ARCTIC PASSION

11th June 2024

Air Quality Forecast for Arctic Communities Service (AURORAE) Model details

WP4 – Task 4.5 - Pilot Service 5

A.Cuzzucoli[2], I.Crotti[1], D.De Marchi[1], F.Eyraud[1], L.Selmi[2], A.Pasini[2], S.Dobricic[1]

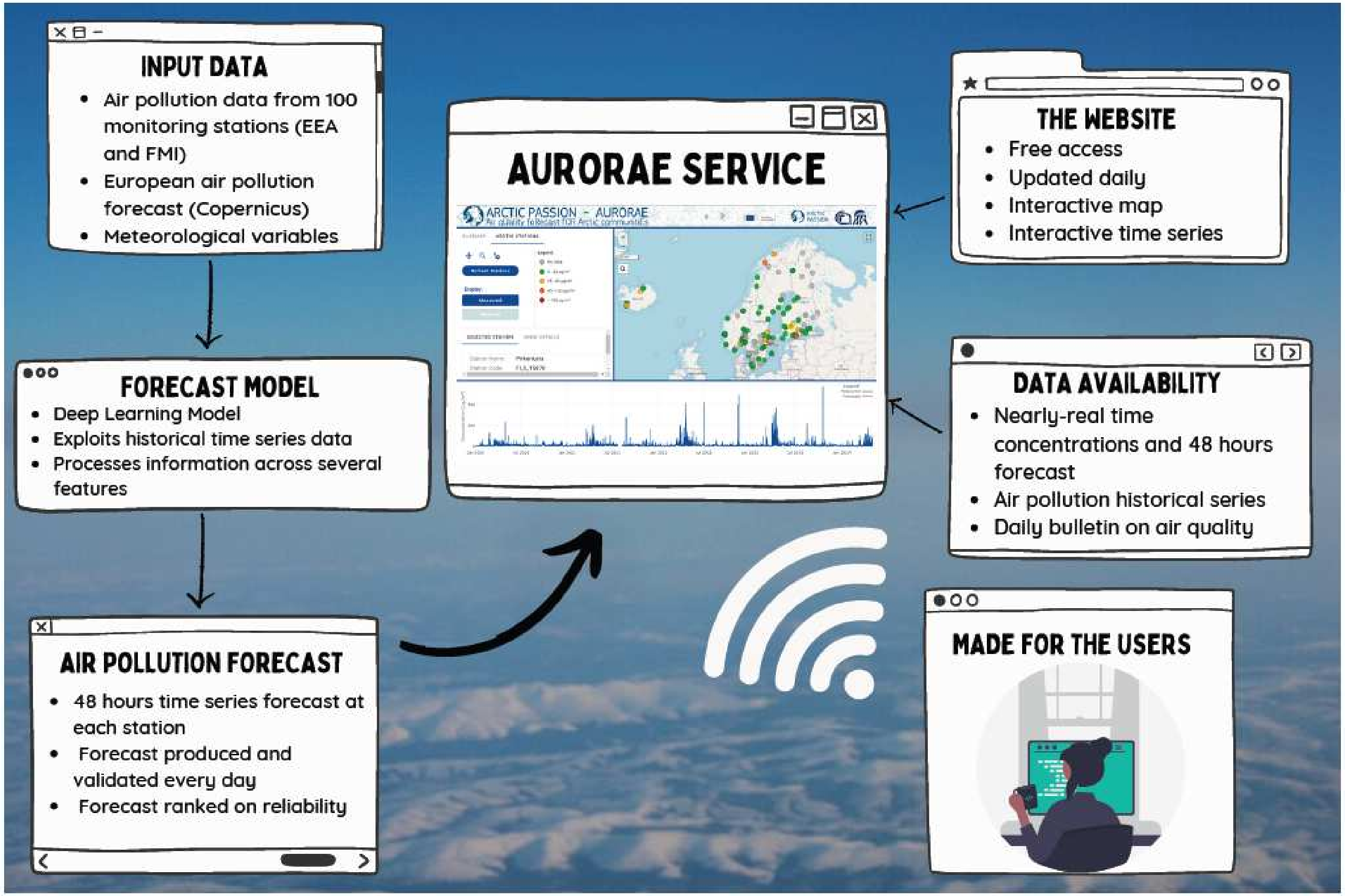
[1]European Commission, Joint Research Centre, Ispra, Italy

