Predicting Daily Ozone Air Pollution With Transformers

Sebastian Hickman,[1] **Paul Griffiths,** [1] **Peer Nowack,** [2] **Alex Archibald** [1]

[1] Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK
[2] School of Environmental Sciences, University of East Anglia, Norwich, NR47 TJ, UK
shmh4@cam.ac.uk

Abstract

Surface ozone is an air pollutant that contributes to hundreds of thousands of premature deaths annually, primarily by causing cardiovascular and respiratory disease. Ozone at the surface also has considerable negative impacts on vegetation and crop yields. Ozone concentrations are affected by environmental factors, including temperature, which means that ozone concentrations are likely to change in future climates, posing increased risks to human health. This effect is known as the ozone climate penalty, and recent work suggests that currently polluted areas are likely to become more polluted by ozone in future climates. In light of recent stricter WHO regulations on surface ozone concentrations, we aim to build a predictive data-driven model for recent ozone concentrations, as a step towards more accurate ozone forecasting tools, which could be used to make predictions of ozone concentrations in future climates. We use observational station data from three European countries to train a transformer-based model to make predictions of daily maximum 8-hour ozone.

Introduction

Ozone at the surface is a secondary air pollutant which is not directly emitted by anthropogenic activities, but formed in the troposphere via a series of photochemical reactions (Finlayson-Pitts and Pitts Jr 1997). Ozone air pollution is estimated to contribute to between 365,000 and 1,100,000 premature deaths worldwide annually (Murray et al. 2020; Anenberg et al. 2010; Malley et al. 2017; Silva et al. 2013), primarily by causing cardiovascular and respiratory diseases (Kim, Kim, and Kim 2020; Filippidou and Koukouliata 2011; Sun et al. 2022). The impacts of ozone pollution have been linked to both long and short-term exposure to high ozone concentrations (Bell et al. 2004; Nuvolone, Petri, and Voller 2018).

Ozone air pollution is both a global and local issue. Background levels of ozone in remote areas routinely exceed guidelines set by the WHO, while local ozone concentrations can far exceed these guidelines. The WHO estimates that 99% of the world's population live in areas where ozone concentrations routinely exceed guidelines (WHO and ECE 2021).

Due to the phytotoxicity of ozone, the negative effects of ozone air pollution on vegetation, ecosystems and crop yields are significant (Fowler et al. 2009; Emberson et al. 2001). Ozone impacts vegetation by oxidative damage of cells (Ainsworth 2017). This damage leads both to considerable economic losses from reduced crop yields (Burney and Ramanathan 2014), and the potential for increased climate change, as damaged vegetation has a reduced capacity to sequester carbon dioxide from the atmosphere (Ainsworth et al. 2012; Sitch et al. 2007).

Effects of climate change on ozone

Derived from the key processes controlling ozone, in-situ photochemical production and transport, we observe a number of relationships between ozone and other variables, such as temperature (Laña et al. 2016) and meteorology. The contribution of each of these factors makes accurate prediction of ozone with numerical forward chemical transport models (CTMs) difficult. For example, large scale meteorological phenomena affect both temperatures and transport. The relative contributions to ozone concentrations from the changes in temperature compared to the changes in transport are difficult to ascertain, particularly at extrema.

Furthermore, increasing temperatures under climate change are also expected to affect ozone concentrations differently across regions (Schnell et al. 2016). This effect is known as the ozone climate penalty (Rasmussen et al. 2013). It has been suggested that increases in temperatures will lead to increases in ozone concentrations in regions polluted with NO_x, while in less polluted regions ozone is expected to decrease (Bloomer et al. 2009; Rasmussen et al. 2013). Recent studies have suggested this effect is significant in tropical, forested regions (Brown et al. 2022).

With higher temperatures we may expect increased concentrations of chemical ozone precursors in the atmosphere, firstly by thermal decomposition of peroxyacyl nitrate (PAN), a reservoir for NO_x (Beine et al. 1997), which may stifle the impact of projected NO_x emission cuts. There is also evidence that there will be increased NO_x emissions from soil (Hall, Matson, and Roth 1996). Emissions of VOCs (also ozone precursors, e.g. isoprene) from vegetation may also increase (Lathiere et al. 2005; Constable et al. 1999) due to changes in stomatal conductance with temperature (Niinemets and Reichstein 2003).

