# Estimating Wind and Emission Parameters in an Atmospheric Transport Model

Andres Yarce Botero[1,2(✉)], Santiago Lopez Restrepo[2], Olga Lucia Quintero[2], and Arnold Heemink[1]

[1] Department of Applied Mathematics, TU Delft, 2628CC Delft, The Netherlands
{a.yarcebotero,A.W.Heemink}@tudelft.nl
[2] Mathematical Modelling Research Group, EAFIT University, 050022 Medellín, Colombia
{ayarceb,slopezr2,oquinte}@eafit.edu.co

**Abstract.** The present study proposes a novel data assimilation (DA) approach for estimating emission and wind direction parameters in an advection-diffusion model. This implementation aims to improve the prediction of a chemical transport model over long distances by updating the emission operator in the model using DA techniques. As a first step, we want to test the method in a small-scale scenario. A low-dimensional advection-diffusion model was utilized to evaluate the effectiveness of the proposed approach under various sampling observation numbers. The model's emission and wind parameters are perturbed as a source of uncertainty. The parameters are sequentially estimated with the adjoint-free Ensemble Kalman filter with an augmented state vector. These sequential DA techniques exploit the ensemble of multiple model realizations to reduce uncertainty in the state and parameter representation. An associated stream function with a divergence-free condition controls the wind fields, and the estimation of this stream function through the assimilation process allows corrections of the wind fields without violating physical laws. The technique's performance was compared against validation observations such as the Root-Mean Square (RMS), and it was found that the number of assimilated observations had a significant impact on the parameter estimations results. This study demonstrates the potential of the proposed DA approach for improving the prediction of transport in the advection-diffusion model through parameter estimation.

**Keywords:** Air Quality modeling · Parameter estimation problem · Chemical Transport Model · 2D advection-diffusion model · Data Assimilation · divergence-free · LOTOS-EUROS · Ensemble Kalman filter information · atmospheric transport model

## 1 Introduction

Chemical Transport Models (CTMs) predict the concentration, transport, and deposition of several components emitted from natural and anthropogenic sources. Although

Supported by TuDelft and EAFIT.


M. Mujica Mota and P. Scala (Eds.): EUROSIM 2023, CCIS 2032, pp. 31–43, 2024.
https://doi.org/10.1007/978-3-031-68435-7_3

these atmospheric transport models have been evolving for years, refining the physical, chemical, and mathematical structure representing reality, uncertainty is always there due to the highly nonlinear phenomena these models try to describe. One of the primary sources of uncertainty in CTMs comes from the emission inventories, which could differ significantly from the technique's location and time against reality. We inspired or contributed to a real-life scenario where we have a CTM model in a region with coarse meteorological input information, so the winds are the secondary source of uncertainty between the model and real directions. Uncertainty in the emission inventories or parameters involved in the emission processes may be reduced but never eliminated [1]. With the help of data assimilation (DA), incorporating information in these models alleviates some difficulties.

Emission parameter estimation has been the center of attention for DA around the globe for CTM implementations, mainly because the emission inventories tend to have high uncertainty values, with some different degrees of uncertainty between regions and source sectors [3,4]. In South America, updating emission inventories can be challenging due to historical institutional limitations, making it difficult to have up-to-date information available. This significant variability in CTMs from the emission parameters [5] has been given the principal attention in DA techniques to incorporate observations into numerical models for determining the most appropriate state and parameter sets that produce a new state that follows this observation [6]. DA method is commonly used to incorporate measurements into the system, estimate its states and parameters, and enhance subsequent simulation forecasts, making it a valuable tool for improving model performance.

In this work, the state and model parameters, emissions and winds, are estimated simultaneously using the Ensemble Kalman Filter (EnKF) to estimate the error covariances among several variables with a small scale model of a CTM, an emission advection-diffusion model. EnKF is the DA technique that generates an update at the first moment of the probability density function found in the analysis once new observations are available [8,9]. An ensemble of model propagation is created by perturbing emission factors and wind direction parameters. The method for estimating the parameters using the data provided by the observations uses covariances that include the relationship between states and initial conditions and compliance with the observed stated. The "augmented state space" vector technique, proposed by Jazwinsky in 1970 [10], is used to estimate parameters. Here both the state variables and the model parameters are part of the state vector and can be estimated through the sequential steps of the filter. This approach has been used in different parameter estimation studies, specifically with the LOTOS-EUROS mode" as CTM [11,12].

This chapter is organized as follows: In Sect. 2, we introduce the specific methodology for the EnKF, including the stochastic model for parameter perturbation and propagation, as well as the technical details of the code and stream function formulation. In Sect. 3, we focus on the results of this technique with the advection-diffusion model. In Sect. 4, we present our discussion and concluding remarks. This study offers valuable information on the effectiveness of stream function DA in improving CTM and can contribute to developing more accurate air quality monitoring systems.

