Source term estimation: variational method versus machine learning applied to urban air pollution

Roman Lopez-Ferber, Sylvain Leirens, Didier Georges

Source Term Estimation: Variational Method Versus Machine Learning Applied to Urban Air Pollution

R. Lopez-Ferber * **S. Leirens** ** **D. Georges** ***

* *Univ. Grenoble Alpes, CEA, Leti, F-38000 Grenoble, France*
** *Univ. Grenoble Alpes, CEA, Leti, F-38000 Grenoble, France*
*** *Univ. Grenoble Alpes, CNRS, Grenoble INP*, GIPSA-lab, F-38000 Grenoble, France*
**Institute of Engineering and Management,Univ. Grenoble Alpes*

Abstract: Source detection is a field of study gaining interest due to environmental concerns about air quality in populated areas. We developed a machine learning framework inspired by previous works on road traffic estimation, and compared it to a classical variational method under a unidimensional and stationary problem. We tested source reconstruction with datasets coming from 12 and 50 sensors with and without noise. Noise was set to follow a gaussian law with a dependent variance from the maximum measured value of a concentration profile. Both methods are reasonably robust to noise. The results reveal that the Neural Network used here, a multilayer perceptron, performs very well compared to the classical 3D-Var method.

Keywords: Source Detection, Source Term Estimation, variational methods, 3D-Var, air pollution, neural network, advection-diffusion

1. INTRODUCTION

To the knowledge of the authors, machine learning is not much used for Source detection. Hutchinson et al. (2017) cite several works where Machine Learning methods were applied, but none of them aims to identify a source of pollution from a set of geolocalised data set.. Another work dealt with building of 2D maps of pollutant concentrations using mobile sensors measurements (Ma et al. (2019)), but no method on source term estimation was proposed. Recently machine learning has been applied on inverse problem dealing with road traffic estimation (Georges (2020)) and has inspired the present work, where a feed forward neural network (NN) is used with a few hidden layers. This study proposes a method to reconstruct a source of pollution in an urban area. In the case of a unidimensional domain and stationary conditions, one could imagine the reconstruction of a unique source of pollution located on a boulevard of a town surrounded by an urban area. The machine learning approach is compared with a model-based variational framework (3D-Var) which acts as a reference.

1.1 Physical model used in machine learning and variational frameworks

In both frameworks, a stationary physical model was used to create source and concentration data of pollutants. In a simple one dimensional model where a stationary source has emitted long enough to reach stationarity, the displacement of pollutants is governed by the following advection-diffusion law,

$$c\partial_x u(x) - b\partial^2_{xx} u(x) = q(x)\ , \tag{1}$$

with c $[m/s]$, b $[m^2/s]$, parameters corresponding to the wind field and turbulent diffusion respectively, both considered as scalars. u denotes pollutant concentrations in g/m^3 or in mol/m^3 depending on the considered source rate. q denotes the source with units of concentration per time and is defined on the physical domain (x in one dimension). The source profile q used in the experiments of this work is defined as follows,

$$q\,(x) = A_s\ \Delta_s\,(x - x_s)\ , \tag{2}$$
$$\text{with}\quad \Delta_s\,(x - x_s) = \exp\left(-\sigma^2_{disp}\,(x - x_s)^2\right),$$

where $A_S,\ x_s, \sigma^2_{disp}$ $[\emptyset, km,\ km^{-2}]$ represent the amplitude, the position and the inverse of the variance of a unique source of pollution. Δ_S denotes the function of dispersion chosen for the source around its centre. The partial differential equation can be solved in three dimensions by a function called *gaussian puff* if a pointwise source is considered. In this study, the equation is solved by finite differences on a domain $\Omega = x \in [0, 1]$. Boundaries are chosen as Dirichlet conditions imposing a null concentration at the edges of x which model the absence of pollution in the background. The chosen unit of length is kilometre $[km]$.

The model $\mathcal{M}$ is solved with spatial finite differences and leads to,

$$\begin{aligned} u &= \mathcal{M}(\theta) \\ &= A^{-1}(B_1 q(\theta) + B_2 C), \end{aligned} \tag{3}$$

where, A and C are the representative matrices of discretisation and boundary conditions respectively. B_1, B_2

⋆ This work is supported by The French Alternative Energies and Atomic Energy Commission

are matrices that map the discretised source profile q and the boundary conditions C towards the target system to inverse.