[2]Institute of Atmospheric Pollution Research, National Research Council, Roma, Italy

Contacts : ilaria.crotti@ec.europa.eu, alice.cuzzucoli@iia.cnr.it, srdan.dobricic@ec.europa.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 101003472

A forecast tailored for the Arctic: the AURORAE service



ARCTIC PASSION

Predicting Air pollution in North Europe

Particulate matter (PM_{10}) is a complex mixture of solids and aerosols with a diameter of 10 microns or less.

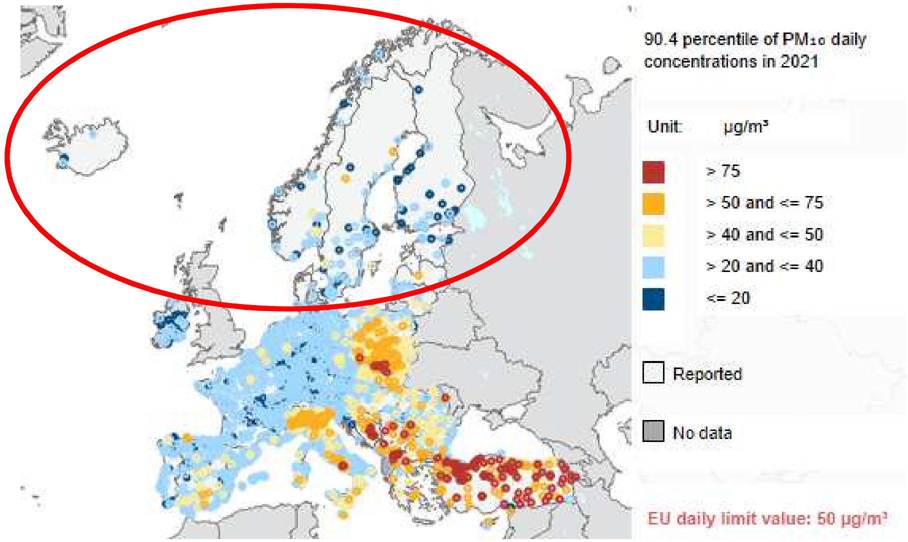
Behaviour of pollutants can be addressed via deterministic or statistical models to:

- Produce long term forecasts
- Flag unexpected behaviour

Numerical Models from Copernicus Atmospheric Monitoring Service (CAMS) can provide hourly forecasts of pollutants, but are not particularly reliable in the Arctic region.

↓

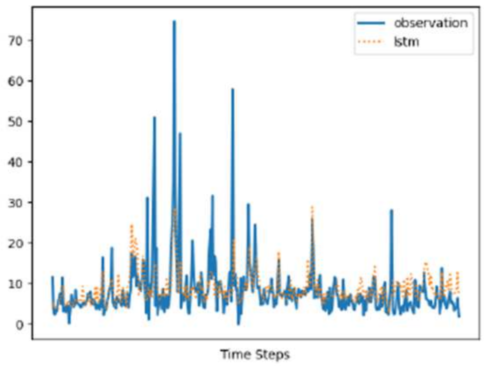
Apply a Deep Learning approach based on Neural Networks to enhance predictions from CAMS



