EXO-SPACETIMEFORMER: TIME SERIES PREDICTION WITH EXTERNAL FORECAST INTEGRATION

Boris Kraychev
GATE Institute
Sofia University St. Kliment Ohridski
Sofia, Bulgaria
boris.kraychev@gate-ai.eu

S. Ensiye Kiyamousavi
GATE Institute
Sofia University St. Kliment Ohridski
Sofia, Bulgaria
ensiye.kiyamousavi@gate-ai.eu

ABSTRACT

This study introduces the Exo-Spacetimeformer, an innovative adaptation of the Spacetimeformer model, designed to enhance PM2.5 air quality forecasting by incorporating external factors such as weather conditions, traffic intensity, and air-water-soil temperature differences into its decoder input. Through attention heatmap analysis and performance comparison across different time intervals, the Exo-Spacetimeformer demonstrates improved predictive accuracy over the original model, offering a promising tool for more reliable and comprehensive air pollution forecasting.

1 INTRODUCTION

Figure 1: Accumulation of particulate matters in urban environments

In urban areas, air pollution is a serious problem, particularly in cities surrounded by mountains, where geographical conditions can increase pollution levels. This issue becomes more pronounced in colder temperatures. During the winter, there is a higher demand for heating homes, which often leads to the increased use of solid fuels like coal and wood. During combustion, these fuels release significant amounts of PM2.5[1] and other pollutants into the air (Glojek et al., 2022). At the same time, vehicle emissions, especially from older, less environmentally friendly vehicles, contribute to air pollution (Wang et al., 2022a). The combined effect of traffic and domestic heating emissions creates a dense concentration of contaminants. The surrounding mountains near cities act as a natural barrier, trapping these particles and leading to a phenomenon known as temperature inversion (Glojek et al., 2022). This inversion layer prevents the release of particles, causing them to accumulate near the ground level, seriously affecting air quality and posing health risks to the city's inhabitants. Forecasting PM2.5 concentrations in mountain-enclosed cities during the winter is crucial for public health planning, enabling timely warnings and interventions to mitigate the adverse effects of air pollution on residents' health.

In this context, deep learning models have emerged as powerful tools to process large datasets and identify complex patterns. CNNs (Kow et al., 2020), RNNs (Athira et al., 2018), LSTM (Gao & Li, 2021) and transformers (Wang et al., 2022b) are examples of common deep learning-based forecast models. Among these forecast models, recent studies have explored various transformer-based models for PM2.5 prediction. In recent years, transformer-based models have been widely used in air quality predictions. Wavelet-Transformer (Xu et al., 2023) is one of the methods to predict PM2.5 and PM10 in Guilin, China. Spatiotemporal Transformer (Yu et al., 2023) has been used to predict PM2.5 in wildfire-prone areas in Los Angeles, U.S. Combining Transformer with mechanisms for sparse attention and temporal embedding led to the development of Informer (Zhou et al., 2021; Al-qaness et al., 2023; Zhou & Yang, 2022). As mentioned in (Feng et al., 2023), air quality data not only has temporal correlation features but also has spatial correlation features. Consider-

[1] Fine particulate with an aerodynamic diameter of $2.5\mu m$ or smaller

Table 1: Loss value on test dataset for 6, 24, and 48 hours intervals

Model	6h	24h	48h
Exo-Spacetimeformer	**0.1974**	**0.2727**	**0.4504**
Spacetimeformer	0.2104	0.3464	0.5796

ing this, Spatiotemporal Informer (Feng et al., 2023) suggests a new spatiotemporal embedding and spatiotemporal attention to improving forecast accuracy.

2 The Exo-Spacetimeformer

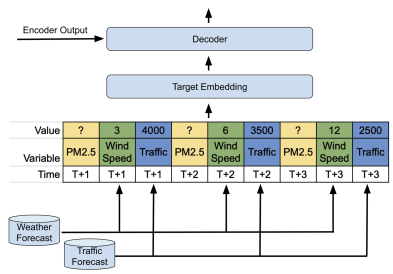

Figure 2: Incorporating exo-forecasts into the decoder

As the introduction explains, the PM2.5 concentration depends on many external factors, like weather, traffic intensity, and air-water-soil temperature differences. Therefore, we cannot effectively predict pollution without having a reliable forecast for these values. To overcome this problem, we propose the Exo-Spacetimeformer: a spacetimeformer-based model incorporating forecast values from external models in its decoder input. We alter the Spacetimeformer model (Grigsby et al., 2023) to include external target values as a known future context, such as wind speed, air temperature, water temperature, and traffic estimation. The values are provided to the decoder input, as shown in figure 2. Refer to Figure 3 from the original paper by (Grigsby et al., 2023) for more information.