1.2 Simulations Parameters

As stated at the beginning of section 1, this study aims to reconstruct a source of pollution in an urban area where the wind field velocity generally does not exceed $\sim 10\,m/s$ as seen in Wang et al. (2019) and Wang et al. (2020). The latter studied a technique to assess wind field without using CFD models in Seoul. The turbulent coefficient is set to $b = 25000\,m^2/s$ as found in Winiarek et al. (2012) for the horizontal components of b in a problem of source identification at a continental scale. In urban scale, this coefficient is often imposed by the chosen wind turbulence model with,

$$b = \left(D_m + \frac{\nu_t}{Sc_t} \right), \tag{4}$$

where D_m is the molecular diffusion of the pollutant, v_t the turbulent viscosity and Sc_t the Schmidt Number. According to Lin et al. (2021) this number is chosen as a constant in atmospheric study with a value usually set to $Sc_t = 0.3$ whereas v_t is determined by the turbulence model of Navier-Stockes which is written in Kikumoto (2020) is for a $k - \epsilon$ RANS model,

$$v_t = C_\mu k T_e \quad \text{where,} \quad v_t = C_\mu \frac{k^2}{\epsilon} \text{ and } Te = \frac{k}{\epsilon}. \tag{5}$$

1.3 Noise considered on sensors

Many studies consider measurements with a dedicated measurement sampling function to take into account that a sensor i, $1 \leq i \leq N_C$ (N_C denoting the number of sensors) measures average concentration in its surrounding, leading to the sensor function for the i^{th} sensor (See Georges (2019), Hammond et al. (2019), Krysta and Bocquet (2007)),

$$\mathcal{H}_i \ : u,\ x_i \mapsto \int_\Omega \phi_i \left(x - x_i \right) u \left(x \right) dx, \quad 1 \leq i \leq N_C \ , \tag{6}$$

where $\phi_i \ : \ x \ \mapsto \ \phi \left(x - x_i \right)$ is a sampling function associated to a sensor i at a position x_i. In this work ϕ is chosen as a Gaussian kernel but studies can also consider a rectangular function as in Kovalets et al. (2018). The overall observation operator $\mathcal{H}$ which maps the outputs of the physical model to measurements is thus an operator which returns,

$$\mathcal{H} \ : u \mapsto [\mathcal{H}_1(u), \ldots, \mathcal{H}_{N_C}(u)] \tag{7}$$

Gaussian noise is also considered for several experiments in this paper. This noise measurements is applied to each of the N_C sensors following the expression,

$$u_i^m = \mathcal{H}_i(u) + \epsilon_i^{obs} \quad \text{with, } \epsilon_i^{obs} \sim \mathcal{N}(0, {\sigma^{obs}}^2) \ , \tag{8}$$

where ϵ_i^{obs} is a gaussian noise corresponding to a i^{th} sensor. Sensor noise is often defined in literature by two contributions: background noise and relative noise. The last depends on the magnitude of the measurement. Thus noise can be written as follows,

$$\epsilon_i^{obs} = n_R + n_b, \tag{9}$$

with $n_b \sim \mathcal{N}(0, \sigma_{base}^2)$ a basal noise which would command the limit of blank of the sensor (according to the definitions of Armbruster and Pry (2008)) and $n_R \sim \mathcal{N}(0, (\sigma_{relative} \times \mathcal{H}_i(u))^2)$ the measure-dependant noise. However this formulation would modify the expression of the covariance matrix of measurement errors written in (13) and hence making the problem non quadratic. For this reason, we apply a constant variance of measurements as written in (8). Consequently, for each measurement of a concentration profile u_s, a constant standard deviation with $\sigma_s^{obs} = 5\% \times max\left(\mathcal{H}(u_s)\right)$ is chosen. It constitutes also a conservative hypothesis on the amount of noise for lower measured values. The basal noise is neglected. The choice of gaussian errors of measurements is motivated by Defforge et al. (2021) and Asch et al. (2016) where gaussian errors are advised to be considered in order to build proven optimal filters such as the Kalman filter.

2. METHODS

The goal of this source identification study is to reconstruct the source term written in (2) given measurements. To carry this study two methods are applied and compared on the stationary problem described in 1.1: a method based on a Neural Network model (NN) trained with a synthetic database and a classical model-based method (3D-Var).

