npj | clean air

Article

https://doi.org/10.1038/s44407-024-00002-5

A neural operator for forecasting carbon monoxide evolution in cities

Check for updates

Sanchit Bedi[1], Karn Tiwari[2], Prathosh A. P.[2] ✉, Sri Harsha Kota[1] ✉ & N. M. Anoop Krishnan[1,3] ✉

Real-time forecasting of carbon monoxide (CO) concentrations is essential for enabling timely interventions to improve urban air quality. Conventional air quality models often require extensive computational resources for accurate, multi-scale predictions, limiting their practicality for rapid, real-time application. To address this challenge, we introduce the Complex Neural Operator for Air Quality (CoNOAir), a machine learning model that forecast CO concentrations efficiently. CoNOAir demonstrates superior performance over state-of-the-art models, such as the Fourier Neural Operator (FNO), in both short-term (hourly) and extended (72-h) forecasts at a national scale. It excels in capturing extreme pollution events and performs consistently across multiple Indian cities, achieving an R^2 above 0.95 for hourly CO predictions across all evaluated locations. CoNOAir equips authorities with an effective tool for issuing early warnings and designing targeted intervention strategies. This work marks a step forward in achieving dependable, real-time CO pollution predictions for densely populated urban centres.

Indian cities, epicentres of economic prosperity and cultural vibrancy, are poised to accommodate over 675 million residents by 2035[1]. The rapid urbanisation, alongside rampant industrialisation and a burgeoning population, has triggered a notable surge in the concentration of airborne pollutants such as Carbon Monoxide (CO), a criteria pollutant extensively studied historically to gauge urban air quality[2–6]. While CO is recognised as an excellent tracer for atmospheric dynamics[7], it plays a pivotal role in tropospheric ozone chemistry[8] and exerts a profound impact on climate dynamics[9]. Furthermore, CO impacts neurodevelopment in children[10,11], and is associated with cardiovascular diseases[12,13].

Incorporating air quality management into policymaking for the development of sustainable urban centres is crucial[14–16]. Traditionally, data from ground-based analysers is utilised to prioritise air quality management scenarios[15]. However, a low-density monitor network fails to capture spatial variations; for instance, India has 544 ground-based continuous monitoring stations measuring CO[17], suggesting that the current state of air quality monitoring in Indian cities is inadequate[18,19]. This challenge can be addressed by supplementing ground-based observations with air quality models to enhance the spatial coverage. Several studies have compared the performance of air quality models such as Model for Ozone and Related Tracers (MOZART)[20–23] and Weather Research and Forecasting-Chem (WRF-Chem)[24–27] with ground-based observations and satellite retrievals for CO. Despite their widespread usage, the performance of these models ranges from poor to moderate, often resulting in underestimation of pollutant concentrations. Additionally, computations become increasingly expensive due to the involvement of the chemistry module and the enhanced spatial and temporal resolutions required[28]. These computations are performed for every instance with new initial boundary conditions, increasing the time required to make new simulations[29].

Data-driven machine learning (ML) models are gaining attraction as substitutes for these conventional models, given the availability of large amounts of data for training. ML models offer the distinct advantage of rapidly generating inferences on new instances with significantly lower computational demands. Various ML algorithms such as Linear Regression, Artificial Neural Networks, K-nearest neighbour, Support Vector Regression, Long short-term memory, and Convolutional Neural Networks (CNNs) were employed by several authors in the past to simulate CO[30–40]. Several approaches were explored towards modelling pollutant concentrations using ML at much finer resolutions for urban scales[41–45] with similar work applied at coarser resolutions for regional and country scales[46–48]. Such models can facilitate the generation of information regarding pollutant concentrations in the absence of monitoring stations.

However, these ML models learn the function map between the finite-dimensional input and output spaces and need to be retrained for each new case i.e. a separate model for every new city. CNN-based models further incorporate local convolution to account for the nearby grid's contribution, whereas modelling air pollution can be better represented using global convolution as it is a continuous function, being a solution of the advection-

[1]Department of Civil Engineering, Indian Institute of Technology Delhi, New Delhi, India. [2]Department of Electrical Communication Engineering, Indian Institute of Science Bengaluru, Bengaluru, India. [3]Yardi School of Artificial Intelligence, Indian Institute of Technology Delhi, New Delhi, India. ✉e-mail: prathosh@iisc.ac.in; harshakota@iitd.ac.in; krishnan@iitd.ac.in

diffusion equation. These issues can be mitigated by adopting the approach from conventional models—learning the infinite-dimensional operator between input and output function spaces using partial differential equations with global convolution filters. Such a data-driven model would also address the constraints related to mesh dependency and resolution invariance[49]. CO, like any other air pollutant, has its concentration influenced by factors such as emissions (primarily from combustion processes), meteorological conditions, and terrain characteristics. However, CO is relatively inert, with dilution predominating chemistry. This characteristic simplifies CO modelling, making it a preferable choice for investigating the potential of operator learning in pollution modelling in urban cities. Some approaches in this direction include the work on Fourier Neural Operator (FNO)[50], which can learn the map between initial boundary conditions to the time evolution of further steps. However, these models rely on stationary signals and do not perform well for modelling pollutants such as CO, which are inherently non-stationary.

To address the challenges of (a) expensive computations, (b) slow inference speeds, (c) inability of previous ML models to learn the underlying dynamics, and (d) issues with mesh dependency and generalisability of the models, a neural operator architecture, namely Complex Neural Operator for Air Quality (CoNOAir)[51], to model the evolution of CO concentrations over the entire country is developed. CoNOAir employs a complex neural network along with the fractional Fourier transform that can capture the non-stationary signals, as in the case of CO concentrations. This model is trained using four years' hourly data from WRF-Chem simulations that work well for regional-scale air quality modelling. Furthermore, the performance of this model was compared to that of state-of-the-art neural operator FNO. The present framework can serve as a useful tool for generating real-time forecasts of CO concentrations in urban cities of a country in an extremely economical fashion, making it suitable for integration into urban planning.

Results

CO concentration in Indian cities

In this work, we considered several major cities of India by modelling the CO evolution at the country level for a period from 2016 to 2019. Note that the CO evolution for this period was simulated using the WRF-Chem model (v3.9.1), henceforth referred to as ground truth (for details, please refer to the "Methods" section). Figure 1a depicts the study area spanning from 7.060° N to 38.004° N and 67.791° E to 97.934° E with the CO concentrations. Note the timestamp (03rd March 2016 at 04:00 AM IST) corresponding to the maximum cumulative CO concentration is used, representing the extreme air pollution event over the entire study area. We have also identified the grid point having the maximum concentration at this timestamp, which is given by the coordinates 23.316° N, 86.416° E. This grid's proximity to several thermal power plants and industrial areas could be a cause of elevated concentrations. Figure 1b shows the distribution of CO concentrations observed over all the grid points and over all the time stamps for the years 2016, 2017, 2018 and 2019.