Considerable damage to human health from ozone comes from short-term exposure to high ozone levels. These extreme ozone episodes are often accompanied by high temperatures, leading to a combination of risks that further increase mortality (Filleul et al. 2006; Dear et al. 2005). It is projected that these compound events will increase under climate change leading to increases in the global risk to human health, despite projected emissions controls (Lei, Wuebbles, and Liang 2012). In particular, compound meteorological events, where both high temperatures and stagnant conditions lead to high ozone concentrations, are likely to become more common under climate change (Zhang et al. 2018). Predicting and understanding these extreme ozone events under climate change is important to accurately quantify the risks of ozone pollution to human and ecosystem health (Schnell et al. 2014).

Machine learning for ozone prediction

Machine learning methods are increasingly being deployed to make short-term predictions of ozone concentrations. As discussed previously, ozone concentrations are controlled by physical and chemical processes. These processes act on varying timescales. For example, extended periods of hot weather and stagnancy affect ozone concentrations, and therefore an ML approach that accounts for these temporal relationships is likely to be more skillful in predicting ozone compared to standard methods. Recently, both convolutional and recurrent neural networks have been implemented to make short-term predictions of ozone (Kleinert, Leufen, and Schultz 2021; Eslami et al. 2020; Biancofiore et al. 2015) (Table 1). However, transformer architectures have been less widely deployed (Chen et al. 2022). Transformers have been shown to be highly effective in sequential domains such as natural language processing (Brown et al. 2020; Ji et al. 2021), in part due to their ability to attend to long-term dependencies in the data. A transformer-based model may provide an intrinsic advantage over standard ML models and convolutional and recurrent neural networks that have been previously explored to predict ozone concentrations.

Therefore, in this work we implement a transformer-based temporal machine learning (ML) model to make short-term predictions of ozone, to evaluate the capacity of a transformer-based architecture to predict ozone, particularly for extreme concentrations. This is a step towards ML models capable of forecasting ozone accurately now, and in future climates.

Data selection and pre-processing

The TOAR dataset (Schultz et al. 2017) was selected as a suitable exploratory dataset for our predictive model, due to its global coverage and high fidelity and quantity of data, with daily measurements stretching back to the 1980s in some locations. The dataset is hosted by the Julich Supercomputing Centre, and provides 2.6 billion individual observations of ozone concentrations. By building a scalable approach with governmental observations we expect that the models would improve as more high fidelity observations are made.

As well as providing data on chemical species such as ozone and NO_x, this dataset also provides environmental variables on a daily scale, drawn from ERA5 reanalysis data (Hersbach et al. 2020), and landcover attributes for each station. This allows us to build supervised ML models with key ozone predictors affected by climate change.

We scraped data from the TOAR database for three European countries: the UK, France and Italy. These were chosen to represent three different environments, in order to test whether a single model could be trained to make accurate predictions across countries. Data from all months of the year and from both urban and rural stations were included in our dataset. This dataset therefore provides a larger sample of different environments than have been studied in previous work (Kleinert, Leufen, and Schultz 2021; Biancofiore et al. 2015), with data from 1997 to 2013, from 1012 individual stations. Our final dataset contains more than 2 million individual days of data.

Due to the large size and relative completeness of our dataset, imputing missing values was deemed unnecessary. By simply removing missing data we removed the risk of bias from data imputation, although future work will examine the impact of replacing missing values. We scaled our features with min-max normalisation (Jayalakshmi and Santhakumaran 2011).

The data include both static and dynamic features. The static features relate to characteristics of a particular station, such as the local population density, while the dynamic features are environmental or chemical variables which change through time, such as temperature. Some static features are given as strings in the TOAR dataset, such as station type (e.g. 'background'). These static categorical features were converted to numerical features.

To train, validate and test our models, we split the data temporally, with the penultimate year used for validation, the final year used for testing and the remainder for training. We used the previous 21 days of observations of ozone and covariates to make ozone predictions up to 4 days ahead.