2.1 Machine Learning

Let us name, $\theta = \left[A_S,\ x_s,\ \sigma^2_{\text{disp}} \right]$ a vector representing source parameters. Alternatively, we could also consider that the source contains as much parameters as grid points of simulation, i.e. $\theta_i = q(i\Delta x)$ with $0 \leq i \leq N_x - 1$, N_x being the total number of discretised points in the x axis. A multilayer perceptron with one or several hidden layers is built to learn source parameters given concentration profiles and wind-field inputs. This type of architecture was shown to be reliable to approximate an inverse function as done in Georges (2020) and as proven by mathematicians (see Hornik et al. (1989)).

We consider three tensors C, c and Θ. C contains n examples of stationary concentration profiles induced by: 1- n wind coefficients contained in the tensor c (here a 1D vector) and 2- by a gaussian stationary source of parameters θ_k; $1 \leq k \leq n$. $[C, c]$ and Θ play the role of X and Y traditionally written in machine-learning framework, where X is the input tensor which is estimated by the Neural Network given a target tensor Y :

$$(X,\ Y) \equiv ([C, c], \Theta). \tag{10}$$

To train the model, a given number $n_{train} \in \mathbb{N}$ of simulations of pollutant dispersion are performed using the model written in (1) with various wind speeds c and source parameters θ. To cover all the space of possible training examples, wind speeds and training source parameters are generated with a Sobol sequence. Sobol sequence has the property to be a pseudo-random generator with low discrepancy so that each training example is not far in Euclidean distance from another set of parameters (see Niederreiter (1988)). Its python implementations used in this work are borrowed from an open-source module written by Ng (2016).

Training data are generated from n_{train} wind and source parameters c_k and θ_k , $1 \leq k \leq n_{train}$, with $c_k \in [-10, 10]$

m/s and $\theta_k \in [1e-2\ ,\ 2] \times\ [0.1, 0.9] \times\ [50,\ 10^4]$. The sources of pollution are considered to be defined by a bell curve of amplitude A_S such as a source is defined by,

$$q_k(x) = \theta_k(1)\, exp\left(-\theta_k(3)^2 (x - \theta_k(2))^2\right),\ 1 \leq k \leq n_{train} \tag{11}$$

The target tensor corresponds to $\Theta_k = [\theta_k(1), \theta_k(2), \theta_k(3)]$, $1 \leq k \leq n_{samples}$, with $n_{samples}$ the number of training or validation examples. Source parameters need to be scaled to the same order of magnitude. To carry this we build the following, adimensional training parameters,

$$\tilde{\theta}(j) = \frac{\theta(j) - min(\Theta(j))}{max(\Theta(j)} \qquad 1 \leq j \leq 3, \tag{12}$$

for each set of the n_{train} parameters used for the training. Thus the model returns an estimation of the adimensional parameters as an output. A rescaling is applied to the output estimate using $min(\Theta(j))$ and $max(\Theta(j))$ the scaling parameters used in training, such that a scaled estimation of source parameters is performed on the validation set. To avoid learning biases, the rescaling still uses the extrema of parameters from training to rescale validation examples.

2.2 Variational methods

Variational methods rely on the minimisation of a cost function using a known model of transport of pollutants. These kinds of methods are widely used and considered as a standard way to estimate parameters given some measurement especially with the use of adjoint estimation (Georges (2019), Nguyen (2016)). Here we propose to apply a variational method to a stationary problem where u can be expressed as an explicit function of θ, thus avoiding the use of the adjoint. The 3D-Var cost function is chosen. A regularisation term (Tikhonov and Arsenin (1977)) is provided by an initial guess such that the entire cost function is written,

$$\begin{aligned} \mathcal{J}(\theta) &= \frac{1}{2}(\theta - \theta^b)^T B^{-1}(\theta - \theta^b) \\ &+ \frac{1}{2}(\mathcal{F}(\theta) - u^m)^T R^{-1}(\mathcal{F}(\theta) - u^m) \end{aligned} \tag{13}$$

with,

$$\begin{aligned} u &= \mathcal{M}(\theta) \\ &= A^{-1}(B_1 q(\theta) + B_2 C), \end{aligned} \tag{14}$$

the concentration profile written by the physical model $\mathcal{M}$ written in (1). And with,