## 2 Methodology

### 2.1 The Ensemble Kalman Filter for Wind and Emission Parameter Estimation

Ensemble Kalman Filter (EnKF) is a sequential DA mathematical procedure used to incorporate data into a dynamic model to improve its representation of reality [19,20]. One benefit of the such sequential technique is the estimation of the current state $\mathbf{x} \in \Re^{n\times 1}$, with $n$ as the state dimension of a dynamical system that evolves according to some numerical model operator $\mathcal{M}$ from the past to the current time.

$$\mathbf{x}_k = \mathcal{M}_{(k-1)\rightarrow k}(\mathbf{x}_{(k-1)}(\lambda_{k-1}, \gamma_{k-1}))\,, \tag{1}$$

where $k$ denotes the time index. It follows here that a discretization in time takes place within the EnKF DA technique. This model operator transforms the model states from time $k-1$ to time $k$. $\lambda_k$ represents the emission parameter and $\gamma_k$ represents the wind direction parameter for the current time $k$, with the following distribution:

$$\gamma \sim \mathcal{N}\left(\gamma_n, Q_1\right)\,, \tag{2}$$

$$\lambda \sim \mathcal{N}\left(\lambda_n, Q_2\right)\,, \tag{3}$$

where the mean corresponds to the nominal values $\gamma_n$ and $\lambda_n$ of the parameters, and the covariances $Q_1$ and $Q_2$ describe the degree of freedom of the magnitude that those to have in the system.

### 2.2 Stochastic Uncertainty Representation for Wind and Emission Parameters

A stochastic representation of the parameter uncertainty is required to implement the DA method. The emissions employed by the model operator are consequently modeled as a stochastic process with a factor of random variation.

$$\hat{\boldsymbol{e}}_k \;=\; \boldsymbol{e}_k\,(\,1\,+\,\boldsymbol{\gamma}_k\,)\,. \tag{4}$$

Here, $\boldsymbol{e}_k$ is the nominal emission from the emission inventory. The emission deviation is modeled as a colored noise process [10] as follows:

$$\boldsymbol{\gamma}_k \;=\; \alpha_k \boldsymbol{\gamma}_{k-1} + \sigma\sqrt{1-\alpha^2}\,\boldsymbol{w}_\gamma\,, \tag{5}$$

where $\boldsymbol{w}_k$ is a white noise process with zero mean and unity standard deviation:

$$\boldsymbol{w}_\gamma \;\sim\; \mathcal{N}(0,1)\,. \tag{6}$$

The stochastic factors are taken from a normal distribution with zero mean and standard deviation $\sigma$. The temporal correlation coefficient $\alpha_k \in [0,1]$ is used to describe the temporal variation, where the value should be set between two extremes: for $\alpha = 0$, the deviation is pure white noise with completely different values for every sample; for $\alpha = 1$ there is no temporal variation at all, and the deviation factor is a single sample

out of the normal distribution. In this study, the correlation parameter is described using a temporal length scale $\tau$ following [21]:

$$\alpha_k \;=\; \exp\left(-|t_k - t_{k-1}|/\tau\right) \;, \tag{7}$$

where the stochastic model state is formed by augmenting the state vector with a similar model for the wind parameter perturbation. The DA estimation is performed based on a first guess or prior estimate $\mathbf{x}^b \in \Re^{n\times 1}$. In ensemble-based methods, an ensemble of $N$ model realizations and $n$ states is generated:

$$\mathbf{X}_k^b = [\mathbf{x}_k^{b[1]}, \mathbf{x}_k^{b[2]}, .., \mathbf{x}_k^{b[N]}] \in \Re^{n\times N} \;, \tag{8}$$

which correspond to the different realization of the model generated by perturbing initial conditions assumed to be normally distributed:

$$\mathbf{x} \sim \mathcal{N}(\mathbf{x}^b, \mathbf{B}) \;. \tag{9}$$

The mean $\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}^i$ and the covariance background error matrix are:

$$\mathbf{B} = \frac{1}{n-1}\sum_{i=1}^{n}(\mathbf{x}^i - \overline{\mathbf{x}})(\mathbf{x}^i - \overline{\mathbf{x}})^T \in \Re^{n\times n} \;. \tag{10}$$

The parameters involved are the emission factor $\gamma$ and the wind direction $\lambda$ factor, which has the following distribution for the ensemble method of N model realizations:

$$\mathit{\Gamma} = [\gamma_{[1,1]}, \gamma_{[1,2]}, ..., \gamma_{[1,N]}] \;, \tag{11}$$

$$\mathit{\Lambda} = [\lambda_{[1,1]}, \lambda_{[1,2]}, ..., \lambda_{[1,N]}] \;. \tag{12}$$

