation (12) over the same set. Right : Normalized bias (13) over the same set.

In figure 9 We see the relative errors of the full model concentration projected onto the reduced basis $\{\psi_n^{AQ}\}_{1\leq\leq N}$, averaged over the set of deterministic simulations for the month of November 2014. We also see the emulated concentration relative error, averaged over the same set of simulations. While the emulation of the projection coefficients is globally responsible for the most error, we can see that the regions with the highest projection error correspond to high errors in the meta-model as well. This is expected, as the emulated solution can only perform as well as the projected solution. We see that larger errors are located on roads, mostly the large highway and outside the dense urban area. Meta-model error remains below 20% over a large portion of the domain, which shows that much of the spatial variation of the concentration is captured by the reduced order solution.

In table 3 we can see statistical scores comparing the meta-modeled concentration to the full concentration model over all hours of November 2014. We compare both the entire grid (here c_m is the concentration at a grid

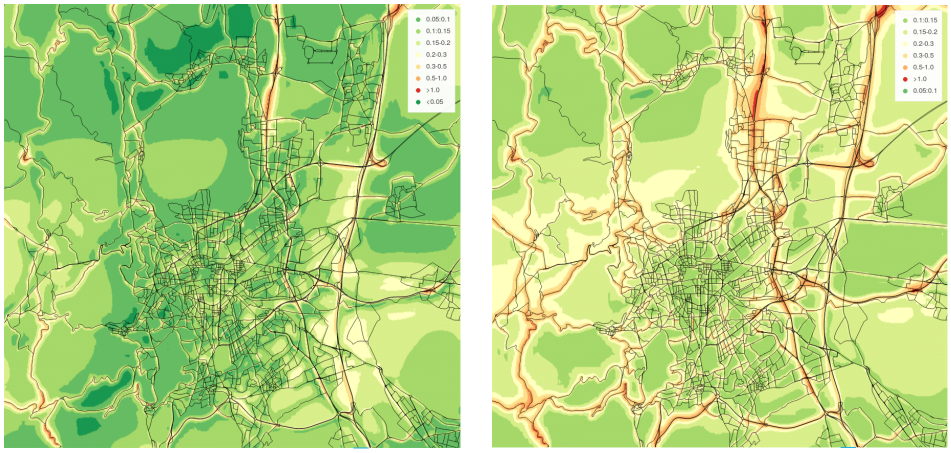

FIGURE 9: Left : Relative mean error of the projected NO_2 concentration field compared to the full model solution, for parameters over the month of November 2014. Right : Relative mean error of the emulated NO_2 concentration field compared to the full model solution over the same parameter set.

point and $M = N_{grid}$ is the total number of grid points) and at the NO_2 sensor locations. While the dimensional reduction means the meta-model does not fully capture spatial variations of the simulated concentration state, we can see that the relative RMSE errors are satisfactorily low, and the correlation between the two is very high.

State Estimation $c(\mathbf{p})$ vs. $\hat{c}(\mathbf{p})$	Statistical scores over 1-month simulations		
Over the full grid	MNRMSE (14)	$\sqrt{\text{NMSE}}$ (11)	Correlation (12)
Meta-model vs. full model chain	0.274	0.25	0.93
At receptor locations	NRMSE (10)	$\sqrt{\text{NMSE}}$ (11)	Correlation (12)
Meta-model vs. full model chain	0.22	0.18	0.96

TABLE 3: Statistical scores of the meta-model approximation results compared to the scores of the model chain using the full air quality model.

In figure 10 we see a visual representation of hourly scores of the meta-model solution compared to the full solution at each grid point for simulations corresponding to the month of November 2014. The NMSE (11) remains globally below 0.4, and the RMSE (10) often below $10\frac{\mu g}{m^3}$. Correlations scores are grouped above 0.75, and the bias distribution is nearly centered around $-2\frac{\mu g}{m^3}$, showing a slightly higher concentration approximation by the meta-model, when averaged over the grid.