$$\mathcal{F}(\theta) = \mathcal{H} \circ \mathcal{M}(\theta), \tag{15}$$

a forward operator which is a composition of the observation operator $\mathcal{H}$ with $\mathcal{M}$ where $\mathcal{H}$ maps the concentrations returned by the model on the whole domain to the space of the measurements. Matrices B and R respectively contain the covariance errors of initialisation and the covariance errors of measurements. We assume here that these errors are uncorrelated so that B and R are diagonal– and with an equal variance. σ^b and σ^{obs} are the standard deviations associated with initial guess and observations. To proceed gradient descent, the following gradient is given to the cost function,

$$\begin{aligned} &\nabla\mathcal{J}(\theta) = \mathbf{B^{-1}}\left(\theta - \theta^b\right) + \\ &\left(\int_\Omega \sum_{i=1}^{N_C} \phi(x - x_i)\, \mathbf{A^{-1}} B_1 \nabla q(\theta) dx\right)^{\mathbf{T}} \mathbf{R^{-1}}\left(\mathcal{F}(\theta) - u^m\right). \end{aligned} \tag{16}$$

The optimisation of the cost function to find the optimal set of source parameters θ^* is carried out by a quasi-Newton descent using the L-BFGS-B algorithm.

3. CASE STUDY : SOURCE IDENTIFICATION IN A STREET

Source detection gains an increasing interest for the estimation sources of pollution in urban areas (e.g. Defforge et al. (2021), Kovalets et al. (2018)). The approach proposed here could consider an emissive vehicle stopped on a boulevard with its motor turned on and where stationarity has been reached.

3.1 Parameters

Database is built by building the set of parameters,

$$\left\{p_k = \left[A_{Sk},\, x_{Sk}, \sigma^2_{disp_k}, c_k\right]\ ;\ 1 \leq k \leq\ n_{train}\right\}, \tag{17}$$

with $\left[A_S,\, x_S, \sigma^2_{disp}, c_k\right] \in [1e-2, 4] \times [0.1, 0.9] \times [50, 2000] \times [-10e-3, 10e-3]$ generated by Sobol sequence, with the units $\left[\emptyset, km, km^{-2}, km/s\right]$ respectively. The validation set is created in the same manner with n_{test} samples, choosing $\left[A_S,\, x_S, \sigma^2_{disp}, c\right] \in [5e-2, 3] \times [0.1, 0.9] \times [70, 1500] \times [-10e-3, 10e-3]$. Source position is purposely generated so that $x_S \in [0.1, 0.9]$ to avoid to generate sources which would be located too close from domain boundaries leading to a non-physical solution of our dispersion model written in (1). We remind that every scaling parameter involved in (12) is based on training only. Our database contains $n_{train} = 15000$ training samples, whilst validation is applied on $n_{test} = 200$ samples. As a metric of performance, the relative root-square error (RRSE) is chosen, defined by,

$$RRSE = \frac{\|\hat{q} - q_T\|_2}{\|q_T\|_2}, \tag{18}$$

where $(q_T, \hat{q}) \in \mathbb{R}^{{N_x}^2}$ are respectively the real source profile and its estimate. To be fair with variational method, a fair initialisation is given with a vector,

$$\theta^b_k = [0,\ argmax(u^m_k),\ 0], \quad 1 \leq k \leq n_{test}; \tag{19}$$

so that the initial guessed source position is located at the highest measured concentration.

The best neural network architecture has been chosen after having performed some experiments. Hence the chosen Neural Network contains 2 dense hidden layers of 50 neurons each, whose *tanh* is the activation function. The optimisation is carried out using the L-BFGS-B algorithm. L2 regularisation is set at a low value : 10^{-12}. The machine learning framework is built using *keras* and *tensorflow* libraries.

Some estimates returned by the neural network and by the variational method are not physical because on one hand, the learning of source parameters in the Neural Network is not strongly constrained by the source model, on the other

hand, the variational method can stop gradient descent on local minima. We will exclude these few examples from our results.

3.2 Comparison without noise

Experiment with 12 sensors Simulations are performed without noise to compare how the machine learning framework performs next to the variational method. Several non-physical estimates returned by the neural network and bad estimates from the variational method are excluded: 16 out of 200, leading to the study of 184 estimates.

On these examples, the neural network is able to estimate a source with an RRSE in the same order of magnitude than with the standard 3D-Var, with a mean RRSE and a standard deviation of $\mu^{NN}_{RRSE} = 0.22$ and $\sigma^{NN}_{RRSE} = 0.12$ and $\mu^{var}_{RRSE} = 0.31$ and $\sigma^{var}_{RRSE} = 0.21$ using neural network and variational method respectively. Overall the Neural Network performs slightly better.