Each one of the model state, emission factor, and wind deviation matrices ensembles are rest by the mean value of the ensembles:

$$\mathbf{X}_k^b = [\mathbf{x}_k^{b[1]} - \overline{x}, \mathbf{x}_k^{b[2]} - \overline{x}, .., \mathbf{x}_k^{b[N]} - \overline{x}] \in \Re^{n\times N} \;, \tag{13}$$

$$\mathit{\Gamma} = [\gamma_{[1,1]} - \overline{\gamma}, \gamma_{[1,2]} - \overline{\gamma}, ..., \gamma_{[1,N]} - \overline{\gamma}] \;, \tag{14}$$

$$\mathit{\Lambda} = [\lambda_{[1,1]} - \overline{\lambda}, \lambda_{[1,2]} - \overline{\lambda}, ..., \lambda_{[1,N]} - \overline{\lambda}] \;. \tag{15}$$

From these deviation arrays, we create the following augmented vector and calculate the approximate covariance matrix of this augmented system:

$$P_a = \begin{bmatrix} \mathbf{X} \\ \cdots \\ \mathit{\Gamma} \\ \cdots \\ \mathit{\Lambda} \end{bmatrix} \;. \tag{16}$$

This matrix can be expanded for each one of the components:

$$\begin{bmatrix} \mathbf{x}_k \\ \boldsymbol{\gamma}_k \\ \boldsymbol{\lambda}_k \end{bmatrix} = \begin{bmatrix} \mathcal{M}_{(k-1)\to k}\left(\mathbf{x}_{(k-1)}(\lambda_{k-1}, \gamma_{k-1})\right) \\ \alpha_k\, \boldsymbol{\gamma}_{k-1} \\ \beta_k\, \boldsymbol{\lambda}_{k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \sigma\sqrt{1-\alpha^2}\boldsymbol{w}_\gamma \\ \eta\sqrt{1-\beta^2}\boldsymbol{w}_\lambda \end{bmatrix}. \tag{17}$$

Using the augmented vector (17), it is possible to estimate both the state and the emission correction factor using a sequential DA approach. The non-linear operator $\mathcal{M}$ propagates the augmented state vector $\boldsymbol{x}$ in time, while the right part of the expression corresponds to the stochastic forcing $\boldsymbol{w}_k$ over the elements of the state. The covariance matrix is calculated as

$$\mathbf{P} = \frac{P_a P_a^T}{N-1}\,. \tag{18}$$

The observations are also assumed to be normally distributed

$$\mathbf{y} \sim \mathcal{N}\left(\mathbf{H}\cdot\mathbf{x}, \mathbf{R}\right)\,, \tag{19}$$

where $\mathbf{H} \in \Re^{m\times n}$ is a linear operator that propagates the state space into the observation space, and $R \in \Re^{m\times m}$ is the observation error covariance matrix. In this way, the deviation observation matrix is:

$$\mathbf{Y} = [\mathbf{y}_{[1,1]} - \overline{\mathbf{y}}, \mathbf{y}_{[1,2]} - \overline{\mathbf{y}}, ..., \mathbf{y}_{[1,N]} - \overline{\mathbf{y}}]\,. \tag{20}$$

The EnKF computes an optimal variance gain with the expression:

$$\mathbf{K} = \mathbf{P}\mathbf{H}^T[\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}]^{-1}\,, \tag{21}$$

which is applied to each member of the ensemble:

$$\mathbf{x}_i^a = \mathbf{x}^b + [\mathbf{P}^{-1} + \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}]^{-1}\mathbf{H}^T\mathbf{R}\mathbf{d}, \tag{22}$$

where the innovation $\mathbf{d}$ deviation space is created with the inclusion of noise $v_i \sim N(0, \mathbf{R})$ is added to the observation $y_i = y + v_i$ for each ensemble that is going to be used to perturb the observations before the difference with the correspondent $\mathbf{H}_i\mathbf{x}$:

$$\mathbf{d}_k = [\mathbf{y}_1 - \mathbf{H}_1\mathbf{x}; \mathbf{y}_2 - \mathbf{H}_2\mathbf{x}, ..., \mathbf{y}_n - \mathbf{H}_n\mathbf{x}]\,. \tag{23}$$

### 2.3 Stream Function Formulation

Consider the scalar stream function

$$\Psi(x, y)\,, \tag{24}$$

where

$$u = \frac{\partial\Psi}{\partial y}\,, \tag{25}$$

$$v = -\frac{\partial \Psi}{\partial x} \,. \tag{26}$$

This scalar stream function constructs the vector field, satisfying the divergence-free property $\nabla \cdot f = 0$.

