https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development Discussions

FastCTM (v1.0): Atmospheric chemical transport modelling with a principle-informed neural network for air quality simulations

Baolei Lyu[1,2,3], Ran Huang[4,5], Xinlu Wang[4], Weiguo Wang[6], Yongtao Hu[7]

[1] Huayun Sounding Meteorological Technology Co. Ltd., Beijing 102299, China
[2] Key Laboratory of Intelligent Meteorological Observation Technology, Beijing 100081, China
[3] China Meteorological Administration Xiong'an Atmospheric Boundary Layer Key Laboratory, Xiong'an, 071000, China
[4] Hangzhou AiMa Technologies, Hangzhou, Zhejiang 311121, P. R. China
[5] Nanjing AiMa Environmental, Nanjing, Jiangsu 210000, P. R. China
[6] SAIC, at Environment Modelling Center, NOAA/National Centers for Environmental Prediction, College Park, Maryland 20740, United States
[7] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, United States

Correspondence to: Baolei Lyu (baoleilv@foxmail.com), Ran Huang (ranhuang2019@163.com)

Abstract. Chemical transport models (CTM) have wide and profound applications in air quality simulations and managements. However, its applications are often constrained by high computational burdens. In this study, we developed a neural network based CTM model (FastCTM) to efficiently simulate ten air pollutant composition variables, including major $PM_{2.5}$ species of SO_4^{2-}, NO_3^-, NH_4^+, organic matters and other inorganic components, coarse part of PM_{10}, SO_2, NO_2, CO and O_3. The FastCTM has a principle-informed structure by explicitly encoding atmospheric physical and chemical processes in a basic simulator. Specifically, in the simulator, five neural network modules are proposed to respectively represent five major atmospheric processes of primary emissions, transport, diffusion, chemical reactions and depositions. Given 1-hour initial condition data, the FastCTM is able to simulate future 24-hour concentrations of the ten air pollutants with corresponding meteorology fields and emissions as input. The FastCTM is trained with operational forecast data from a numerical CTM model named Community Multiscale Air Quality (CMAQ) in 2018-2022. The well-trained FastCTM is evaluated comparing to the long-term CMAQ forecast in an independent year 2023, and achieves high agreements with mean RMSE values of 9.1, 11.9, 4.4, 4.0, 48.9 and 10.9 μg/m^3 and R^2 values of 0.8, 0.81, 0.8, 0.83, 0.9 and 0.7 for $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3. Besides, assessed against hourly site observations of six criteria pollutants, the RMSE values of FastCTM have small relative differences of 4.3%, 4.2%, -2.8%, -1.7%, -0.3% and -3.2% compared to that of CMAQ. The FastCTM model also exhibited reasonable responses of air quality to meteorological variables of air temperature, wind speed and planetary boundary layer height, as well as to input pollutant emissions. Furthermore, due to the principles-oriented structure, internal process analysis could be performed by FastCTM to quantify the specific contribution from each of the five processes for hourly air pollutant concentration changes. In a nutshell, FastCTM has multi-functional advantages in air pollutant concentration simulations, sensitivity analysis and internal process analysis with high computation efficiencies on GPU and accuracy.

1 Introduction

Effective air quality management requires an accurate understanding of air pollution conditions in current time and future to take targeted emission cut and control measures (Wang et al., 2010; Council, 2004). Driven by this demand, deterministic air quality numeric models have been developed to simulate spatiotemporal variances and evolutions of ambient air pollutants in the atmosphere (Hakami et al., 2003; Eder et al., 2006). In these models, such as the Community Multiscale Air Quality (CMAQ) model, atmospheric physical and chemical processes (e.g., emissions, chemical reaction, horizontal advection, and diffusion etc.) are mathematically defined by partial differential equations. The air pollutant and species concentrations can be then calculated by solving these complicated equations with numeric methods (Byun and Schere, 2006), which is often time-consuming and requires intense computational resources.

Recent developments in deep learning models provide promising alternative pathways to build fast and data-driven deep leaning-based CTM models, owing to the strong capabilities of neural networks in encoding and representing complex features, patterns and relationships that could be learned from long-term and large-size data (Lecun et al., 2015; He et al., 2016; Liao et al., 2020). Such deep learning-based CTM models are expected to provide accurate simulations that are comparable to the current deterministic numeric CTMs but with much higher computational efficiency and better learnable capabilities. However, related advances have been limited due to difficulties in designing proper neural network structures to simultaneously achieve the goals of high accuracies, structural interpretations, and long-term simulations (Reichstein et al., 2019; Irrgang et al., 2021). In the constructions of deep learning-based CTM models, air quality simulations and predictions were always viewed as sequence-to-sequence prediction problems (Shi et al., 2015; Zhang et al., 2024) to model the spatiotemporal correlations among multiple variables. Therefore, previous studies mainly focused on refining the representation capabilities of the neural network by proposing new neural-network operations and structures to improve error back-propagation efficiencies and model encoding capabilities (Wang et al., 2018; Huang et al., 2021; Mao et al., 2021). For example, Xing et al. (2022) developed a deep learning-based module named deepCTM through mimicking atmospheric photochemical modeling to simulate ozone concentrations. However, these deep learning-based CTMs are often structured in an uninterpretable black-box style to generate simulations that reflect the cumulative effect of all physical and chemical processes. These black-box models have limitations in modelling error attribution, internal processes inspection and knowledge findings etc. (Reichstein et al., 2019). Besides, current deep learning-based CTMS are generally dedicated to specific one function, i.e. either forecast, or sensitivity analysis and transport analysis, while the deterministic numeric CTM models like CMAQ are multifunctional to conduct species concentration simulation, sensitivity analysis and internal process analysis at the same time. Quantifying the contributions of individual processes would provide fundamental explanations for a model's predictions, and therefore is also useful in identifying potential sources of error in the model formulation or its inputs (Liu et al., 2010).

In this study, we proposed a principles-oriented neural network model (FastCTM), which has explicit structures comparable to the traditional numeric CTMs to ensure model explanations, inspections, and revisions. The well-trained FastCTM model is capable of achieving multi-functionalities similar to a traditional numeric CTM, such as air quality simulations (forecasts), process analysis, emission evaluations, etc. Interpretations of the FastCTM are also widely vowed to improve deep learning model applications in earth system science and climate studies. The FastCTM model would bring many benefits with their high computation speed, efficient data assimilation and fast model updates. The FastCTM is currently configured to simulate hourly concentrations of 10 pollutant variables, including and major species of $PM_{2.5}$ (SO_4^{2-}, NO_3^-, NH_4^+, organic matters and other inorganic components, coarse part in PM_{10}, CO, NO_2, SO_2 and O_3.

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development Discussions
Open Access
EGU

2 Data and Methods

2.1 CTM Model Simulations

In this study, the FastCTM model was designed to replicate the CMAQ structures, trained by learning CMAQ's underlying physical and chemical processes among multiple air pollutants including the complicated chemical reaction, transport, diffusion and deposition. The weather and air quality simulations in 2018~2023 were conducted using a WRF-CMAQ modelling system that consists of three major components: The meteorology component of Weather Research and Forecast model (WRF, v3.4.1)(Michalakes et al., 2005; Skamarock et al., 2008) provides meteorological fields, the emission component provides gridded estimates of hourly emissions rates of primary pollutants that matched to model species, and the CTM component (CMAQ v5.0.2 (Byun and Schere, 2006)) solves the governing physical and chemical equations to obtain 3-D pollutant concentrations fields. We used hourly average concentrations of dominant $PM_{2.5}$ components of sulfate (SO_4), nitrate (NO_3), ammonium (NH_4), organic carbon (OC) and other components (EC and soil, etc.) and CO, SO_2, NO_2 and O_3 in the surface layer. Meteorological variables used in this study include relative humidity (RH), air temperature (T), wind components (U, V) at surface 10 meters height, precipitation (RN), cloud fraction (CFRAC) and planetary boundary layer height (PBLH). Wind speed (WS) was calculated from U and V. The data covered the whole China at a horizontal resolution of 12 km with 372×426 grid cells. The simulation data of 2018~2022 is used as the training dataset, while the remaining simulation data in 2023 is used for independent evaluation. The surface topographic data (HGT, Figure S1 in the supplementary material, obtained from https://lta.cr.usgs.gov/GTOPO30) and land cover data (Zhang et al., 2020) of urban and tree fraction (LULC) are also used to reflect the effects of land surface conditions in this study.