The root squared errors obtained from machine learning and variational framework shew to follow a log-normal law confirmed by the use of a Kolmogorov-Smirnov test ($p - value = 0.39 > 0.05$ and $0.44 > 0.05$ respectively). The distributions of RRSE are shown in Fig. 2.

The machine learning framework method is data-driven and thus does not contain any physical model whereas variational framework relies on the knowledge of the physics and on a good guess. The guess written in (19) specifies the position of the sensor which measures the maximum concentration. We notice that the RRSE metric is sensitive to a slightly error of estimated position. For example in the case of the 65^{th} sample of validation illustrated in Fig. 1, the neural network estimates well the amplitude and the dispersion, but a slight error of position leads to an $RRSE^{NN} = 0.16$ against $RRSE^{var} = 0.04$ for the neural network and the variational framework respectively. We will get back to the metric choice at the end of the paper.

Overall, without noise and applied to 12 sensors in total (10 regularly spaced in the domain and 2 which measure the boundary conditions), the machine learning framework performs slightly better than the variational framework regarding the Relative Root Square Error. The neural network gave sometimes non-physical estimates such as negative sources or non-Gaussian sources. Fig. 3 show that the performance of the two methods are close.

Experiment with 50 sensors: Simulations with 50 sensors are carried out to observe a sensitivity of the estimations towards measurement density. Several non-physical estimates returned by the neural network or by the variational framework are excluded of the network: 9 out of 200 (but not the same as with 12 sensors), leading to the study of 191 estimates. Both methods return similar results as previously, with mean and standard deviation of $\mu^{NN}_{RRSE} = 0.13$ and $\sigma^{NN}_{RRSE} = 0.08$ compared to $\mu^{var}_{RRSE} = 0.24$ and $\sigma^{var}_{RRSE} = 0.15$ for neural network and variational frameworks respectively. This means that the error decreased by about a half for the NN and by one third for 3D-Var comparing to the use of 12 sensors. Again the Neural Network performs better than 3D-Var. As with 12 sensors RRSE distributions shew to follow log-normal law with a

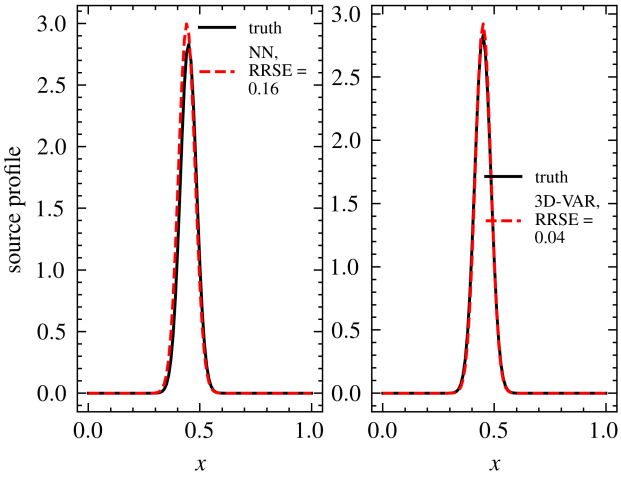

Fig. 1. Example of source reconstruction : left, with a neural network; right, with a variational method

Kolmogorov-Smirnov test ($p - value = 0.19 > 0.05$ for the NN and $p - value = 0.90 > 0.05$ for 3D-Var). Overall, source reconstruction using 50 noiseless sensors (48 regularly spaced in the domain and 2 which measure the boundary conditions) achieves 30 to 50% improvements compared to the case with 12 noiseless sensors. Fig. 3 shows that the performance of the neural network is clearly better than the variational framework, its distribution being narrower and closer to 0 on RRSE axis. On the other hand, both methods improve with the use of 50 sensors instead of 12.

3.3 Comparison with noise

In this experience, a gaussian noise n_s is applied to each sample s of data-set of concentration profiles after having applied a characteristic sensor function as described in 1.3, with n_s following the law :

$$n_s \sim \mathcal{N}(0, \sigma^{obs} = 5\% \times max(\mathcal{H}(u_s)), \quad 1 \leq s \leq n_{test} \text{ or } n_{train}. \tag{20}$$

Thus the gaussian noise of a set of measurement has a constant variance, which depends of the maximum of the noiseless-measured concentration profile. This hypothesis is reasonably conservative because the lower measured values of a concentration profile will be relatively noisier than its maximum value.