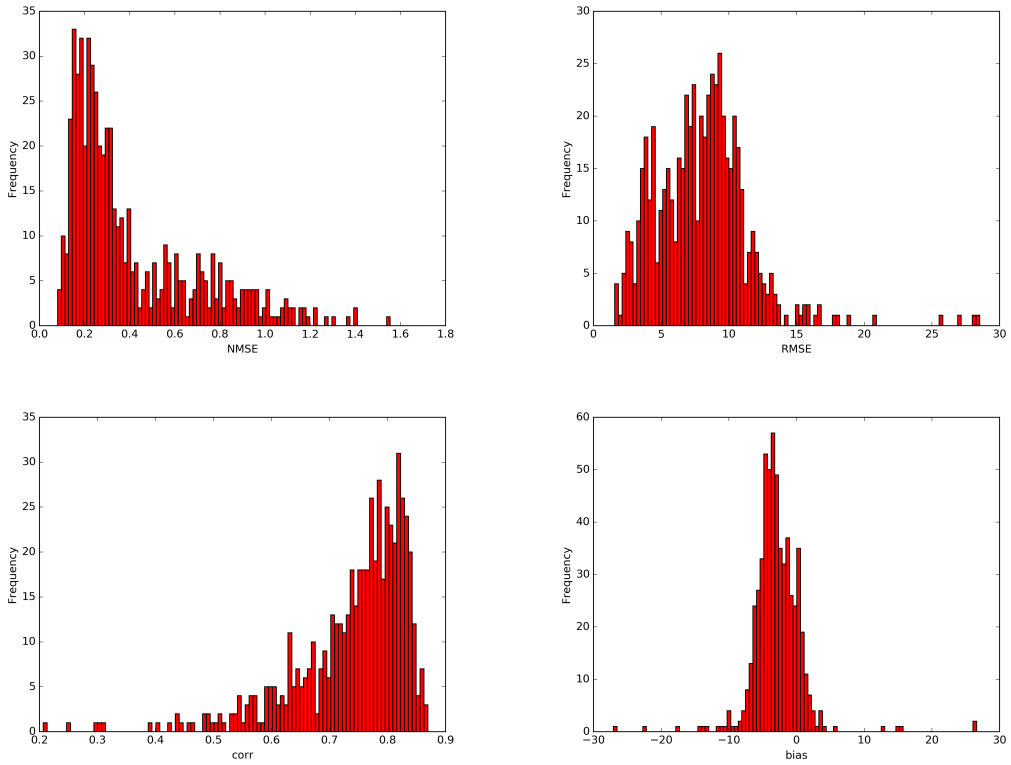

FIGURE 10: Scores of the meta-modeled NO_2 concentration field compared to the full model solution, for parameters over the month of November 2014. Top left : NMSE (11). Top right : RMSE (10). Bottom left : correlation (12). Bottom right : bias (13).

4.1.2. Comparison with observational data

We next analyze the accuracy of the full model and meta-model compared to observational data on NO_2 concentrations. In figure 11, we see the temporal profile of average NO_2 concentrations at $M = 4$ sensor locations : $\frac{1}{M}\sum_{m=1}^{M} c_m$. We compare observed, emulated, projected and Sirane modeled concentrations of all weekdays in November 2014. We see that the bias in the modeled concentrations underestimating peak concentrations, notably during heavy traffic periods in the mornings and evenings. We also notice a seemingly delayed reaction of the model chain to the pollution increase during the evening peak hour. In [17] this delay was less evident, suggesting that factors such as the dispersion and reaction parametrizations in the AQ model or the averaging of time scales from 15 minutes to one hour may have an effect. The exploration of this question will require more study of uncertainties in the model chain. We notice the the temporal trend representing morning and evening peak hours in traffic is reproduced by the model chain. We also note that the emulated concentrations are closer to the observations than the full model. This is likely due to the "smoothing" effect of the dimensional reduction causing less sharp concentration variations, as small parts of the modeled concentration fields are not reproduced by the reduced basis.