Predictive results

In order to provide a complement to existing numerical CTMs for ozone prediction, we deploy a state-of-the-art temporal deep learning architecture, the temporal fusion transformer (TFT) (Lim et al. 2021). Transformers have been used widely and successfully in natural language processing and on other sequential data (Brown et al. 2020; Ji et al. 2021). The TFT combines gated residual networks, variable selection networks, an LSTM encoder-decoder layer, and multi-head attention.

The TFT is able to ingest both static and dynamic features to make predictions of ozone. When making predictions operationally, it is important to quantify the uncertainty in predictions. In order to extract prediction intervals from the TFT, we used a quantile loss function.

Despite being a relatively computationally expensive ML method, training the TFT on our dataset took 2 hours using 2 Tesla V100 GPUs. Once trained, making predictions

across 1012 individual stations took 25 μs. This illustrates the vast speed-ups compared to CTMs possible with ML models. Hyperparameters were optimised manually on the validation data.

Performance of the TFT

When predicting ozone concentrations using concurrently observed covariate data (infilling a dataset, or forecasting ozone with a meteorological forecast), the model was skillful (MAE = 4.1 ppb, R^2 = 0.86, RMSE = 5.6 ppb, r = 0.92). These predictions rely on previous ozone observations and concurrent covariate data, and therefore these predictions are suitable for making short-term future predictions with meteorological forecasts as input, and infilling missing ozone values in historical data. The TFT was also used to make short-term future predictions without concurrent covariates, up to four days ahead, which also yielded good performance (MAE = 5.7 ppb, R^2 = 0.73, RMSE = 7.8 ppb, r = 0.82) which was 20% more accurate in terms of MAE compared to a persistence model. While we cannot make direct comparisons due to differing test datasets, the skill of our method compares favourably to other ML methods and numerical air quality forecasting models such as AQUM (Neal et al. 2014; Im et al. 2015), especially given the size and variety of our test dataset (Table 1).

A correlation plot of TFT predictions on the test set, using concurrent covariates, against observations, is given in Figure 1. The model was more accurate than other standard ML approaches such as random forests and LSTMs, and approximately 40% more accurate compared to a persistence model in terms of MAE.

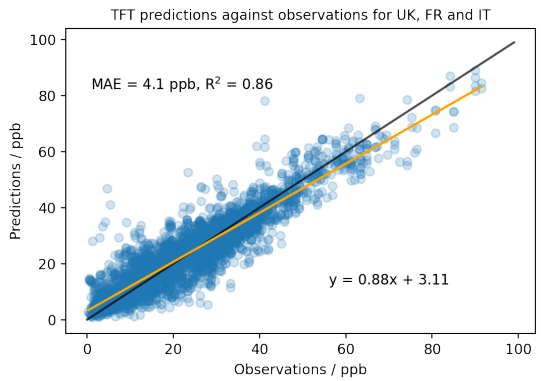

Figure 1: Correlation plot of predictions against observations on the test data for forecasting ozone concentrations with the TFT using concurrent covariates. Overall prediction skill was good (MAE = 4.1 ppb, R^2 = 0.86).

We can further visualise the skill of the TFT by looking at predictions and observations at individual stations in our dataset, as illustrated in Figure 2. Figure 2 also shows days from the past which the attention mechanism in the model used to inform the predictions, shown by the grey line. The model pays attention to previous high ozone days to make future predictions of high ozone concentrations.

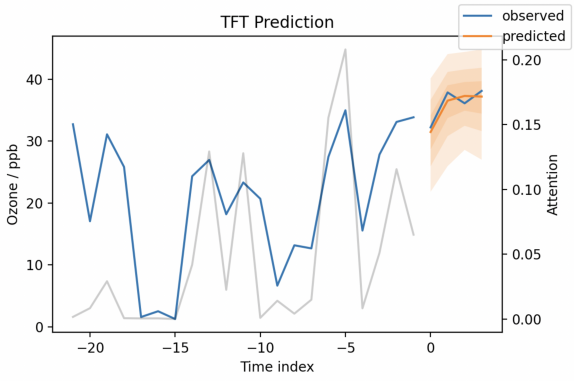

Figure 2: Example predictions of the TFT at an individual station in the UK, accompanied by prediction intervals derived from the quantile predictions of the model. The 21 days of previous ozone is also shown with a negative time index. The plot also shows the attention that the model pays to each particular day of data from the past, illustrating which information the model uses to make predictions.

