AN ADVANCED HYBRID MODEL BASED ON STOCHASTIC-EULERIAN NUMERICAL APPROACH: APPLICATION TO ATMOSPHERIC POLLUTION

AMINE AJDOUR*, BRAHIM YDIR, JAMAL CHAOUFI, RADOUANE LEGHRIB

Laboratory of Materials, Signals, Systems and Physical Modeling, Physics Department, Faculty of Sciences, Ibn Zohr University, Agadir, Morocco
* *Corresponding author, E-mails*: amine.ajdour@edu.uiz.ac.ma; aajdour@gmail.com

Received September 9, 2024

Abstract. In this paper, we propose for the first time to the best of our knowledge, extend the application of a stochastic Eulerian numerical approach based on the Extended Kalman Filter ($EKF_{E.N.M.}$) to address the limitations of the Eulerian air pollution model CHIMERE. This approach integrates a comprehensive set of processes, including advection, turbulence, chemical reactions, emissions, and deposition, to model the dynamics of pollutant mass concentration. The EKF technique is employed to transform nonlinear dynamic problems into a succession of locally linearized ones, which are then used to estimate system states and adjust pollutant concentrations based on measured data. This stochastic approach is tested through two scenarios: one without external forces or control terms, and another that incorporates external factors like temperature, wind speed, and nitrogen dioxide as ozone precursors. A comparison of the obtained results with those from the standard CHIMERE model and studies from the literature demonstrates the accuracy and effectiveness of the proposed method.

Key words: Eulerian numerical model; Stochastic model; Extended Kalman filter; atmospheric pollution dynamics; mass conservation equation.

DOI: https://doi.org/10.59277/RomJPhys.2024.69.808

1. INTRODUCTION

Air pollution modeling is crucial in civil and industrial engineering and is a significant focus for many researchers today [1–4]. Mathematical models are essential for predicting air pollution, making robust software necessary for simulating the dispersion of atmospheric pollutants [5–7]. This need has driven the development of numerical methods that can address the challenges associated with various formulations.

Currently, one of the most effective approaches for modeling air pollution from Eulerian numerical model (E.N.M.) is the use of models such as CHIMERE. These models translate the processes identified in the conceptual model into mathematical equations, such as the atmospheric diffusion equation or the mass conservation equation, to represent the evolution of pollutant concentrations over space and time. In the vast majority of cases, there is no analytical solution to these mathematical equations, necessitating the development of numerical (approximate) solutions [8, 9]. CHIMERE utilizes three-dimensional (3D) grid-based models to simulate pollutant dispersion and transformations. However, there are inherent uncertainties associated with deterministic models [10, 11]. The conceptual model may be incomplete or partially incorrect, and the mathematical equations do not always precisely capture the physical and atmospheric processes. Numerical algorithms can introduce errors, and input data, such as meteorological conditions and emission sources, are often uncertain [12]. Additionally, boundary conditions and initial conditions are not perfectly known, and the spatial and temporal resolutions of the model may not capture all phenomena, with values averaged over model grid cells and time intervals [13–15]. Despite these limitations, CHIMERE provides valuable insights into air quality by employing fluid mechanics principles and advanced numerical techniques to model complex atmospheric processes. However, it requires continual adjustments and corrections to account for uncertainties and inaccuracies, ensuring that the model remains as accurate and reliable as possible.

Recently, several methods have been developed for the numerical solution of atmospheric pollution equations. For instance, the Eulerian approach used in CHIMERE is complemented by advanced numerical techniques to handle complex processes such as pollutant dispersion and atmospheric reactions [16, 17]. Researchers have explored various methodologies to improve the accuracy and efficiency of these models. For example, the use of numerical method such as, finite element methods coupled with has been investigated to enhance resolution in areas of high pollutant concentration [18, 19]. Additionally, the implementation of ensemble forecasting techniques helps address uncertainties in meteorological data and emission sources, providing a range of possible outcomes for better risk assessment [20]. In recent studies, the integration of machine learning algorithms with Eulerian numerical models has shown promise in refining predictions and optimizing model performance [21, 22]. These developments illustrate the ongoing efforts to advance atmospheric pollution modeling, addressing both the challenges of accuracy and the need for reliable predictions in atmospheric pollution management.

In this work, we adapt the CHIMERE model, a Eulerian numerical model, to predict mass concentration. By integrating CHIMERE with the Weather Research and Forecasting (WRF) model and using EmiSurf emissions data, we enhance our ability to simulate air pollution dynamics. WRF provides meteorological data, such as energy transfer, surface processes, and atmospheric conditions, while EmiSurf handles detailed emissions modeling. This combined data is then input into CHIMERE to generate accurate predictions of mass concentration. To address uncertainties and improve model accuracy, we incorporate a stochastic approach using the Extended Kalman Filter (EKF) for adjustment. This adjustment process is tested under two scenarios: one without external forces or control terms, and one including external forces that account for temperature, wind speed, and nitrogen dioxide as a precursor to ozone mass concentration. The coupling of the Eulerian numerical model with EKF helps overcome the limitations of traditional Eulerian models, offering a more robust and precise representation of atmospheric pollution dynamics. Based on the results of this study and others, the proposed models effectively address the shortcomings of conventional models and methods.

The structure of this paper is as follows: Section 2 introduces the mathematical and physical modeling for the Eulerian numerical model and the stochastic model, along with the relevant equations, schemes, and models. Section 3 provides a detailed explanation of the numerical applications and discusses the results from the Eulerian model, including the adjustments made by the stochastic model. Two scenarios are analyzed to validate the approach. Finally, Section 4 presents the conclusion, summarizing the main findings.

2. MATHEMATICAL AND PHYSICAL MODELING

2.1. EULERIAN NUMERICAL MODEL

The temporal evolution of pollutant mass concentration (γ_S) is described by a system of partial differential equations (PDEs) (Equation 1) that explicitly integrate various physical and chemical processes. Key contributions include advection, modeling transport through fluid motion; turbulence, introducing additional mixing and diffusion through a turbulent diffusion term; and chemistry, adjusting mass concentration based on specific chemical reactions. Emissions are included as sources, while deposition accounts for losses due to processes like dry deposition. This detailed formulation, implemented in the CHIMERE model as a Eulerian numerical model (E.N.M.), allows for a comprehensive understanding of the dynamics influencing pollutant mass concentration over time

$$\frac{\partial \gamma_S}{\partial t} = \left(\frac{\partial \gamma_S}{\partial t}\right)_{Advection} + \left(\frac{\partial \gamma_S}{\partial t}\right)_{Turbulence} + \left(\frac{\partial \gamma_S}{\partial t}\right)_{Chemistry} + \left(\frac{\partial \gamma_S}{\partial t}\right)_{Emissions} + \left(\frac{\partial \gamma_S}{\partial t}\right)_{Depot} \tag{1}$$

Table 1 shows the physical processes and modeling approaches used in the CHIMERE model. The numerical solver uses the TwoStep solver, which enhances the accuracy and stability of numerical solutions. Deep Convection Fluxes are modeled using the Tiedtke scheme, which simulates heat and moisture exchanges in deep convection. Advection is managed by the Upwind scheme, optimizing variable transport and minimizing numerical errors. Photolytic rates are calculated with the radiative transfer model, ensuring precise photolysis rate estimation. Boundary-Layer Turbulence is modeled by the Troen and Mahrt scheme, which describes turbulent processes in the boundary layer. Soil Moisture is handled by the Fecan scheme, providing detailed soil moisture simulations. Lastly, the Dry threshold friction velocity is determined by the Shao and Lu scheme, which evaluates the friction velocity needed to initiate dry soil friction. These options enable detailed and realistic modeling of atmospheric processes. It is important to note that CHIMERE relies on input data from WRF for meteorological conditions and Emi-Surf for emissions to model pollutant dispersion accurately.