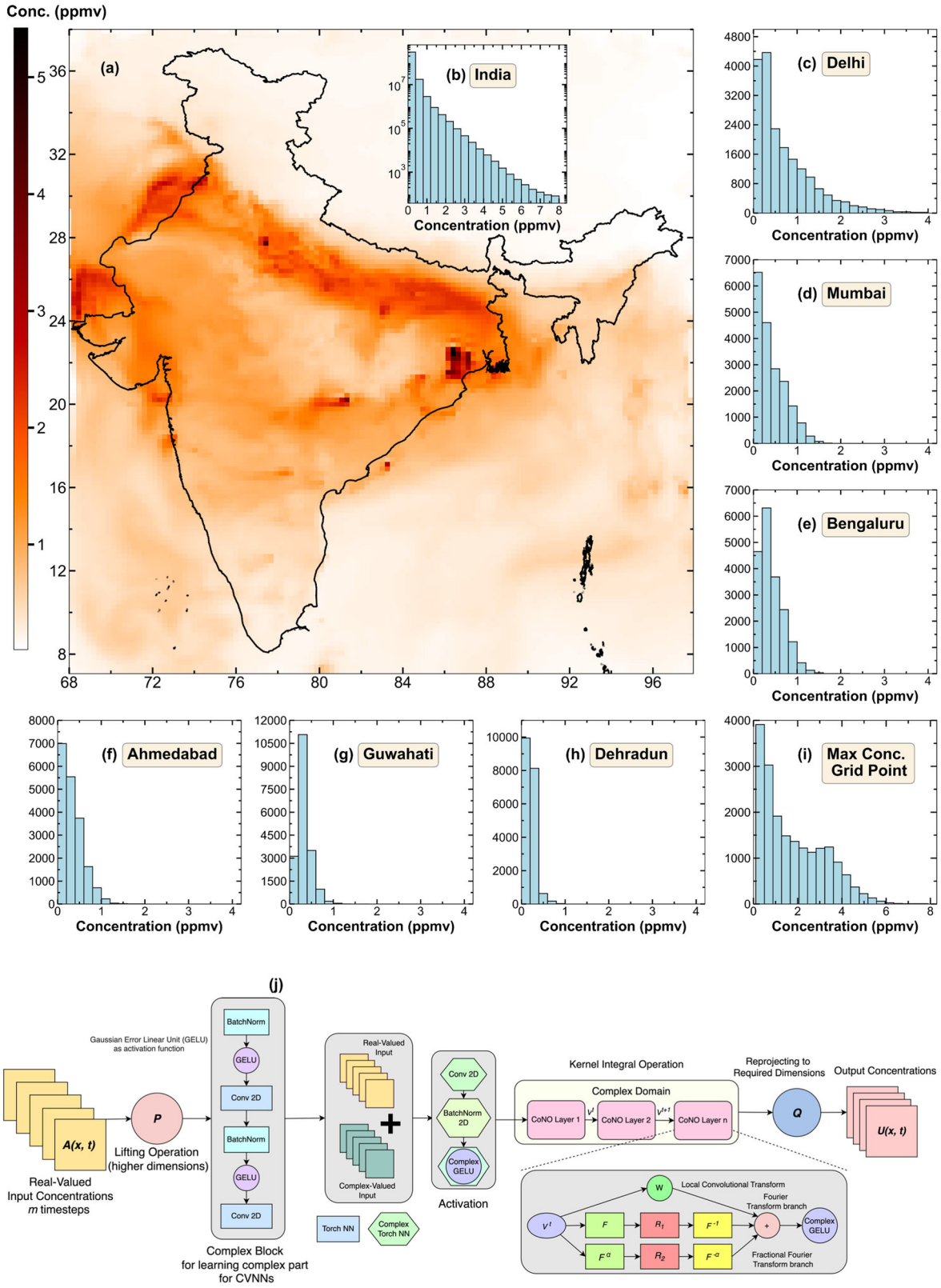
For analysing urban CO concentration, India is divided into six major zones to group the airsheds covering the 131 cities under the National Clean Air Programme (NCAP) namely South, Central, Northeast, Indo-Gangetic Plain (IGP), Northwest, and Himalayan[19]. Although these airsheds have been defined based on Particulate Matter (PM), we use the same zones for CO pollution as well and have identified six major cities in these zones for detailed analysis. The identified cities are designated as non-attainment cities in the NCAP plan developed by the Indian government and are targeted for reducing their $PM_{2.5}$ concentrations. Each city represents one of the six airsheds and is the largest city belonging to that airshed. The cities are namely Delhi (28.602° N, 77.103° E) in IGP, Mumbai (18.975° N, 72.937° E) in Central, Bengaluru (13.089° N, 77.593° E) in South, Ahmedabad (23.091° N, 72.692° E) in Northwest, Guwahati (26.209° N, 91.808° E) in Northeast, and Dehradun (30.309° N, 78.084° E) in Himalayan. Figure 1c–h depicts the distribution of CO concentration over these cities for the entire simulated data. In comparison to the 'good' air quality index thresholds for CO[52], Delhi transgressed the threshold 30.22% of the time, followed by Mumbai (10.5%), followed by Bengaluru (6.7%) and Ahmedabad (3.5%). While Guwahati and Dehradun exceeded the thresholds only 0.71% and 0.05% of the time, respectively, the grid point registering the highest concentration surpassed the limit in 63.86% of instances. Figure 1i depicts the distribution of CO concentrations over the maximum concentration grid point.

Complex Neural Operator for air quality

To predict the evolution of CO, we present our neural operator-based framework namely, Complex Neural Operator for Air Quality (CoNOAir). Figure 1j shows the CoNOAir architecture for forecasting CO. The model uses the CO concentration from the previous k timesteps to forecast the concentrations in the next timestep, which is then used in an autoregressive fashion to forecast the CO concentration in further timesteps. Here, we consider k as 10 timesteps. The model consists of three primary operations, namely, lifting, kernel integration, and projection. First, the model projects input to a high-dimensional latent space using a fully dense pointwise neural network. Then, the model employs a series of affine operations along with kernel integration, that are passed through a non-linear activation function. Finally, the output is obtained by projecting the final embedding in the latent space through a dense pointwise neural network. In contrast to classical neural operators such as FNO, CoNOAir employs a fractional Fourier transform for the kernel integration along with complex neural networks for the affine transformations. As demonstrated later, these features are particularly effective for capturing non-stationary signals such as the CO evolution. The models are trained by minimising the Relative L^2 norm (RL^2 Norm) loss of the predicted CO concentration over n future timesteps. Here, we present three models with n values of 4, 6, and 16. For comparison, we also train FNO models under the same setting. FNO outperforms traditional DL methods for learning PDEs, hence it was considered as the baseline for CoNOAir. Note that the models are trained on the data from 2016 to 2018 and the data from 2019 is kept as the test set for evaluating the model. Further details of the models and the hyperparameters are provided in the "Methods" section.

Forecasting CO concentrations

First, we evaluate the performance of CoNOAir in comparison with FNO for forecasting CO concentrations. Specifically, we consider three models trained on 4, 6, and 16 forward timesteps as outlined in the "Methods" section (refer to Supplementary Fig. 1a for the train and 1b for test loss curves). To evaluate the performance of these models, we compute the error metrics for three different cases: 16, 20, and 72 h forecasting. Table 1 shows the performance of CoNOAir and FNO models, trained on 4, 6, and 16 forward timesteps, for 16, 20, and 72 h forecasting. CoNOAir outperforms FNO in all the cases for a similar setting. Interestingly, while the CoNOAir(4) performs best for the 1-h forecast (Supplementary Table 1), CoNOAir(16) performs best for the longer-term forecast of 16 h. This is due to the fact that CoNOAir(16), being trained on 16 forward steps, exhibits better longer-term stability in predictions, whereas CoNOAir(4) exhibits better short-term predictions with the errors accumulating for the long-term predictions. Similar performance is observed for 20th and 72nd-h predictions, with CoNOAir(16) giving the best results in both cases. More detailed discussions on these, along with the plots, are included in Supplementary Note 2. Altogether, it is evident that a given variation of the CoNOAir model performs better than the corresponding variation of the FNO model (Supplementary Fig. 2). Moreover, while CoNOAir(16) gives the best performance for forecasting 16, 20, and 72 h, CoNOAir(4) gives the best performance for the next few hours prediction. Thus, depending on the fidelity and timescale of interest, the respective CoNOAir models could be used for forecasting the CO concentrations. Since, CoNOAir consistently outperforms FNO both in short-term and long-term forecasts, henceforth, our discussion is focused only on CoNOAir with the rest of the figures being included in Supplementary Material.



Predicting extreme pollution event

Predicting extreme events accurately is important for developing early warning systems in cities. Moreover, since ML models are statistical in nature and most of the training data points are generally closer to the mean, these models may not perform well on extrema. Figure 2 shows the ability of CoNOAir(4) to predict the three worst-case scenarios, namely, 12th February 2019, 07:00 IST, 6th February 2019, 19:00 IST and the midnight hour on 6th February 2019 IST. These three timestamps had the highest summed-up concentrations over the entire test data for the corresponding hour of the day, and represent morning rush hour, evening rush hour, and night-time radiation inversion (which restricts the potential dispersion of air pollutants), respectively. Figure 2a–c shows the ground truth, CoNOAir

Fig. 1 | Data visualisation and model architecture. a Indian subcontinent for the timestamp of 03rd March 2016 at 04:00 AM IST having the maximum cumulative CO concentration. **b** Histogram of CO concentration over the entire subcontinent for the entire study period considered from 2016 to 2019. Histograms of CO concentration for the same timestamp in six major cities, namely, **c** Delhi, **d** Mumbai, **e** Bengaluru, **f** Ahmedabad, **g** Guwahati, and **h** Dehradun, **i** Histogram of CO concentration corresponding to the maximum concentration grid point (unit: ppmv). **j** Proposed CoNOAir architecture: In CoNOAir, real-valued time-series data as a 3D array of shapes (x, y, time) for CO undergoes the lifting operation, projecting the input to higher dimensions. This goes through a complex block consisting of three operations namely batch normalisation, GELU activation and convolution layer two times. The model generates a complex-valued representation of the signal, which passes an activation block to introduce non-linearity in the complex domain. The input passes four CoNOAir layers for kernel integration operation, undergoing three transformations namely local linear transform W, Fourier transform, and Fractional Fourier transform for global convolution operation. These are combined to pass through complex GELU activation. Finally, the outputs are projected back to the required dimensions and result in CO concentration for the next hours.