Figure 2 also illustrates the prediction intervals generated by the TFT, which are a great benefit to evaluate trust in the model. The uncertainty estimates correlate reasonably well with the accuracy of predictions at different stations.

We also tested whether the TFT was able to make skillful predictions of ozone concentrations at both urban and rural stations. The most comparable recent work to ours, Kleinert et al. (Kleinert, Leufen, and Schultz 2021), only trained and evaluated their model on rural stations, and therefore it is unclear whether a single deep learning model is able to generalise across these two distinct environments. Our model performed similarly on urban and rural data (MAE = 4.2, R^2 = 0.86 and MAE = 4.0, R^2 = 0.85 respectively), which suggests that architectures of this type are able to generalise across the two environments given sufficient training data.

Predicting extreme ozone

Ozone concentrations in Europe tend to peak in the spring and summer months, typically between April and June (Lewis et al. 2021). Making accurate forecasts of high ozone is important, as these high ozone concentrations pose a great threat to health, and are likely to occur more frequently in future climates. We therefore evaluated the skill of the TFT using concurrent covariates during these high ozone periods. We found that the TFT was able to make reasonably skillful forecasts on spring and summertime ozone concentrations (MAE = 6.2 ppb, R^2 = 0.77). However, the performance was poor compared to forecasting on data from the rest of the year (MAE = 3.9 ppb, R^2 = 0.87). The performance is illustrated in Figure 3.

Method (and paper)	r (Pearson)	RMSE / ppb
Persistence	0.42	10.16
(Ivatt and Evans 2020), Geos-CHEM	0.48	16.2
Ridge regression	0.50	9.59
(Neal et al. 2014), AQUM	0.64	20.8
Random forest	0.68	7.51
(Debry and Mallet 2014), DRR	0.70	6.3
(Neal et al. 2014), bias-corrected AQUM	0.76	16.4
(Sayeed et al. 2020), CNN	0.77	8.8
(Eslami et al. 2020), CNN	0.79	12.0
(Ivatt and Evans 2020), bias-corrected Geos-CHEM	0.84	7.5
LSTM	0.85	6.11
(Biancofiore et al. 2015), RNN	0.86	12.5
(Chen et al. 2022), CNN-Transformer	*NA*	7.8
TFT	0.90	5.6

Table 1: The relative performance of different ML and numerical approaches to ozone forecasting. Methods in italics were evaluated on our dataset, while the other methods used different datasets. The difficulty of comparing methods evaluated on different datasets is shown by the varying RMSE values.

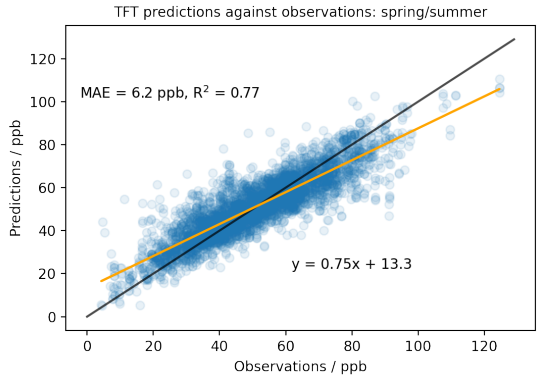

Figure 3: Comparison of the TFT when making forecasts with concurrent covariates during the spring/summer and the non-spring/summer months. When predicting on spring/summertime ozone, the performance was considerably worse (MAE = 6.2 ppb and R^2 = 0.77) compared to predicting during the rest of the year (MAE = 3.9 ppb and R^2 = 0.87). The root cause of this appears to be the underprediction of ozone at high values.

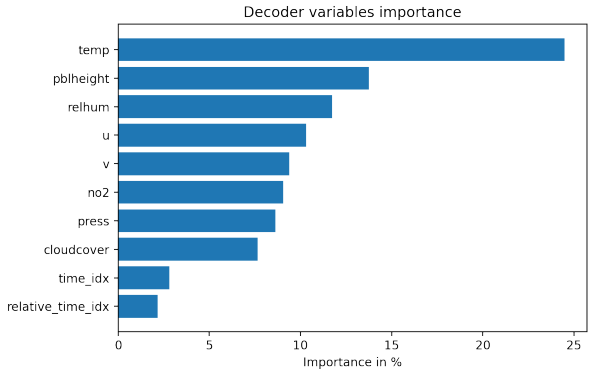

Figure 4: The variable importances of the TFT when making infilling predictions, derived from the weights of the attention mechanism. These are largely in line with expected physical and chemical relationships.