$$f(x, y) = \begin{pmatrix} u \\ v \end{pmatrix} \tag{27}$$

The stream function $\Psi(x, y)$ is a scalar function describing fluid flow in a two-dimensional space. The stream function divergence-free property $\nabla \cdot f = 0$ means that the total flow across any closed curve in the fluid domain is zero. This feature is beneficial because it allows physical constraints, such as continuity or conservation laws. The proposed approach involves perturbing the stream function instead of the two wind velocities. The stream function's divergence-free property is characteristic of an auxiliary space in DA used for parameter estimation, which can have some properties that help describe the magnitudes related to the model space. Section 3 provides an illustrative example of this method, where we select a suitable stream function to demonstrate its effectiveness. Moreover, the stream function divergence-free property reduces the dimensionality of the DA problem. The stream function can transform the state vector into a new space with a smaller dimension, reducing computational costs. The stream function can be used as an auxiliary space for DA by transforming the state vector.

## 3 Results

### 3.1 Advection-Diffusion Model

Testing an implementation with a simplified model is widely used in the DA field to check the behavior of new techniques designed. This section considers a 2D advection-diffusion model to which synthetic data is fed through data assimilation (DA). The advection-diffusion model for a gas concentration $C$ is described by the following equation in the directions $x$ and $y$ with a constant diffusion coefficients $D_x$ and $D_y$, and velocities in the two directions $U_x$ and $U_y$:

$$\frac{\partial C}{\partial t} = -U_x \frac{\partial C}{\partial x} - U_y \frac{\partial C}{\partial y} + D_x \frac{\partial^2 C}{\partial x^2} + D_y \frac{\partial^2 C}{\partial y^2} + \hat{\boldsymbol{e}}_k \,. \tag{28}$$

Synthetic data in this scenario correspond to the data sampled from the true model states of the advection-diffusion model propagation with different initial conditions. The true model propagation is one propagation of the model that is taken as the reality. This model propagation is perturbed, and then this synthetic data sampled from here is the observation dataset to assimilate. The change in parameters presents a challenge for DA and predictability. The size of this simple test model is 3600 states ($60 \times 60$ grid) for the concentrations. The emission parameter depends on the number of sources to be updated, and the wind direction parameter the number is two times the number of states because the horizontal wind directions are $u$ and $v$. In this case, one advantage of the stream function is reducing the number of wind direction parameters in half because

once the stream function parameters are estimated, the wind direction is calculated in the model space through the gradient.

For the discretization form, the central step second order finite difference [2] was used to generate the propagating following expression from an initial state. The boundary condition used for solving the experiment was the Dirichlet homogeneous zero or null value fixed in the contour [27].

### 3.2 Results Advection-Diffusion Model

Figure 1 provides a schematic normally-distributed perturbation for the wind direction along with a comparison of various standard deviation values of the wind direction perturbation for the advection-diffusion model as a grey shade region that describes a set of variations in the direction of the wind from the ideal or true value. In comparison, the different standard deviation values of the wind direction perturbation in the advection-diffusion model Eq. 28 determine the region of uncertainty expanded where the variation in the modeled wind direction occurs. The Figure provides insights into how perturbations in the wind direction can impact the accuracy of the advection-diffusion model. It highlights the importance of accurately estimating wind direction in improving model accuracy. Furthermore, comparing the different standard deviation values of the wind direction perturbation is a useful guide for determining the maximum level of uncertainty or variability that should be considered when modeling wind direction for various applications. The wind fields $V_w$ were homogeneously perturbated in direction assuming that the magnitudes remain constant and equal to $|V_w| = 1$. The noise distribution for generating the distribution is $\mathcal{N}_y \sim \mathcal{N}(\frac{1}{\sqrt{2}}, R)$

$$V\hat{y}_k = V\hat{y}_{k-1} + \boldsymbol{\lambda}_{k-1}\,, \tag{29}$$

$$V\hat{x}_k = \sqrt{1 - V^2\hat{y}_k}\,, \tag{30}$$

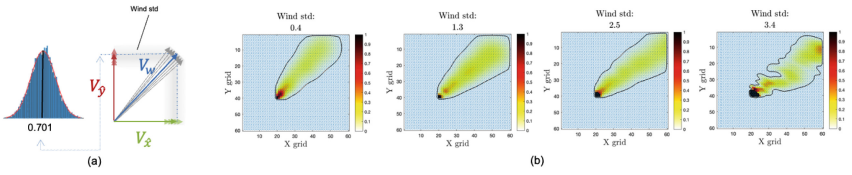


**Fig. 1.** (a) Schematic perturbation for the wind direction according to the model in Eq. (30) both in one and in two dimensions (b) Comparison for different standard deviation values of the wind direction perturbation for the advection-diffusion model. Colorbar indicates concentration. Different panels correspond to different time steps as indicated.

