

# Supplementary Materials: Deep Sequence Learning for Prediction of Daily NO<sub>2</sub> Concentration in Coastal Cities of Northern China

Xingbin Jia <sup>1</sup>, Xiang Gong <sup>1,\*</sup>, Xiaohuan Liu <sup>2,3</sup>, Xianzhi Zhao <sup>1</sup>, He Meng <sup>4</sup>, Quanyue Dong <sup>5</sup>, Guangliang Liu <sup>6</sup> and Huiwang Gao <sup>2,3</sup>

<sup>1</sup> School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao, 266061, China

<sup>2</sup> Key Laboratory of Marine Environment and Ecology, Ministry of Education of China, Ocean University of China, Qingdao 266071, China

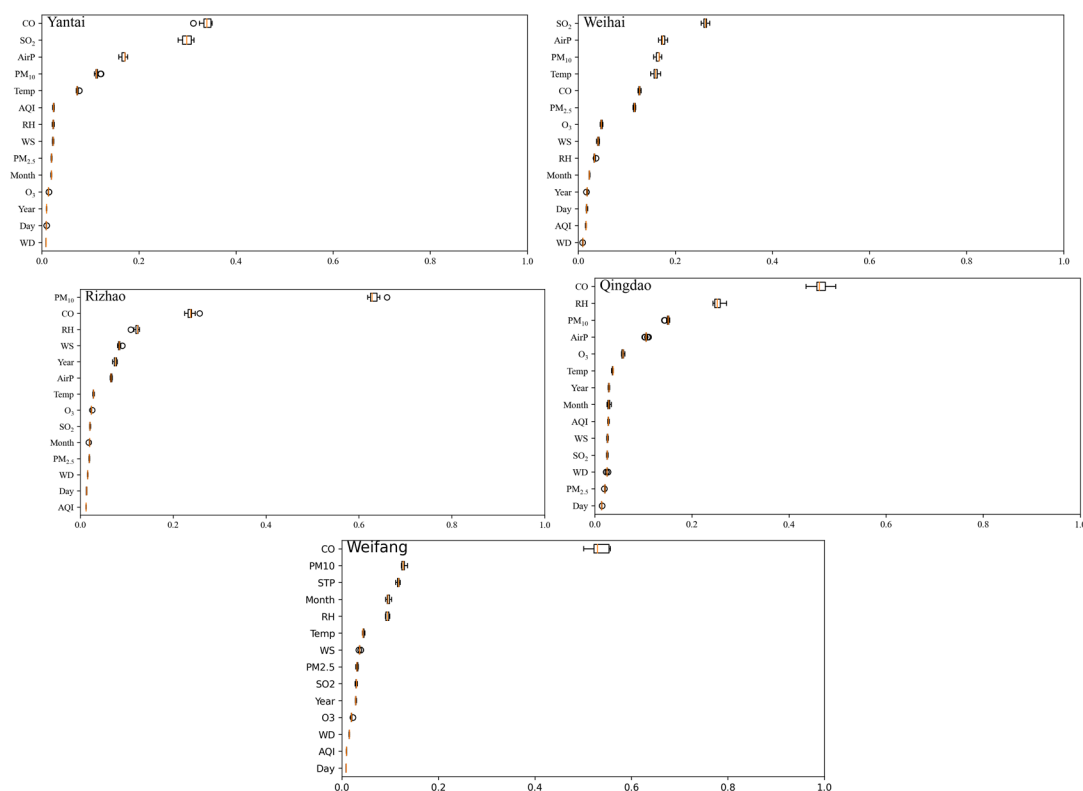
<sup>3</sup> Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China

<sup>4</sup> Qingdao Eco-Environment Monitoring Center of Shandong Province, Qingdao 266003, China

<sup>5</sup> College English Office, Qingdao Vocational and Technical College of Hotel Management, Qingdao 266100, China

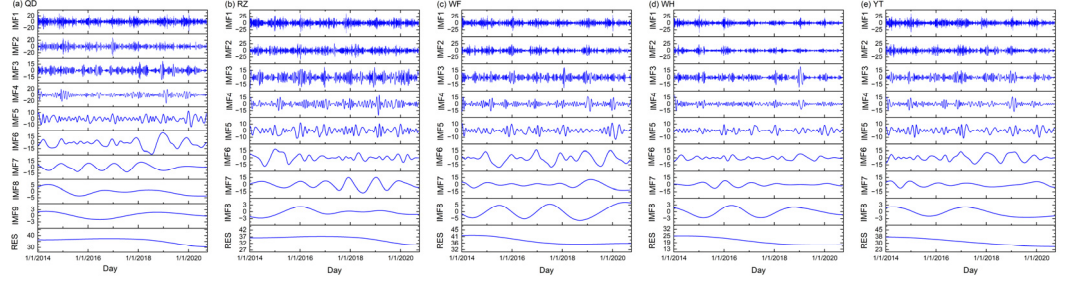
<sup>6</sup> Shandong Provincial Key Laboratory of Computer Networks, Qilu University of Technology (Shandong Academy of Sciences), Jinan 250101, China

\* Correspondence: gongxiang@qust.edu.cn



**Figure S1.** Boxplots of random forest experiments with in a random order of 13 candidate predictors.

**Text S1:** Empirical Mode Decomposition (EMD) is a non-parametric signal processing technique for decomposing a signal into a set of intrinsic oscillatory modes. It is based on the empirical observation that a given signal can be decomposed into a finite number of intrinsic modes which are simple oscillatory components of the signal. Each of these intrinsic modes is then further decomposed into component sub-modes using a method of sifting. The resulting set of intrinsic modes and sub-modes are then used to reconstruct the original signal. EMD is often used in time-series analysis and has been successfully applied to a variety of problems in signal processing and analysis.



**Figure S2.** Temporal time-series signal of intrinsic mode functions (IMFs) and residuals (RES) from the EMD transformation of NO<sub>2</sub> concentration at 5 coastal cities of the northern China.

To extract time series feature of the output predictor, we employed the EMD method. The EMD provides a set of intrinsic mode functions (IMFs) decomposed from a signal, which allows users to decompose singular values and avoid being trapped in a local optimum (Zhou et al., 2021). Each IMF presents the time scale of data. The cycle ( $T$ ) of data variation in each IMF can be estimated by using Eq. (S1) (Flandrin et al, 2004).

$$T = \frac{L}{N-1} \quad (\text{S1})$$

where  $N$  is the number of intersection points between the IMF and X-axis and  $L$  is the distance between the first intersection point to the last one. The final data that can no longer be decomposed is called residual (RES). In general, residuals represent trends in the original time series data.

After decomposition, all IMFs at each coastal city were used in conjunction with 13 input data as the training, validation and test dataset. We trained and tested the EMD-S2S model with the same training and test set applicable to the RF-S2S model.