What is the TFT paying attention to?

We extracted feature importances, derived from the weights of attention mechanism in our model, to examine which dynamic features are most important when making predictions with our data. These importances are largely in line with what is expected physically: both temperature and planetary boundary layer height are key variables. Interestingly, the concentration of chemical species are less important, which agrees somewhat with the findings of model-based studies (Porter and Heald 2019). However, our data does not contain all the variables available in model-based studies.

Conclusions

Ozone is difficult to model with existing numerical methods. A transformer-based ML model, the TFT, makes skillful predictions of ozone concentrations at stations across Europe. The model is able to make accurate predictions across environments, and performs reasonably well when predicting extrema. The TFT pays attention to relevant physical processes driving the ozone concentrations. This model provides a promising, computationally cheap method to make short-term forecasts of ozone concentrations, which could be deployed to make accurate forecasts in future climates.

Acknowledgments

SH acknowledges funding from EPSRC via the AI4ER CDT at the University of Cambridge (EP/S022961/1). PG and AA thank NCAS and the Met Office. PN is supported through an Imperial College Research Fellowship.

References

Ainsworth, E. A. 2017. Understanding and improving global crop response to ozone pollution. *The Plant Journal*, 90(5): 886–897.

Ainsworth, E. A.; Yendrek, C. R.; Sitch, S.; Collins, W. J.; Emberson, L. D.; et al. 2012. The effects of tropospheric ozone on net primary productivity and implications for climate change. *Annual Review of Plant Biology*, 63(1): 637–661.

Anenberg, S. C.; Horowitz, L. W.; Tong, D. Q.; and West, J. J. 2010. An estimate of the global burden of anthropogenic ozone and fine particulate matter on premature human mortality using atmospheric modeling. *Environmental Health Perspectives*, 118(9): 1189–1195.

Beine, H. J.; Krognes, T.; Stordal, F.; et al. 1997. High-latitude springtime photochemistry. Part I: NOx, PAN and ozone relationships. *Journal of Atmospheric Chemistry*, 27(2): 127–153.

Bell, M. L.; McDermott, A.; Zeger, S. L.; Samet, J. M.; and Dominici, F. 2004. Ozone and short-term mortality in 95 US urban communities, 1987-2000. *Jama*, 292(19): 2372–2378.

Biancofiore, F.; Verdecchia, M.; Di Carlo, P.; Tomassetti, B.; Aruffo, E.; Busilacchio, M.; Bianco, S.; Di Tommaso, S.; and Colangeli, C. 2015. Analysis of surface ozone using a recurrent neural network. *Science of the Total Environment*, 514: 379–387.

Bloomer, B. J.; Stehr, J. W.; Piety, C. A.; Salawitch, R. J.; and Dickerson, R. R. 2009. Observed relationships of ozone air pollution with temperature and emissions. *Geophysical Research Letters*, 36(9).

Brown, F.; Folberth, G. A.; Sitch, S.; Bauer, S.; Bauters, M.; Boeckx, P.; Cheesman, A. W.; Deushi, M.; Dos Santos, I.; Galy-Lacaux, C.; et al. 2022. The ozone–climate penalty over South America and Africa by 2100. *EGUsphere*, 1–33.

Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J. D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. 2020. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33: 1877–1901.

Burney, J.; and Ramanathan, V. 2014. Recent climate and air pollution impacts on Indian agriculture. *Proceedings of the National Academy of Sciences*, 111(46): 16319–16324.

Chen, Y.; Chen, X.; Xu, A.; Sun, Q.; and Peng, X. 2022. A hybrid CNN-Transformer model for ozone concentration prediction. *Air Quality, Atmosphere & Health*, 1–14.

Constable, J. V. H.; Guenther, A. B.; Schimel, D. S.; and Monson, R. K. 1999. Modelling changes in VOC emission in response to climate change in the continental United States. *Global Change Biology*, 5(7): 791–806.