Table 1

Physical Processes and Modeling Schemes Used in the CHIMERE Model

Physical Process	**Modeling Approach**	**Reference**
Numerical solver	TwoStep solver	[23]
Deep convection fluxes	Tiedtke scheme	[24]
Advection	Upwind scheme	[25]
Photolytic rates	Radiative transfer model	[26]
Boundary-layer turbulence	Troen and Mahrt scheme	[27]
Soil moisture	Fecan scheme	[28]
Dry threshold friction velocity	Shao and Lu scheme	

Table 2 shows the physical processes and modeling approaches used in the WRF (Weather Research and Forecasting) to model energy, transfer and surface processes. Radiative transfer is managed by the rapid radiative transfer model, known for its precision in atmospheric radiative calculations. Microphysics are handled by the single-moment 3 class scheme, which simulates the formation and evolution of hydrometeors like rain and ice crystals. Surface processes are modeled using the Noah model, which simulates energy and moisture exchanges between the surface and atmosphere. The MM5 similarity scheme ensures accurate modeling of heat and momentum fluxes in the surface layer. The Kain-Fritsch scheme is used for cumulus parameterization, effectively simulating atmospheric convection. The Yonsei University model captures turbulent processes in the planetary boundary layer. The diffusion process is simplified with a basic diffusion model, and turbulent diffusion is handled using the first-order Smagorinski coefficient for horizontal scales, a common method for addressing subgrid turbulence. In emissions modeling with Emi-Surf, each parameter is carefully defined to capture various types and sources of atmospheric emissions. Parameters include anthropogenic emissions, chemical mechanisms, and natural emissions like mineral dust and marine aerosols [29]. Specific approaches are used for each emission type, such as the MEGAN model for biogenic emissions [30] and Menut's guidelines for temporal distribution [31]. Edge conditions, whether related to dust or gases, are managed using methods like global GOCART and LMDz-INCA [32, 33]. This detailed integration of parameters and methods produces a comprehensive simulation of atmospheric emissions.

Table 2

Physical Processes and Modeling Schemes Used in the CHIMERE Model

Physical Process	Modeling Approach	Reference
Microphysics	WRF Single-Moment 3 class scheme	[34]
Long Wave Radiation	Rapid Radiation Transfer Model	[26]
Shortwave Radiation	Dudhia scheme	[35]
Cloud fraction option	Xu-Randall method	
Surface layer	MM5 similarity	[36]
Land surface	Noah Land Surface Model	[37]
Boundary layer	Yonsei University scheme	[38]
Cumulus option	Kain-Fritsch scheme	[39]
Diffusion	Simple diffusion	[40]
Eddy coefficient	Horizontal Smagorinski 1st order	[41]

2.2. STOCHASTIC MODEL

The Extended Kalman Filter (EKF) is a powerful technique for estimating the states of a dynamic system with non-linear evolution and observation models. While these models are not linear functions of the state, they can be differentiable functions. In EKF, the equations describing the evolution of the state and the observations, as given in Eqs. 2 and 3, are represented by non-linear functions. Equation 2 describes how the state $\boldsymbol{x}_k$ of the system evolves from the previous state $\boldsymbol{x}_{k-1}$, based on the inputs $\boldsymbol{u}_k$ and process noise $\boldsymbol{w}_k$. Equation 3 relates the system state $\boldsymbol{x}_k$ to the observations $\boldsymbol{z}_k$, including measurement noise $\boldsymbol{v}_{\mathrm{k}}$. Note that for the second scenario with external forces, $\boldsymbol{B_k}$ is the control matrix that models how each external force (such as temperature, wind, and NO_2 mass concentration) affects the system's dynamics. $\boldsymbol{u}_k$ is the vector of external forces, including temperature, wind speed, and NO_2 mass concentration.

$$\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}, \boldsymbol{B_k}\boldsymbol{u}_k, \boldsymbol{w}_k) \tag{2}$$

$$\boldsymbol{z}_k = h(\boldsymbol{x}_k, \boldsymbol{v}_{\mathrm{k}}) \tag{3}$$

Since the functions f and h can be non-linear, they cannot be used directly to calculate the state covariance. To address this, the Extended Kalman Filter (EKF) uses local approximations with partial derivative matrices known as Jacobians, as described in Eqs. 4 and 5. Equation 4 represents the matrix of partial derivatives of the evolution function with respect to the state, evaluated at the current estimate. The Jacobian of evolution in Eq. 4 is the matrix of partial derivatives of the evolution function with respect to the state, evaluated at the current estimate. The Jacobian of the observation in Eq. 5 is a matrix that shows how changes in the state affect the observations.

$$\boldsymbol{F}_k = \left.\frac{\partial f}{\partial \boldsymbol{x}}\right|_{\widehat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{u}_k} \tag{4}$$

$$\boldsymbol{H}_k = \left.\frac{\partial h}{\partial \boldsymbol{x}}\right|_{\widehat{\boldsymbol{x}}_{k|k}} \tag{5}$$

The Extended Kalman Filter (EKF) operates in two main processes: prediction and update. Equations 6 and 7 describe the prediction process. Equation 6 calculates the predictive state $\widehat{\boldsymbol{x}}_{k|k-1}$ at time k using the non-linear evolution function f, the estimated state at the previous time $\hat{x}_{k-1|k-1}$, and the inputs. The term 0 represents the process noise, which is assumed to be zero for simplicity. Equation 7 calculates the predicted covariance $\boldsymbol{P}_{k|k-1}$ of the state at time k. This is done using the Jacobian $\boldsymbol{F}_k$ of the evolution function f, which linearizes the function around the current estimate. The previous covariance $\boldsymbol{P}_{k-1|k-1}$ is updated to reflect the evolution of state uncertainty, and the term $\boldsymbol{Q}_k$ represents the covariance of the process noise.

$$\widehat{\boldsymbol{x}}_{k|k-1} = f(\widehat{\boldsymbol{x}}_{k-1|k-1}, \boldsymbol{B}_{\boldsymbol{k}}\boldsymbol{u}_k, 0) \tag{6}$$

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}_k \boldsymbol{P}_{k-1|k-1} \boldsymbol{F}_k^T + \boldsymbol{Q}_k \tag{7}$$

The objective of the update process is to correct state estimates by incorporating new observations and adjusting predictions based on observation errors. This is achieved through several steps: Equation 8 calculates the innovation $\widetilde{\boldsymbol{y}}_k$ which is the difference between the actual measurement $\boldsymbol{z}_k$ and the predicted observation $h(\widehat{\boldsymbol{x}}_{k|k-1}, 0)$. Equation 9 computes the innovation covariance matrix $\boldsymbol{S}_k$, representing the uncertainty in the prediction of the observation. Equation 10 determines the innovation gain $\boldsymbol{K}_k$, indicating how much the state estimate needs to be adjusted for innovation. Equation 11 updates the state estimate $\widehat{\boldsymbol{x}}_{k|k}$, and Eq. 12 revises the state covariance $\boldsymbol{P}_{k|k}$

$$\widetilde{\boldsymbol{y}}_k = \boldsymbol{z}_k - h(\widehat{\boldsymbol{x}}_{k|k-1}, 0) \tag{8}$$

$$\boldsymbol{S}_k = \boldsymbol{H}_k \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T + \boldsymbol{R}_k \tag{9}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_{k|k-1} \boldsymbol{H}_k^T + \boldsymbol{S}_k^{-1} \tag{10}$$

$$\widehat{\boldsymbol{x}}_{k|k} = \widehat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K}_k \widetilde{\boldsymbol{y}}_k \tag{11}$$

$$\boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k) \boldsymbol{P}_{k|k-1} \tag{12}$$

2.3. EXPERIMENTATION AND EVALUATION CRITERIA

The O342M is a continuous ozone analyzer designed for low mass concentrations. It detects ozone using ultraviolet light absorption and benefits from advanced electronic and optical technologies, ensuring minimal maintenance. The sample is drawn through a 6 mm external diameter Teflon tube connected to the back of the instrument and is pulled in by an internal pump. Results are displayed on a graphic screen on the front panel. The measurements were taken at the Mediouna station in Casablanca, Morocco, located at –7.51° longitude and 33.45° latitude, from February 1st to March 27th, 2021. This data was then used to determine the ozone concentration based on the principle of ultraviolet (UV) light absorption, typically at a wavelength of 254 nm. Ozone absorbs the UV light, reducing its intensity in proportion to the ozone concentration. This relationship follows the Beer-Lambert law, which links the change in light intensity to the ozone concentration (Equation 13). The analyzer measures the light intensities before and after absorption and uses the Beer-Lambert law to calculate the ozone concentration by volume, which is then converted to mass concentration by multiplying by ozone's molar mass (Equation 14). Note that I is the intensity of UV light after passing through the air sample, while I_0 is the initial intensity of UV light (without ozone). α is the specific absorption coefficient of ozone at the wavelength used, C is the ozone concentration by volume, and L is the optical path length through the sample. γ is the ozone mass concentration, and M_{O3} is the molar mass of ozone (48 g/mol)

$$I = I_0 e^{-\alpha C L} \tag{13}$$

$$\gamma = C \times M_{O3} \tag{14}$$

Three evaluation criteria were used to compare our models with experimental measurements and results from other studies: correlation (Equation 15), RMSE (Equation 16), and the logarithmic norm of the residual (Equation 17). Correlation ensures that the model accurately reflects data trends, while RMSE measures the average error between predictions γ and measurements $\hat{\gamma}$, assessing accuracy and stability. The logarithmic norm of the residual offers a detailed view of model convergence, capturing small variations, especially during the final optimization stages. Together, these criteria help verify the model's overall performance, stability (low error variability), and convergence toward an optimal solution. Comparing our results with other studies using these indicators helps confirm the robustness and reliability of our models against established benchmarks. Note that n is sample size. $\mathrm{Cov}(\gamma, \hat{\gamma})$ and $\sigma_\gamma \sigma_{\hat{\gamma}}$ denote the covariance and standard deviations of the predictions γ and measurements $\hat{\gamma}$, respectively.