In table 4 we compute statistical scores over the month of November 2014, comparing the full model simulations

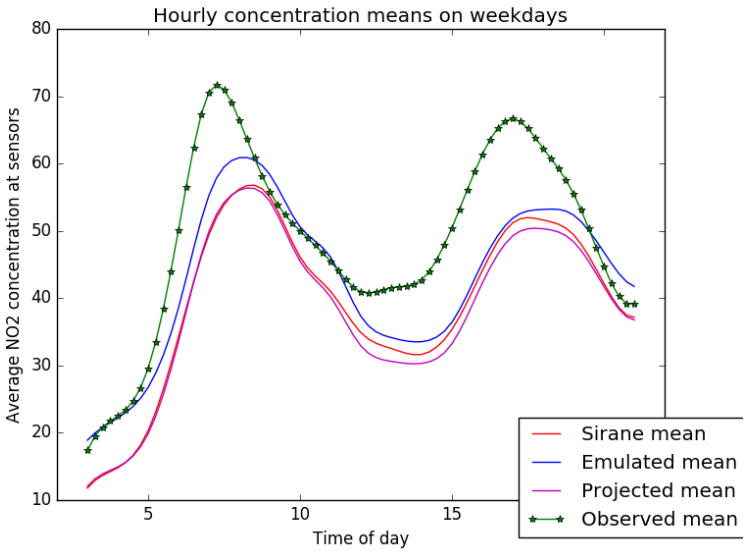

FIGURE 11: Mean NO_2 concentrations at 5 sensor locations over weekdays in November 2014. Curves show observations, full model simulations, projected simulations onto the reduced basis, and emulated solutions.

and the meta-modeled simulations to the observation data at $M = 4$ sensor locations. We again see that the emulated solutions are slightly more accurate than the full model. The stations at which both the model and meta-model perform best are those found in dense urban areas, excepting the station Gare, where heavy traffic induces high NO_2 concentrations, which the model fails to reproduce. We see the highest bias at this location. Finally, the station Chamalières is located outside the city center, where the model exhibits a higher level of bias. The performance of the meta-model with respect to observation data is highly satisfactory.

Statistical scores over	Sirane solution $c(\mathbf{p})$				Meta-model solution $\hat{c}(\mathbf{p})$			
one-month simulations	NRMSE	$\sqrt{\text{NMSE}}$	Bias	Correlation	NRMSE	$\sqrt{\text{NMSE}}$	Bias	Correlation
All 4 stations	0.479	0.499	12.88	0.746	0.461	0.467	10.28	0.728
Lecoq	0.368	0.351	6.81	0.847	0.364	0.347	6.5	0.846
Montferrand	0.366	0.33	2.04	0.833	0.377	0.325	−1.6	0.836
Gare	0.655	0.587	33.13	0.776	0.469	0.52	20.43	0.783
Chamalières	0.474	0.573	16.05	0.719	0.498	0.55	15.8	0.648

TABLE 4: Statistical scores of the meta-model approximation results compared to observation data. NRMSE (10), $\sqrt{\text{NMSE}}$ (11), Correlation (12).

In figure 12 we see a visual representation of daily scores of the meta-model solution and the full solution compared to NO_2 observations over the month of November 2014. The meta-model shows similar score distributions

to the full model, excepting the occasional outlier. The RMSE (10) is below $25\frac{\mu g}{m^3}$ on the majority of days for both the full and reduced simulations, with the bias distribution nearly centered around $10 - 15\frac{\mu g}{m^3}$, showing an underestimation of concentrations by the simulations.