Dear, K.; Ranmuthugala, G.; Kjellström, T.; Skinner, C.; and Hanigan, I. 2005. Effects of temperature and ozone on daily mortality during the August 2003 heat wave in France. *Archives of Environmental & Occupational Health*, 60(4): 205–212.

Debry, E.; and Mallet, V. 2014. Ensemble forecasting with machine learning algorithms for ozone, nitrogen dioxide and PM10 on the Prev'Air platform. *Atmospheric Environment*, 91: 71–84.

Emberson, L.; Ashmore, M.; Simpson, D.; Tuovinen, J.-P.; and Cambridge, H. 2001. Modelling and mapping ozone deposition in Europe. *Water, Air, and Soil Pollution*, 130(1): 577–582.

Eslami, E.; Choi, Y.; Lops, Y.; and Sayeed, A. 2020. A real-time hourly ozone prediction system using deep convolutional neural network. *Neural Computing and Applications*, 32(13): 8783–8797.

Filippidou, E.; and Koukouliata, A. 2011. Ozone effects on the respiratory system. *Prog Health Sci*, 1(2).

Filleul, L.; Cassadou, S.; Médina, S.; Fabres, P.; Lefranc, A.; Eilstein, D.; Le Tertre, A.; Pascal, L.; Chardon, B.; Blanchard, M.; et al. 2006. The relation between temperature, ozone, and mortality in nine French cities during the heat wave of 2003. *Environmental Health Perspectives*, 114(9): 1344–1347.

Finlayson-Pitts, B. J.; and Pitts Jr, J. N. 1997. Tropospheric air pollution: ozone, airborne toxics, polycyclic aromatic hydrocarbons, and particles. *Science*, 276(5315): 1045–1051.

Fowler, D.; Pilegaard, K.; Sutton, M.; Ambus, P.; Raivonen, M.; Duyzer, J.; Simpson, D.; Fagerli, H.; Fuzzi, S.; Schjoerring, J.; et al. 2009. Atmospheric composition change: ecosystems–atmosphere interactions. *Atmospheric Environment*, 43(33): 5193–5267.

Hall, S. J.; Matson, P. A.; and Roth, P. M. 1996. NOx emissions from soil: implications for air quality modeling in agricultural regions. *Annual Review of Energy and the Environment*, 21(1): 311–346.

Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. 2020. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730): 1999–2049.

Im, U.; Bianconi, R.; Solazzo, E.; Kioutsioukis, I.; Badia, A.; Balzarini, A.; Baró, R.; Bellasio, R.; Brunner, D.; Chemel, C.; et al. 2015. Evaluation of operational on-line-coupled regional air quality models over Europe and North America in the context of AQMEII phase 2. Part I: Ozone. *Atmospheric Environment*, 115: 404–420.

Ivatt, P. D.; and Evans, M. J. 2020. Improving the prediction of an atmospheric chemistry transport model using gradient-boosted regression trees. *Atmospheric Chemistry and Physics*, 20(13): 8063–8082.

Jayalakshmi, T.; and Santhakumaran, A. 2011. Statistical normalization and back propagation for classification. *International Journal of Computer Theory and Engineering*, 3(1): 1793–8201.

Ji, Y.; Zhou, Z.; Liu, H.; and Davuluri, R. V. 2021. DNABERT: pre-trained Bidirectional Encoder Representations from Transformers model for DNA-language in genome. *Bioinformatics*, 37(15): 2112–2120.

Kim, S.-Y.; Kim, E.; and Kim, W. J. 2020. Health effects of ozone on respiratory diseases. *Tuberculosis and Respiratory Diseases*, 83(Supple 1): S6.

Kleinert, F.; Leufen, L. H.; and Schultz, M. G. 2021. IntelliO3-ts v1. 0: a neural network approach to predict near-surface ozone concentrations in Germany. *Geoscientific Model Development*, 14(1): 1–25.

Laña, I.; Del Ser, J.; Padró, A.; Vélez, M.; and Casanova-Mateo, C. 2016. The role of local urban traffic and meteorological conditions in air pollution: A data-based case study in Madrid, Spain. *Atmospheric Environment*, 145: 424–438.