Table 1 | RMSE, MAE, RL^2 Norm Error for the evaluation of different datasets by the 6 models trained in this study

Error → Model ↓	FNO(4)	FNO(6)	FNO(16)	CoNOAir(4)	CoNOAir(6)	CoNOAir(16)
Validation Data (Errors for 16th-h forecasts)						
RMSE	0.064 (0.041)	0.056 (0.034)	0.043 (0.022)	0.058 (0.036)	0.047 (0.028)	0.031 (0.017)
MAE	0.034 (0.020)	0.030 (0.016)	0.024 (0.012)	0.029 (0.017)	0.025 (0.014)	0.018 (0.009)
RL^2 Norm	0.284 (0.099)	0.258 (0.088)	0.195 (0.057)	0.247 (0.083)	0.205 (0.067)	0.138 (0.042)
Test Data (Errors for 20th-h forecasts)						
RMSE	0.084 (0.046)	0.078 (0.044)	0.073 (0.031)	0.084 (0.039)	0.074 (0.031)	0.072 (0.030)
MAE	0.041 (0.017)	0.036 (0.014)	0.034 (0.009)	0.038 (0.011)	0.034 (0.009)	0.032 (0.009)
RL^2 Norm	0.382 (0.146)	0.371 (0.169)	0.343 (0.135)	0.377 (0.142)	0.339 (0.122)	0.334 (0.125)
Elevated Pollution Days Subset from Test Data (Errors for 72nd-h forecasts)						
RMSE	0.145 (0.046)	0.104 (0.027)	0.090 (0.027)	0.115 (0.030)	0.111 (0.032)	0.082 (0.024)
MAE	0.089 (0.020)	0.063 (0.012)	0.053 (0.013)	0.066 (0.012)	0.064 (0.013)	0.047 (0.010)
RL^2 Norm	0.468 (0.073)	0.415 (0.086)	0.330 (0.055)	0.419 (0.100)	0.395 (0.150)	0.300 (0.063)

The values in each cell depict the mean (standard deviation) obtained over all the samples over which the errors are calculated. The first part of the table represents the errors over the validation data, followed by the errors over the test data. The third part of the table shows the high concentration days from test data over which the long-term evaluation (72 h) is performed (units: RMSE, MAE in ppmv, RL^2: unitless).

output and absolute error between these two for 12th February 2019, 07:00 h IST. Similarly, Fig. 2d–i panels represent the same quantities for 6th February 2019, 19:00 h IST and 6th February 2019, 00:00 h IST, respectively. We observe that CoNOAir is able to capture the hotspots (grids of high concentrations) and the overall dispersion pattern of CO over the study area very well for all three scenarios. For both the second and third cases, the maximum absolute error is around 10%, while for the first scenario, it is around 17%. The efficiency of the model to learn the trend and identify the hotspots make them suitable for use in policy formulations. Supplementary Fig. 3 depicts that the FNO(4) configuration has more grids having high absolute errors, evident by a larger spread of dark regions along with a greater maximum absolute error than CoNOAir(4).

Short-term forecasting for six cities

Figure 3 shows the performance of CoNOAir(4) over the six cities listed earlier and the maximum concentration grid point for short-term forecasts. Specifically, we focus on the next hour forecasting capability for the month of February 2019 (from the test set) for these locations. Interestingly, we observe that CoNOAir(4) captures the fluctuations, including the spurious ones with sudden increases, in CO concentration both quantitatively and qualitatively. To demonstrate this further, we analyse the parity plot of predicted with respect to ground truth CO concentrations for these locations (see Fig. 3). We observe that the R^2 values for all the locations are more than or equal to 0.95 suggesting a high accuracy in the prediction. From Supplementary Fig. 4, we can infer that the R^2 values for the seven grids for FNO(4) are less than or equal to CoNOAir(4). Supplementary Tables 2 and 3 suggest that CoNOAir(4) has the best performance both in terms of RMSE and MAE for every city and the maximum concentration grid point. The RMSE and MAE, even for the CoNOAir(16), are comparable to FNO(4). These results reinforce the superior performance of CoNOAir both in terms of short-term and extreme-value forecasting. Performance evaluation of WRF-Chem simulations with ground-based observation stations for these cities for the year 2019 can be found in the "Methods" section and Supplementary Material (Supplementary Table 4 and Supplementary Fig. 5).

Long-term forecasting

Long-term forecasting is an important aspect of pollution for strategic planning of the cities and policymaking. We now analyse the capability of CoNOAir to perform a reliable long-term forecast of CO concentration. Figure 4 shows one instance of 72 h forecast using CoNOAir(16) starting 10th February 2019 at 12:00 IST in an autoregressive fashion (see Supplementary Materials for the results of the remaining models, Supplementary Figs. 6–10). The ground truth and the corresponding prediction using CoNOAir for the 1st, 12th, 24th, 48th, 60th and 72nd-h forecasts are shown for comparison (see Fig. 4). We observe that CoNOAir can provide reasonable forecasts both qualitatively and quantitatively for the 72nd hour as well. Specifically, CoNOAir is able to reproduce the overall patterns of dispersion and the distribution of the CO concentrations over the country with sufficient accuracy. On a closer examination, CoNOAir appears to be underestimating the concentrations over the country on longer timescales. The selection of this timestamp, i.e., noon hour when the potential of the dispersion of pollution is high, also explains the ability of our proposed model to build over relatively lower concentrations autoregressively to determine the future concentrations. The evolution of error metrics over all samples is given in Supplementary Fig. 2(g–i).

Finally, we analyse the long-term forecast of CoNOAir(16) for the selected cities and the maximum concentration location for 72 h for the same instance as above (see Fig. 5). CoNOAir(16) performs reasonably well for all the cities except Dehradun where the concentrations of CO are low. This shows the model's competence to efficiently replicate a greater range of concentration values autoregressively (Supplementary Table 5 depicts the RMSE for forecasting 72 h over these cities). It is impressive that the model's

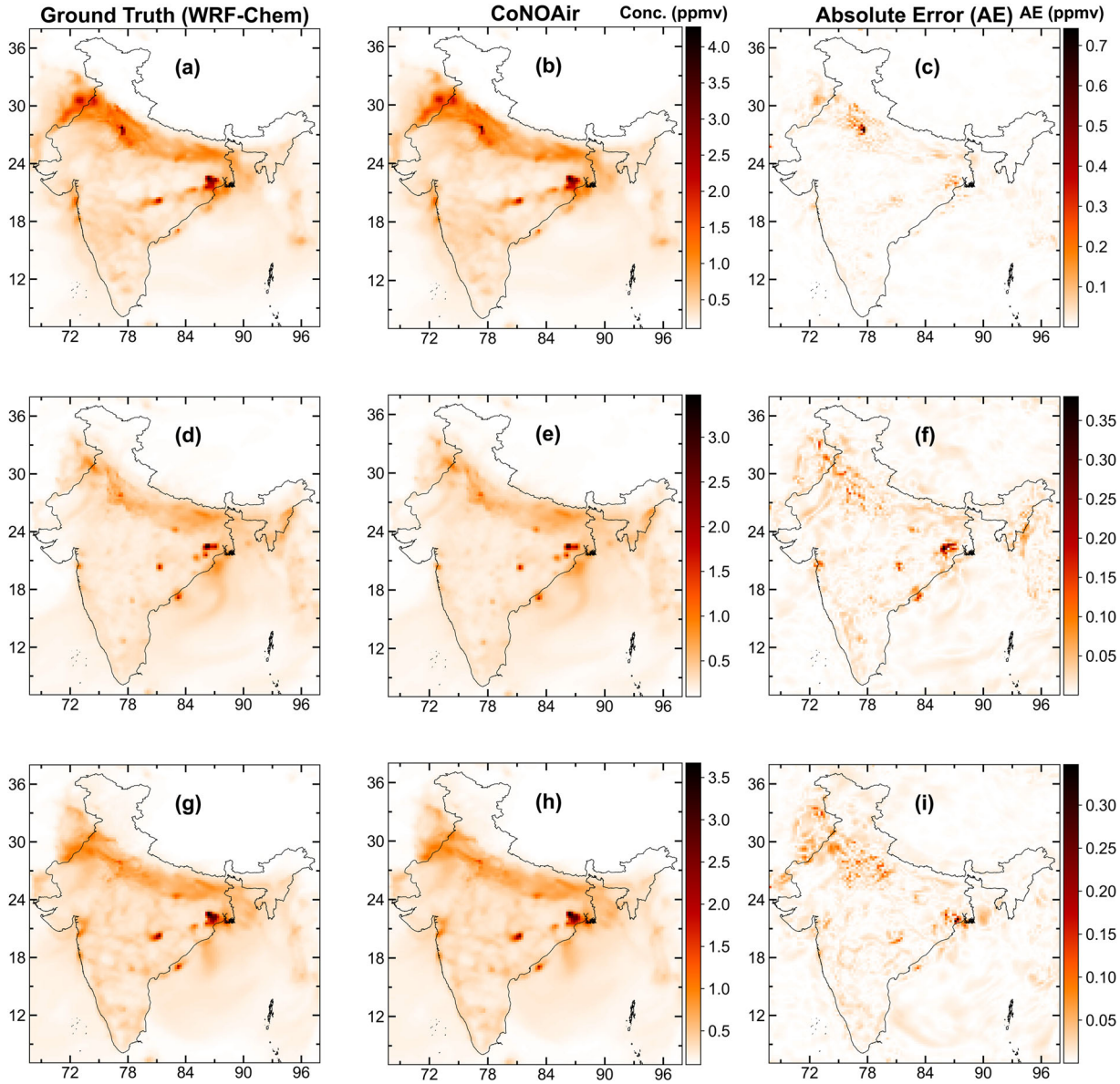


Fig. 2 | Prediction of CoNOAir(4) on the maximum concentration hours. Three maximum concentration hours are identified to assess the performance of CoNOAir. Both the ground truth and CoNOAir maps are normalised to the same range for the purpose of visualisation. Ground truth for 12th February 2019 07:00 IST (**a**), 6th February 19:00 IST (**d**) and 6th February 00:00 IST (**g**). CO concentrations predicted using CoNOAir(4) for the corresponding hours (**b**, **e** and **h**). Absolute error between the ground truth and CoNOAir(4) predictions (**c**, **f** and **i**) (unit: ppmv).

performance over the maximum concentration grid point is reasonably accurate even for a 72nd-h forecast. This explains the superiority of the CoNOAir(16) model in forecasting CO levels at grids with high concentrations, that too autoregressively, making it a reliable model to be used for developing warning systems for predicting future episodes and preparing for them.