$$\boldsymbol{\lambda}_k \;=\; \beta_k\,\boldsymbol{\lambda}_{k-1} + \eta\sqrt{1-\beta^2}\,\boldsymbol{w}_\lambda. \tag{31}$$

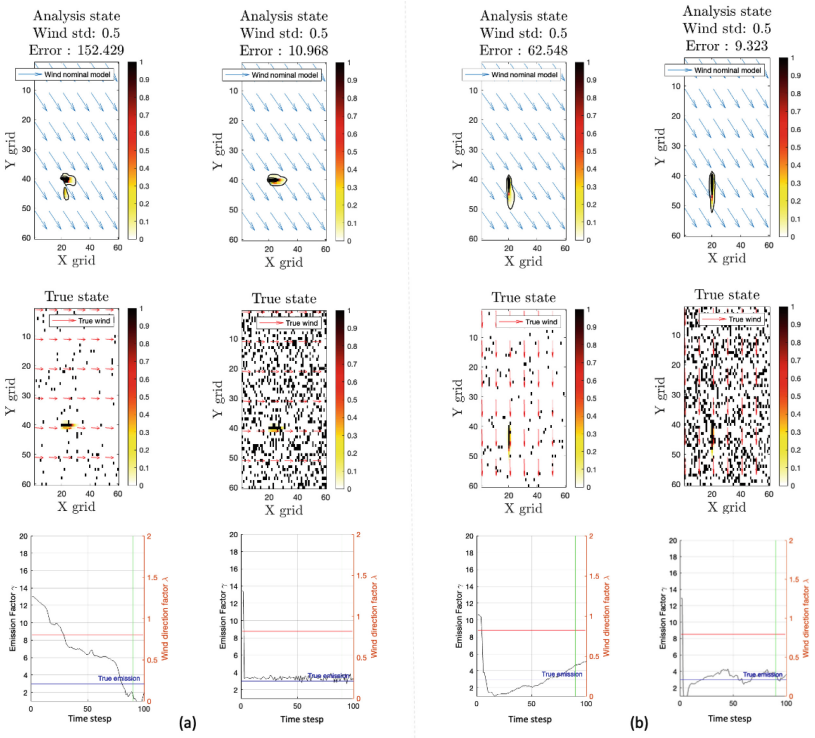


**Fig. 2.** Panels (a) and (b) show the result of the advection-diffusion emission parameter perturbation and estimation with wind direction not being estimated for the true wind direction pointing to the east (a) and the true wind pointing to the south (b). For each one of these cases, two situations for the dimension of the observation operator $\mathcal{H}$ are shown, one where the number of observations is low and the other where they are high. Colorbars represent the concentration levels. In the last row, the emission and wind direction parameters in time are presented.

The following Figures compare the analysis output between the truth state and three scenarios: emission perturbed, winds perturbed, or emissions and winds perturbed. The number of observations was increased for each experiment to study the discrepancy between the analysis state and the true change. Figure 2 presents the results of an Advection Diffusion emission parameter estimation DA experiment.

The simulation was conducted under two wind directions, one pointing towards the east, and the other towards the south. The top image shows the analysis output, indicating that the wind directions were not updated. Two situations for the dimension of the observation operator $H$ are presented for each case, which has a varying impact on the error reduction. This Figure highlights how fundamental it is to accurately estimate wind directions and how large is the impact of the number of observations on the analysis output.

Figure 3 describes an Advection Diffusion model DA for estimating emission parameters and wind direction. The top image shows the analysis wind field, which

results from the DA process. The middle image shows the true wind direction, pointing east in the first case and south in the second case. For each case, two situations with different observation dimensions are shown, where the graphic below shows the estimated parameters. This Figure helps to demonstrate the impact of DA on estimating parameters and wind direction in the Advection Diffusion model. The results suggest that the DA process can improve the accuracy of the estimated parameters, even with fewer observations, and that the true wind direction can be better estimated with more observations.