Lathiere, J.; Hauglustaine, D.; De Noblet-Ducoudré, N.; Krinner, G.; and Folberth, G. 2005. Past and future changes in biogenic volatile organic compound emissions simulated with a global dynamic vegetation model. *Geophysical Research Letters*, 32(20).

Lei, H.; Wuebbles, D. J.; and Liang, X.-Z. 2012. Projected risk of high ozone episodes in 2050. *Atmospheric Environment*, 59: 567–577.

Lewis, A.; Allan, J.; Carruthers, D.; Carslaw, D.; Fuller, G.; Harrison, R.; Heal, M.; Nemitz, E.; and Reeves, C. 2021. Ozone in the UK – recent trends and future projections. *Department for Environment, Food and Rural Affairs Report*.

Lim, B.; Arık, S. Ö.; Loeff, N.; and Pfister, T. 2021. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4): 1748–1764.

Malley, C. S.; Henze, D. K.; Kuylenstierna, J. C.; Vallack, H. W.; Davila, Y.; Anenberg, S. C.; Turner, M. C.; and Ashmore, M. R. 2017. Updated global estimates of respiratory mortality in adults 30 years of age attributable to long-term ozone exposure. *Environmental Health Perspectives*, 125(8): 087021.

Murray, C. J.; Aravkin, A. Y.; Zheng, P.; Abbafati, C.; Abbas, K. M.; Abbasi-Kangevari, M.; Abd-Allah, F.; Abdelalim, A.; Abdollahi, M.; Abdollahpour, I.; et al. 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet*, 396(10258): 1223–1249.

Neal, L.; Agnew, P.; Moseley, S.; Ordóñez, C.; Savage, N.; and Tilbee, M. 2014. Application of a statistical post-processing technique to a gridded, operational, air quality forecast. *Atmospheric Environment*, 98: 385–393.

Niinemets, Ü.; and Reichstein, M. 2003. Controls on the emission of plant volatiles through stomata: Differential sensitivity of emission rates to stomatal closure explained. *Journal of Geophysical Research: Atmospheres*, 108(D7).

Nuvolone, D.; Petri, D.; and Voller, F. 2018. The effects of ozone on human health. *Environmental Science and Pollution Research*, 25(9): 8074–8088.

Porter, W. C.; and Heald, C. L. 2019. The mechanisms and meteorological drivers of the summertime ozone–temperature relationship. *Atmospheric Chemistry and Physics*, 19(21): 13367–13381.

Rasmussen, D.; Hu, J.; Mahmud, A.; and Kleeman, M. J. 2013. The ozone–climate penalty: past, present, and future. *Environmental Science & Technology*, 47(24): 14258–14266.

Sayeed, A.; Choi, Y.; Eslami, E.; Lops, Y.; Roy, A.; and Jung, J. 2020. Using a deep convolutional neural network to predict 2017 ozone concentrations, 24 hours in advance. *Neural Networks*, 121: 396–408.

Schnell, J.; Holmes, C.; Jangam, A.; and Prather, M. 2014. Skill in forecasting extreme ozone pollution episodes with a global atmospheric chemistry model. *Atmospheric Chemistry and Physics*, 14(15): 7721–7739.

Schnell, J. L.; Prather, M. J.; Josse, B.; Naik, V.; Horowitz, L. W.; Zeng, G.; Shindell, D. T.; and Faluvegi, G. 2016. Effect of climate change on surface ozone over North America, Europe, and East Asia. *Geophysical Research Letters*, 43(7): 3509–3518.

Schultz, M. G.; Schröder, S.; Lyapina, O.; Cooper, O. R.; Galbally, I.; Petropavlovskikh, I.; Von Schneidemesser, E.; Tanimoto, H.; Elshorbany, Y.; Naja, M.; et al. 2017. Tropospheric Ozone Assessment Report: Database and metrics data of global surface ozone observations. *Elementa: Science of the Anthropocene*, 5.

Silva, R. A.; West, J. J.; Zhang, Y.; Anenberg, S. C.; Lamarque, J.-F.; Shindell, D. T.; Collins, W. J.; Dalsoren, S.; Faluvegi, G.; Folberth, G.; et al. 2013. Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change. *Environmental Research Letters*, 8(3): 034005.