These rollout predictions can be made for timesteps beyond 72 h as well for every model. Supplementary Figs 11–16 show this. However, our previous experience suggests that errors accumulate for subsequent rollout predictions and 72 h becomes a longer period for the accumulation of errors[53]. Further, previous studies[54] have highlighted the challenges in longer-term forecasting (24 h), [55] predicted 48-h in advance for long-term hourly forecasting, while Chiang et al. [56] have made predictions for 72 future timesteps. The latest study by Zhang et al. [57] also developed ML models for hourly 3-days' forecasts. Therefore, we have limited our analysis to 72 h rollout forecasts as long-term for relative comparison to the 1-h predictions, which is the shortest timescale for making predictions.

Discussion

Air quality modelling using ML algorithms can adeptly address challenges inherent in conventional models. This study advances ML models to simulate CO concentrations across India through operator learning. The operator learning approach for PDE inferencing offers the advantage of learning continuous functions using global convolution with Fourier transformation, whereas traditional DL methods employ local convolutions that are better suited to learning edges and shapes. DL methods are further limited by the discretization of data as they learn the mapping between finite-dimensional spaces. Neural operators, due to the usage of global convolution can produce resolution-invariant solutions for operators, whose parameters can be reused for different mesh sizes. Neural-FEM methods model one instance of a known PDE using a neural network, however, Neural Operator learns the solution operator without any information about the PDEs[49]. Moreover, CoNOAir performed better for out-of-distribution data than FNO models proving that operator learning is a better approach to handle the problem of pollution prediction[51].

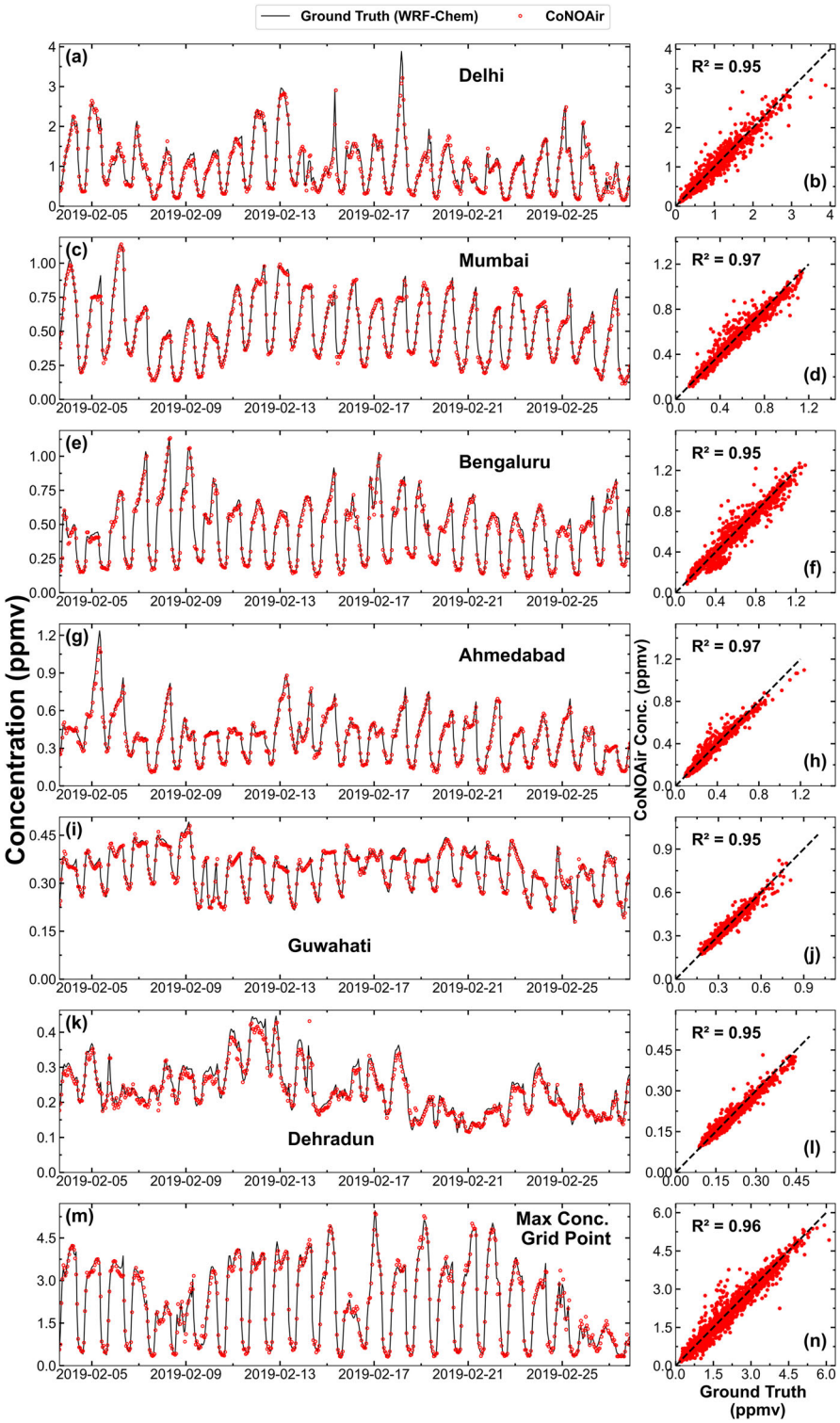


Fig. 3 | Short-term forecast for Indian cities. Time series 1-h forecast for CoNOAir(4) for 6 cities and maximum concentration grid point to evaluate the ability of CoNOAir for grid-level short-term forecasts. A time-series comparison is shown for February month in the left column **a** Delhi, **c** Mumbai, **e** Bengaluru, **g** Ahmedabad, **i** Guwahati, **k** Dehradun, **m** Max conc. grid point. The column on the right shows the predicted values with respect to the ground truth for the whole test dataset. The R^2 values corresponding to each grid point are shown in the inset **b** Delhi, **d** Mumbai, **f** Bengaluru, **h** Ahmedabad, **j** Guwahati, **l** Dehradun, **n** Max conc. grid point. A solid black line represents WRF-Chem simulated CO concentrations, and red markers represent CoNOAir(4)'s output for the corresponding timestamp (unit: ppmv).

We trained the existing FNO model and introduced an enhanced architecture, CoNOAir, utilising the fractional Fourier transform to better represent non-stationary signals and optimise parameters in complex space for learning the time evolution of the advection-diffusion equation. Training the FNO and CoNOAir models ranged from 4 to 80 h on a single NVIDIA V100 GPU with 32 GB memory (see Supplementary Note 3 for more information). Further, generating inferences with these ML models took mere seconds for 72-h forecasts, highlighting their potential to address computational time-related constraints of conventional models.