$$\text{Correlation} = \frac{\text{Cov}(\gamma, \hat{\gamma})}{\sigma_\gamma \sigma_{\hat{\gamma}}} \tag{15}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^{n} (\gamma_k - \hat{\gamma}_k)^2}{n}} \tag{16}$$

$$\text{Logarithmic residual norm} = \text{Log}_{10}(\gamma_k - \hat{\gamma}_k). \tag{17}$$

3. NUMERICAL APPLICATIONS AND DISCUSSIONS

3.1. EULERIAN NUMERICAL MODEL

Figure 1a, b and c compares statistical results for temperature (*T*), wind speed (WS), and mass concentration (γ) measured and modeled by the Eulerian numerical model (E.N.M.). The temperature data show a correlation coefficient of 86%, a mean bias of 0.9°C, indicating a slight overestimation, and an RMSE of 2.0°C. The model tends to overestimate wind speed and shows a moderate correlation of 61% with actual measurements, along with significant prediction errors of 3.1 m/s. The Eulerian numerical model has difficulty accurately representing measured mass concentrations. It shows a significant bias (overestimation of 26.8 $\mu g/m^3$), a moderate correlation indicating an imperfect match between measured and E.N.M. data, and a high RMSE, suggesting notable errors in the model's predictions. These statistical results are consistent with the conclusions of other studies that have also used Eulerian models such as CMAQ and WRF-Chem [42–50]. Although the E.N.M. results are acceptable, we still need to understand the origin of their weaknesses, especially regarding wind speed and concentration.

Figure 1d shows the maximum and minimum values for each variable, both measured and modeled. The Eulerian model (E.N.M.) accurately captured the maximum temperature, but had a 2.6°C bias in the minimum temperature. Overall, the model effectively represented the temperature. The maximum measured wind speed is 3.3 m/s, while the model predicts 6.6 m/s, indicating that the model overestimates the wind speed. The flow components show a meridional flow of –6.2 m/s and a zonal flow of –2.6 m/s, as shown in Fig. 2b, suggesting the model predicts significant north-to-south and east-to-west winds, with a dominant meridional component. In contrast, the minimum measured wind speed is 0 m/s, while the model predicts 1 m/s. The flux components, with a meridional flux of 0.93 m/s and a zonal flux of 0.45 m/s, shown in Fig. 2a, indicate weak south-to-north and west-to-east winds, with values closer to equilibrium. The minimum mass concentration of 32.8 $\mu g/m^3$, compared to an actual 0 $\mu g/m^3$, is highly exaggerated, highlighting a significant weakness. However, the model accurately captured the maximum mass concentration. The overestimations in minimum temperature, minimum mass concentration, and maximum wind speed likely account for the high mean and RMSE of the Eulerian model for mass concentration and wind speed shown in Fig. 1a and b. However, analyzing the dispersion maps in cases that are not well captured is essential to fully understand the behavior of the Eulerian model.

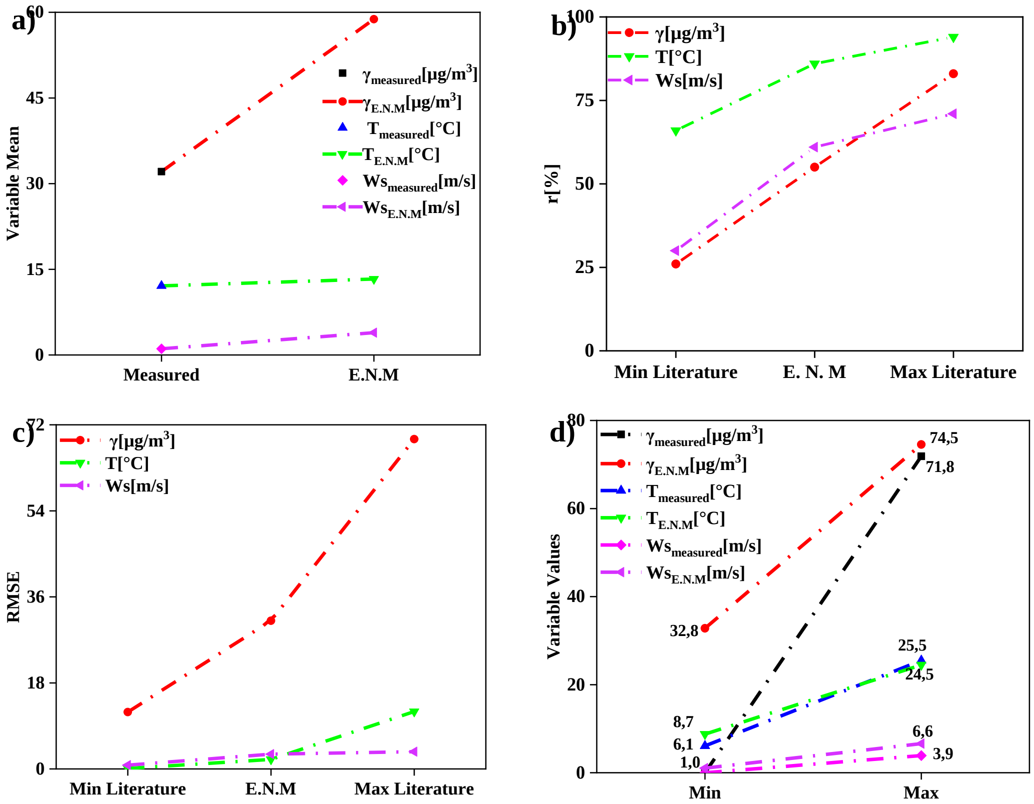

Fig. 1 – Comparison of Statistical Results for: temperature (*T*), wind speed (WS), and ozone concentration (γ) measured and modeled: a) variable mean; b) correlation; c) RMSE; (d) max-min.

Figure 3 presents the dispersion maps for minimum temperature, minimum mass concentration, and the meridional flow of maximum and minimum wind speed. On February 22nd at 6 am, the measured minimum temperature was 6.1°C, while the model showed 8.7°C, as depicted in Fig. 1d. The dispersion map highlights the model's effectiveness in representing the region's thermal dynamics. The ocean, with its high heat capacity, absorbs and releases heat more slowly than land, stabilizing coastal temperatures. At night, land cools faster than the sea, leading to cooler air over land (6°C) and warmer air over the sea (13°C). The model combines these temperature differences to calculate a moderate coastal temperature. This approach, accounting for the attenuation of temperature variations, results in a modeled mean temperature of 8.3°C. This can be explained by the principle of conservation of thermal energy, represented by Eq. 18, which considers mass (m) and specific heat capacity (c). This equation summarizes the energy balance between land and ocean and explains the observed temperature variations.