To be fair with variational method, the initial guess error was modified from $\sigma^b = 1 \times 10^6$ to $\sigma^b = 5 \times 10^4$ to to help to make the problem more convex.

Experiment with 12 sensors: Simulations are performed with noise in the same manner as previously. Several non-physical estimates returned by the NN or 3D-Var are excluded of the results: 9 out of 200. Leading to the study of 191 estimates. Both methods are close in the presence of noise with mean and standard deviation of $\mu^{NN}_{RRSE} = 0.48$ and $\sigma^{NN}_{RRSE} = 0.26$ using the NN and $\mu^{var}_{RRSE} = 0.52$ and $\sigma^{var}_{RRSE} = 0.23$ using 3D-Var. Fig. 4 shows that the distribution of Relative Root Squared Errors of both methods are close to each other.

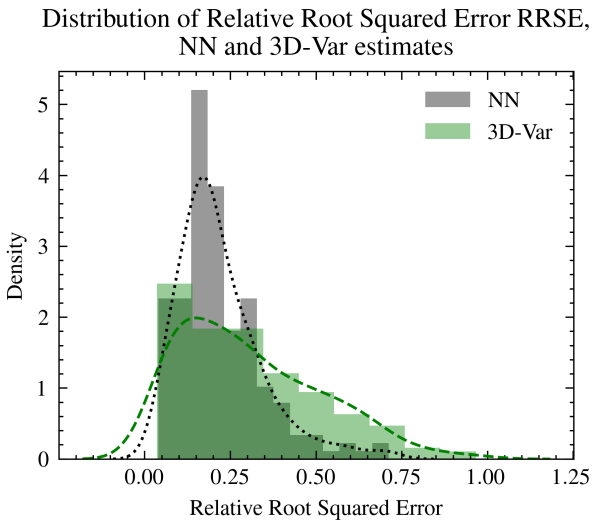

Fig. 2. Compared distribution of Relative Root Squared Error for obtained by Variational and Neural Network Estimates. No noise, 12 sensors

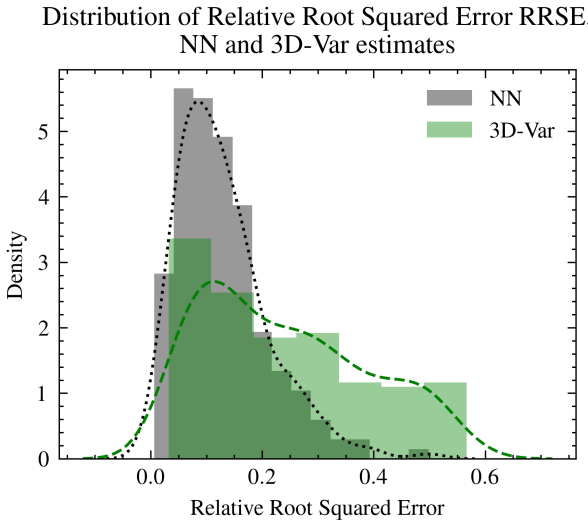

Fig. 3. Compared Distribution of Relative Root Squared Errors obtained by Variational and Neural Network Estimates. No noise, 50 sensors

Experiment with 50 sensors: With 50 sensors, 48 regularly spaced and two measured concentration values at boundaries, the results improve compared to the use of 12 sensors. This lead to mean and standard deviation metrics of $\mu^{NN}_{RRSE} = 0.40$ and $\sigma^{NN} = 0.21$ using NN and $\mu_{RRSE} = 0.45$ and $\mu^{var}_{RRSE} = 0.22$ using 3D-Var. Thus μ^{var}_{NN} and μ^{var}_{RRSE} decreased by 17 and 13 % respectively with the use of 50 sensors instead of 12. Hence the Neural Network performs slightly better the variational framework. However it can be seen in Fig. 5 that the distribution of errors are still fairly close between the two methods. Again error distributions were proven to follow log-normal law.