The original primary emissions used in the aforementioned WRF-CMAQ modelling system are used as input to the FastCTM. The large amount of emission data is grouped according to the simulated 10 pollutant variables. Specifically, the primary $PM_{2.5}$ emissions of SO_4, NO_3, NH_4, OC and other components, and gaseous emissions including sulfur oxide (SO_2), nitrogen oxides (NO_x, including HONO, NO, and NO_2), ammonia (NH_3), volatile organic species (VOCs, including isoprene (ISOP), terpene (TERP), and other species of VOC) are used in the FastCTM. Annual average emission of NOx, SO_2, and VOC are respectively depicted in Figure S2-4 in the supplementary material.

2.2 Guiding Principles in Designing the FastCTM Model

The deterministic CTM models simulate emissions, transport, deposition, diffusion, and chemical transformations of gases and particles in the troposphere through numerically solving the governing equations as follows,

$$\frac{\partial C_i}{\partial t} = -\nabla \cdot (\vec{u} C_i) + \nabla(K \nabla C_i) + R_i + E_i + D_i \quad (1)$$

where C_i is the concentration of species i, u is the air fluid velocity, K is the eddy diffusivity tensor, R_i is the net rate of chemical generation of species i, E_i is the rate of direct addition of the species from primary emissions, and D_i is the deposition rate caused by both dry and wet depositions. A detailed description of CMAQ principles is available elsewhere (Byun and Schere, 2006) Inspired by the traditional numeric CTMs principles and equations, the guiding framework of FastCTM was also structured in a similar formulation to represent the dominant processes in order to simulate air pollutant spatiotemporal variations.

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development
Discussions
Open Access

2.3 FastCTM Model Formulations

2.3.1 General Model Structure

In the context of deep learning, hourly air quality simulation is a spatiotemporal sequence-to-sequence learning problem to predict the most probable future length-*K* sequence given the previous length-*J* sequence as in the following Eq.2,

$$\hat{Y}_{t+1}, \ldots, \hat{Y}_{t+K} = arg \max p \ \left(\left[Y_{t-J+1}, \ldots, Y_t\right], \left[X_{t-J+1}, \ldots, X_t, X_{t+1}, \ldots, X_{t+K}\right]\right) \quad (2)$$

Where the *arg* max (short for "argument of the maximum") function is used to find the *p* class with the highest predicted probability. The $X_t \in \boldsymbol{R}^{M \times N \times V_X}$ is the data of V_X input variables at the spatial grid of $M \times N$ at time *t*. The $Y_t \in \boldsymbol{R}^{M \times N \times V_Y}$ is the data of V_Y predictive variables at time *t*. Specifically, the FastCTM simulates future *K*-hour air pollutant concentrations, given *J*-hour air pollutant concentrations $\left[Y_{t-J+1}, \ldots, Y_t\right]$ as initial fields and (*K*+*J*)-hour meteorological and emission conditions $\left[X_{t-J+1}, \ldots, X_t, X_{t+1}, \ldots, X_{t+K}\right]$. Previous studies generally used multiple-step input data with *J*>1 to ensure sufficient spatial-temporal correlations contained in the training data (Sum et al., 2022; Xing et al., 2022). Instead, we use 1-hour initial pollutant concentration (*J*=1) to simulate 24-hour air quality pollutants (*K*=24), to ensure FastCTM is dedicated to learn air quality changes between neighboring two hours.

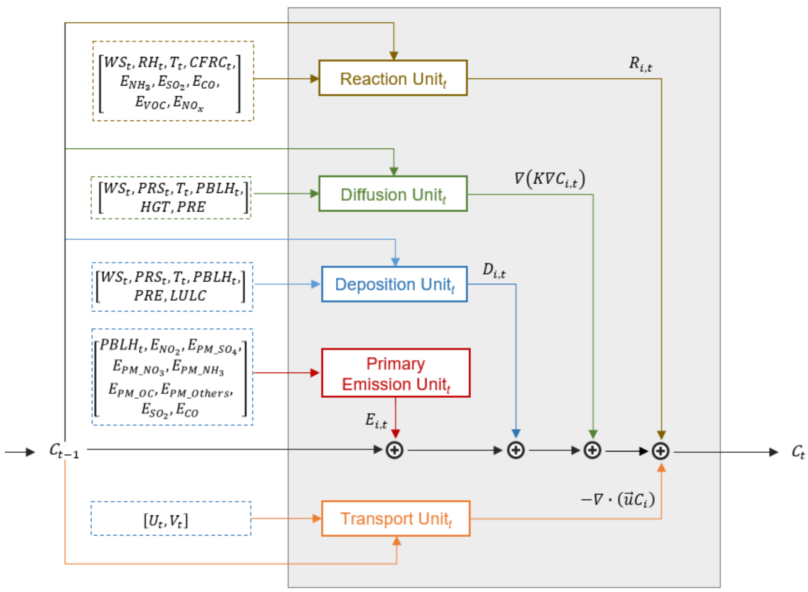


Figure 1: The basic simulator module structure at the time step *t* of deep learning simulation model FastCTM designed according to Eq.1. Arrows and boxes with different colours represent calculation modules of different atmospheric physical and chemical processes.

The FastCTM model uses the basic simulator module (Figure 1) recursively for hourly simulations. In contrast to directly learning spatiotemporal correlations of predictand itself as in most previous studies (Wang et al., 2018; Shi et al., 2017), the basic simulator is formulated following the atmospheric physical and chemical equations and constraints shown in Eq.1, and was composed of five modules to respectively represent the physics-chemical processes to improve the model performance. The modules for each of the five processes in the basic simulator are described in the following section. The time step used in FastCTM was 60 seconds.

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development
Open Access
Discussions
EGU

2.3.2 Primary Emissions Module

Primary pollutants are assumed to be directly emitted into the atmosphere and instantly well-mixed within the PBL. Therefore, hourly air pollutant concentrations enhancement caused by primary emissions could be described in the following Eq.3.

$$E_{m,n,i,t} = \frac{PE_{m,n,i,t}}{PBLH \times \mathrm{d}x \times \mathrm{d}y} \qquad (3)$$

Where $E_{m,n,k,t}$ refers to the concentration changes contributed by primary emissions at spatial coordinate (m, n) for species *i* at time *t*. The $PE_{m,n,i,t}$ is the corresponding total primary emissions within the grid cell per second. Considering that the cell size in the FastCTM is 12 km by 12 km, we have $\mathrm{d}x = 12000$ and $\mathrm{d}y$ are 12000 in this study.

2.3.3 Horizontal Transport Module

In the FastCTM, horizontal transports usually have significant influences on air quality variations (Lang, 2013). In CMAQ, the regional transport was in general represented as the divergence of the product of wind field and air pollutant species as in Eq.1, inferred from continuity equations and convection equations (Michalakes et al., 2001; Byun and Schere, 2006). By decomposing the air mass movement into two orthogonal directions of east-west (*x*) and north-south (*y*), they could be re-written in the form as shown in Eq. 4,

$$\nabla \cdot (\vec{u} C_i) = \frac{\partial (C_i U)}{\partial x} + \frac{\partial (C_i V)}{\partial y} \quad (4)$$

Where the wind field was represented as $\vec{u}$, which was then decomposed into U and V, respectively in the *x* and *y* directions. In the deep learning framework, the partial equation in Eq. 4 could be rewritten in a discrete form as convolution operations and inner product calculations as shown in Eq. 5 with a finite difference method. The convolutional kernels of W_x and W_y were defined in an up-wind scheme as shown in Eq. 6 and Eq. 7.

$$\nabla \cdot (\vec{u} C_i) = \frac{W_x * (C_i \times U)}{\mathrm{d}x} + \frac{W_y * (C_i \times V)}{\mathrm{d}y} \quad (5)$$

$$W_x = \begin{cases} [-1 \quad 1 \quad 0] \; if \; U < 0 \\ [0 \quad -1 \quad 1] \; if \; U \geq 0 \end{cases} \quad (6)$$

$$W_y = \begin{cases} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \; if \; V < 0 \\ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \; if \; V \geq 0 \end{cases} \quad (7)$$

2.3.4 Diffusion Module

The turbulence diffusion process $\nabla(K \nabla C_i)$ in Eq.1 helps the spread of pollutants in the atmosphere. It is expressed as the second-order deviation of species concentrations as shown in Eq. 8. They could also be discretized to convolutional operations with finite difference method as shown in Eq. 9, just like that in the horizontal transport process module.

$$\nabla(K \nabla C_i) = \frac{\partial}{\partial x}\left(K \frac{\partial C_i}{\partial x}\right) + \frac{\partial}{\partial y}\left(K \frac{\partial C_i}{\partial y}\right) \quad (8)$$

$$\nabla(K \nabla C_i) = \frac{W_x * (K \times W_x * C_i)}{\mathrm{d}x \times \mathrm{d}x} + \frac{W_y * (K \times W_y * C_i)}{\mathrm{d}y \times \mathrm{d}y} \quad (9)$$

$$K = Encoder_K([T, RH, PRS, PBLH]) \quad (10)$$

The turbulent diffusivity K is closely related to the meteorological conditions of the atmosphere and is simulated with an encoder module $Encoder_K$ (Eq. 10). The input variables of the $Encoder_K$ include temperature T, humidity RH, surface

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development
Discussions
Open Access
EGU

pressure PRS, and boundary layer height $PBLH$. The $Encoder_K$ is determined to be a grid-to-grid regression model based on the Unet++ model with a nested structure (Zhou et al., 2018; Ronneberger et al., 2015). The $Encoder_K$ model consists of 5 layers with each layer respectively composed of 16, 32, 64, 128 and 256 filters.