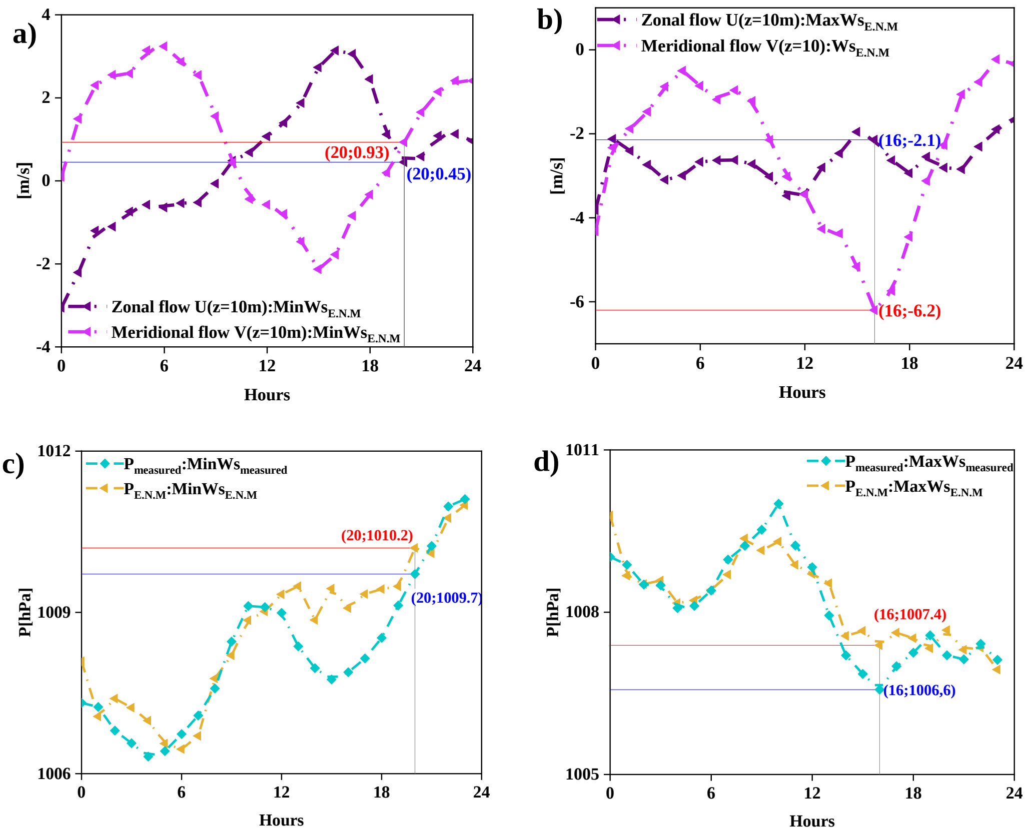

Fig. 2 – Zonal and meridional flow for minimum wind speed (a) and maximum wind speed (b), with pressure measured and modeled; (c) minimum wind speed and (d) maximum wind speed.

The maximum wind speed shown in Fig. 1d is recorded on February 13 at 16:00. From Fig. 2b, we can deduce that the general wind direction is from northeast to southwest, with a dominant meridional component, as supported by the dispersion map of the maximum Meridional Flow. The minimum wind speed from

Fig. 1d is recorded on February 15 at 20:00. From Fig. 2a, we can determine that the general wind direction is from southwest to northeast, with a dominant meridional component, consistent with the dispersion map of the minimum meridional flow. Equation 19, a simplified version of the Navier-Stokes equations known as the atmospheric primitive equation, may help explain the discrepancies between the numerical Eulerian model and measurements. Note that $(-\frac{1}{\rho}\nabla P)$ is a pressure term, $(\boldsymbol{f_{Coriolis}})$ is a Coriolis force, $((\boldsymbol{V}\cdot\nabla)\cdot\boldsymbol{V})$ is an advective term, and $(\boldsymbol{F_{Friction}})$ is a friction force. Misrepresentation of the pressure term could result in either overestimation or underestimation of wind speed. From Figure 2c and d, the discrepancies between measured and modeled pressures suggest that the model may be overestimating the pressure gradient. It can be noted that the overestimation of pressure in the maximum wind speed is significant compared to the minimum wind speed. This overestimation could explain the higher wind speeds predicted by the model, as an exaggerated pressure gradient results in a greater pressure force, which artificially accelerates the wind in the simulations.

From Fig. 3, we see that vertical dispersion (south-north) is more significant than horizontal dispersion (west-east) for minimum mass concentration. Based on the local form of the continuity equation (Equation 20), which describes mass conservation, and Fig. 1d, which shows that the model effectively captures maximum mass concentration, we will examine the spatial correlation effects of temperature, zonal flow, and meridional flow in both cases of maximum and minimum mass concentration, as presented by Fig. 4. This figure shows two configurations: in the first, the cell is fixed along the y-axis (north-south) and varied along the x-axis (west-east), noted as Configuration 1($\gamma_{min}(i_n, j = 15)$ and $\gamma_{max}(i_n, j = 15)$). In the second configuration, the x-axis is fixed while the y-axis is varied, noted as Configuration 2 ($\gamma_{min}(i = 15, j_n)$ and $\gamma_{max}(i = 15, j_n)$). In each case, the effects of T, U, and V on γ_{min} and γ_{max} are compared. Note that *n* varies from 14 to 16. In configuration 1, the E.N.M. model uses the same mechanisms to estimate both minimum and maximum mass concentrations. This consistency may lead to an overestimation of γ_{min}, as the same forces that drive high mass concentrations (strong advection with U and V) are also applied to γ_{min}, which is not appropriate. Physically, this suggests that the model does not sufficiently distinguish between the processes that should limit minimum mass concentration and those that increase maximum mass concentration. In configuration 2, the model correctly distinguishes between the mechanisms for γ_{min} and γ_{max}. It applies a dispersing effect from V (and partially from U) on γ_{min} while capturing the concentration effect of V and T on γ_{max}. This distinction allows the model to estimate γ_{max} accurately while more realistically overestimating γ_{min}. Given the limitations of Eulerian numerical modeling, especially in predicting minimum mass concentration, we considered improving it with a stochastic model based on a linear-quadratic state estimator (LQE), which estimates the states of a dynamic system from incomplete or noisy measurements.

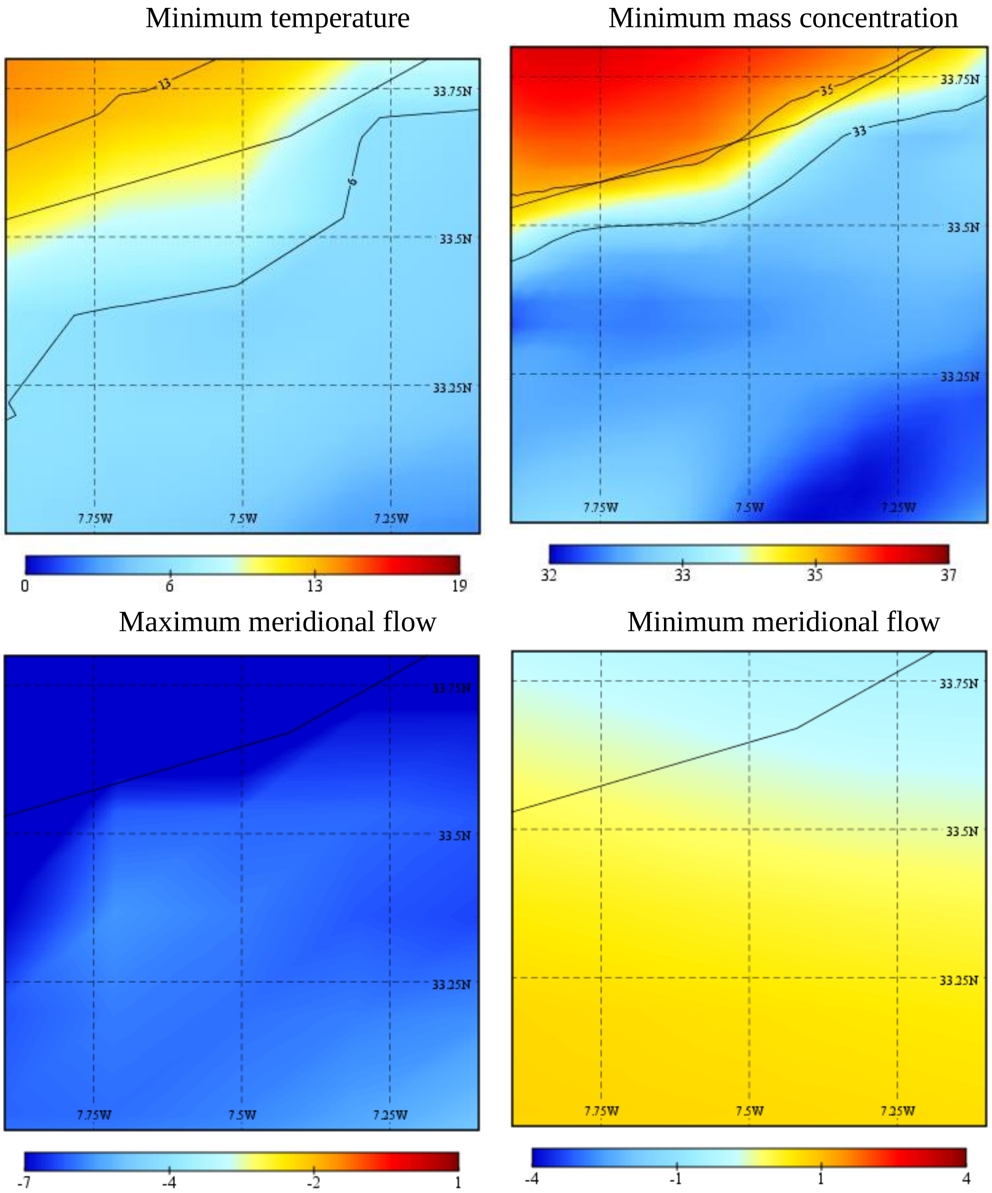

Fig. 3 – Dispersion maps for minimum temperature, minimum mass concentration, and meridional flow of maximum and minimum wind speed.

