Towards a general forecasting framework for Particulate Matter within urban environments

Fotios K. Anagnostopoulos,[1,*] Spyros Rigas,[2] Michalis Papachristou,[3]
Christos Tryfonopoulos,[1] Ioannis Chaniotis,[3] and Paraskevi Raftopoulou[1]

[1]*Department of Informatics & Telecommunications, University of Peloponnese, Karaiskaki 70, Athens, Greece*
[2]*School of Electrical and Computer Engineering, National & Technical University of Athens, Polytechnioupoli Zografou, 15772, Athens, Greece*
[3]*Department of Physics, National & Kapodistrian University of Athens, Panepistimioupoli Zografou 15773, Athens, Greece*

According to the World Health Organization, for more than 6000 cities in 117 countries, the level of particulate matter (PM) pollution is beyond healthy limits [1]. Furthermore, the effect's intensity appears to be inversely proportional to the population's average income. Consequently, in parallel with mitigation strategies, the development of accurate early-warning systems is imperative [2].

Of the two main approaches, namely transport models and simulations within the Computational Fluid Dynamics framework, the former suffers from low accuracy within the complex urban agglomeration, while the latter is very costly in terms of computational resources [3]. Moreover, both approaches require a detailed, city-specific modelling of the pollution sources. In this work we propose an alternative direction, towards a general framework for an early-warning system: by employing *Machine Learning* algorithms and specifically Artificial Neural Networks (ANN) [4–6], we deploy high-accuracy models with fast inference on real-time data. In addition, the city-specific information is encoded in each model during its training, thus resolving all of the aforementioned shortcomings of "traditional" approaches.

With the aid of a state-of-the-art ANN, we utilize publicly available PM measurements from low cost sensors and open meteorological data. Specifically, we train a Long Short-Term Memory (LSTM) Neural Network that provides a level of interpretability through an attention-like mechanism [7]. A novel characteristic of our work lies in capturing the spatial dependence of the phenomenon through urban planning features, such as Mean Floor Ratio [8] and Population Density. This allows for easily adding or removing sensors *and* employing moving sensors, for example in drones [9] or city bus [10], which is not trivial in the context of previous works in the field [4]. Importantly, our approach is applicable to any kind of sensors.

As a case study, we apply our method to a number of Greek cities with installed PurpleAir [11] sensor networks, e.g. Patras and Athens. Such cities are excellent candidates for the evaluation of our model, since they provide distinct environments with different meteorological conditions and emission budgets. The selected features are UV high, Solar Radiation High, Wind Gust Average, Average Wind Direction, Average Wind Speed, Average Temperature, Mean Floor Area Ratio, Dew point Average, Mean Population Density, Pressure, Precipitation Rate and Average Relative Humidity, and are extracted from meteorological data drawn from local stations[1]. Consecutive measurements of PM during the previous 24h correspond to our last feature.

We report that the LSTM network shows a forecasting accuracy that is comparable to the sensor's resolution, combined with meaningful interpretations of its results, which provide insight into the Physics of the problem. That said, in parallel with a re-training schedule, our model can be used as an accurate, low-cost, early-warning system, in order to attenuate the health hazards of particulate matter pollution, by providing automated, real-time alerts to the public.

[1] Data and code will be publicly available under Creative Commons licence upon publication.

*Electronic address: fotisanagn@uop.gr

Acknowledgments

FA wishes to thank Ioannis Anastasiou, UoP for valuable help in obtaining meteorological data. This work was supported in part by project ENIRISST+ under grant agreement No. MIS 5047041 from the General Secretary for ERDF & CF, under Operational Programme Competitiveness, Entrepreneurship and Innovation 2014-2020 (EPAnEK) of the Greek Ministry of Economy and Development (co-financed by Greece and the EU through the European Regional Development Fund).

ENIRISST PLUS

[1] World Health Organization, *Billions of people still breathe unhealthy air: new who data* (2022), URL https://www.who.int/news/item/04-04-2022-billions-of-people-still-breathe-unhealthy-air-new-who-data.
[2] F. J. Kelly, G. W. Fuller, H. A. Walton, and J. C. Fussell, Respirology **17**, 7 (2012).
[3] D. Toscano, M. Marro, B. Mele, F. Murena, and P. Salizzoni, Building and Environment **196**, 107812 (2021).
[4] X. Li, L. Peng, X. Yao, S. Cui, Y. Hu, C. You, and T. Chi, Environmental pollution **231**, 997 (2017).
[5] J. Zhao, F. Deng, Y. Cai, and J. Chen, Chemosphere **220**, 486 (2019).
[6] L. Li, R. Zhang, J. Sun, Q. He, L. Kong, and X. Liu, Journal of Environmental Health Science and Engineering **19**, 401 (2021).
[7] T. Guo, T. Lin, and N. Antulov-Fantulin, in *International conference on machine learning* (PMLR, 2019), pp. 2494–2504.
[8] A. Faludi, *A reader in planning theory*, vol. 5 (Elsevier, 2013).
[9] H. A. Hedworth, T. Sayahi, K. E. Kelly, and T. Saad, Journal of Aerosol Science **152**, 105702 (2021).
[10] S. Kaivonen and E. C.-H. Ngai, Digital Communications and Networks **6**, 23 (2020).
[11] *Purpleair, inc.* (2022), URL https://www2.purpleair.com/collections/air-quality-sensors/products/purpleair-pa-ii.