2.3.5 Chemical Reaction Module

The air pollutant concentration changes caused by chemical reactions are represented in the following Eq. 11. In the equation, the rate of chemical reaction of species i is expressed as the product of a rate constant k and a term that is dependent on the concentrations of its reactants j (Carter, 1990; Carter and Atkinson, 1996).

$$R_{m,n,i,t} = k_{m,n,i,t} \times f\left(C_{m,n,j,t}\right) \quad (11)$$

$$k_i = Encoder_k([T, RH, PRS, WS, PRE, CFRAC]) \quad (12)$$

The reaction kinetics constant k is generally temperature-dependent. They could also be related to atmospheric pressures and moisture humidity in some reaction processes. Therefore, the reaction rate constant k is simulated using a spatial encoder function $Encoder$ as shown in Eq. 12, which has the same structure as that of reaction and deposition encoder modules (Eq. 10). There are 6 input variables of the $Encoder_k$ including T, RH, PRS, WS, RN and $CFRAC$. The concentration processor f is designed as a simple multi-layer convolutional network with a kernel size of 1 to represent high-order and complex relations among different reactants.

2.3.6 Deposition Module

Air quality changes due to the deposition process are expressed linearly as the product of the deposition rate d and the corresponding air pollutants concentrations C, as shown in Eq. 13. The constant d is closely related to the current and previous meteorological conditions, terrains and underlying land cover types. Therefore, they are all simulated with an *Encoder* module as shown in Eq. 14.

$$D_{m,n,i,t} = d_{m,n,i,t} \times C_{m,n,i,t} \quad (13)$$

$$d = Encoder_d([WS, RH, RN, HGT, LULC]) \quad (14)$$

The model structure and parameter configurations are also the same as that of $Encoder_k$. The input data variables of $Encoder_d$ include WS, RH, RN, HGT and LULC.

2.4 Model Training

The FastCTM was programmed with Python3 on the deep learning framework TensorFlow (Abadi et al., 2016). The model was trained with the WRF-CMAQ operational forecast data in China for 2018~2022. Considering that on each day we had 120-hour forecasts with a spatial coverage of 426×372 grid cells (each with a size of 12×12 km^2) for 9 meteorological variables and $\boldsymbol{I}$=10 air pollutant variables, the total training data have a size of $\boldsymbol{TD} = \boldsymbol{R}^{1826,120,426,372,19}$, where 1826 represents the total counting days from 2018 to 2022. Since the model was set to predict 24-hour $PM_{2.5}$ concentrations from input 1-hour data, the total input sequence length was 25 hours in each training step. Besides, the size $M \times N$ of input data X_t to FastCTM was decided to be 150×150, equal to an area of 1800×1800 km^2 in 12-km resolution. Therefore, the input data for FastCTM in each step should be in the size of $\boldsymbol{BD} = \boldsymbol{R}^{b,25,150,150,19}$, where b is the batch size (determined as 1 in this study). In the training process, the input data $\boldsymbol{BD}$ are randomly sliced from the whole training dataset $\boldsymbol{TD}$ in each training iteration. We did not use the fixed area as that in the previous studies (Xing et al., 2022) to ensure that the model learns inherent physical and chemical principles rather than just statistical spatiotemporal autocorrelations in a fixed area.

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development Discussions

Besides, the spatio-temporal random samples contain varied emissions which would improve FastCTM adaption to changing emission levels.

The loss function was determined to be L2 loss (Bühlmann and Yu, 2003) of the regularized mean squared error (MSE) as shown in Eq. 15. The model was optimized with the Adam optimizer (Kingma and Ba, 2014).

$$\mathcal{L} = \frac{1}{J \times N \times M \times I} \sum_{t=1}^{J} \sum_{m=1}^{M} \sum_{n=1}^{N} \sum_{i=1}^{I} \left(C_{m,n,i,t} - \tilde{C}_{m,n,i,t} \right)^2 \quad (15)$$

The FastCTM model was trained on one entry-level professional acceleration card of NVIDIA A40 with a running time of 10 hours for every 10000 iterations. A total of 300, 000 iterations were performed before the remaining model loss becoming stable.

2.5 Model Evaluation

The main objective of our study is to build and validate a principles-guided neural network based FastCTM that could simulate spatial-temporal fields of hourly concentrations of major air pollutant species like a traditional CTM. Besides, the FastCTM could model individual contributions from each of the atmospheric processes of transport, diffusion, deposition, reaction and emission. Therefore, the FastCTM simulations were first assessed against CMAQ simulations using the same input emission data and meteorological fields. The CMAQ model simulated 120-hour forecasts from 0:00 local time on each day of 2023, while the FastCTM model generated 119-hour forecasts with 1-hour initial input data. The 119-hour forecasts are achieved by iteratively using an initialized condition from the previous step. The 119-hour forecast data by the two models were compared hour-by-hour at each corresponding time. For example, when we had 120-hour forecast starting at 0:00 on January 1, 2023 at Beijing Local Time (BLT), the data of 0:00 on January 1, 2023 were fed into FastCTM to get the 119-hour forecasts until 23:00 on January 5. The 10 species forecasts by FastCTM were compared against the CMAQ forecasts at each corresponding hour. Furthermore, CMAQ and FastCTM forecasts were both evaluated by hourly observations from national monitoring sites (as shown in Figure S5 in the supplementary material) for six criteria pollutants ($PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3). The metrics of root mean square error (RMSE) and coefficient of determination (R^2) were calculated.

Besides, the FastCTM was also assessed from the aspects of sensitivity analysis to emission inputs and meteorological fields. For meteorological variables, responses of six criteria pollutant concentrations to T, WS and PBLH were calculated. For emissions, responses to paired variables of SO_2/NH_4 and NOx/VOC emissions were calculated. Finally, the contributions by five internal processes of transport, diffusion, emission, reaction, and deposition were also analyzed and discussed for an example pollution episode.

3 Results

3.1 Forecast Performance by FastCTM

3.1.1 Evaluation of FastCTM forecasts against CMAQ forecasts

The FastCTM has exhibited strong and stable strengths in reproducing CMAQ forecasts for a long-lasting forecast period of 119 hours evaluated in 2023 (Figure 2). The average RMSE values for six criteria pollutants of $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3 are respectively 9.1, 11.9, 4.4, 4.0, 48.9 and 10.9 μg/m^3. For R^2 values, they are 0.8, 0.81, 0.8, 0.83, 0.9 and 0.7. As for $PM_{2.5}$ components, RMSE values are 1.68, 2.68, 1.52, 1.98 and 4.25 μg/m^3 respectively for SO_4^{2-}, NO_3^-, NH_4^+, organic matters and other inorganic components, while the R^2 values are 0.72, 0.6, 0.3, 0.83 and 0.68. The low R^2 value of NH_4^+ could be caused by insufficient chemical reactions represent in FastCTM as not enough chemicals considered in the model. Compared to the ~5ppb (~10.5 μg/m³) in the previous study by Xing et al. (2022), the FastCTM model has similar RMSE values. Hourly RMSE values have apparent diurnal variations with lower RMSE values in the nighttime than that in the daytime. This is probably due to more active physical and chemical processes in the daytime, which is the header to simulate for FastCTM. Besides, since the FastCTM is a 2-D model only considering atmospheric processes within the boundary layer, lower consistency with the CMAQ model during daytime could be due to more active vertical turbulence which is not fully represented. It is important to note that the relatively low R^2 values observed for NH_4^+ can be attributed to the fact that it is the sole cation included in the FastCTM model without a corresponding acid-base balance, which may affect the model's predictive accuracy.