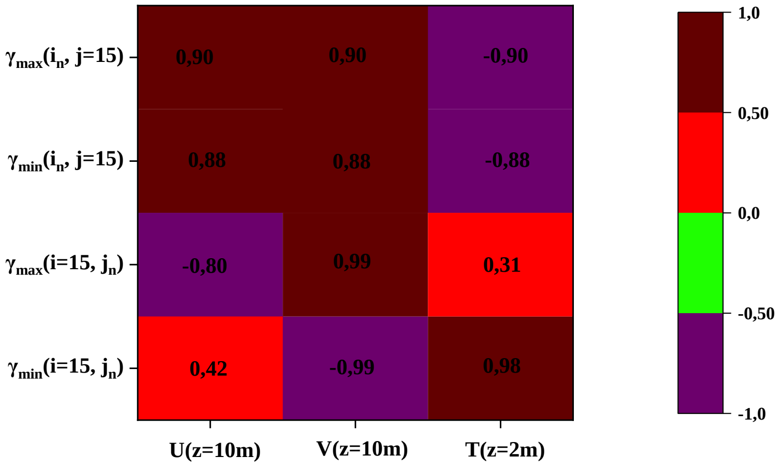

Fig. 4 – Spatial correlation effects of temperature, zonal flow, and meridional flow for maximum and minimum mass concentration.

$$m_{ocean}C_{ocean}\frac{\mathrm{d}T}{\mathrm{d}t}+m_{land}C_{land}\frac{\mathrm{d}T}{\mathrm{d}t}=0 \tag{18}$$

$$\frac{\partial \boldsymbol{V}}{\partial t}+(\boldsymbol{V}\cdot\nabla)\cdot\boldsymbol{V}=-\frac{1}{\rho}\nabla P+\boldsymbol{f}_{\boldsymbol{Coriolis}}+\boldsymbol{F}_{\boldsymbol{Friction}} \tag{19}$$

$$\left(\frac{\partial \gamma_s}{\partial t}\right)_{Advection}=-U\cdot\frac{\partial \gamma_s}{\partial x}-V\cdot\frac{\partial \gamma_s}{\partial y} \tag{20}$$

$$\left(\frac{\partial \gamma_s}{\partial t}\right)_{Chemistry}=P_s-L_s.\gamma_s \tag{21}$$

$$\gamma_{NO2}\rightarrow\gamma_{NO}+\gamma_{O3}. \tag{22}$$

3.2. STOCHASTIC MODEL ADJUSTMENT

To adjust the Eulerian numerical modeling, a stochastic algorithm which is the Extended Kalman Filter (EKF) is used, as detailed in Eqs. 2 to 12. To achieve a more efficient solution and a more stable model, the stochastic model was applied under extreme conditions on a day during the study period when the Eulerian numerical model severely overestimated the minimum, maximum, and average mass concentrations, as shown in Fig. 7b (Minγ_{O3}= 53.9 µg/m^3; Meanγ_{O3}= 65.1 µg/m^3; Maxγ_{O3}= 71.4 µg/m^3). In the adjustment process, we tested two scenarios: one without external forces or control term, and one with the external forces considering the effects of temperature, wind speed, and nitrogen dioxide as an ozone mass concentration precursor. Uncertainty levels are shown with these abbreviations: VLU (Very Low Uncertainty) = 0.05, LU (Low Uncertainty) = 0.1, MU (Moderate Uncertainty) = 0.5, and HU (High Uncertainty) = 1.0. To evaluate the convergence and stability of the numerical method, the logarithmic residual norm (Equation 17) measures convergence by tracking the error reduction between adjustment and measurement, while the correlation coefficient (Eq. 15) ensures stability by confirming that the methods accurately reflect the trends in the measured data.

Figure 5a shows the logarithmic residual norm for various levels of process noise (Q) and measurement noise (R). Combining HU for Q (0.340) with VLU for R (0.077) achieves a balance where the process model is fairly accurate but with some uncertainty, and the observational data is very reliable. This setup maximizes the EKF filter's ability to provide precise estimates by leveraging high-quality measurements and managing model uncertainty effectively. In practice, this leads to more accurate adjustments to mass concentration dynamics while maintaining filter stability. Figure 5b shows that the very high correlation values indicate the model fits the measurements well, even with varying levels of uncertainty. Figure 6a illustrates how different state functions influence the accuracy of the E.N.M. model. Physically, the choice of a non-linear function directly impacts the EKF-adjusted E.N.M.'s ability to capture the complexity of atmospheric dynamics, as expressed by Eq. 1. The cubic function is the most effective at modeling the complex interactions and nonlinear phenomena necessary for a realistic representation of mass concentration, represented by Eq. 21. P_i is a vector and L_i is a diagonal matrix representing respectively, production and loss terms. The components $P_i(t,\gamma)$ and $L_i(t,\gamma)\cdot\gamma_i$ are nonnegative. These results demonstrate that to achieve accurate and reliable outcomes, it's crucial to use high-quality measured data along with a process model that can effectively manage its inherent uncertainty. Additionally, employing suitable state functions, like non-linear ones, is key to modeling complex phenomena and enhancing the performance of the Extended Kalman Filter (EKF), which, in this case, operates without external forces or control terms.

To increase the complexity of our stochastic model, we have included an external forces term to account for the effects of temperature, wind speed, and NO_2 mass concentration on the system. This addition makes the model more realistic in representing state evolution. Figure 6b shows the logarithmic residual norms for various functions representing external forces, with a cubic state function used for consistency, as justified by Fig. 6a. Given its better overall performance and most negative residual values, the linear model is the most suitable for hourly data in this case. It offers a more accurate adjustment of the effects of temperature, wind speed, and NO_2 mass concentration on ozone mass concentration. Real physical phenomena like pollutant production and dispersion usually follow relatively simple dynamics under normal conditions. Ozone mass concentration is primarily influenced by reaction and meteorological processes, which can often be approximated by linear relationships within a typical range, as expressed by Eq. 22. While significant non-linear effects might occur at extreme thresholds or special conditions, a linear model generally captures most variations under typical conditions.

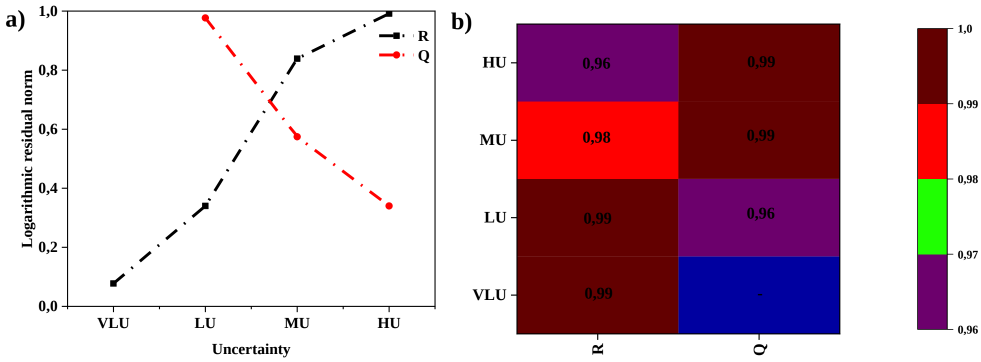

Fig. 5 – a) Logarithmic residual norm for various levels of process noise (Q) and measurement noise (R); b) correlation for different levels of process noise (Q) and measurement noise (R).

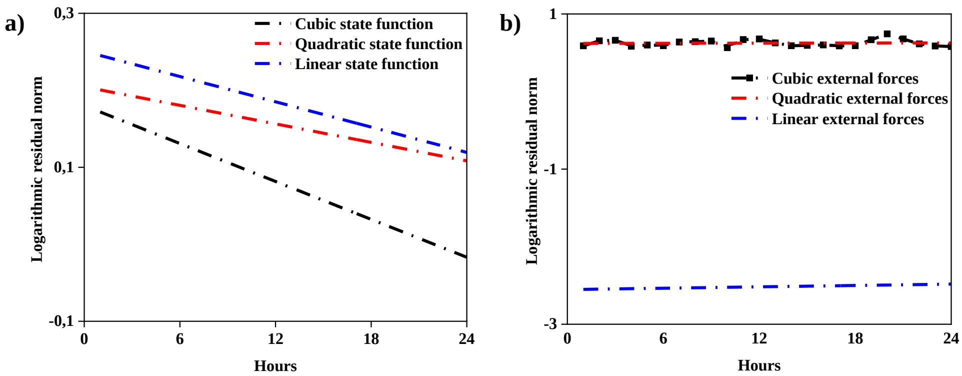

Fig. 6 – Hourly logarithmic residual norms for: a) state functions; b) external forces functions.