3 Experiments

Once we perform a complete training of the Exo-Spacetimeformer, we can visualize the cross-attention of all four heads in a heatmap to discover which input features are the most important for pollution prediction. An example visualization is shown in Figure 3 where, (un)surprisingly, we see that the attention heads highlight the primary mechanisms for pollution accumulation: difference in air and water temperature (Attn h0), wind speed (Attn h1), traffic and humidity (Attn h2), and precipitation and wind gusts (Attn h3). The attention given to the PM 2.5 value is minimal because this is our hidden column that needs to be forecasted.

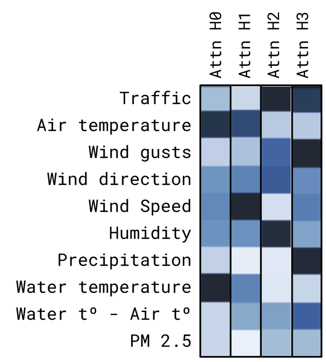

Figure 3: Attention heatmap for PM 2.5 forecasting

As a next step, we compare the performance of our method to the original Spacetimeformer model for predicting the PM 2.5 concentration for various long intervals, namely 6, 24, and 48 hours. Table 3 presents the loss values over the test set after performing identical training iterations and using the means square error for loss calculation. More details about the experiments' reproducibility can be found in the appendix.

4 Conclusion

The proposed Exo-Spacetimeformer model significantly enhances PM2.5 concentration predictions by integrating key external factors into its decoder input. This approach, modifying the original Spacetimeformer model, highlights crucial environmental and anthropogenic influences on air quality. Our findings, particularly through attention heatmap analysis, demonstrate the model's effectiveness in capturing vital variables like temperature differences and traffic. Compared to the original model, the Exo-Spacetimeformer shows superior performance in various forecasting intervals, promising more accurate air quality forecasting.

5 Acknowledgment

This research was supported by the GATE project, funded by the H2020 WIDESPREAD-2018-2020 TEAMING Phase 2 programme, under grant agreement no. 857155, FLEdge projects, funded by the Driving Urban Transitions (DUT) Partnership programme, under agreement no. KP-06-D002/5 and INTEND project, funded by European Union's Horizon 2020 research and innovation programme under grant agreement mo.101135576.

URM Statement

The authors acknowledge that at least all key authors of this work meet the URM criteria of ICLR 2024 Tiny Papers Track.

References

Mohammed AA Al-qaness, Abdelghani Dahou, Ahmed A Ewees, Laith Abualigah, Jianzhu Huai, Mohamed Abd Elaziz, and Ahmed M Helmi. Resinformer: residual transformer-based artificial time-series forecasting model for pm2. 5 concentration in three major chinese cities. *Mathematics*, 11(2):476, 2023.

V Athira, P Geetha, Rab Vinayakumar, and KP Soman. Deepairnet: Applying recurrent networks for air quality prediction. *Procedia computer science*, 132:1394–1403, 2018.

Yang Feng, Ju-Song Kim, Jin-Won Yu, Kuk-Chol Ri, Song-Jun Yun, Il-Nam Han, Zhanfeng Qi, and Xiaoli Wang. Spatiotemporal informer: A new approach based on spatiotemporal embedding and attention for air quality forecasting. *Environmental Pollution*, 336:122402, 2023.

Xi Gao and Weide Li. A graph-based lstm model for pm2. 5 forecasting. *Atmospheric Pollution Research*, 12(9):101150, 2021.

Kristina Glojek, Griša Močnik, Honey Dawn C Alas, Andrea Cuesta-Mosquera, Luka Drinovec, Asta Gregorič, Matej Ogrin, Kay Weinhold, Irena Ježek, Thomas Müller, et al. The impact of temperature inversions on black carbon and particle mass concentrations in a mountainous area. *Atmospheric Chemistry and Physics*, 22(8):5577–5601, 2022.

Jake Grigsby, Zhe Wang, Nam Nguyen, and Yanjun Qi. Long-range transformers for dynamic spatiotemporal forecasting, 2023.

Pu-Yun Kow, Yi-Shin Wang, Yanlai Zhou, I-Feng Kao, Maikel Issermann, Li-Chiu Chang, and Fi-John Chang. Seamless integration of convolutional and back-propagation neural networks for regional multi-step-ahead pm2. 5 forecasting. *Journal of Cleaner Production*, 261:121285, 2020.

Meng Wang, Yusen Duan, Zhuozhi Zhang, Juntao Huo, Yu Huang, Qingyan Fu, Tao Wang, Junji Cao, and Shun-cheng Lee. Increased contribution to pm2. 5 from traffic-influenced road dust in shanghai over recent years and predictable future. *Environmental Pollution*, 313:120119, 2022a.