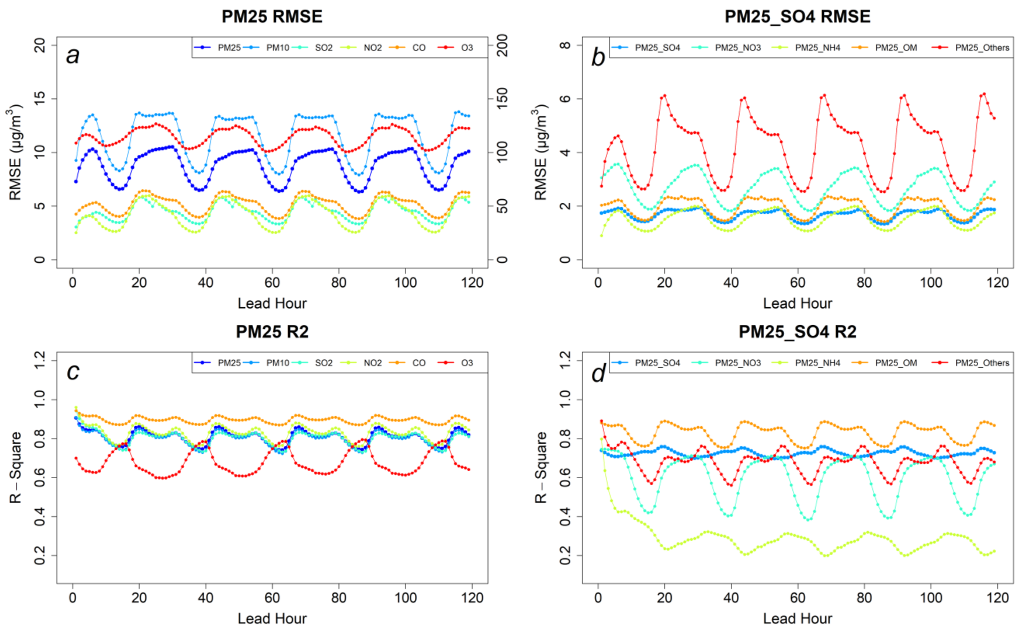
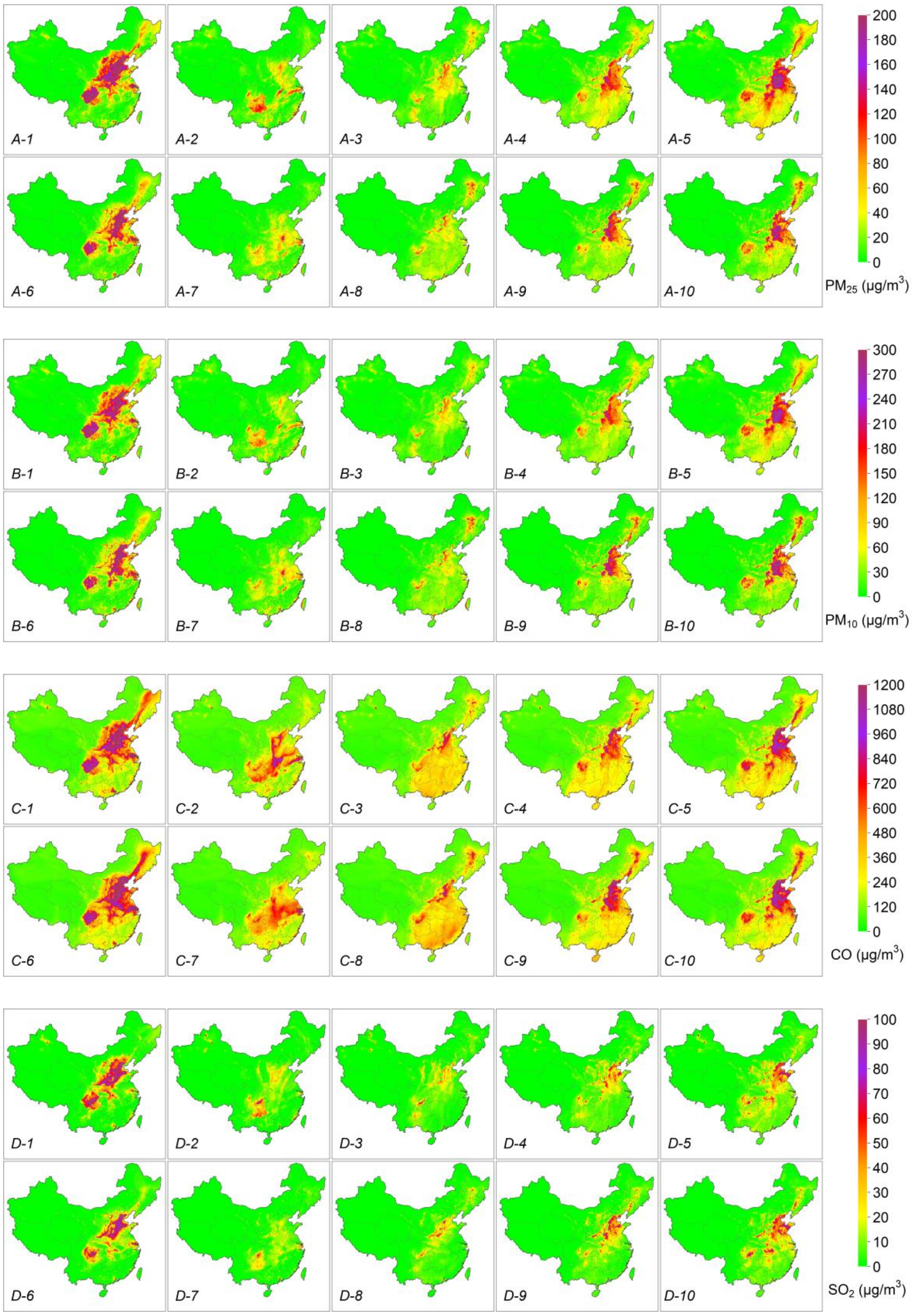


Figure 2: The evaluation performances of FastCTM forecasts against CMAQ forecasts in 2023. Panel (a) and (b) respectively show RMSE values of criteria pollutants and the $PM_{2.5}$ components of. Panel (c) and (d) respectively show R^2 values. It should be noted that RMSE value of CO corresponds to the right axis in panel (a).

Furthermore, we tested the influences of initial condition on FastCTM long-term simulations. As shown in Figure S6 in the SI, FastCTM forecasts using zero values as input air quality data were almost the same as that using ordinary input in the long leading hours, indicating that FastCTM simulations in long leading hours are not affected by initial conditions, just like deterministic numeric CTMs (such as CMAQ). In other words, the insensitivities of FastCTM to initial conditions indicate that it has well learned and encoded the most physical and chemical principles in CMAQ CTM, rather than just spatio-temporal correlations among air quality sequences.

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development
Discussions



https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

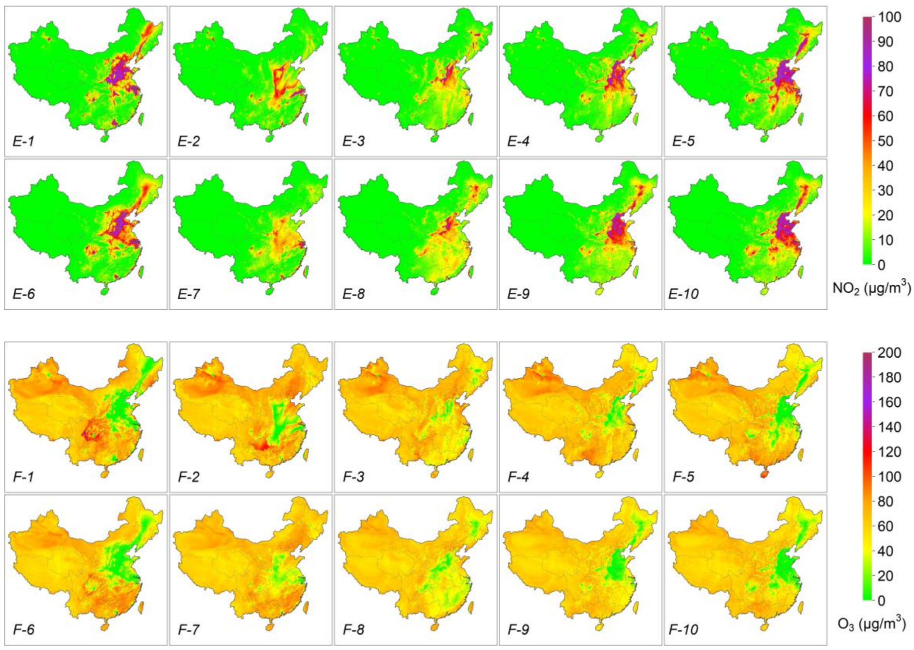


Figure 3: Air quality forecast examples of CMAQ and FastCTM at leading time of 24, 48, 72, 94 and 120 hours starting from 0:00 on March 4th, 2023. Panel A-F respectively refers to $PM_{2.5}$, PM_{10}, CO, SO_2, NO_2 and O_3. The 1-5 sub-panels in the first row (1-5) in each panel are the CMAQ forecasts, while the 6-10 sub-panels in the second row are FastCTM forecasts.