To fully understand the impact of adding external forces to the stochastic model, we compared the mass concentration measured by the Eulerian numerical modeling with and without external forces, as shown in Fig. 7a. When external forces are not considered, models tend to overestimate ozone mass concentration. This is because dynamic factors like wind dispersion and thermal degradation of ozone are not accounted for. The model assumes a scenario where ozone formation processes dominate without adequate dissipation mechanisms, leading to an overestimation. Integrating temperature, wind speed, and NO_2 into the stochastic model significantly improves forecast accuracy. This highlights the crucial role these factors play in ozone formation, degradation, and dispersion processes. Figure 7b presents summary statistics (minimum, mean, and maximum) for measured ozone mass concentrations, as well as for the adjusted E.N.M. with and without external forces. The E.N.M. model adjusted without external forces significantly overestimates the minimum ozone mass concentration. This overestimation may occur due to the absence of ozone dissipation mechanisms, like the reaction with nitrogen monoxide (NO_2). In contrast, the model with external forces predicts a minimum concentration much closer to the measured value, indicating that factors like temperature and wind are crucial for accurately modeling low ozone concentrations. The model's mean with external forces, which is almost identical to the measured mean, indicates that these forces correct for modeling errors and lead to a more accurate estimate. For the maximum mass concentration, the adjustment has little effect, as the E.N.M. without adjustment already captures the maximum concentration well, as shown in Fig. 1d.

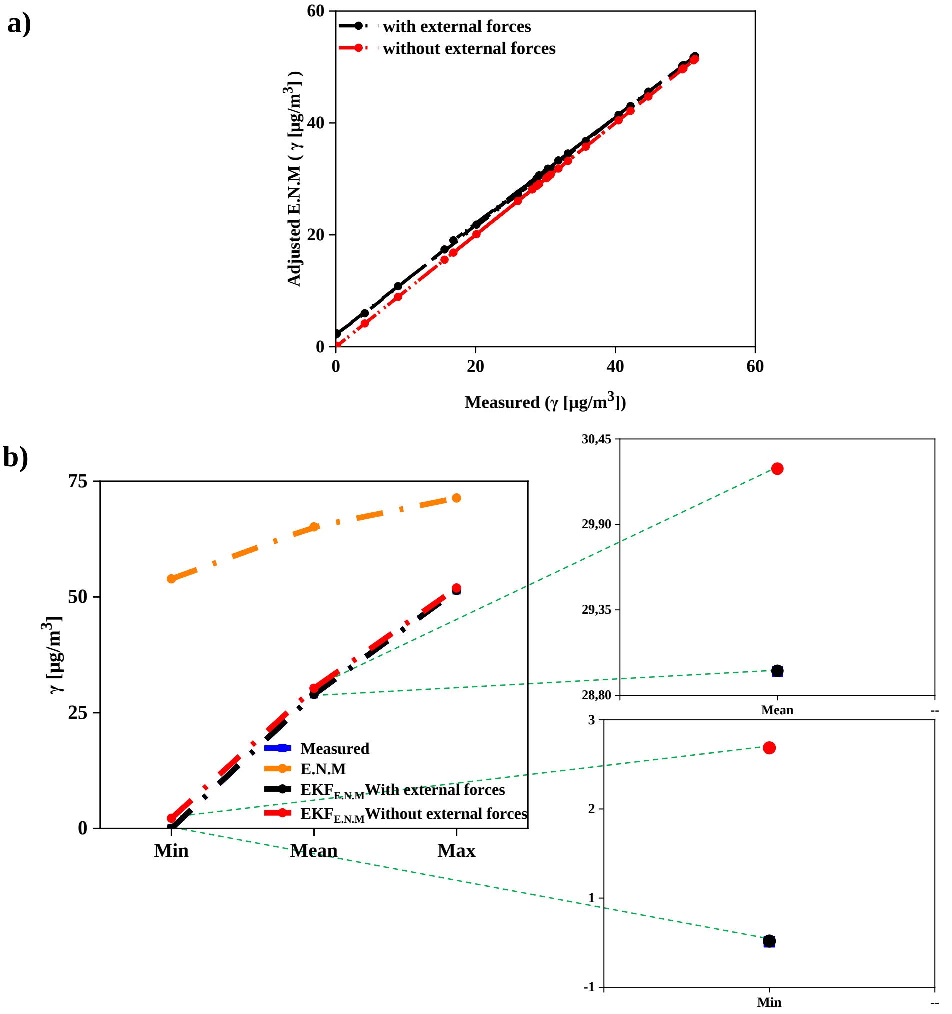

Fig. 7 – a) Comparison of mass concentration measured with adjusted E.N.M. with and without external forces; b) comparison of min, mean, and max values between measured, ENM, and adjusted ENM ($EKF_{E.N.M.}$) with and without external forces.

The results demonstrate the success of combining a numerical Eulerian model with a stochastic model, specifically an Extended Kalman Filter, addressing the limitations of using the Eulerian model alone for predicting pollutant mass concentration. This improved model effectively corrected the overestimation of the minimum mass concentration by the E.N.M. The findings also highlight that accurate modeling requires including external forces to capture the complexity of atmospheric processes affecting mass concentration. A correlation coefficient very close to 1, along with a steadily decreasing logarithmic residual norm, suggests good convergence. Stability in the model is indicated by a consistently high correlation across different uncertainty levels and a low, stable residual norm. Our proposed $EKF_{E.N.M.}$ model, when applied with external forces over a one-month period (about 672 hours), outperformed other studies – Shams *et al.*, Xiong *et al.*, Zhou *et al.*, and Ajdour *et al.* [10, 45, 51, 52] achieving the lowest RMSE of 3.34 µg/m³ and the highest correlation of 99%, as shown in Fig. 8a and b. This demonstrates that our model is not only highly accurate in its predictions but also effectively captures underlying trends over an extended period.

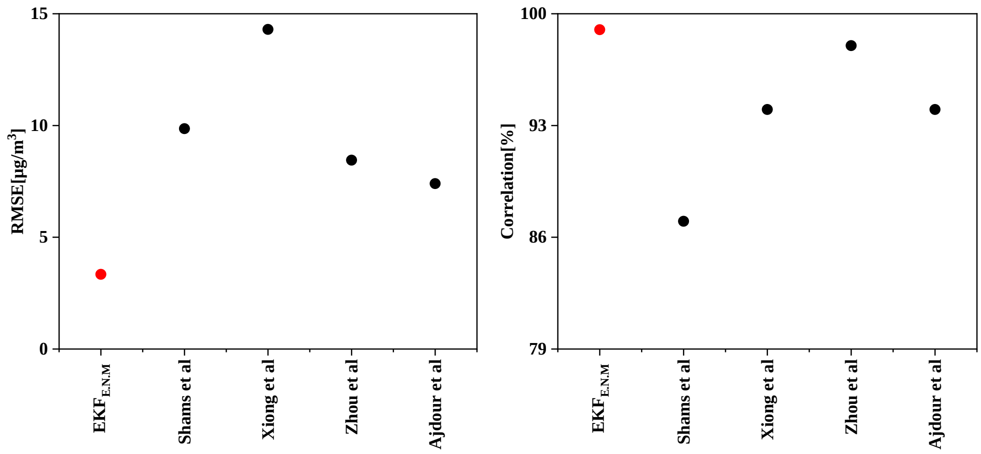

Fig. 8 – Comparison of: a) RMSE; b) correlation between the $EKF_{E.N.M.}$ approach and other studies.

4. CONCLUSION

For the first time to the best of our knowledge, a stochastic Eulerian numerical approach based on $EKF_{E.N.M.}$ is used to address the limitations of the Eulerian air pollution model CHIMERE. This approach integrates advection, turbulence, chemical reactions, emissions, and deposition for a comprehensive model of pollutant mass concentration dynamics. The Extended Kalman Filter (EKF) employs nonlinear dynamic and measurement equations to model the system, linearizing them locally to estimate system states and adjust mass concentrations based on available measurements. The approach's accuracy is demonstrated through two scenarios: one without external forces or control terms and another considering external forces such as temperature, wind speed, and nitrogen dioxide. Results are compared with those from the CHIMERE (E.N.M.) model and the work of Shams *et al.*, Xiong *et al.*, Zhou *et al.*, and Ajdour *et al.*, showing that this approach is competitive in accuracy and adaptable to other nonlinear problems. Ongoing work includes applying this method to other areas, such as modeling pollutant flow from high-speed stacks and studying atmospheric oscillations' effects on pollutant dispersion.