We evaluated the performances of three CoNOAir and FNO configurations based on short-term (1 h) and long-term (3 days) forecasting. CoNOAir configurations consistently outperformed their FNO counterparts, indicating that FNO would require more parameters to achieve a similar level of performance. Detailed performance analysis over six major cities and the location with the highest CO concentrations revealed that CoNOAir(4) excelled in one-hour forecasts, while CoNOAir(16) outperformed or matched FNO(16) across all grids for forecasting 72 h in an autoregressive fashion. CoNOAir has several architectural features that

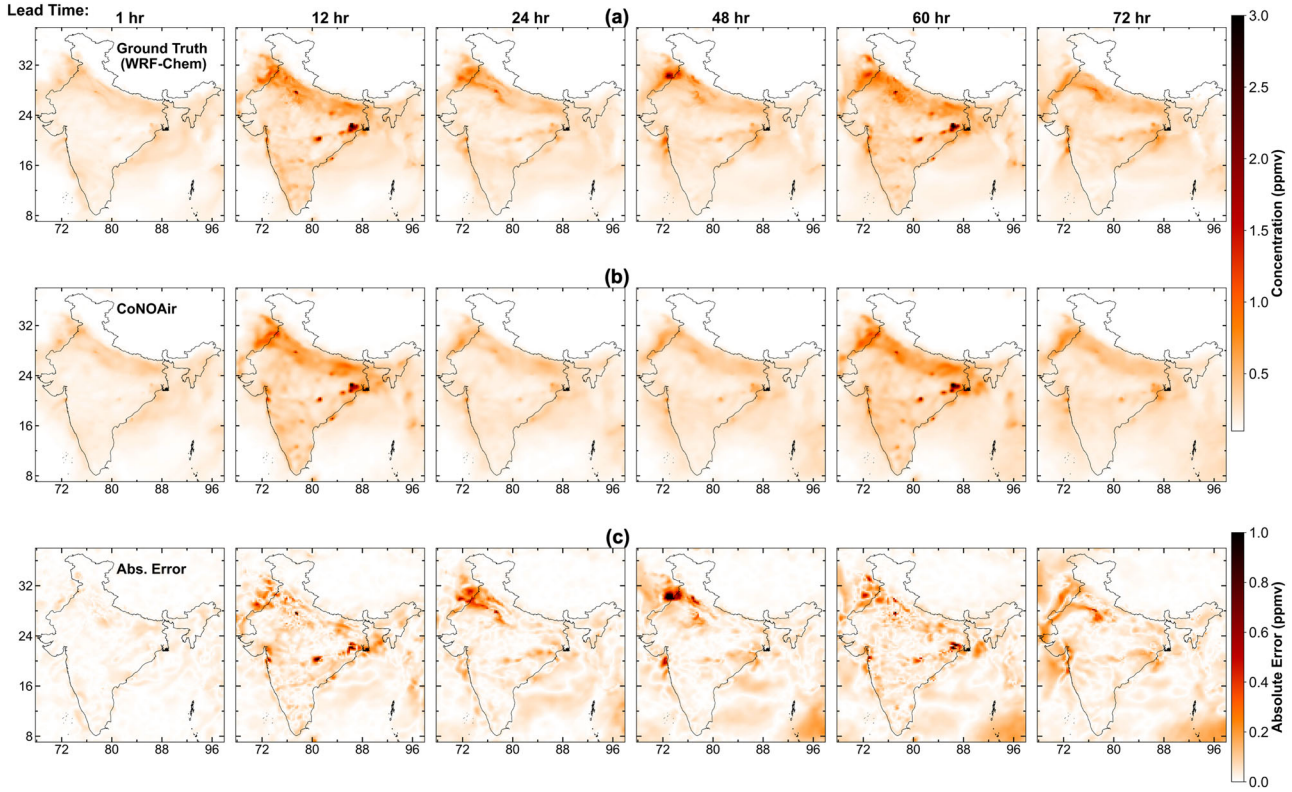


Fig. 4 | Country-wide evolution of long-term forecast. 72 h autoregressive forecasts for one instance for the whole country for evaluating the long-term forecasting capability of the model. The top row **a** represents the ground truth, while the middle row **b** shows the CoNOAir(16)'s predictions. All the maps are normalised between 0 and 3 ppmv for the purpose of easy comparison. The absolute error between these two sets of plots is also shown in the bottom row (**c**) which is normalised between 0 to 1 ppmv (unit: ppmv).

provide superior performance in comparison to FNO. One key feature of CoNOAir is the use of fractional FT (FrFT) in lieu of FFT, which provides both time and frequency representation. In contrast, FFT uses just frequency and hence, may perform poorly on non-stationary signals. FrFT can capture non-stationary signals (changes in signal characteristics with time) better than FFT and identify the instantaneous change in frequency better than FFT, a typical feature exhibited by our datasets. Additionally, CoNOAir utilises Complex-Valued Neural Networks to further improve the representation of input signals in terms of both magnitudes and phases, providing greater information for the model to learn. FFT has limited spatial granularity of the order ($n\log(n)$) on a uniform grid and struggles to capture local variations. FrFT has a learnable parameter alpha to transform the input and preserve the information of the time variation of frequency in the input signal to capture local variations effectively.

The super-resolution capability of the operators allows predictions at multiple resolutions for the micromanagement of hotspots and cities, facilitating the development of advanced intervention strategies. Future directions include integrating physics and chemistry-based information for precise CO and other pollutant predictions and developing real-time early warning systems. This framework could potentially be scaled globally to forecast CO concentrations for any city worldwide, offering enhanced management scenarios and improving living standards.

It is imperative to discuss the limitations and future direction of the current work. The current methodology incorporates only past concentration over India as input to output future concentration, whereas the concentrations are a function of emissions, meteorology, terrain, and background concentrations for instance. This limits the ability of CoNOAir to forecast when these inputs face sudden variations. The inclusion of these input parameters would further constrain CoNOAir to learn the complex relationships between them and make it robust. Uncertainties in emission inventory and meteorological simulations during WRF-Chem simulations generate uncertainties in the training data. Including ground-based observation data from monitoring stations during model training can reduce uncertainties which is our future direction for this work.

Methods

Data generation and preprocessing

In this study, we used the Weather Research and Forecasting coupled with the chemistry module (WRF-Chem) model (v3.9.1) for simulating air pollutant concentrations from 2016 to 2019 across India. More details about the configuration and validation of WRF-Chem are available elsewhere and are only briefly discussed here[25,58–62]. Simulations were performed at a spatial resolution of 25 × 25 km^2 comprising 140 × 124 grids covering India. Model for Ozone and Related Chemical tracers' version 4 (MOZART-4) along with gas-phase chemistry coupled with Model for Simulating Aerosol Interactions and Chemistry (MOSAIC-4 bins) was employed. The boundary conditions for the chemical fields were taken from the Community Atmosphere Model with Chemistry (CAM-Chem) Global Chemical Transport Model at 6 h temporal resolution. The Emissions Database for Global Atmospheric Research (EDGARv5.0) was used for the anthropogenic emissions of various pollutants at a resolution of 0.1°. These emissions from 26 sectors were combined into 5 major sectors and scaling factors were applied on the base year 2015 for every sector for the subsequent years. The WRF-Chem output, generated in the form of a NetCDF file, was extracted as NumPy arrays and saved into *.mat file format for training the models. The ground-level CO concentrations (available at multiple elevation levels from WRF-Chem) as 2D arrays for the entire country with a temporal resolution of 1 h (considered as ground truth in this study) and the corresponding timestamp were extracted. Subsequently, for the purpose of