Air quality forecasts (Figure 3) starting from 00:00 a.m. on March 4th, 2023 demonstrated the strong capabilities of FastCTM in modeling the complex spatio-temporal changes in a large spatial domain and over a relatively long period. In this period, air quality experienced rapid deterioration. For the pollutants except for O_3, both CMAQ and FastCTM simulations have predicted very high concentrations at the 24th-hour forecast in the areas of the North China and Sichuan Basin area. During the next four days, the air quality was first cleaned up but then became worse, which was reflected both in the CMAQ and FastCTM. Generally, in this complicated process, the FastCTM generated very similar forecasts to that of the CMAQ forecasts in a long-term period over a large area. The O_3 generally has a close relationship with the ratio of VOCs/NOx, the increased NO_2 could lead to decreased O_3 due to titration effect (Ren and Xie, 2022). The results have indicated that, with FastCTM, hourly ground-level concentrations of major air pollutants can be generated fast with high reducibility to CMAQ.

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development Discussions

3.1.2 Evaluation of FastCTM forecasts against station observations

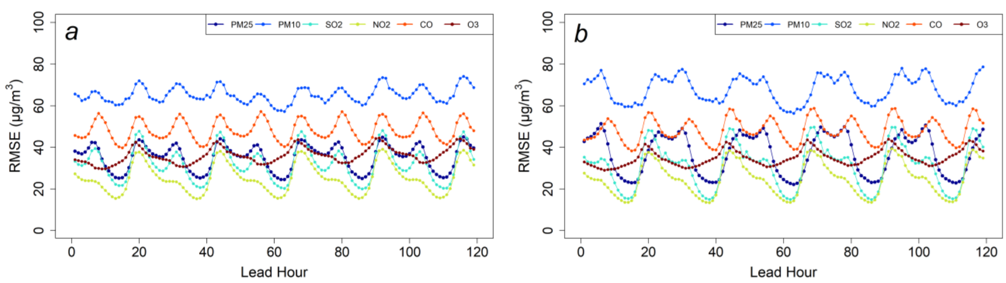


Figure 4: The evaluation performances of $PM_{2.5}$ forecasts by FastCTM (a) and CMAQ (b) against observations in 2023.

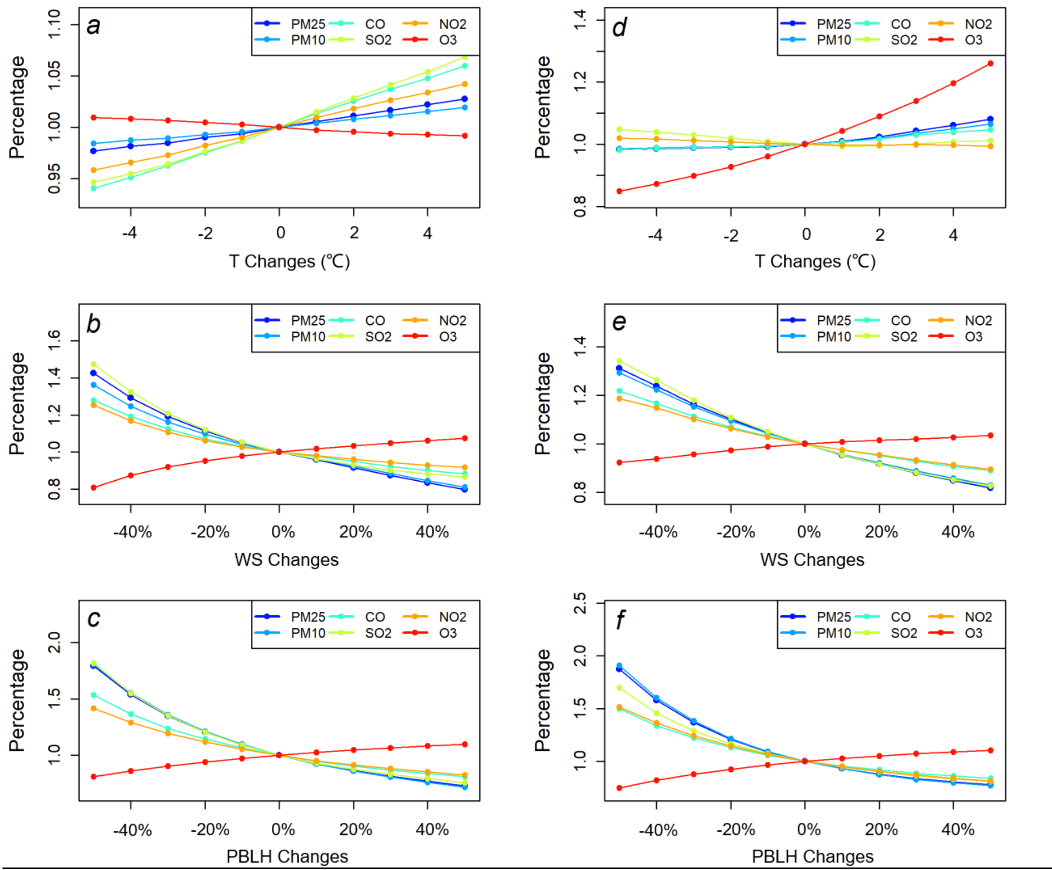
The FastCTM forecasts also have comparable performances to CMAQ forecasts in the evaluation against observations at national monitoring sites as shown in Figure 4. Generally, both the FastCTM and CMAQ forecasts have lower accuracies in the daytime than that during the night. For FastCTM, average RMSE values for $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO, and O_3 are respectively 36.7, 67.9, 31.05, 24.7, 482.1 and 36.2 μg/m^3, compared to that of 35.2, 65.2, 31.9, 25.2, 483.4 and 35.0 μg/m^3 for CMAQ. The relative difference for the RMSE values of FastCTM to CMAQ are respectively 4.3%, 4.2%, -2.8%, -1.7%, -0.3% and -3.2%. The differences between FastCTM and CMAQ are within a small range of ±5%. In consideration that the FastCTM model was trained with CMAQ simulations, their close evaluation performances are well within expectations.

3.2 Sensitivity Analysis with FastCTM

The FastCTM model was trained with 5-year meteorological and air quality simulations by WRF-CMAQ. These simulations used the same annual emission inventory data for each year. In this condition, the FastCTM model has learned the relationships between the air quality and varied meteorology with fixed emissions input. Considering that the FastCTM model has exhibited high accuracy at an independent evaluation year 2023 when new meteorological fields are fed into FastCTM, the deep learning model should be able to simulate responses of air pollutant concentrations to meteorological variables. However, for the response of air pollutant concentrations to emissions, the training data do not contain relationships between inter-annual varied emissions and air quality under the condition of same annual meteorological fields. Therefore, it is less expected for FastCTM to simulate reliable and correct response relationships between emissions and air quality. To validate these analyses, we calculated the sensitivities of simulated air pollutant concentrations to changes in meteorological variables and emissions.