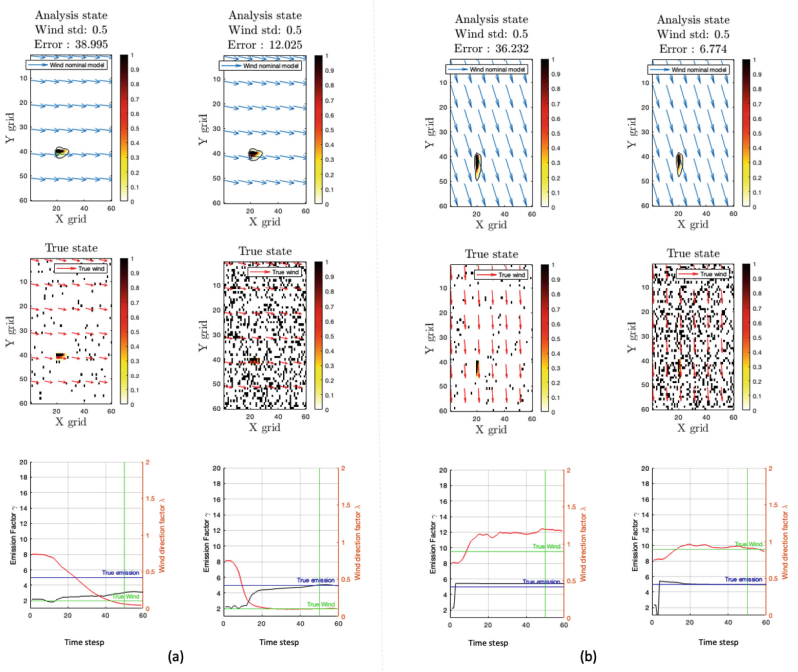


**Fig. 3.** Panels (a) and (b) show the result of the Advection-diffusion emission parameter perturbation and estimation with wind direction now being estimated for the true wind direction pointing to the east (a) and the true wind pointing to the south. For each one of these cases, two situations for the dimension of the observation operator $\mathcal{H}$ are shown, one where the number of observations is low and the other where they are high. Colorbars represent the concentration levels. In the last row, the emission and wind direction parameters in time are presented.

Figure 4 compares the error between the true state and the analysis state after assimilation and how it relates to the sampled space size (the number of observations as the black pixels in the true state of the images above). The experiment was conducted 50 times under two scenarios: one with a small (2.7%) and another with a large (55%) percent of sampled space size. The Figure shows the density of the group of dots, which confirms the experiment's success in reducing error consistently. The analysis suggests

that a larger sample size leads to a more accurate analysis state and lower error. This finding emphasizes the importance of adequately sampling the state space to improve the accuracy of the assimilation process.

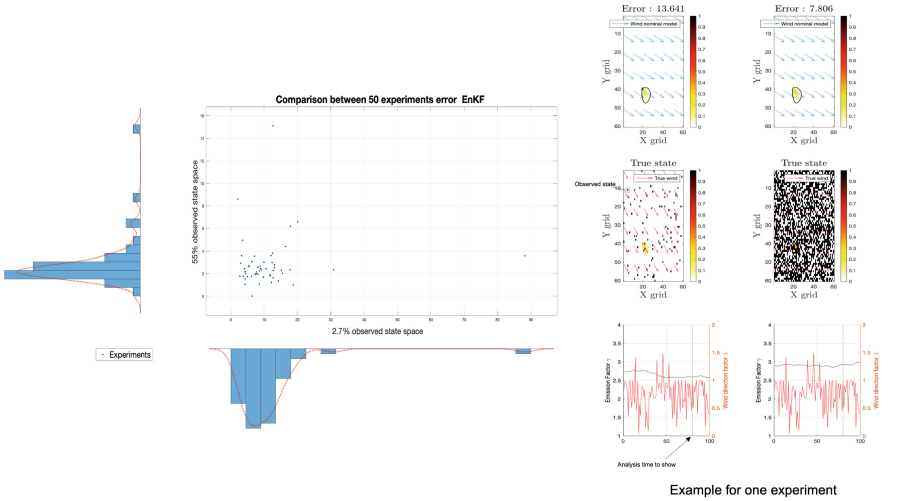


**Fig. 4.** Statistical comparison for the error for 50 experiments between one scenario with a small (2.7%) and large (55%) size, compared to the state space size, for the observation matrix $\mathcal{H}$.

To test the case with the scalar stream function as an auxiliary space to perform the DA for the parameter estimations, we use the following linear function for a test case.

$$\Psi_b(x, y) = u_b y - v_b x \,, \tag{32}$$

where

$$u_b = \frac{\partial \Psi_b}{\partial y} \,, \tag{33}$$

$$v_b = -\frac{\partial \Psi_b}{\partial x} \,. \tag{34}$$

Figure 5 compares the stream function correction after sequential DA, which involves estimating emission and wind direction parameters. This Figure illustrates the true and analysis wind direction corresponding to the gradient direction, depicted with the red and blue arrows, respectively. Over time, the ends of the arrows align, indicating that the DA process successfully corrected the stream function. This result highlights the potential of DA techniques to improve model accuracy and the importance of accurately estimating emission and wind direction parameters. Overall, the insights from this Figure provide a valuable contribution to the literature on DA and offer guidance to researchers in optimizing their modeling approaches.