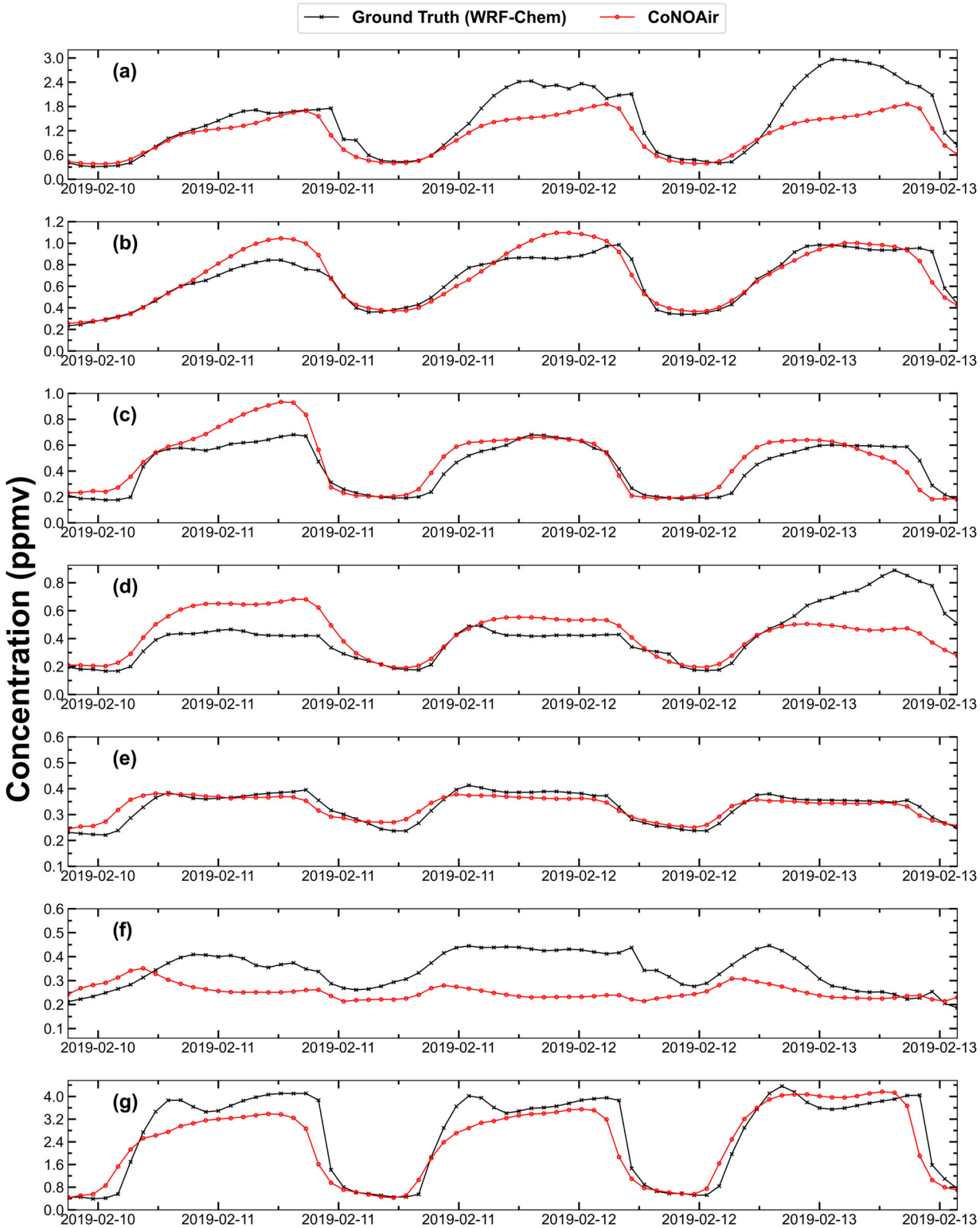


Fig. 5 | Grid-level evolution of long-term forecast. 72 h forecasts for the same instance as Fig. 4 using CoNOAir(16) for 6 cities **a** Delhi, **b** Mumbai, **c** Bengaluru, **d** Ahmedabad, **e** Guwahati, **f** Dehradun and **g** Maximum concentration grid point for evaluating the long-term ability of the model to generate grid-level forecasts. Red line represents our model CoNOAir's outputs, black line represents WRF-Chem simulated CO concentrations (unit: ppmv).

training and validation, respective datasets were prepared using the concentrations simulated for the years 2016, 2017, and 2018. To generate the training and validation set, we used the moving window approach to generate smaller time series each of 26 h from the ground-truth data. All the samples were then shuffled randomly to obtain 3000 samples for training and 500 samples were obtained for testing the model's performance during the training as validation set. The simulated data for the year 2019 was kept separately as a holdout (test) set to assess our models' performances. The training, validation and test datasets were then normalised to the [0, 1] range using the absolute minimum and maximum values from the training

dataset, considering both temporal and spatial dimensions to accelerate the training process.

Model architecture

A neural operator, CoNOAir is developed in this study to model CO concentrations and the performances are evaluated against the ground truth. Further, we compare the performance of CoNOAir with one of the most widely used neural operators, namely, FNO, which gives state-of-the-art performance in several problems related to the time evolution of partial differential equations[50,63,64]. The problem formulation of the Neural Operator and its theory is presented subsequently.

Let $d \in \mathbb{N}, \Omega \in \mathbb{R}^d$ denotes bounded, open set with $A = A\left(\Omega; \mathbb{R}^{d_a}\right)$ and $U = U\left(\Omega; \mathbb{R}^{d_u}\right)$ denoting separable Banach spaces of functions, representing elements in $\mathbb{R}^{d_a}$ and $\mathbb{R}^{d_u}$, respectively. Suppose $G^\dagger : A \to U$ represents the underlying non-linear surrogate mapping derived from a solution operator of parametric PDE. It is assumed that independent and identically distributed (i.i.d.) observations $(a_j, u_j)_{j=1}^N$ are available, where $a_j \sim \mu$ are sampled from the underlying probability measure μ supported on A, and $u_j = G^\dagger(a_j)$.

The aim of operator learning is to approximate $G^\dagger$ using a parametric mapping $G : A \times \Theta \to U$ or equivalently, $G_\theta : A \to U$, $\theta \to \Theta$, belongs to a finite-dimensional parameter space Θ. The objective is to identify $\theta^\dagger \in \Theta$ such that $G(., \theta^\dagger) = G_\theta^\dagger \approx G^\dagger$. It enables learning in infinite-dimensional spaces by solving the optimisation problem in Eq. 1 constructed with a loss function $\mathcal{L} : U \times U \to \mathbb{R}$.

$$\min_{\theta \in \Theta} E_{a \sim u}[\mathcal{L}\left(G(a, \theta),\ G^\dagger(a)\right)] \tag{1}$$

In neural operator learning, G_θ is parameterised using deep neural networks. Equation 1 is optimised using a data-driven approximation of the loss function. G_θ consists of three main components as follows:

a. Lifting Operation ($P : \mathbb{R}^{d_a} \to \mathbb{R}^{v_0}$): projects input to high-dimensional latent space using a fully dense pointwise neural network along the channel dimension.
b. Kernel Integration Operation (K): represented iterative kernel operation that maps through a series of linear operators and non-linear activation. It can be represented as follows:

$$V_{l+1} = \sigma(W v_l + b_l + \mathrm{K}(a;\phi)v_l) \tag{2}$$

where σ represents a non-linear activation function and ϕ denotes the kernel parameterised using neural network and W represents the bias term.

c. Projection Operation ($Q: \mathbb{R}^{v_t} \to \mathbb{R}^{d_u}$): projects high-dimensional latent space to output using a fully dense pointwise neural network along the channel dimension,

Overall, G_θ can be represented as follows:

$$G_\theta = Q \circ K_l \circ \ldots \circ K_1 \circ P \tag{3}$$

where $\circ$ denotes the composition operation.

Kernel Integral Operator: For the iterative updates in neural operators, the linear transformation is represented by a nonlocal kernel integral operator. It can be characterised using a kernel function κ:

$$\mathrm{K}(a;\phi)v_l(x) = \int_D \kappa(x, y, a(x), a(y);\phi)v_l(y)\mathrm{d}y \tag{4}$$

This kernel integral operation is formulated as a convolution operation which becomes a multiplication operation in the Fourier Domain. Let $\mathcal{F}$ represent the Fourier transform:

$$\mathrm{K}(a;\phi)v_l(x) = \mathcal{F}^{-1}(\mathcal{F}(\kappa_\phi).\mathcal{F}(\nu_l))(x) \tag{5}$$

Hence, the FNO model parameterises the Fourier transform of this kernel function directly in Fourier space and the outputs are again converted to the spatial domain as follows:

$$\mathrm{K}(a;\phi)v_l(x) = \mathcal{F}^{-1}(R_\phi.\mathcal{F}(\nu_l))(x) \tag{6}$$

In CoNOAir, the kernel integral is formulated using the Fractional Fourier Transform (FrFT) with a learnable order. The kernel integral is carried out within the complex-valued domain using Complex-Valued Neural Networks (CVNNs). Unlike, the FNO employs the Fourier Transform to handle the integral kernel with real-valued inputs and outputs. In CoNOAir, the integral kernel can be mathematically formulated as:

$$\mathrm{K}(a;\phi)v_l(x) = \mathcal{F}^{-\alpha}(R_\phi.\mathcal{F}^{\alpha}(\nu_l))(x) \tag{7}$$

where $\mathcal{F}^\alpha$ denotes the order α fractional Fourier transform and $\mathcal{F}^{-\alpha}$ denotes inverse fractional Fourier transform. The kernel is parameterised as a learnable weight $R_\phi \in \theta$. In CoNOAir, α also becomes a learnable parameter.

Model training and hyperparameters

In this study, we used the variation of FNO, which performs convolution in the 2D spatial domain only and the input is fed into the model as a time series, just like a Recurrent Neural Network. A similar architecture for CoNOAir was used to compare their ability to make forecasts. The number of Fourier modes used was 12, with 20 hidden channels for each transformation unit. All the variations of these models were run for 500 epochs with an initial learning rate of 0.001 which was halved after every 100 epochs. RL^2 Norm, as given in Eq. 8, was used as the loss function with Adam as the optimiser.

$$\mathrm{RL}^2\ \mathrm{Norm} = \frac{\sqrt{\sum_{u=1}^{n \times m}\left(Y_u - \widehat{Y_u}\right)^2}}{\sqrt{\sum_{u=1}^{n \times m}\left(Y_u\right)^2}} \tag{8}$$

Y_u is Ground Truth (WRF-Chem) and $\hat{Y}_u$ is Predicted Value (FNO or CoNOAir). $n \times m$ represents the total number of grids after converting it into a 1D vector.

For our ablation studies, we used 8, 10, 12 and 16 previous timesteps as inputs to our models, from which we concluded that 10 and 12 configurations worked the best. Finally, we selected the configuration of utilising 10 previous timesteps as inputs to predict one timestep at a time. Subsequently, we can use the forward pass in an autoregressive fashion to output the required number of future steps. To update the models' parameters during the backpropagation of loss, we calculated the cumulative loss over the autoregressive forecasting for 4, 6 and 16 timesteps. The rationale behind this step was to develop models for both short-term and long-term forecasting by accumulating losses over smaller and longer time frames and to compare their abilities to make forecasts on both timescales. We established a uniform nomenclature for the rest of the paper to identify the different configurations of the models trained in this study. The name will be a combination of the type of architecture employed, whether it is CoNOAir or FNO and the number of timesteps (4, 6, or 16) used to propagate the loss backwards. For example, FNO(6) will represent the FNO architecture where 6-step forecasts were used to propagate the loss backwards and so on.