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

3.2.1 Response of Air Pollutant Concentration to Meteorology



295 **Figure 5: The FastCTM predicted air pollutant percentage changes responding to changes of T, WS, and PBLH in Beijing on January 2nd (a-c respectively in the left column) and August 1st (d-f respectively in the right column) , 2023. The air pollutant concentrations are relative to those at the baseline meteorological conditions.**

The responses of six criteria pollutants to meteorological changes simulated by FastCTM are evaluated as exhibited in Figure 5. For ground-level temperature T, O_3 concentrations have distinct response curvature compared to the other five
300 criteria pollutants. O_3 concentrations have slight negative responses to T in January as shown in Figure 6a, which is probably due to stronger dilution effects with increased NOx emissions with higher air temperature. O_3 concentrations had the strongest positive responses in August among six pollutants, which is consistent with previous observation-based studies (Flaum et al., 1996). The O_3 had larger sensitivities when the air temperature was higher. The gaseous pollutants of CO, NO_2 and SO_2 have the most significant positive responses to air temperature, which could be caused by the shift of
305 chemical equilibrium towards to the higher release of these gaseous pollutants (Bassett and Seinfeld, 1983; Cox, 1982). The particulate matter pollutants, especially PM_{10}, have the weakest responses in six pollutants. Considering that there are

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

dominating proportions of chemically inert species in particulates, the weak responses of $PM_{2.5}$ and PM_{10} are expected. For the wind speed and PBLH, the responses of pollutants are having similar patterns for the same pollutant. First, O_3 concentrations exhibited adverse patterns contrast to other pollutants both in January and August. Higher wind speed would increase the dispersion and transport of air pollutants (Feng et al., 2015; Lv et al., 2017) resulting in lower pollution levels, which is the reason for decreased concentrations along increased with increasing wind speed, except for O_3. The contradictory response of ozone and particulate matter concentrations to PBLH is consistent with the analysis results of multiple-year observations (Liu and Tang, 2024). Theoretically, the air pollutant concentrations should exhibit an inverse relationship between air pollution concentrations and PBLH. The actual air pollutant concentration changes simulated by FastCTM generally fit the theory that there are negative non-linear effects with increasing PBLH. Meanwhile, the sensitivity is stronger when the PBLH is lower (Figures 6e and 6f), which is consistent with previous observation-based analysis (Wang et al., 2019; Su et al., 2020). The totally different relationship of O_3 to wind speed and PBLH compared to other pollutants could be due to its high dependence on chemical precursors, such as NO_x and VOC. Concentrations of these precursors could have an inverse relationship with O_3 at specific locations. FastCTM model itself is trained with multi-year CMAQ simulations, indicating that it is preconditioned on varied meteorological fields with the same atmospheric physical and chemical rules. Therefore, the sensitivity of air quality simulations to meteorology variations could be well learned, especially with the disciplinary-based model FastCTM.

3.2.1 Response of Air Pollutant Concentration to Emission

The sensitivity analysis with a "brute force" method can be carried out with the FastCTM model quickly due to its high computational efficiency on GPU. The responses of $PM_{2.5}$ concentrations to doubled emissions of SO_2, NOx were explored in a winter month of January 2023 (Figure 6). For doubled NOx, the $PM_{2.5}$ concentrations exhibited positive responses in most areas of China as shown in Figure 6a. The most significant increase occurred in regions like North China, Heilongjiang province in Northeast China, Yangtze River Delta and Sichuan province. In these places, the NOx emission are relatively large. For doubled SO_2, $PM_{2.5}$ concentrations increased in almost all China as shown in Figure 6b. The response was larger in places of North China, Northeast China and Sichuan basin. The responses results of $PM_{2.5}$ simulated by the FastCTM was generally consistent to previous studies (Li et al., 2022).

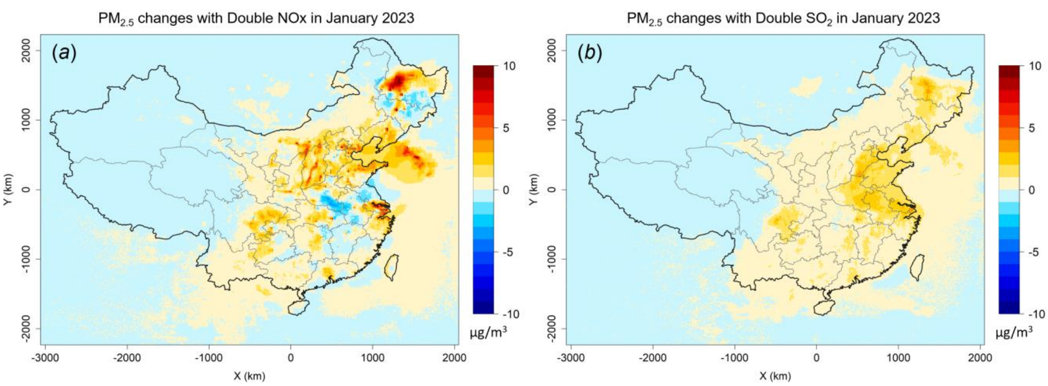


Figure 6: Average predictions of $PM_{2.5}$ concentrations in 5 lead-days with doubled emissions in January 2023. Panel (*a*) refers to predictions with doubled NOx and panel (*b*) refers to double SO_2.

As for ozone, its responses to doubled NOx and VOC are explored as shown in Figure 7. For NOx emission, decreased O_3

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

concentrations in polluted regions like North China, Yangtze River Delta and other highly industrial regions are well simulated by FastCTM (Figure 7a). The response is reasonable considering that these regions are generally abundant with NOx emissions and at VOC-limited conditions. Doubled VOC emission lead to significant decrease of O_3 (Figure S7 in the supplementary material), which could be caused by the reason that increased VOC could consume O_3 in these regions. The spatial patterns of O_3 responses to NOx and VOC are similar to previous deep learning study trained by emission-controlled simulation data (Xing et al., 2022). However, due to complex speciation of VOC emissions that's simplified in the FastCTM, uncertainties for responses of O_3 to VOC should be noted.

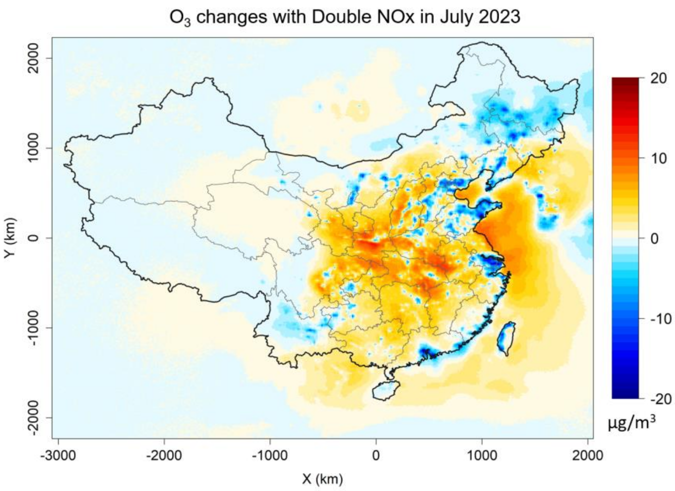


Figure 7: Average predictions of hourly O_3 concentrations in 5 lead-days with doubled NOx emissions in July 2023.

We used a principles-constrained formulation approach in designing the FastCTM model framework. As shown in Eq.4, atmospheric chemical reactions are in the Atkinson form which independently estimate the reaction rate from meteorological conditions and polynomials of reactants concentrations in multiple powers. The principle-based formulation should be the reason for the relatively significant and reasonable response simulations of $PM_{2.5}$ and O_3 to precursor emissions, even though the FastCTM itself is not trained by emission-controlled CMAQ scenario simulations. The remaining uncertainties should be attributed to the reason that FastCTM only considered environmental chemical reactants in part comparing to that of CMAQ model (Binkowski and Roselle, 2003).

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Geoscientific Model Development
Discussions

3.3 Internal Processes Analysis with FastCTM

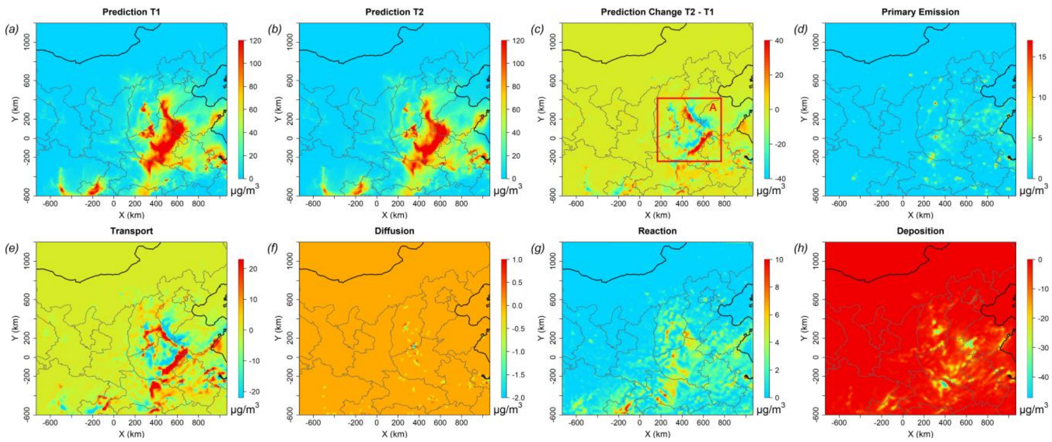


Figure 8: An example of the $PM_{2.5}$ concentration at T1 (18:00, panel *a*) and T2 (19:00, panel *b*) on January 13, 2023 (with the forecast leading time of 42 hours) and hourly changes (panel *c*). Changes caused by each of the five dominant processes are depicted in panels *d-h*.