Acknowledgements. The Faculty of Science at Ibn Zohr University supported this research at every stage. We also thank the Dynamic Meteorology Laboratory (LMD) for their valuable assistance. The simulations were carried out using the national HPC resources managed by the National Center for Scientific and Technological Research (CNRST) in Morocco. We extend special thanks to the Scientific Computing team for their availability and support.

REFERENCES

1. R. Yauri, B. Loayza, A. Yauri, A. Aquino, *Air quality monitoring system based on low power wide area network technology at public transport stops*, Int. J. Reconfigurable Embed. Syst. **13**, 699 (2024). https://doi.org/10.11591/ijres.v13.i3.pp699-707.
2. M.U. Şimsek, İ. Kök, S. Özdemir, *DeepFogAQ: A fog-assisted decentralized air quality prediction and event detection system*, Expert Syst. Appl. **251**, 123920 (2024). https://doi.org/10.1016/j.eswa.2024.123920.
3. C. Stihi, O.C. Bute, *Indoor Air Quality Monitoring in Educational Environments: a Case Study*, Rom. Rep. Phys. **75**, 706 (2023). https://doi.org/10.59277/RomRepPhys.2023.75.706.
4. R. V. Chiritescu, E. Luca, G. Iorga, *Observational Study of Major Air Pollutants Over Urban Romania in 2020 in Comparison With 2019*, Rom. Rep. Phys. **76**, 702 (2024) https://doi.org/10.59277/RomRepPhys.2024.76.702.
5. Q. Zhou, X. Li, J. Hu, Q. Zhang, *Dynamics and optimal control for a spatial heterogeneity model describing respiratory infectious diseases affected by air pollution*, Math. Comput. Simul. **220**, 276–295 (2024). https://doi.org/10.1016/j.matcom.2024.01.024.
6. G. Chirosca, M.E. Mihailov, M.I. Tomescu-Chivu, A.V. Chirosca, *Enhanced Machine Learning Model for Meteo-Oceanographic Time-Series Prediction*, Rom. J. Phys. **67**, 815 (2022). https://rjp.nipne.ro/2022_67_9-10/RomJPhys.67.815.pdf
7. A. Dumitru, E.A. Olaru, M. Dumitru, G. Iorga, *Assessment of Air Pollution By Aerosols Over a Coal Open-Mine Influenced Region in Southwestern Romania*, Rom. J. Phys. **69**, 801 (2024). https://doi.org/10.59277/RomJPhys.2024.69.801.
8. A. Ajdour, R. Leghrib, J. Chaoufi, A. Chirmata, *Numerical Method Combinations Assessment for Transport-Dominated Problems in the CHIMERE Model : A Case Study of Agadir (Morocco)*, **1**, 88–100 (2022). https://doi.org/10.1007/978-3-031-12416-7_9.
9. L. Adenis, S. Mailler, L. Menut, P. Achim, S. Generoso, *Lagrangian and Eulerian modelling of 106Ru atmospheric transport in 2017 over northern hemisphere*, J. Environ. Radioact. **275**, 107416 (2024). https://doi.org/10.1016/j.jenvrad.2024.107416.
10. A. Ajdour, B. Ydir, H. Bouzghiba, I.D. Sulaymon, A. Adnane, D. Ben Hmamou, K. Khomsi, J. Chaoufi, G. Géczi, R. Leghrib, *Investigating Two-dimensional Horizontal Mesh Grid Effects on the Eulerian Atmospheric Transport Model Using Artificial Neural Network*, Aerosol Air Qual. Res. **24**, 230309 (2024). https://doi.org/10.4209/aaqr.230309.

11. L. Menut, A. Cholakian, R. Pennel, G. Siour, S. Mailler, M. Valari, L. Lugon, Y. Meurdesoif, *The CHIMERE chemistry-transport model v2023r1*, Geosci. Model Dev. Discuss. **2024**, 1–44 (2024). https://doi.org/10.5194/gmd-2024-20.
12. Z. Ding, H. Chen, L. Zhou, Z. Wang, *A forecasting system for deterministic and uncertain prediction of air pollution data*, Expert Syst. Appl. **208**, 118123 (2022). https://doi.org/10.1016/j.eswa.2022.118123.
13. A. Ajdour, A. Adnane, B. Ydir, D. Ben hmamou, K. Khomsi, H. Amghar, Y. Chelhaoui, J. Chaoufi, R. Leghrib, *A new hybrid models based on the neural network and discrete wavelet transform to identify the CHIMERE model limitation*, Environ. Sci. Pollut. Res. **30**, 13141-13161 (2023). https://doi.org/10.1007/s11356-022-23084-8.
14. M. Zhang, P. Tian, Y. Zhao, X. Song, J. Liang, J. Li, Z. Zhang, X. Guan, X. Cao, Y. Ren, J. Shi, L. Zhang, *Impact of aerosol-boundary layer interactions on PM2.5 pollution during cold air pool events in a semi-arid urban basin*, Sci. Total Environ. **922**, 171225 (2024). https://doi.org/10.1016/j.scitotenv.2024.171225.
15. G.B. Stewart, D. Dajnak, J. Davison, D.C. Carslaw, A. V Beddows, N. Phantawesak, M.E.J. Stettler, M.J. Hollaway, S.D. Beevers, *New NO_x and NO_2 vehicle emission curves, and their implications for emissions inventories and air pollution modelling*, Urban Clim. **57**, 102103 (2024). https://doi.org/10.1016/j.uclim.2024.102103.
16. A. Maison, L. Lugon, S.-J. Park, C. Boissard, A. Faucheux, V. Gros, C. Kalalian, Y. Kim, J. Leymarie, J.-E. Petit, Y. Roustan, O. Sanchez, A. Squarcioni, M. Valari, C. Viatte, J. Vigneron, A. Tuzet, K. Sartelet, *Contrasting effects of urban trees on air quality: From the aerodynamic effects in streets to impacts of biogenic emissions in cities*, Sci. Total Environ. **946**, 174116 (2024). https://doi.org/10.1016/j.scitotenv.2024.174116.
17. B. Bessagnet, F. Couvidat, V. Lemaire, *A statistical physics approach to perform fast highly-resolved air quality simulations – A new step towards the meta-modelling of chemistry transport models*, Environ. Model. Softw. **116**, 100–109 (2019). https://doi.org/10.1016/j.envsoft.2019.02.017.
18. A. Puigferrat, I. de-Pouplana, F. Amato, E. Oñate, *Numerical prediction of the distribution of black carbon in a street canyon using a semi-Lagrangian finite element formulation*, Build. Environ. **199**, 107910 (2021). https://doi.org/10.1016/j.buildenv.2021.107910.
19. R.A.S. Albani, V.V.L. Albani, *Tikhonov-type regularization and the finite element method applied to point source estimation in the atmosphere*, Atmos. Environ. **211**, 69–78 (2019). https://doi.org/10.1016/j.atmosenv.2019.04.063.
20. X. Yang, Q. Wang, L. Liu, J. Tian, H. Xie, L. Wang, Y. Cao, S.S.H. Ho, *Impacts of emission reduction and meteorological conditions on air quality improvement from 2016 to 2020 in the Northeast Plain, China*, J. Environ. Sci. **151**, 484–496 (2025). https://doi.org/10.1016/j.jes.2024.04.017.
21. F. Lemmouchi, J. Cuesta, M. Lachatre, J. Brajard, A. Coman, M. Beekmann, C. Derognat, *Machine Learning-Based Improvement of Aerosol Optical Depth from CHIMERE Simulations Using MODIS Satellite Observations*, Remote Sens. **15**, 1510 (2023). https://doi.org/10.3390/rs15061510.
22. E. Hajiparvaneh, H. Alizadeh, V. Hosseini, C.R. Koch, *Machine Learning Classification of Air Quality Monitoring Stations to Achieve Ambient NO_2 Objectives Using Emission Scenarios and Chemical Transport Model*, Atmos. Environ. **331**, 120582 (2024). https://doi.org/10.1016/j.atmosenv.2024.120582.
23. J.G. Verwer, J.G. Blom, M. Van Loon, E.J. Spee, *A comparison of stiff ode solvers for atmospheric chemistry problems*, Atmos. Environ. **30**, 49–58 (1996). https://doi.org/10.1016/1352-2310(95)00283-5.
24. M. Tiedtke, *A comprehensive mass flux scheme for cumulus parameterization in large-scale models*, Mon. Weather Rev. **117**, 1779–1800 (1989). https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.
25. W. Zhang, T. Wang, J.S. Bai, P. Li, Z.H. Wan, D.J. Sun, *The piecewise parabolic method for Riemann problems in nonlinear elasticity*, Sci. Rep. **7**, 1–17 (2017). https://doi.org/10.1038/s41598-017-13484-z.