4. CONCLUSION

With the use of few and noisy sensors, both methods are close even if the NN returns more often accurate sources estimates. However, the machine learning framework showed greater improvement than the variational framework when increasing sensor density. We have shown that the machine learning approach is able to reconstruct sources even in the presence of noisy measurements. Sensor noise decreases severely the performance of Source Term Estimation for both methods. Overall, the performance of the Neural Network is as good or superior to the variational approach as demonstrated by table 1. These results revealed that machine learning approaches are promising for source detection in the context of urban air pollution. Future work will consider another metric such as Wasserstein distance to evaluate the performance of source reconstruction. Indeed, Wasserstein distance is more adapted to compare distributions than RRSE. The former is far less punishing with slight errors of source position than the latter. Nevertheless, the behavior of error distribution for NN and 3D-Var remain similar across all the cases adressed in this study using RRSE or either Wasserstein metric.

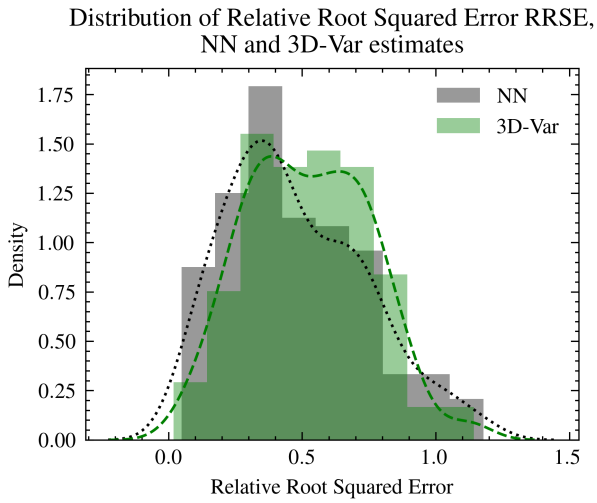

Fig. 4. Compared Distribution of Relative Root Squared Errors obtained by Variational and Neural Network Estimates. Noise, 12 sensors

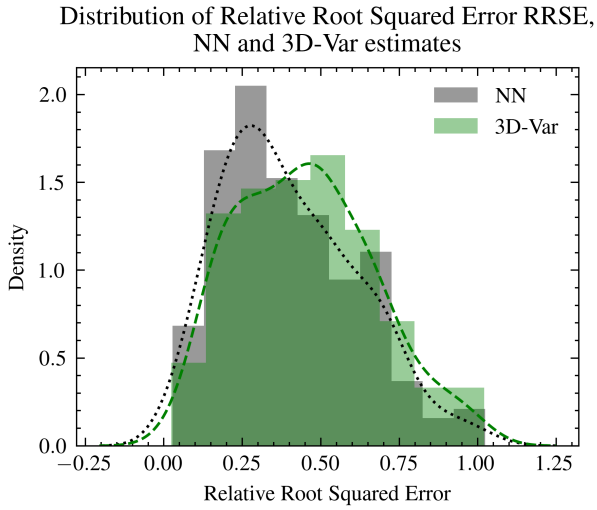

Fig. 5. Compared Distribution of Relative Root Squared Errors obtained by Variational and Neural Network Estimates. Noise, 50 sensors

It would be interesting to develop this approach for real world cases in urban area. Indeed, once trained, a neural network model could be useful to monitor online urban pollution given an urban network of sensors. We could also imagine a framework where a neural network could provide a proper initial guess to a variational method such as 3D-Var or 4D-Var.

Table 1. Measured mean and standard deviation of Relative Root Square Error of validation samples for each case of study

Case			**NN**		**3D-Var**	
N	N° sensors	noise	mean	std	mean	std
184	12	no	0.22	0.12	0.31	0.21
191	50	no	0.13	0.08	0.24	0.15
191	12	yes	0.48	0.26	0.52	0.23
191	50	yes	0.40	0.21	0.45	0.22

ACKNOWLEDGEMENTS

The authors acknowledge the french nuclear and alternative energies commission for the funding of this work and John D. Garrett (2021) for providing IEEE plot styles thanks to his open-source module.

REFERENCES

Armbruster, D.A. and Pry, T. (2008). Limit of Blank, Limit of Detection and Limit of Quantitation. *The Clinical Biochemist Reviews*, 29(Suppl 1), S49. Publisher: Australasian Association for Clinical Biochemistry and Laboratory Medicine.

Asch, M., Bocquet, M., and Nodet, M. (2016). *Data assimilation: methods, algorithms, and applications*. SIAM. Pages: xviii + 306.