The FastCTM is a principles-guided deep neural network to individually simulate the dominant atmospheric physical and chemical processes as defined in Eq.1. The processes are calculated numerically with critical parameters describing the processes being estimated by deep learning encoders. The hourly variations are equal to the sums of air pollutants' concentration changes in each process. Therefore, the contributions of these processes to air pollutant concentrations changes could be elaborately calculated. Figure 8 depicts an example during the night-time of January 13, 2023 when hourly $PM_{2.5}$ concentration changes significantly. Between the two hours of 18:00 and 19:00, hourly $PM_{2.5}$ concentrations have significantly changed in neighbouring areas of Shandong, Hebei and Henan provinces as shown in the red rectangle (denoted as Area A hereafter) in Figure 8c. In this example, strong northern wind prevails leading pollutants moving southward.

For $PM_{2.5}$ concentration changes caused by primary emissions (Figure 8d), it's determined by the primary emission and the mixing volumes determined by PBLH. In this episode, the hourly $PM_{2.5}$ changes are mostly determined by the transport process (Figure 8e) since its spatial pattern has the most resemblance to the total $PM_{2.5}$ concentration changes. In the transport process, air pollutants move from one area to another determined by the wind fields as shown in Eq.4. When the northern clean air prevails as in the Area A, changes should be negative in the upstream direction and positive in the downstream direction. The transport process simulated by FastCTM sticks to this pattern. As known to us, the diffusion process will bring pollutants from a region of high concentration to one of low concentration. Its contribution is low as shown in Figure 8f, which is reasonable considering the relatively large grid cell size of 12 km and short simulation period of 1 hour. $PM_{2.5}$ concentration changes caused by the diffusion process constituted a small proportion compared to other processes. The activities of chemical reactions are determined by both meteorological conditions and related precursor concentrations. $PM_{2.5}$ contribution changes between T1 and T2 caused by chemical reactions are lower in the areas to the north of Area A because the cold and clean air in this area is not favourable for chemical reactions. The deposition is the dominant process that led to $PM_{2.5}$ concentration reductions where regional transport was not significant. In general, depositions were proportional to $PM_{2.5}$ concentrations as shown in Figure 8h (Davis and Swall, 2006). It should be noted that FastCTM simulated air quality in a 2-D domain rather than in 3-D, the deposition could also include the vertical transport of air pollutants to the upper air above PBL (Zhao et al., 2020).

4 Discussions

The FastCTM was a neural network-based CTM model for speeding up air quality simulations and forecasts. Comparing to the previous deep learning based CTMs, the FastCTM has more functionalities like a traditional CTM. First, it is able to simulate 10 air pollutants including criteria gas pollutants, coarse particulate matter, and five species concentrations of $PM_{2.5}$. The FastCTM has relatively high agreements in long-term forecasts with the conventional CTM. Besides, FastCTM simulations are not related to its initial condition of input air quality fields after around two-day simulation, which indicates that it has well learned the inherent physical and chemical processes in CTM rather than only the spatial-temporal auto-correlations of input time-series data. Meanwhile, it has exhibited reasonable responses to precursor emission changes and meteorological condition changes in the sensitivity analysis. Furthermore, the internal processes in the FastCTM model were checkable and interpretable by analyzing the contributions of dominant atmospheric chemical and physical processes separately. These processes are encoded within FastCTM by designing dedicated neural network modules.

Previous deep learning-based models for emission sensitivity analysis were generally trained by simulations with a group of different emission scenarios, whereas the FastCTM model was trained by CMAQ simulations of unvaried annual emissions. The relative reasonable responses simulations to emissions and meteorological data revealed that the principles used in formulating the FastCTM have helped the model to better learn inherent physical and chemical processes within the training data. Considering the high computation consumption by conventional CTM, FastCTM would reduce substantial computational resources.

The FastCTM has the capabilities to generate hourly pollutant simulations with nearly equal accuracies to that by CMAQ CTM, evaluated by observations at national monitoring sites. There are still differences and potential errors within the FastCTM, arising from inadequate representations of actual atmospheric processes and mechanisms. First, there are missing processes were considered within the FastCTM. The chemical reactions in traditional CMAQ are very complex and involves many organic and inorganic species in gaseous and aqueous phases. The FastCTM has just modeled potential chemical reactions among several atmosphere compositions. Besides, long-range air pollutant transport in the upper atmosphere above the planetary boundary layer was not considered within the FastCTM model. The remaining uncertainties of FastCTM compared to CMAQ could be further reduced after carefully detailing atmospheric processes with properly designed neural network modules.

It should also be noted that atmospheric physical and chemical processes are defined in principles-guided neural network modules in FastCTM. Their specific formulation was learned and optimized to minimize the sum loss errors of all species concentrations, rather than being supervised by data of actual internal processes in CMAQ. The actual contributions of air pollutant concentration changes by each of these processes could be calculated with the integrated process rate (IPR) analysis and integrated reaction rate (IRR) analysis tools within CMAQ. Future studies could use these IPR and IRR results to supervise the simulated processes in FastCTM to further improve its simulation accuracies and robustness.

Data availability. The land use and land cover data are available at the Data Sharing and Service Portal of Chinese Academy of Science (http://data.casearth.cn/en/sdo/detail/5ebe2a9908415d14083a4c24). The CTM simulation data and source code files of the exact version used to produce the results used in this paper is available at https://doi.org/10.5281/zenodo.13757211 on Zenodo (Lyu, 2024). The configuration files for running models of WRF v3.4.1 and CAMQ v5.0.2 are also available at https://doi.org/10.5281/zenodo.5152621 (Hu, 2021).

Author contributions. BL and YH conceived the study. BL developed the model and codes. RH and XW contributed the CTM simulation data. BL and RH collected the observation data. BL analyzed data and wrote the paper with contributions from YH, RH, WW and XW. RH managed the project.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research has been in part supported by the AiMa R&D Project (R#2016-004) of Hangzhou AiMa Technologies. The findings in this research do not necessarily reflect the views of the sponsors.

Reference

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., and Devin, M.: Tensorflow: Large-scale machine learning on heterogeneous distributed systems, arXiv preprint arXiv:1603.04467, 2016.

Bassett, M. and Seinfeld, J. H.: Atmospheric equilibrium model of sulfate and nitrate aerosols, Atmospheric Environment (1967), 17, 2237-2252, https://doi.org/10.1016/0004-6981(83)90221-4, 1983.

Binkowski, F. S. and Roselle, S. J.: Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description, Journal of Geophysical Research Atmospheres, 108, 335-346, 2003.

Bühlmann, P. and Yu, B.: Boosting With the L2 Loss, Publications of the American Statistical Association, 98, 324-339, 2003.

Byun, D. and Schere, K. L.: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system, Applied Mechanics Reviews, 59, 51-77, 2006.

Carter, W. P. L.: A detailed mechanism for the gas-phase atmospheric reactions of organic compounds, Atmospheric Environment. Part A. General Topics, 24, 481-518, https://doi.org/10.1016/0960-1686(90)90005-8, 1990.

Carter, W. P. L. and Atkinson, R.: Development and evaluation of a detailed mechanism for the atmospheric reactions of isoprene and NOx, International Journal of Chemical Kinetics, 28, 497-530, https://doi.org/10.1002/(SICI)1097-4601(1996)28:7<497::AID-KIN4>3.0.CO;2-Q, 1996.

Council, N.: Air quality management in the United States, National Academies Press2004.

Cox, R. A.: Chemical Transformation Processes for NoX Species in the Atmosphere, in: Studies in Environmental Science, edited by: Schneider, T., and Grant, L., Elsevier, 249-261, https://doi.org/10.1016/B978-0-444-42127-2.50027-0, 1982.