Model evaluation metrics

To evaluate our models' performance over short-term forecasting as well as long-term forecasting we have chosen three commonly used metrics namely Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and RL^2 Norm. We calculated the RMSE, MAE and RL^2 Norm between every predicted and corresponding observed image pair over all the samples and compared the mean and standard deviations. To calculate every metric, we converted the 2D arrays into 1D vectors and performed the respective

operations. RMSE between one observed and predicted image was calculated as given in Eq. 9 follows

$$\mathrm{RMSE} = \sqrt{\frac{1}{n \times m} \times \sum_{u=1}^{n \times m} (Y_u - \widehat{Y_u})^2} \quad (9)$$

Similarly, MAE and RL^2 Norm were calculated as shown in Eqs. 10 and 8, respectively.

$$\mathrm{MAE} = \frac{1}{n \times m} \times \sum_{u=1}^{n \times m} |Y_u - \widehat{Y_u}| \quad (10)$$

where RMSE is Root Mean Squared Error, MAE is Mean Absolute Error, Y_u is Ground Truth (WRF-Chem) and $\hat{Y}_u$ is Predicted Value (FNO or CoNOAir). n × m represents the total number of grids after converting it into a 1D vector.

In this study, we have evaluated our models' performance on the test data that was used for the validation during the training process, which comes from a similar distribution as the training data. Subsequently, we tested our models on out-of-sample data which comes from the observed data of the 2019 year. For the test data, we forecasted in an autoregressive fashion for the next 16 h and obtained the mean and standard deviation of the metrics defined above corresponding to every timestep's forecast. Similarly, for the whole out-of-sample data we have made forecasts for the next 20 h. We also took 168 samples spanning over an entire week from the month of February 2019 to evaluate our models' capacity to forecast longer timeframes, i.e., the next 72 h over the entire country whose results are shown in Table 1, and Supplementary Fig. 2. We further analysed the short-term forecasting ability, i.e. the next one hour, of our models on the grids representing the six cities listed previously, along with the maximum concentration grid point. For this evaluation, we have used the R^2 metric, RMSE and MAE to compare our model performance. The complete results are shown in Fig. 3, Supplementary Fig. 4, Supplementary Tables 2 and 3.

We also analysed the comparison between the time-series evolution for 72 h using autoregressive forecasting with observed time series for these 7 grid locations, whose results are shown in Fig. 5 and Supplementary Table 5. Finally, we assessed our models' performance for the worst CO concentration events obtained using cumulative maximum concentration for the three crucial hours (i) morning rush, (ii) evening rush and (iii) midnight inversion illustrated in Fig. 2 and Supplementary Fig. 3.

Validation of WRF-Chem simulations with ground-based observation data

We obtained the AQ data available from the Central Pollution Control Board's monitoring stations corresponding to the grids representing Delhi, Mumbai, Guwahati, Bengaluru and Ahmedabad for 2019, the year for which our models are evaluated. We averaged the data from all the stations wherever multiple stations were available in the grid for comparison. Delhi had 16 stations, Mumbai had 5 stations, Bengaluru had 4 stations and Ahmedabad and Guwahati had 1 station each. There was no station data for Dehradun and max concentration grid point for this year. The metrics for the evaluation of performance for each city (Mean Bias, Root Mean Squared Error) are shown in Supplementary Table 4 and time-series plots are shown in Supplementary Fig. 5.

Data availability

Data used for training and testing of models will be made available upon request. The trained models' parameters are available at https://github.com/M3RG-IITD/CoNO_for_CO.

Code availability

Codes for training and evaluation of the models can be found at https://github.com/M3RG-IITD/CoNO_for_CO.

Received: 7 September 2024; Accepted: 5 December 2024;
Published online: 12 March 2025

References

1. UN-Habitat, World cities report 2022: envisaging the future of cities. UN-Habitat. xxxi, 387 p. : (2022).
2. Jaffe, L. S. Ambient carbon monoxide and its fate in the atmosphere. *J. Air Pollut. Control Assoc.* **18**, 534–540 (1968).
3. Gokhale, S. & Khare, M. A theoretical framework for the episodic-urban air quality management plan (e-UAQMP). *Atmos. Environ.* **41**, 7887–7894 (2007).
4. Aneja, V. P. et al. Measurements and analysis of criteria pollutants in New Delhi, India. *Environ. Int.* **27**, 35–42 (2001).
5. Kwiecień, J. & Szopińska, K. Mapping carbon monoxide pollution of residential areas in a Polish city. *Remote Sens.* **12**, 2885 (2020).
6. Liñán-Abanto, R. N. et al. Temporal variations of black carbon, carbon monoxide, and carbon dioxide in Mexico City: mutual correlations and evaluation of emissions inventories. *Urban Clim.* **37**, 100855 (2021).
7. Pani, S. K. et al. Relationship between long-range transported atmospheric black carbon and carbon monoxide at a high-altitude background station in East Asia. *Atmos. Environ.* **210**, 86–99 (2019).
8. Novelli, P. C., Masarie, K. A., Tans, P. P. & Lang, P. M. Recent changes in atmospheric carbon monoxide. *Science* **263**, 1587–1590 (1994).
9. Holloway, T., Levy, H. & Kasibhatla, P. Global distribution of carbon monoxide. *J. Geophys. Res. Atmospheres* **105**, 12123–12147 (2000).
10. Calderon-Garciduenas, L. et al. Air pollution, cognitive deficits and brain abnormalities: a pilot study with children and dogs. *Brain Cogn.* **68**, 117–127 (2008).
11. Levy, R. J. Carbon monoxide pollution and neurodevelopment: A public health concern. *Neurotoxicol Teratol.* **49**, 31–40 (2015).
12. Liu, C. et al. Ambient carbon monoxide and cardiovascular mortality: a nationwide time-series analysis in 272 cities in China. *Lancet Planet Health* **2**, e12–e18 (2018).
13. Satran, D. et al. Cardiovascular manifestations of moderate to severe carbon monoxide poisoning. *J. Am. Coll. Cardiol.* **45**, 1513–1516 (2005).
14. Kura, B., Verma, S., Ajdari, E. & Iyer, A. Growing public health concerns from poor urban air quality: strategies for sustainable urban living. *Comput. Water Energy Environ. Eng.* **2**, 1–9 (2013).
15. Gulia, S., Shiva Nagendra, S. M., Khare, M. & Khanna, I. Urban air quality management-a review. *Atmos. Pollut. Res.* **6**, 286–304 (2015).
16. Zhang, X. et al. Linking urbanization and air quality together: a review and a perspective on the future sustainable urban development. *J. Clean. Prod.* **346**, 130988 (2022).
17. Central Pollution Control Board. Air Quality Monitoring Dashboard. https://airquality.cpcb.gov.in/ccr/#/caaqm-dashboard-all/caaqm-landing (accessed 10 August 2024).
18. Brauer, M. et al. Examination of monitoring approaches for ambient air pollution: a case study for India. *Atmos. Environ*. https://doi.org/10.1016/j.atmosenv.2019.116940 (2019).
19. Guttikunda, S., Ka, N., Ganguly, T. & Jawahar, P. Plugging the ambient air monitoring gaps in India's national clean air programme (NCAP) airsheds. *Atmos. Environ.* **301**, 119712 (2023).
20. Chandra, N., Venkataramani, S., Lal, S., Sheel, V. & Pozzer, A. Effects of convection and long-range transport on the distribution of carbon monoxide in the troposphere over India. *Atmos. Pollut. Res.* **7**, 775–785 (2016).
21. Yarragunta, Y., Srivastava, S. & Mitra, D. Validation of lower tropospheric carbon monoxide inferred from MOZART model simulation over India. *Atmos. Res.* **184**, 35–47 (2017).
22. Elguindi, N. et al. Current status of the ability of the GEMS/MACC models to reproduce the tropospheric CO vertical distribution as measured by MOZAIC. *Geosci. Model Dev.* **3**, 501–518 (2010).
23. Surendran, D. E. et al. Air quality simulation over South Asia using Hemispheric Transport of Air Pollution version-2 (HTAP-v2) emission