Sitch, S.; Cox, P.; Collins, W.; and Huntingford, C. 2007. Indirect radiative forcing of climate change through ozone effects on the land-carbon sink. *Nature*, 448(7155): 791–794.

Sun, H. Z.; Yu, P.; Lan, C.; Wan, M. W.; Hickman, S.; Murulitharan, J.; Shen, H.; Yuan, L.; Guo, Y.; and Archibald, A. T. 2022. Cohort-based long-term ozone exposure-associated mortality risks with adjusted metrics: A systematic review and meta-analysis. *The Innovation*, 100246.

WHO; and ECE. 2021. *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization.

Zhang, J.; Gao, Y.; Luo, K.; Leung, L. R.; Zhang, Y.; Wang, K.; and Fan, J. 2018. Impacts of compound extreme weather events on ozone in the present and future. *Atmospheric Chemistry and Physics*, 18(13): 9861–9877.

Appendices

Features from the TOAR dataset

Table 2 describes the data used as features for the machine learning model. The features are split into static and dynamic features. Static features describe the characteristics of a particular station, while dynamic features vary through time.

Model hyperparameters

Table 3 details the hyperparameters used for the TFT model. These hyperparameters were selected with manual optimisation, however more principled methods such a random search or Bayesian optimisation will be implemented in future work.

Variable Name	Description
Static	
station type	Characterisation of site, e.g. "background", "industrial", "traffic".
landcover	The dominant IGBP landcover classification at the station location extracted from the MODIS MCD12C1 dataset (original resolution: 0.05 degrees).
toar category	A station classification for the Tropospheric Ozone Assessment Report based on the station proxy data that are stored in the database. One of unclassified, low elevation rural, high elevation rural or urban.
pop density	Year 2010 human population per square km from CIESIN GPW v3 (original horizontal resolution: 2.5 arc minutes).
max 5km pop density	Maximum population density in a radius of 5 km around the station location.
max 25km pop density	Maximum population density in a radius of 25 km around the station location.
nightlight 1km	Year 2013 Nighttime lights brightness values from NOAA DMSP (original horizontal resolution: 0.925 km).
nightlight max 25km	Year 2013 Nighttime lights brightness values (original horizontal resolution: 5 km).
alt	Altitude of station (in m above sea level). Best estimate of the station altitude, which frequently uses the elevation from Google Earth.
station etopo alt	Terrain elevation at the station location from the 1 km resolution ETOPO1 dataset.
nox emi	Year 2010 NOx emissions from EDGAR HTAP inventory V2 in units of g m^{-2} yr^{-1} (original resolution: 0.1 degrees)
omi nox	Average 2011-2015 tropospheric NO_2 columns from OMI at 0.1 degree resolution (Env. Canada) in units of 10^{15} molecules cm^{-2}.
Dynamic	
o3	Ozone concentration, daily maximum 8-hour average statistics according to the using the EU definition of the daily 8-hour window starting from 17 h of the previous day. Measured at the station, with UV absorption.
cloudcover	Daily average cloud cover from ERA5 reanalysis for the grid cell containing a particular station.
relhum	Daily average relative humidity from ERA5 reanalysis for the grid cell containing a particular station.
press	Daily average pressure from ERA5 reanalysis for the grid cell containing a particular station.
temp	Daily average temperature from ERA5 reanalysis for the grid cell containing a particular station.
v	Daily average meridional wind speed from ERA5 reanalysis for the grid cell containing a particular station.
u	Daily average zonal wind speed from ERA5 reanalysis for the grid cell containing a particular station.
pblheight	Daily average planetary boundary layer height from ERA5 reanalysis for the grid cell containing a particular station.
no2	NO_2 concentration, as the daily maximum 8-hour average statistics according to the using the EU definition of the daily 8-hour window starting from 17 h of the previous day. Measured at the station, with chemiluminescence.

Table 2: Table giving the relevant data extracted from the TOAR database.

Model	Hyperparameter value
TFT	
attention head size	4
dropout	0.2
hidden continuous size	32
hidden size	64
learning rate	0.0135
lstm layers	4
optimizer	ranger

Table 3: Table giving the hyperparameters for the final TFT used for model evaluation.