Zhenyu Wang, Yingdong Yang, and Shaolong Yue. Air quality classification and measurement based on double output vision transformer. *IEEE Internet of Things Journal*, 9(21):20975–20984, 2022b.

Rui Xu, Deke Wang, Jian Li, Hang Wan, Shiming Shen, and Xin Guo. A hybrid deep learning model for air quality prediction based on the time–frequency domain relationship. *Atmosphere*, 14(2): 405, 2023.

Manzhu Yu, Arif Masrur, and Christopher Blaszczak-Boxe. Predicting hourly pm2. 5 concentrations in wildfire-prone areas using a spatiotemporal transformer model. *Science of The Total Environment*, 860:160446, 2023.

Haoyi Zhou, Shanghang Zhang, Jieqi Peng, Shuai Zhang, Jianxin Li, Hui Xiong, and Wancai Zhang. Informer: Beyond efficient transformer for long sequence time-series forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, volume 35, pp. 11106–11115, 2021.

Yuhe Zhou and Guangfei Yang. A predictive model of indoor pm2. 5 considering occupancy level in a hospital outpatient hall. *Science of The Total Environment*, 844:157233, 2022.

A Dataset Construction

The training and validation dataset for our model was sourced from *Sensor.Community*, a well-known contributor-driven network (https://sensor.community/) and was paired to weather data from *stormglass.io*. The dataset comprises filtered and aggregated data from Sofia, the capital of Bulgaria: a city with more than 1.2M inhabitants, notably impacted by air pollution, a condition exacerbated by its unique geographical setting. Large and encircled by towering mountains, it faces additional environmental challenges due to the widespread use of hard fuels for household heating and the prevalence of older vehicles on its roads. This combination of factors makes it an ideal case study for our research. The dataset was constructed from over 500 million records gathered over three years, ensuring a comprehensive and robust data foundation for analysis.

B Experimental Setup

To complete the experiments, we used the standard code for the Spacetimeformer (Grigsby et al., 2023) and the proposed transformer: Exo-Spacetimeformer. The training has been executed with a batch size of 512 records, 50 epoch iterations, and a learning rate of 0.001. The loss function has been fixed to mean square error. As an input to the model, we have used a context of 12 records and a forecasted output of 6, 24, and 48 records, respectively. Both models have an encoder and decoder with one NN layer. The dimensions of the model and the feed-forward network are 100, resulting in a transformer with a total of 2.1M trainable parameters.

C Experiment Insights

To further visualize our experiments figures 4 - 7 provide the validation loss during the training process as well as prediction results for 24h and 48h predictions.

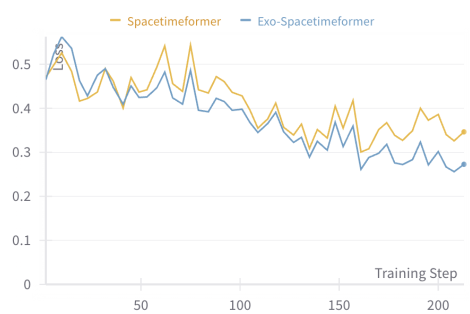

Figure 4: Validation loss for 24h predictions

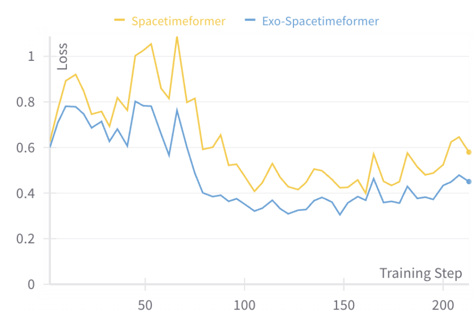

Figure 5: Validation loss for 48h predictions

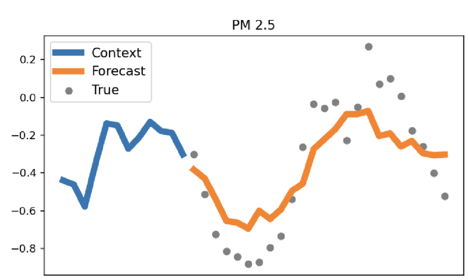

Figure 6: Sample prediction of PM 2.5 values for 24h by Exo-Spacetimeformer

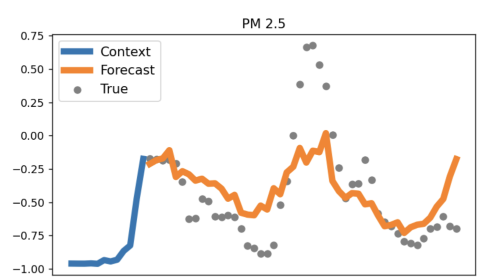

Figure 7: Sample prediction of PM 2.5 values for 48h by Exo-Spacetimeformer