Davis, J. M. and Swall, J. L.: An examination of the CMAQ simulations of the wet deposition of ammonium from a Bayesian perspective, Atmospheric Environment, 40, 4562-4573, 2006.

Eder, B., Kang, D., Mathur, R., Yu, S., and Schere, K.: An operational evaluation of the Eta–CMAQ air quality forecast model, Atmospheric Environment, 40, 4894-4905, 2006.

Feng, X., Li, Q., Zhu, Y., Hou, J., Jin, L., and Wang, J.: Artificial neural networks forecasting of $PM_{2.5}$ pollution using air mass trajectory based geographic model and wavelet transformation, Atmospheric Environment, 107, 118-128, http://dx.doi.org/10.1016/j.atmosenv.2015.02.030, 2015.

Flaum, J. B., Rao, S. T., and Zurbenko, I. G.: Moderating the Influence of Meteorological Conditions on Ambient Ozone Concentrations, Journal of the Air & Waste Management Association (1995), 46, 35-46, 1996.

Hakami, A., Odman, M. T., and Russell, A. G.: High-Order, Direct Sensitivity Analysis of Multidimensional Air Quality Models, Environmental Science & Technology, 37, 2442-2452, 10.1021/es020677h, 2003.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, IEEE, 2016.

Huang, L., Liu, S., Yang, Z., Xing, J., and Liu, T. Y.: Exploring Deep Learning for Air Pollutant Emission Estimation, Geoscientific Model Development, 14, 4641–4654, 2021.

Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., and Saynisch-Wagner, J.: Towards neural Earth system modelling by integrating artificial intelligence in Earth system science, Nature Machine Intelligence, 3, 667-674, 10.1038/s42256-021-00374-3, 2021.

Kingma, D. and Ba, J.: Adam: A Method for Stochastic Optimization, Computer Science, 2014.

Lang, J.: A Monitoring and Modeling Study to Investigate Regional Transport and Characteristics of PM2.5 Pollution, Aerosol & Air Quality Research, 13, 943-956, 2013.

LeCun, Y., Bengio, Y., and Hinton, G.: Deep learning, Nature, 521, 436-444, 10.1038/nature14539, 2015.

Li, J., Dai, Y., Zhu, Y., Tang, X., Wang, S., Xing, J., Zhao, B., Fan, S., Long, S., and Fang, T.: Improvements of response surface modeling with self-adaptive machine learning method for $PM_{2.5}$ and O_3 predictions, Journal of Environmental Management, 303, 114210, https://doi.org/10.1016/j.jenvman.2021.114210, 2022.

Liao, Q., Zhu, M., Wu, L., Pan, X., Tang, X., and Wang, Z.: Deep Learning for Air Quality Forecasts: a Review, Current Pollution Reports, 6, 399-409, 10.1007/s40726-020-00159-z, 2020.

Liu, X.-H., Zhang, Y., Xing, J., Zhang, Q., Wang, K., Streets, D., Jang, C., Wang, W.-X., and Hao, J.-M.: Understanding of regional air pollution over China using CMAQ, Part II. Process analysis and sensitivity of ozone and particulate matter to precursor emissions, Atmospheric Environment, 44, 3719-3727, 10.1016/j.atmosenv.2010.03.036, 2010.

Liu, Y. and Tang, G.: Contradictory response of ozone and particulate matter concentrations to boundary layer meteorology, Environmental Pollution, 343, 123209, https://doi.org/10.1016/j.envpol.2023.123209, 2024.

Lv, B., Cai, J., Xu, B., and Bai, Y.: Understanding the Rising Phase of the $PM_{2.5}$ Concentration Evolution in Large China Cities, Scientific Reports, 7, 46456, 10.1038/srep46456

Mao, W., Wang, W., Jiao, L., Zhao, S., and Liu, A.: Modeling air quality prediction using a deep learning approach: Method optimization and evaluation, Sustainable Cities and Society, 65, 102567, , 2021.

Michalakes, J., Chen, S., Dudhia, J., Hart, L., Klemp, J., Middlecoff, J., and Skamarock, W.: Development of a next-generation regional weather research and forecast model, IEEE International Conference on High Performance Computing, Data, and Analytics, 11/1/2001, 10.1142/9789812799685_0024, 2001.

Michalakes, J., Dudhia, J., Gill, D., Henderson, T., Klemp, J., Skamarock, W., and Wang, W.: The Weather Research and Forecast Model: Software Architecture and Performance, 10.1142/9789812701831_0012, 2005.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195-204, 10.1038/s41586-019-0912-1, 2019.

Ren, J. and Xie, S.: Diagnosing ozone-NOx-VOC sensitivity and revealing causes of ozone increases in China based on 2013–2021 satellite retrievals, 10.5194/acp-2022-347, 2022.

Ronneberger, O., Fischer, P., and Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, Medical

https://doi.org/10.5194/gmd-2024-198
Preprint. Discussion started: 20 November 2024

Image Computing and Computer-Assisted Intervention – MICCAI 2015, Cham, 2015//, 234-241,

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-k., and Woo, W.-c.: Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, Computer Science, 2015.

Shi, X., Gao, Z., Lausen, L., Wang, H., Yeung, D. Y., Wong, W., and Woo, W.: Deep Learning for Precipitation Nowcasting: A Benchmark and A New Model, 2017.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., and Powers, J. G.: A description of the advanced research WRF version 3, NCAR Tech. Note 2008., 2008.

Su, T., Li, Z., Zheng, Y., Luan, Q., and Guo, J.: Abnormally Shallow Boundary Layer Associated With Severe Air Pollution During the COVID-19 Lockdown in China, Geophysical Research Letters, 10.1029/2020GL090041, 2020.

Sum, H., Fung, J. C. H., Chen, Y., Li, Z., Yuan, D., Chen, W., and Lu, a. X.: Development of an LSTM broadcasting deep-learning framework for regional air pollution forecast improvement, Geosci. Model Dev., 15, 8439–8452, 2022.

Wang, C., Jia, M., Xia, H., Wu, Y., Wei, T., Shang, X., Yang, C., Xue, X., and Dou, X.: Relationship analysis of $PM_{2.5}$ and boundary layer height using an aerosol and turbulence detection lidar, Atmospheric Measurement Techniques, 12, 3303-3315, 10.5194/amt-12-3303-2019, 2019.

Wang, L., Jang, C., Zhang, Y., Wang, K., Zhang, Q., Streets, D., Fu, J., Lei, Y., Schreifels, J., He, K., Hao, J., Lam, Y.-F., Lin, J., Meskhidze, N., Voorhees, S., Evarts, D., and Phillips, S.: Assessment of air quality benefits from national air pollution control policies in China. Part II: Evaluation of air quality predictions and air quality benefits assessment, Atmospheric Environment, 44, 3449-3457, http://dx.doi.org/10.1016/j.atmosenv.2010.05.058, 2010.

Wang, Y., Gao, Z., Long, M., Wang, J., and Yu, P. S.: PredRNN++: Towards A Resolution of the Deep-in-Time Dilemma in Spatiotemporal Predictive Learning, 2018.

Xing, J., Zheng, S., Li, S., Huang, L., Wang, X., Kelly, J. T., Wang, S., Liu, C., Jang, C., Zhu, Y., Zhang, J., Bian, J., Liu, T.-Y., and Hao, J.: Mimicking atmospheric photochemical modeling with a deep neural network, Atmospheric Research, 265, 105919, https://doi.org/10.1016/j.atmosres.2021.105919, 2022.

Zhang, X., Liu, L., Chen, X., Gao, Y., Xie, S., and Mi, J.: GLC_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery, Earth System Science Data Discussion, https://doi.org/10.5194/essd-2020-182, 2020.

Zhang, Z., Zhang, S., Chen, C., and Yuan, J.: A systematic survey of air quality prediction based on deep learning, Alexandria Engineering Journal, 93, 128-141, https://doi.org/10.1016/j.aej.2024.03.031, 2024.

Zhao, J., Ma, X., Wu, S., and Sha, T.: Dust emission and transport in Northwest China: WRF-Chem simulation and comparisons with multi-sensor observations, Atmospheric Research, 241, 104978, https://doi.org/10.1016/j.atmosres.2020.104978, 2020.

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., and Liang, J.: UNet++: A Nested U-Net Architecture for Medical Image Segmentation, Cham, 3-11,