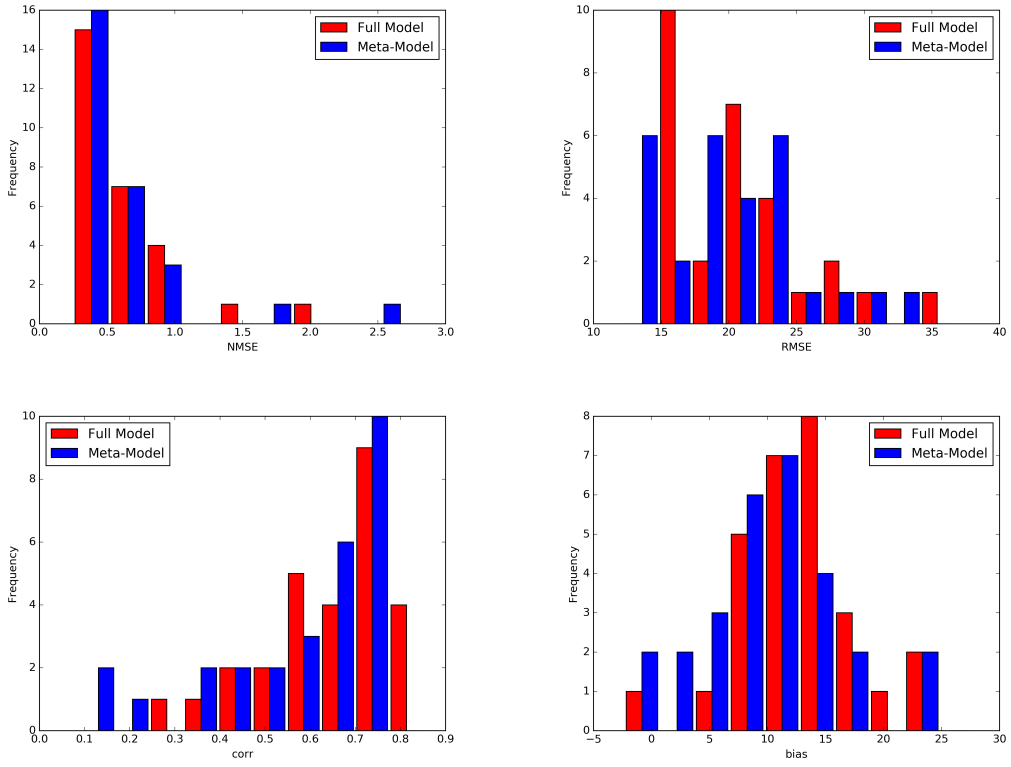

FIGURE 12: Scores of the Sirane NO_2 concentration field compared to the observation data over the month of November 2014. Top left : NMSE (11). Top right : RMSE (10). Bottom left : correlation (12). Bottom right : bias (13).

While we have seen that the model reduction by statistical emulation causes loss of precision, and the meta-model simulations contain error with respect to the full model, comparing to observation data suggests that this error is not significant with respect to the model error inherent to operational models for urban air quality, and does not reduce the accuracy of the predicted concentrations at sensor locations.

4.2. Computational savings

We have seen that the meta-model chain produces satisfactory results when compared to observational data, and determined that the loss of precision due to the dimensional reduction is not higher than the error committed by the full model. Now we will show the computational savings afforded by the meta-model chain. In table 5 we can see the computational times required for a single simulation of the chain by the meta-models or the full models. The meta-models depend on three reduced basis, representing traffic for the traffic assignment meta-model, road emissions for the reduction of pollution model input dimension, and concentration fields for the pollution meta-model. The initialization of the meta-model chain requires loading these basis and building the RBF emulators.

Once the chain is initialized, it can be run for any number of simulations at very low cost, under 0.1 seconds for a simulation representing a one-hour period. In comparison, the full model chain requires nearly three hours for a single simulation.