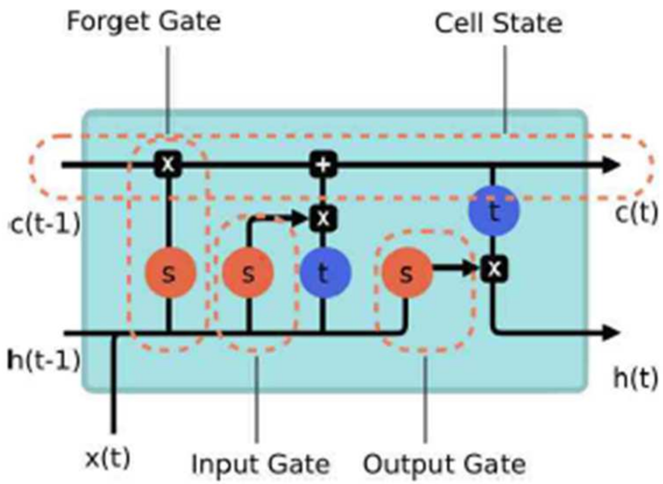
- PM10 observations from around 100 stations in IS, FI, NO, SE
- PM10 forecasts for 48 consecutive hours from CAMS Models
- Meteorological variables forecasts for 48 consecutive hours (temperature, wind magnitude and direction, boundary layer height, total precipitations and pressure)

Source: European Environmental Agency (www.eea.europa.eu/)

Time series forecasting



- Sequence-to-sequence task
- Models need to store information from previous states of the variables
- Deep models involving Recurrent Neural Networks have been state of the art for a long time



The motivation for preferring a Deep Learning approach to a statistical method (such as autoregressive models) comes from the potential to:

- Better process a large amount of data
- Better deal with correlations between several variables
- Forecasting accuracy

Long Short-Term Memory Network (LSTM) Architecture, as used in Fazzini, P.; Montuori, M.; Pasini, A.; Cuzzucoli, A.; Crotti, I.; Campana, E.F.; Petracchini, F.; Dobricic, S., Forecasting PM10 Levels Using Machine Learning Models in the Arctic: A Comparative Study. Remote Sens. 2023, 15, 3348.

ARCTIC PASSION

Transformer Architecture

In recent developments, time series forecasting is turning away from RNNs as new models coming from Natural Language Processing are being adapted to time series analysis tasks as they:

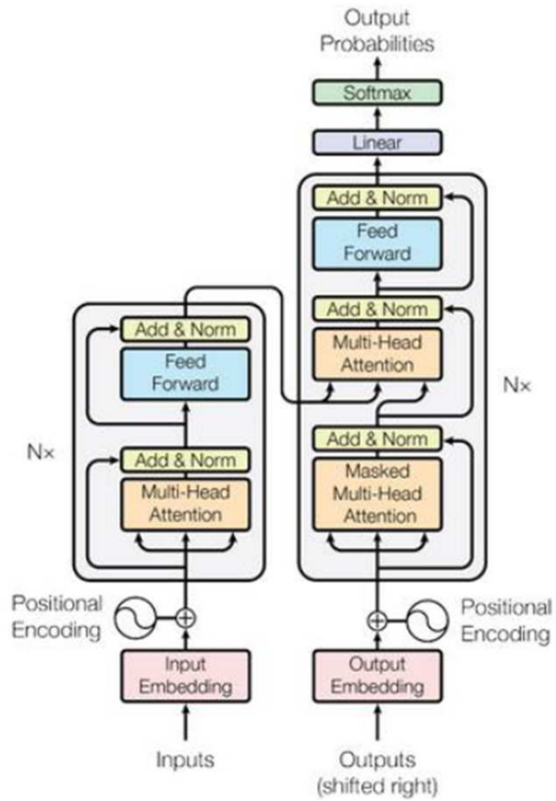
- better incorporate sequential data and thus historical information
- prove to be more suited for long sequence forecasting
- Better encode temporal and cross dimensional information