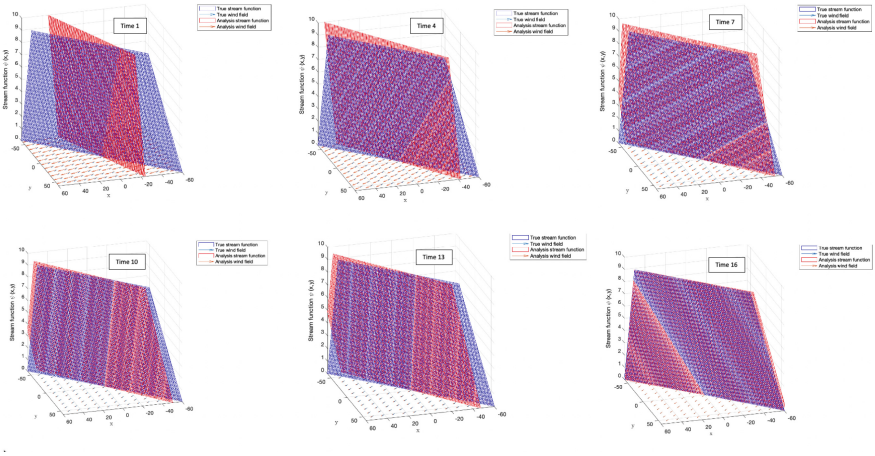


**Fig. 5.** Comparison of the stream function correction after time through the sequential DA, which estimates emission and wind direction parameters, the true and analysis wind direction correspond to the gradient direction of the stream function, which ends align through time.

## 4 Discussion

DA techniques have been increasingly incorporated into CTM to improve their accuracy. Most of these models use DA to improve emission estimates by ingesting new data. The next step in our research is to test the stream function DA technique with advection diffusion in a larger model, such as the LOTOS-EUROS, to evaluate its effectiveness. The results of this work can provide valuable insights into the application of DA techniques in different domains and contribute to the ongoing efforts to improve the accuracy of CTMs. Atmospheric trace gasses measurements retrieved from Satellite instruments can perceive different plumes of pollutants emitted from sources and recently different works propose techniques to detect them [22–24], thus estimating the wind direction through the incorporation to the model of concentration plumes can act as a virtual sensor of wind direction.

## 5 Conclusions

In conclusion, this study presents a novel DA approach that addresses the challenge of estimating emission and wind direction parameters in an advection-diffusion model for improving the accuracy of CTM over long distances. The proposed approach utilizes the ensemble Kalman filter with an augmented state vector to sequentially estimate both parameters while considering additional uncertainties. By perturbing both emission and wind parameters, the proposed approach provides a comprehensive understanding of both parameters' impact on the chemical transport prediction.

The effectiveness of the proposed approach was evaluated using a 2D low-dimensional advection-diffusion model under various sampling observation numbers.

The technique's performance was compared against validation observations using different metrics such as Mean Factorial Bias (MFB) and Root-Mean-Square (RMS). The results indicate that the number of assimilated observations significantly impacts the accuracy of the predicted chemical transport.

One of the significant advantages of the proposed approach is the estimation of the associated stream function that controls the wind fields. The assimilation of this stream function allows for the correction of the wind fields while preserving the physical laws. Considering the uncertainties associated with wind direction, the proposed approach provides a more accurate representation of chemical transport in advection-diffusion models. This study demonstrates the potential of the proposed DA approach for improving the accuracy of CTM in advection-diffusion models. The next crucial step towards advancing our research is introducing spatially variable perturbations to the stream function, thereby generating a diverse range of spatially varying velocity patterns to introduce divergence-free structures on the wind fields to be estimated.