Computational times				
Meta-model simulation	Traffic Emissions	Compute $\hat{\alpha}^{lin}(\mathbf{p}_m)$	Compute $\hat{\alpha}^{AQ}(\mathbf{p}_m)$	Total CPU
Initialize meta-model chain	–	–	–	**24** min
Emulating $\alpha^{AQ}(\mathbf{p}_m)$	0.05 sec	0.006 sec	0.02 sec	**0.076** sec
Full model simulation	Traffic Emissions	–	Compute $c(\mathbf{p}_m)$	Total CPU
Simulation of $c(\mathbf{p}_m)$	117 min	–	23 min	**140** min

TABLE 5: Computation times using the meta-model or full model chain.

The offline construction of the meta-models required 6000 traffic model simulations [17] and 10000 pollution model simulations, which represents a significant computational investment. However, these meta-models are trained over training points $\{\mathbf{p}_i\}_{1\leq i\leq N_{train}} \in \mathcal{D}$ representing two years of data, and once constructed are useful for study over multiple years. In the absence of high performance computing machines or clusters, the simulations can be run using a pseudo-parallel technique running one simulation per core on desktop calculation machines. The Sirane simulations described in section 3 took around one day using this method on multiple machines of $64GB$ RAM or less. Once the meta-model chain is constructed, the online phase for the simulation given any parameter $\mathbf{p} \in \mathcal{D}$ is very cheap, which makes real-time or many-query contexts possible, for example for use in uncertainty quantification study.

5. Conclusions

In this work we constructed a meta-model chain by statistical emulation of reduced basis projection coefficients for an urban air quality modeling chain over the agglomeration of Clermont-Ferrand. We used the road traffic meta-model constructed in [17], built a reduced basis representing road traffic emissions, and constructed a second meta-model of NO_2 concentration fields over the agglomeration, substituting thus a low-cost chain of meta-models for a computationally costly modeling chain over a large urban area. This required the dimensional reduction of the inputs to the atmospheric pollution model, and the appropriate sampling of the parameter spaces to construct a satisfactory reduced basis and reduced order modeling scheme. We reduced computation time from over two computational hours per simulation representing an hourly concentration field to under 0.1 second. Results show good precision of the meta-model simulations with respect to the full model chain, and similar accuracy when compared to measurement data. The meta-model can be used in various applications requiring numerous solutions

to the model chain, rendering studies requiring many solutions to the model chain computationally feasible. In future work, we will use this low-cost modeling chain in the study of uncertainty quantification and the propagation of uncertainties throughout the meta-model chain.

Acknowledgments

This research is supported by the French National Research Agency (ANR, the Agence Nationale de la Recherche), project ANR-13-MONU-0001, ESTIMAIR. The City of Clermont-Ferrand provided traffic flow measurement data. SMTC (Syndicat Mixte des Transports en Commun de l'agglomération clermontoise) provided the traffic network geometry for the agglomeration of Clermont-Ferrand and the static O-D matrix representing spatial traffic demand in the traffic assignment model. The SME NUMTECH provided data necessary for the model SIRANE, including geometrical features of the traffic network, meteorological data, emissions data, and background concentrations. We particularly thank David Poulet for his contributions to the data treatment. The SIRANE model was developed by "Laboratoire de Mécanique des Fluides et d'Acoustique", UMR CNRS 5509, University of Lyon, Ecole Centrale de Lyon, INSA Lyon, Université Claude Bernard Lyon I. We thank Lionel Soulhac for providing it and for his guidance. Atmo Auvergne-Rhone-Alpes (https ://www.atmo-auvergnerhonealpes.fr/) provided pollutant concentration measurements used in the calculation of statistical scores.