Defforge, C.L., Carissimo, B., Bocquet, M., Bresson, R., and Armand, P. (2021). Improving Numerical Dispersion Modelling in Built Environments with Data Assimilation Using the Iterative Ensemble Kalman Smoother. *Boundary-Layer Meteorology*, 179(2), 209–240. doi: 10.1007/s10546-020-00588-9.

Garrett, J. (2021). SciencePlots (v1.0.9). doi: 10.5281/zenodo.5512926.

Georges, D. (2019). A Variational Calculus Approach to Wildfire Monitoring Using a Low-Discrepancy Sequence-Based Deployment of Sensors. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, 5912–5917. IEEE, Nice, France. doi: 10.1109/CDC40024.2019.9029962.

Georges, D. (2020). Machine Learning for Receding Horizon Observer Design: Application to Traffic Density Estimation. *IFAC*, 2020 - 21st IFAC World Congress 2020, 7.

Hammond, J., Chakir, R., Bourquin, F., and Maday, Y. (2019). PBDW: A non-intrusive Reduced Basis Data Assimilation method and its application to an urban dispersion modeling framework. *Applied Mathematical Modelling*, 76, 1–25. doi:10.1016/j.apm.2019.05.012.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366. doi:10.1016/0893-6080(89)90020-8.

Hutchinson, M., Oh, H., and Chen, W.H. (2017). A review of source term estimation methods for atmospheric dispersion events using static or mobile sensors. *Information Fusion*, 36, 130–148. doi: 10.1016/j.inffus.2016.11.010.

Kikumoto, H. (2020). Turbulent diffusivity limiter with travel time for eulerian-CFD eulerian analysis of point-source pollutant dispersion. *Wind Engineering Research*, 26, 9.

Kovalets, I.V., Efthimiou, G.C., Andronopoulos, S., Venetsanos, A.G., Argyropoulos, C.D., and Kakosimos, K.E. (2018). Inverse identification of unknown finite-duration air pollutant release from a point source in urban environment. *Atmospheric Environment*, 181, 82–96. doi: 10.1016/j.atmosenv.2018.03.028.

Krysta, M. and Bocquet, M. (2007). Source reconstruction of an accidental radionuclide release at European scale. 16.

Lin, C., Ooka, R., Kikumoto, H., and Jia, H. (2021). Eulerian RANS simulations of near-field pollutant dispersion around buildings using concentration diffusivity limiter with travel time. *Building and Environment*, 202, 108047. doi:10.1016/j.buildenv.2021.108047.

Ma, R., Liu, N., Xu, X., Wang, Y., Noh, H.Y., Zhang, P., and Zhang, L. (2019). A deep autoencoder model for pollution map recovery with mobile sensing networks. In *Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers - UbiComp/ISWC '19*, 577–583. ACM Press, London, United Kingdom. doi:10.1145/3341162.3349327.

Ng, V. (2016). crankycoder/sobol_seq. Original-date: 2015-04-28T23:36:48Z.

Nguyen, v.T. (2016). *Adjoint-based approach for estimation & sensor location on 1D hyperbolic systems with applications in hydrology & traffic*. Automatic, Université Grenoble Alpes.

Niederreiter, H. (1988). Low-discrepancy and low-dispersion sequences. *Journal of Number Theory*, 30(1), 51–70. doi:10.1016/0022-314X(88)90025-X.

Tikhonov, A.N. and Arsenin, V.Y. (1977). *Solution of Ill-posed Problems*.

Wang, B., Qian, F., and Zhong, W. (2019). Wind field reconstruction for the dispersion modeling of accidental chemical spills on complex geometry. *Chinese Journal of Chemical Engineering*, 27(11), 2712–2724. doi: 10.1016/j.cjche.2019.02.029.

Wang, J.W., Yang, H.J., and Kim, J.J. (2020). Wind speed estimation in urban areas based on the relationships between background wind speeds and morphological parameters. *Journal of Wind Engineering and Industrial Aerodynamics*, 205, 104324. doi: 10.1016/j.jweia.2020.104324.

Winiarek, V., Bocquet, M., Saunier, O., and Mathieu, A. (2012). Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant: ESTIMATION OF ERRORS IN INVERSE MODELING. *Journal of Geophysical Research: Atmospheres*, 117(D5), n/a–n/a. doi:10.1029/2011JD016932.