**Acknowledgements.** Universidad EAFIT and TuDelft University.

## References

1. Montoya, O.L.Q., Niño-Ruiz, E.D., Pinel, N.: On the mathematical modeling and data assimilation for air pollution assessment in the Tropical Andes. J. Environ. Sci. Pollut. Res. **27**, 1–20 (2020)
2. Laplace, P.-S., Courant, R.: The finite difference method
3. Park, S., Dash, U., Yu, J., Yumimoto, K., Uno, I., Song, C.: Implementation of an ensemble kalman filter in the community multiscale air quality model (CMAQ model v5.1) for data assimilation of ground-level $PM_{2.5}$. Geoscientific Model Dev. **15**, 2773–2790 (2022). https://gmd.copernicus.org/articles/15/2773/2022/
4. Solazzo, E., Crippa, M., Guizzardi, D., Muntean, M., Choulga, M., Janssens-Maenhout, G.: Uncertainties in the emissions database for global atmospheric research (EDGAR) emission inventory of greenhouse gases. Atmos. Chem. Phys. **21**, 5655–5683 (2021)
5. Bocquet, M., et al.: Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models. Atmos. Chem. Phys. **15**, 5325–5358 (2015)
6. Mo, J., Gong, S., He, J., Zhang, L., Ke, H., An, X.: Quantification of SO2 emission variations and the corresponding prediction improvements made by assimilating ground-based observations. Atmosphere **13**, 470 (2022)
7. Sekiyama, T., Kajino, M., Kunii, M.: Ensemble dispersion simulation of a point-source radioactive aerosol using perturbed meteorological fields over eastern Japan. Atmosphere **12**, 662 (2021)
8. Evensen, G., Van Leeuwen, P.: An ensemble Kalman smoother for nonlinear dynamics. Mon. Weather Rev. **128**, 1852–1867 (2000)
9. Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. **53**, 343–367 (2003)
10. Jazwinski, A.: Stochastic Processes and Filtering Theory. Academic Press, Cambridge (1970)
11. Skoulidou, I., et al.: Changes in power plant NOx emissions over northwest Greece using a data assimilation technique. Atmosphere **12**, 900 (2021)

12. Lopez-Restrepo, S., et al.: An efficient ensemble Kalman Filter implementation via shrinkage covariance matrix estimation: exploiting prior knowledge. Comput. Geosci. **25**, 985–1003 (2021)
13. Menut, L., Bessagnet, B.: What can we expect from data assimilation for air quality forecast? Part I: quantification with academic test cases. J. Atmos. Oceanic Technol. **36**, 269–279 (2019)
14. Hanea, R., Velders, G., Heemink, A.: Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport model. J. Geophys. Res. Atmos. **109** (2004)
15. Lu, S., Lin, H., Heemink, A., Fu, G., Segers, A.: Estimation of volcanic ash emissions using trajectory-based 4D-Var data assimilation. Mon. Weather Rev. **144**, 575–589 (2016)
16. Jin, J., Lin, H., Heemink, A., Segers, A.: Spatially varying parameter estimation for dust emissions using reduced-tangent-linearization 4DVar. Atmos. Environ. **187**, 358–373 (2018)
17. Lopez-Restrepo, S., Yarce, A., Pinel, N., Quintero, O., Segers, A., Heemink, A.: Forecasting PM10 and PM2.5 in the Aburrá Valley (Medellin, Colombia) via EnKF based data assimilation. Atmos. Environ. **232**, 117507 (2020)
18. Milewski, T., Bourqui, M.: Potential of an ensemble Kalman smoother for stratospheric chemical-dynamical data assimilation. Tellus A Dyn. Meteorol. Oceanogr. **65**, 18541 (2013)
19. Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res. Oceans **99**, 10143–10162 (1994)
20. Anderson, J., Anderson, S.: A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. Mon. Weather Rev. **127**(12), 2741–2758 (1999)
21. Barbu, A., Segers, A., Schaap, M., Heemink, A., Builtjes, P.: A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe. Atmos. Environ. **43**, 1622–1631 (2009). https://doi.org/10.1016/j.atmosenv.2008.12.005
22. Finch, D., Palmer, P., Zhang, T.: Automated detection of atmospheric NO 2 plumes from satellite data: a tool to help infer anthropogenic combustion emissions. Atmos. Meas. Tech. **15**, 721–733 (2022)
23. Georgoulias, A., Boersma, K., Van Vliet, J., Zhang, X., Zanis, P., Laat, J.: Detection of NO2 pollution plumes from individual ships with the TROPOMI/S5P satellite sensor. Environ. Res. Lett. **15**, 124037 (2020)
24. Kuhlmann, G., et al.: Detectability of CO 2 emission plumes of cities and power plants with the Copernicus Anthropogenic CO 2 Monitoring (CO2M) mission. Atmos. Measur. Tech. **12**, 6695–6719 (2019)
25. Brandis, A., Centurelli, G., Schmidt, J., Vollmer, L., Djath, B., Dörenkämper, M.: An investigation of spatial wind direction variability and its consideration in engineering models. Wind Energy Sci. **8**, 589–606 (2023)
26. Potts, D., Timmis, R., Ferranti, E., Vande Hey, J.: Identifying and accounting for the Coriolis effect in satellite NO 2 observations and emission estimates. Atmos. Chem. Phys. **23**, 4577–4593 (2023)
27. Cheng, A., Cheng, D.: Heritage and early history of the boundary element method. Eng. Anal. Boundary Elem. **29**, 268–302 (2005)