26. E.J. Mlawer, S.J. Taubman, P.D. Brown, M.J. Iacono, S.A. Clough, *Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave*, J. Geophys. Res. Atmos. **102**, 16663–16682 (1997). https://doi.org/10.1029/97jd00237.
27. A.K. Betts, J.H. Ball, A.C.M. Beljaars, M.J. Miller, P.A. Viterbo, *The land surface‐atmosphere interaction: A review based on observational and global modeling perspectives*, J. Geophys. Res. Atmos. **101**, 7209–7225 (1996). https://doi.org/10.1029/95JD02135
28. F. Fécan, B. Marticorena, G. Bergametti, *Parametrization of the increase of the aeolian erosion threshold wind friction velocity due to soil moisture for arid and semi-arid areas*, Ann. Geophys. **17**, 149–157 (1998). https://doi.org/10.1007/s00585-999-0149-7.
29. Ulas Im, *Impact of sea-salt emissions on the model performance and aerosol chemical composition and deposition in the East Mediterranean coastal regions*, Atmos. Environ. **75**, 329–340 (2013). https://doi.org/10.1016/j.atmosenv.2013.04.034.
30. A.B. Guenther, X. Jiang, C.L. Heald, T. Sakulyanontvittaya, T. Duhl, L.K. Emmons, X. Wang, *The model of emissions of gases and aerosols from nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions*, Geosci. Model Dev. **5**, 1471–1492 (2012). https://doi.org/10.5194/gmd-5-1471-2012.
31. G. K. Georgiou, T. Christoudias, Y. Proestos, J. Kushta, P. Hadjinicolaou, and J. Lelieveld, *Air quality modelling in the summer over the eastern Mediterranean using WRF-Chem: Chemistry and aerosol mechanism intercomparison*, Atmos. Chem. Phys. **18**, 1555-1571 (2018). https://doi.org/10.5194/acp-18-1555-2018.
32. P. Ginoux, J.M. Prospero, O. Torres, M. Chin, *Long-term simulation of global dust distribution with the GOCART model: correlation with North Atlantic Oscillation*, Environ. Model. Softw. **19**, 113–128 (2004). https://doi.org/https://doi.org/10.1016/S1364-8152(03)00114-2.
33. G.A. Folberth, D.A. Hauglustaine, J. Lathièrel, F. Brocheton, *Interactive chemistry in the Laboratoire de Météorologie Dynamique general circulation model: Model description and impact analysis of biogenic hydrocarbons on tropospheric chemistry*, Atmos. Chem. Phys. **6**, 2273–2319 (2006). https://doi.org/10.5194/acp-6-2273-2006.
34. S.Y. Hong, J. Dudhia, S.H. Chen, *A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation*, Mon. Weather Rev. **132**, 103–120 (2004). https://doi.org/10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.
35. J. Dudhia, *Numerical Study of Convection Observed during the Winter Monsoon Experiment Using a Mesoscale Two-Dimensional Model*, J. Atmos. Sci. **46**, 3077–3107 (1989). https://doi.org/https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.
36. C.W. Fairall, E.F. Bradley, J.E. Hare, A.A. Grachev, J.B. Edson, *Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm*, J. Clim. **16**, 571–591 (2003). https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2.
37. F. Chen, H. Kusaka, M. Tewari, J.W. Bao, H. Hirakuchi, *Utilizing the Coupled WRF/LSM/Urban Modeling System with Detailed Urban Classification to Simulate the Urban Heat Island Phenomena over the Greater Houston Area*, Fifth Conf. Urban Environ. **25**, 9–11 (2004). https://ams.confex.com/ams/AFAPURBBIO/techprogram/paper_79765.htm.
38. P.A. Jiménez, J. Dudhia, *Improving the representation of resolved and unresolved topographic effects on surface wind in the wrf model*, J. Appl. Meteorol. Climatol. **51**, 300–316 (2012). https://doi.org/10.1175/JAMC-D-11-084.1.
39. O.R. Bullock, K. Alapaty, J.A. Herwehe, J.S. Kain, *A Dynamically Computed Convective Time Scale for the Kain–Fritsch Convective Parameterization Scheme*, Mon. Weather Rev. **143**, 2105–2120 (2015). https://doi.org/10.1175/mwr-d-14-00251.1.
40. B. R. Reddy, C. V Srinivas, S. S. R. Shekhar, R. Baskaran, and B. Venkatraman, *Impact of land surface physics in WRF on the simulation of sea breeze circulation over southeast coast of India*, Meteorol. Atmos. Phys. **132**, 925-943 (2020). https://doi.org/10.1007/s00703-020-00726-5.
41. J.S. Simon, F.K. Chow, *Alternative Anisotropic Formulations for Eddy-Viscosity Models in the Weather Research and Forecasting Model*, Boundary-Layer Meteorol. **181**, 11–37 (2021). https://doi.org/10.1007/s10546-021-00642-0.

42. H. He, Z. Li, R.R. Dickerson, Ozone *Pollution in the North China Plain during the 2016 Air Chemistry Research in Asia (ARIAs) Campaign: Observations and a Modeling Study*, Air. **2**, 178–208 (2024). https://doi.org/10.3390/air2020011.
43. Y. Guo, C. Roychoudhury, M.A. Mirrezaei, R. Kumar, A. Sorooshian, A.F. Arellano, *Investigating ground-level ozone pollution in semi-arid and arid regions of Arizona using WRF-Chem v4.4 modeling*, Geosci. Model Dev. **17**, 4331–4353 (2024). https://doi.org/10.5194/gmd-17-4331-2024.
44. H. Lu, M. Xie, X. Liu, B. Liu, M. Jiang, Y. Gao, X. Zhao, *Adjusting prediction of ozone concentration based on CMAQ model and machine learning methods in Sichuan-Chongqing region, China*, Atmos. Pollut. Res. **12**, 101066 (2021). https://doi.org/10.1016/j.apr.2021.101066.
45. S.R. Shams, Y. Choi, D. Singh, M. Ghahremanloo, M. Momeni, J. Park, *Innovative approaches for accurate ozone prediction and health risk analysis in South Korea: The combined effectiveness of deep learning and AirQ+*, Sci. Total Environ. **946**, 174158 (2024). https://doi.org/10.1016/j.scitotenv.2024.174158.
46. D. Blanco-Ward, A. Rocha, C. Viceto, A.C. Ribeiro, M. Feliciano, E. Paoletti, A.I. Miranda, *Validation of meteorological and ground-level ozone WRF-CHIMERE simulations in a mountainous grapevine growing area for phytotoxic risk assessment*, Atmos. Environ. **259,** 118507 (2021). https://doi.org/10.1016/j.atmosenv.2021.118507.
47. P. Gupta, S. Payra, R. Bhatla, S. Verma, *WRF-Chem modeling study of heat wave driven ozone over southeast region, India*, Environ. Pollut. **340**, 122744 (2024). https://doi.org/10.1016/j.envpol.2023.122744.
48. S. Liu, *Sensitivity of WRF-Simulated 2-m Temperature and Precipitation to Physics Options over the Loess Plateau*, Adv. Meteorol. **2024**, 6633255 (2024). https://doi.org/https://doi.org/10.1155/2024/6633255.
49. D. Yan, T. Liu, W. Dong, X. Liao, S. Luo, K. Wu, X. Zhu, Z. Zheng, X. Wen, *Integrating remote sensing data with WRF model for improved 2-m temperature and humidity simulations in China*, Dyn. Atmos. Ocean. **89**, 101127 (2020). https://doi.org/10.1016/j.dynatmoce.2019.101127.
50. M.H. Mohammadi, A.H. Meshkatee, S. Ghader, M. Azadi, *Wind gust forecasting by post-processing the WRF model outputs using ANN*, Dyn. Atmos. Ocean. **105**, 101425 (2024). https://doi.org/https://doi.org/10.1016/j.dynatmoce.2023.101425.
51. K. Xiong, X. Xie, J. Mao, K. Wang, L. Huang, J. Li, J. Hu, *Improving the accuracy of O3 prediction from a chemical transport model with a random forest model in the Yangtze River Delta region, China*, Environ. Pollut. **319**, 120926 (2023). https://doi.org/10.1016/j.envpol.2022.120926.
52. H. Zhou, T. Wang, H. Zhao, Z. Wang, *Updated Prediction of Air Quality Based on Kalman-Attention-LSTM Network*, Sustain. **15**, 356 (2023). https://doi.org/10.3390/su15010356.