Encoder + Decoder structure

Temporal embedding

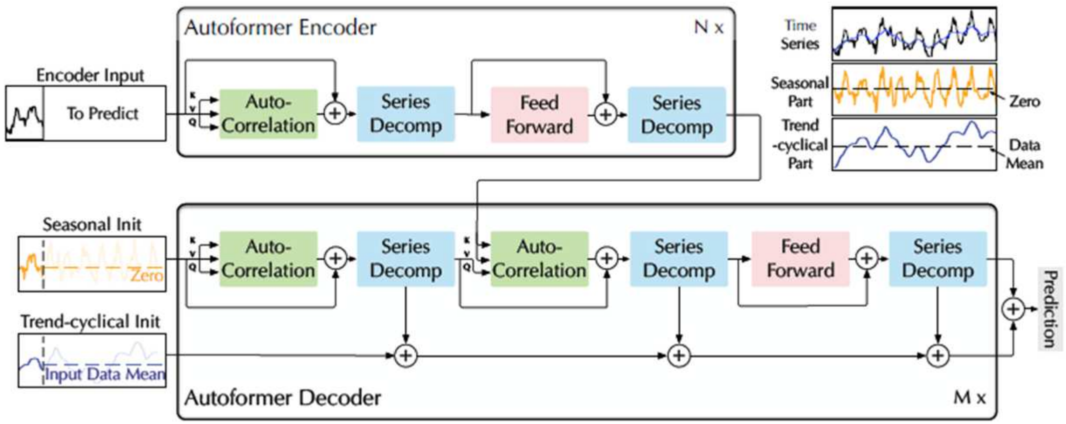
Attention mechanism

Feed-Forward Network



Original Transformer Architecture as defined in Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. Attention is all you need. CoRR, abs/1706.03762, 2017

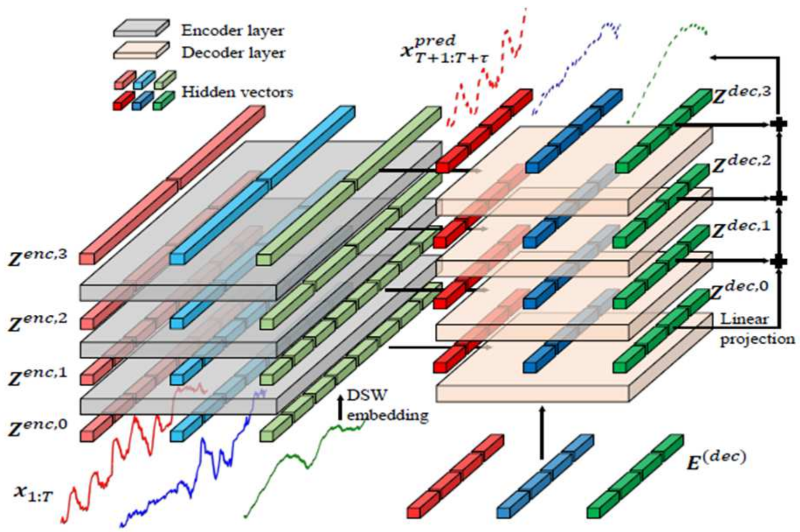
Variation: Incorporating Time Series Decomposition



Original Autoformer Architecture from Wu, H., Xu, J., Wang, J., and Long, M. Autoformer: Decomposition transformers with auto-correlation for longterm series forecasting. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), pp. 101–112, 2021

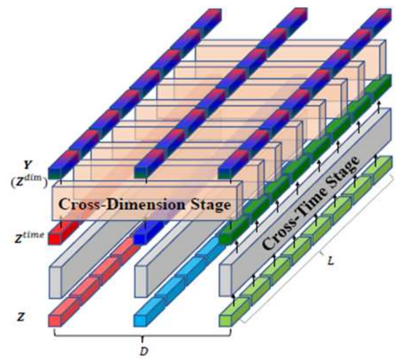
- Instead of working with the entire time series, one idea is to process season and trend parts separately
- Encoder determines context of season part of the timeseries
- Decoder translates both season and trend part from context
- Attention is determined from series autocorrelation in time-domain (Autoformer) or frequency domain (FEDformer)

Variation: Incorporating Cross-Dependencies



Original Crossformer Architecture as represented in Zhang, Y. and Yan J., Crossformer: Transformer Utilizing Cross-Dimension Dependency for Multivariate Time Series Forecasting, International Conference on Learning Representations, 2023

- Most Transformer architectures try to capture cross dependencies amongst features within embedding and forward stages, while it is possible to incorporate a cross-dimension stage within an encoder layer as a two-step attention block



- The embedding also considers segment-wise decomposition in order to track correlations between subsequences