inventory and Model for Ozone and Related chemical Tracers (MOZART-4). *Atmos. Environ.* **122**, 357–372 (2015).
24. I, N., Srivastava, S., Yarragunta, Y., Kumar, R. & Mitra, D. Distribution of surface carbon monoxide over the Indian subcontinent: investigation of source contributions using WRF-Chem. *Atmos. Environ.* https://doi.org/10.1016/j.atmosenv.2020.117838 (2020).
25. Kumar, R. et al. Simulations over South Asia using the Weather Research and Forecasting model with Chemistry (WRF-Chem): chemistry evaluation and initial results. *Geosci. Model Dev.* **5**, 619–648 (2012).
26. Tie, X. et al. Characterizations of chemical oxidants in Mexico City: a regional chemical dynamical model (WRF-Chem) study. *Atmos. Environ.* **41**, 1989–2008 (2007).
27. Benavente, N. R. et al. Air quality simulation with WRF-Chem over southeastern Brazil, part I: model description and evaluation using ground-based and satellite data. *Urban Clim.* **52**, 101703 (2023).
28. Carmichael, G. R. et al. Predicting air quality: improvements through advanced methods to integrate models and measurements. *J. Comput. Phys.* **227**, 3540–3571 (2008).
29. Werhahn, M., Xie, Y., Chu, M. & Thuerey, N. A multi-pass GAN for fluid flow super-resolution. *Proc. ACM Comput. Graph. Interact. Tech.* **2**, 1–21 (2019).
30. Almubaidin, M. A. et al. Machine learning predictions for carbon monoxide levels in urban environments. *Results Eng.* https://doi.org/10.1016/j.rineng.2024.102114 (2024).
31. Wong, P.-Y. et al. Incorporating land-use regression into machine learning algorithms in estimating the spatial-temporal variation of carbon monoxide in Taiwan. *Environ. Model. Softw.* https://doi.org/10.1016/j.envsoft.2021.104996 (2021).
32. Feizi, H., Sattari, M. T., Prasad, R. & Apaydin, H. Comparative analysis of deep and machine learning approaches for daily carbon monoxide pollutant concentration estimation. *Int. J. Environ. Sci. Technol.* **20**, 1753–1768 (2023).
33. Bougoudis, I., Demertzis, K. & Iliadis, L. HISYCOL a hybrid computational intelligence system for combined machine learning: the case of air pollution modeling in Athens. *Neural Comput. Appl.* **27**, 1191–1206 (2015).
34. Suárez Sánchez, A., García Nieto, P. J., Riesgo Fernández, P., del Coz Díaz, J. J. & Iglesias-Rodríguez, F. J. Application of an SVM-based regression model to the air quality study at local scale in the Avilés urban area (Spain). *Math. Comput. Model.* **54**, 1453–1466 (2011).
35. Barthwal, A. & Goel, A. K. Advancing air quality prediction models in urban India: a deep learning approach integrating DCNN and LSTM architectures for AQI time-series classification. *Model. Earth Syst. Environ.* **10**, 2935–2955 (2024).
36. Moazami, S. et al. Reliable prediction of carbon monoxide using developed support vector machine. *Atmos. Pollut. Res.* **7**, 412–418 (2016).
37. Navares, R. & Aznarte, J. L. Predicting air quality with deep learning LSTM: towards comprehensive models. *Ecol. Inform.* https://doi.org/10.1016/j.ecoinf.2019.101019 (2020).
38. Xu, S., Li, W., Zhu, Y. & Xu, A. A novel hybrid model for six main pollutant concentrations forecasting based on improved LSTM neural networks. *Sci. Rep.* **12**, 14434 (2022).
39. Alimissis, A., Philippopoulos, K., Tzanis, C. G. & Deligiorgi, D. Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmos. Environ.* **191**, 205–213 (2018).
40. Li, C. & Zhu, Z. Research and application of a novel hybrid air quality early-warning system: a case study in China. *Sci. Total Environ.* **626**, 1421–1438 (2018).
41. Song, J. & Stettler, M. E. J. A novel multi-pollutant space-time learning network for air pollution inference. *Sci. Total Environ.* **811**, 152254 (2022).
42. Song, J. et al. Toward high-performance map-recovery of air pollution using machine learning. *ACS EST Eng.* **3**, 73–85 (2023).
43. Song, J., Han, K. & Stettler, M. Deep-MAPS: machine learning based mobile air pollution sensing. *IEEE Internet Things J.* https://doi.org/10.1109/JIOT.2020.3041047 (2020).
44. Guo, R. et al. High-resolution urban air quality mapping for multiple pollutants based on dense monitoring data and machine learning. *Int. J. Environ. Res. Public Health* **19**, 8005 (2022).
45. Xu, J., Liu, M., Chao, Y. & Chen, H. A novel prediction framework for estimating high spatial resolution near-ground PM2.5 and O3 concentrations at street-level in urban areas. *Build Environ.* **267**, 112141 (2025).
46. Wang, S. et al. Reconstructing long-term (1980–2022) daily ground particulate matter concentrations in India (LongPMInd). *Earth Syst. Sci. Data* **16**, 3565–3577 (2024).
47. Zhong, J. et al. Reconstructing 6-hourly PM2.5 datasets from 1960 to 2020 in China. *Earth Syst. Sci. Data* **14**, 3197–3211 (2022).
48. Tian, H. et al. Estimating PM2.5 from multisource data: a comparison of different machine learning models in the Pearl River Delta of China. *Urban Clim.* **35**, 100740 (2021).
49. Kovachki, N. et al. Neural operator: learning maps between function spaces with applications to PDEs. *J. Mach. Learn. Res.* **24**, 89 (2023).
50. Li, Z. et al. Fourier neural operator for parametric partial differential equations. https://doi.org/10.48550/arXiv.2010.08895 (2020).
51. Tiwari, K., Krishnan, N. & Prathosh, A.P. CoNO: Complex neural operator for continuous dynamical physical systems. *arXiv preprint arXiv:2406.02597* (2024).
52. CPCB. National Air Quality Index. Central Pollution Control Board (CPCB). http://app.cpcbccr.com/ccr_docs/FINAL-REPORT_AQI_.pdf (2014).
53. Bedi, S., Katiyar, A., Krishnan, N. M. A. & Kota, S. H. Utilizing LSTM models to predict PM2. 5 levels during critical episodes in Delhi, the world as most polluted capital city. *Urban Climate* **53**, 101835 (2024).
54. Zhang, Y. et al. In *Twenty-Eighth International Joint Conference on Artificial Intelligence {IJCAI-19}* 4341–4347 (International Joint Conferences on Artificial Intelligence Organization, Macao, China, 2019).
55. Yi, X. et al. Predicting fine-grained air quality based on deep neural networks. *IEEE Trans. Big Data* **8**, 1326–1339 (2022).
56. Chiang, P. W. & Horng, S. J. Hybrid time-series framework for daily-based PM2.5 forecasting. *IEEE Access* **9**, 104162–104176 (2021).
57. Zhang, Z., Johansson, C., Engardt, M., Stafoggia, M. & Ma, X. Improving 3-day deterministic air pollution forecasts using machine learning algorithms. *Atmos. Chem. Phys.* **24**, 807–851 (2024).
58. Sharma, S., Chandra, M. & Harsha Kota, S. Four year long simulation of carbonaceous aerosols in India: Seasonality, sources and associated health effects. *Environ. Res.* **213**, 113676 (2022).
59. Nayak, D. K., Habib, G. & Kota, S. H. Can Landuse Landcover changes influence the success of India's national clean air plans? *Atmos. Environ X* **22**, 100251 (2024).
60. Gupta, M., Nayak, D. K. & Harsha Kota, S. Impact of particulate matter-centric clean air action plans on ozone concentrations in India. *ACS Earth Space Chem.* **7**, 1038–1048 (2023).
61. Jat, R., Gurjar, B. R. & Lowe, D. Regional pollution loading in winter months over India using high resolution WRF-Chem simulation. *Atmos. Res.* https://doi.org/10.1016/j.atmosres.2020.105326 (2021).
62. Ritter, M., Müller, M. D., Jorba, O., Parlow, E. & Liu, L. J. S. Impact of chemical and meteorological boundary and initial conditions on air quality modeling: WRF-Chem sensitivity evaluation for a European domain. *Meteorol. Atmos. Phys.* **119**, 59–70 (2012).
63. Burark, P., Tiwari, K., Rashid, M. M., Prathosh, A. P. & Krishnan, N. M. A. CoDBench: a critical evaluation of data-driven models for continuous dynamical systems. *Digit. Discov.* **3**, 1172–1181 (2024).
64. Takamoto, M. et al. *Proceedings of the 36th International Conference on Neural Information Processing Systems.* Article 117 (Curran Associates Inc., New Orleans, LA, USA, 2022).

Acknowledgements
We acknowledge the supercomputing facility at the Indian Institute of Technology, Delhi for providing computational resources. The authors thank the Central Pollution Control Board for its publicly available ground-based observation data to validate WRF-Chem simulations. We acknowledge the support of the IHUB-ANUBHUTI-IIITD Foundation set up under the NM-ICPS scheme of the Department of Science and Technology, India. N.M.A.K. acknowledges the funding support from the Alexander von Humboldt Foundation.

Author contributions
S.H.K. and N.M.A.K. conceived the work. S.B. wrote the codes and performed the experiments and K.T. contributed the codes. S.B. analysed the results with support from all the authors. S.B. wrote the first draft of the manuscript, which was reviewed by all the authors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44407-024-00002-5.

Correspondence and requests for materials should be addressed to Prathosh A. P., Sri Harsha Kota or N. M. Anoop Krishnan.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2025