[1] M. Milliez, B. Carissimo, Computational Fluid Dynamical Modelling of Concentration Fluctuations in an Idealized Urban Area, Boundary-Layer Meteorology 127 (2) (2008) 241–259. doi:10.1007/s10546-008-9266-1.
URL http://link.springer.com/10.1007/s10546-008-9266-1

[2] Y. Tominaga, T. Stathopoulos, CFD simulation of near-field pollutant dispersion in the urban environment: A review of current modeling techniques, Atmospheric Environment 79 (2013) 716–730. doi:10.1016/j.atmosenv.2013.07.028.
URL http://linkinghub.elsevier.com/retrieve/pii/S1352231013005499

[3] V. Mallet, A. Tilloy, D. Poulet, S. Girard, F. Brocheton, Meta-modeling of ADMS-Urban by dimension reduction and emulation, Atmospheric Environment 184 (2018) 37–46. doi:10.1016/j.atmosenv.2018.04.009.
URL http://www.sciencedirect.com/science/article/pii/S1352231018302346

[4] D. J. Carruthers, H. A. Edmunds, C. A. McHugh, R. J. Singles, Development of Adms-Urban and Comparison with Data for Urban Areas in the UK, in : Air Pollution Modeling and Its Application XII, Springer, 1998, pp. 467–475.

[5] L. A. Lee, K. S. Carslaw, K. J. Pringle, G. W. Mann, D. V. Spracklen, Emulation of a complex global aerosol model to quantify sensitivity to uncertain parameters, Atmospheric Chemistry and Physics 11 (23) (2011)

12253–12273. doi:10.5194/acp-11-12253-2011.
URL https://www.atmos-chem-phys.net/11/12253/2011/

[6] P. Armand, F. Brocheton, D. Poulet, F. Vendel, V. Dubourg, T. Yalamas, Probabilistic safety analysis for urgent situations following the accidental release of a pollutant in the atmosphere, Atmospheric Environment 96 (2014) 1–10. doi:10.1016/j.atmosenv.2014.07.022.
URL http://www.sciencedirect.com/science/article/pii/S1352231014005433

[7] S. Girard, V. Mallet, I. Korsakissok, A. Mathieu, Emulation and Sobol' sensitivity analysis of an atmospheric dispersion model applied to the Fukushima nuclear accident, Journal of Geophysical Research : Atmospheres 121 (7) (2016) 3484–3496.

[8] M. Fallah Shorshani, M. André, C. Bonhomme, C. Seigneur, Modelling chain for the effect of road traffic on air and water quality: Techniques, current status and future prospects, Environmental Modelling & Software 64 (2015) 102–123. doi:10.1016/j.envsoft.2014.11.020.
URL http://linkinghub.elsevier.com/retrieve/pii/S1364815214003466

[9] A. Russell, R. Dennis, NARSTO critical review of photochemical models and modeling, Atmospheric Environment 34 (12) (2000) 2283–2324.
URL http://www.sciencedirect.com/science/article/pii/S1352231099004689

[10] Y. Zhang, M. Bocquet, V. Mallet, C. Seigneur, A. Baklanov, Real-time air quality forecasting, part II: State of the science, current research needs, and future prospects, Atmospheric Environment 60 (2012) 656–676. doi:10.1016/j.atmosenv.2012.02.041.
URL http://linkinghub.elsevier.com/retrieve/pii/S1352231012001562

[11] F. LEURENT, On network assignment and demand-supply equilibrium : an analysis framework and a simple dynamic model, in : PROCEEDINGS OF THE EUROPEAN TRANSPORT CONFERENCE (ETC) 2003 HELD 8-10 OCTOBER 2003, STRASBOURG, FRANCE, 2003.

[12] F. Leurent, V. Aguiléra, Large problems of dynamic network assignment and traffic equilibrium : Computational principles and application to Paris road network, Transportation Research Record 2132 (1) (2009) 122–132.

[13] R. Chen, V. Mallet, Pollemission software computing traffic emissions of atmospheric pollutants with copert-iv formulations.
URL https://github.com/pollemission.

[14] L. Ntziachristos, D. Gkatzoflias, C. Kouridis, Z. Samaras, COPERT: A European Road Transport Emission Inventory Model, in : D. I. N. Athanasiadis, P. A. E. Rizzoli, P. A. Mitkas, P. D.-I. J. M. Gómez (Eds.), Information Technologies in Environmental Engineering, Environmental Science and Engineering, Springer

Berlin Heidelberg, 2009, pp. 491–504.
URL http://link.springer.com/chapter/10.1007/978-3-540-88351-7_37

[15] L. Soulhac, P. Salizzoni, F.-X. Cierco, R. Perkins, The model SIRANE for atmospheric urban pollutant dispersion; part I, presentation of the model, Atmospheric Environment 45 (39) (2011) 7379–7395. doi:10.1016/j.atmosenv.2011.07.008.
URL http://linkinghub.elsevier.com/retrieve/pii/S1352231011007096

[16] C. Prud'homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. T. Patera, G. Turinici, Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods, Journal of Fluids Engineering 124 (1) (2002) 70–80.
URL http://fluidsengineering.asmedigitalcollection.asme.org/article.aspx?articleid=1429475

[17] R. Chen, V. Mallet, V. Aguilera, F. Cohn, D. Poulet, Metamodeling of a dynamic traffic assignment model at metropolitan scale 43.

[18] Ruiwei Chen, Vivien Mallet, Pollemission software computing traffic emissions of atmospheric pollutants with COPERT-IV formulations, original-date : 2016-01-21T17 :19 :00Z (2016).
URL https://github.com/pollemission

[19] D. Gkatzoflias, C. Kouridis, L. Ntziachristos, Z. Samaras, COPERT 4 : Computer programme to calculate emissions from road transport, European Environment Agency.

[20] EEA, EMEP/EEA air pollutant emission inventory guidebook - Part B.1.A.3.b.iiv Road transport (2016).
URL https://www.eea.europa.eu/publications/emep-eea-guidebook-2016

[21] D. C. Carslaw, Evidence of an increasing NO2/NOX emissions ratio from road traffic emissions, Atmospheric Environment 39 (26) (2005) 4793–4802. doi:10.1016/j.atmosenv.2005.06.023.
URL http://www.sciencedirect.com/science/article/pii/S1352231005005443

[22] S. D. Beevers, E. Westmoreland, M. C. de Jong, M. L. Williams, D. C. Carslaw, Trends in NOx and NO2 emissions from road traffic in Great Britain, Atmospheric Environment 54 (2012) 107–116. doi:10.1016/j.atmosenv.2012.02.028.
URL http://www.sciencedirect.com/science/article/pii/S1352231012001434

[23] R. Kurtenbach, J. Kleffmann, A. Niedojadlo, P. Wiesen, Primary NO2 emissions and their impact on air quality in traffic environments in Germany, Environmental Sciences Europe 24 (1) (2012) 21. doi:10.1186/2190-4715-24-21.
URL https://doi.org/10.1186/2190-4715-24-21

[24] J. A. Gillies, A. W. Gertler, J. C. Sagebiel, W. A. Dippel, On-Road Particulate Matter (PM2.5 and PM10) Emissions in the Sepulveda Tunnel, Los Angeles, California, Environmental Science & Technology 35 (6) (2001) 1054–1063. doi:10.1021/es991320p.
URL https://doi.org/10.1021/es991320p

[25] X. Querol, A. Alastuey, C. R. Ruiz, B. Artiñano, H. C. Hansson, R. M. Harrison, E. Buringh, H. M. ten Brink, M. Lutz, P. Bruckmann, P. Straehl, J. Schneider, Speciation and origin of PM10 and PM2.5 in selected European cities, Atmospheric Environment 38 (38) (2004) 6547–6555. doi:10.1016/j.atmosenv.2004.08.037.
URL http://www.sciencedirect.com/science/article/pii/S1352231004008143

[26] L. Soulhac, P. Salizzoni, P. Mejean, D. Didier, I. Rios, The model SIRANE for atmospheric urban pollutant dispersion; PART II, validation of the model on a real case study, Atmospheric Environment 49 (2011) 320–337. doi:10.1016/j.atmosenv.2011.11.031.
URL http://linkinghub.elsevier.com/retrieve/pii/S1352231011012143