Transformers for Fast Emulation of Atmospheric Chemistry Box Models

Herbie Bradley (University of Cambridge), Luke Abraham (University of Cambridge), Peer Nowack (Imperial College London),
Doug McNeall (UK Met Office)

Introduction

- Atmospheric chemistry models are important scientific tools to understand the role of aerosols and ozone in climate change and their effect on human health
- Atmospheric chemistry models = chemical equations + reactions -> system of ODEs.
- Input: initial concentrations for each species, initial reaction rates, and environmental parameters.
- Output: time series of concentrations for each species.

A Challenge

- However, atmospheric chemistry models are computationally expensive to run at global scale.
- The chemistry numerical solver alone makes up 10% of the running time of the UKESM1 Earth system model.
- As a step towards more efficient atmospheric chemistry models, we develop a **neural network emulator with a transformer-based architecture**.
- Challenges of emulation:
  - High dimensionality of data.
  - Long term stability and error growth over time.
  - Chemically realistic output.

The MOZART-4 Box Model

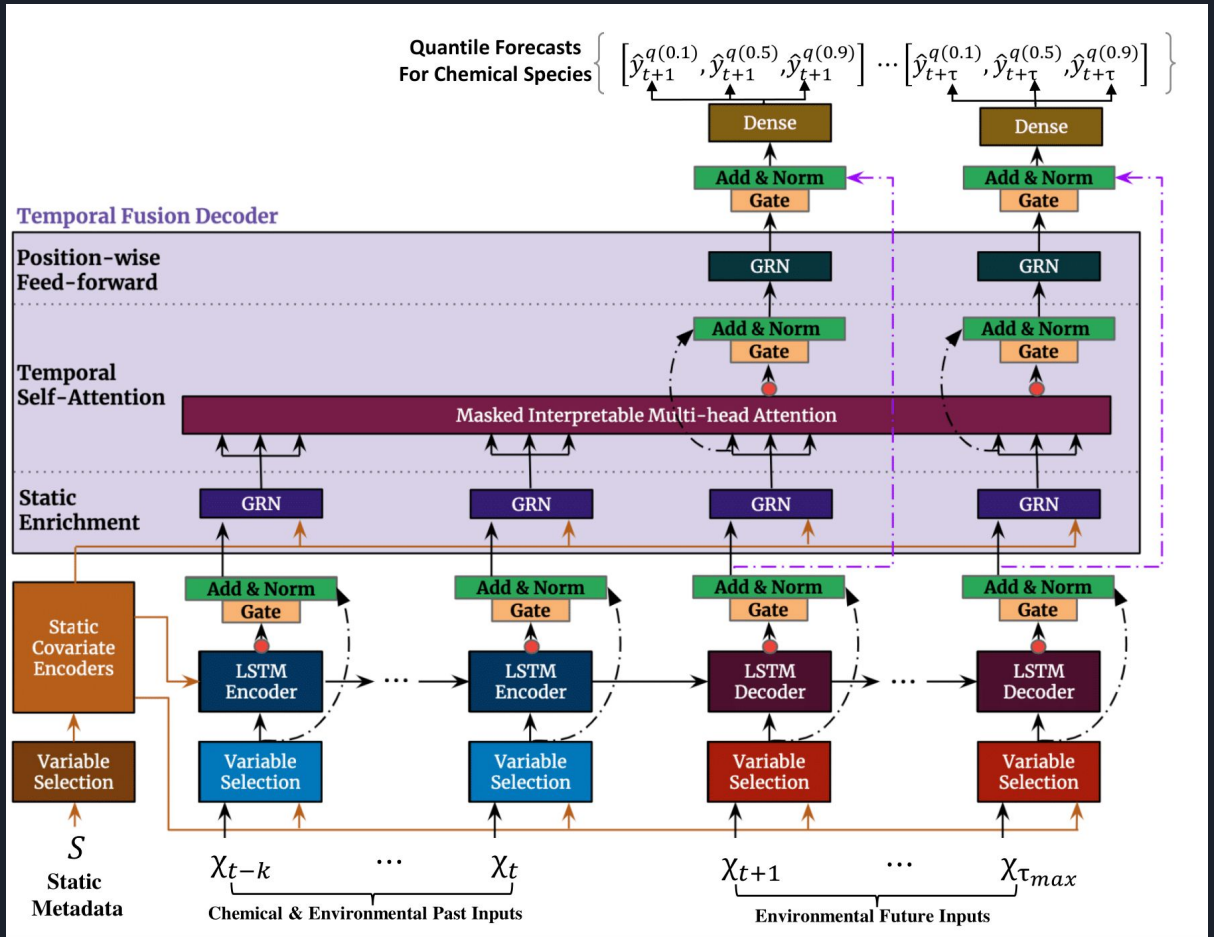
- The long term goal of our project is to develop an emulator for the numerical solver in the UKCA atmospheric chemistry model, to replace it within the UKESM Earth system model.
- The simulation we use for this work is MOZART-4, a tropospheric box model (representing a single grid cell in a global model) of similar complexity to UKCA.
  - ~80 species and ~200 reactions.
- We use BOXMOX to sample 5000 runs of MOZART-4 across many different initial conditions and environmental parameters.
  - Each run is 3 days of simulation time.

Transformer Emulation

- Our model is a Temporal Fusion Transformer [1] — an attention-based neural network capable of multi horizon time series modeling.
- Consists of a multi-head attention module and variable selection networks to select the important features at each time step.
- Model outputs forecasts across different percentiles (10th, 50th, 90th, etc) and minimises loss across all.

[1] Lim, Bryan, et al. "Temporal fusion transformers for interpretable multi-horizon time series forecasting." International Journal of Forecasting 37.4 (2021)

Temporal Fusion Transformer

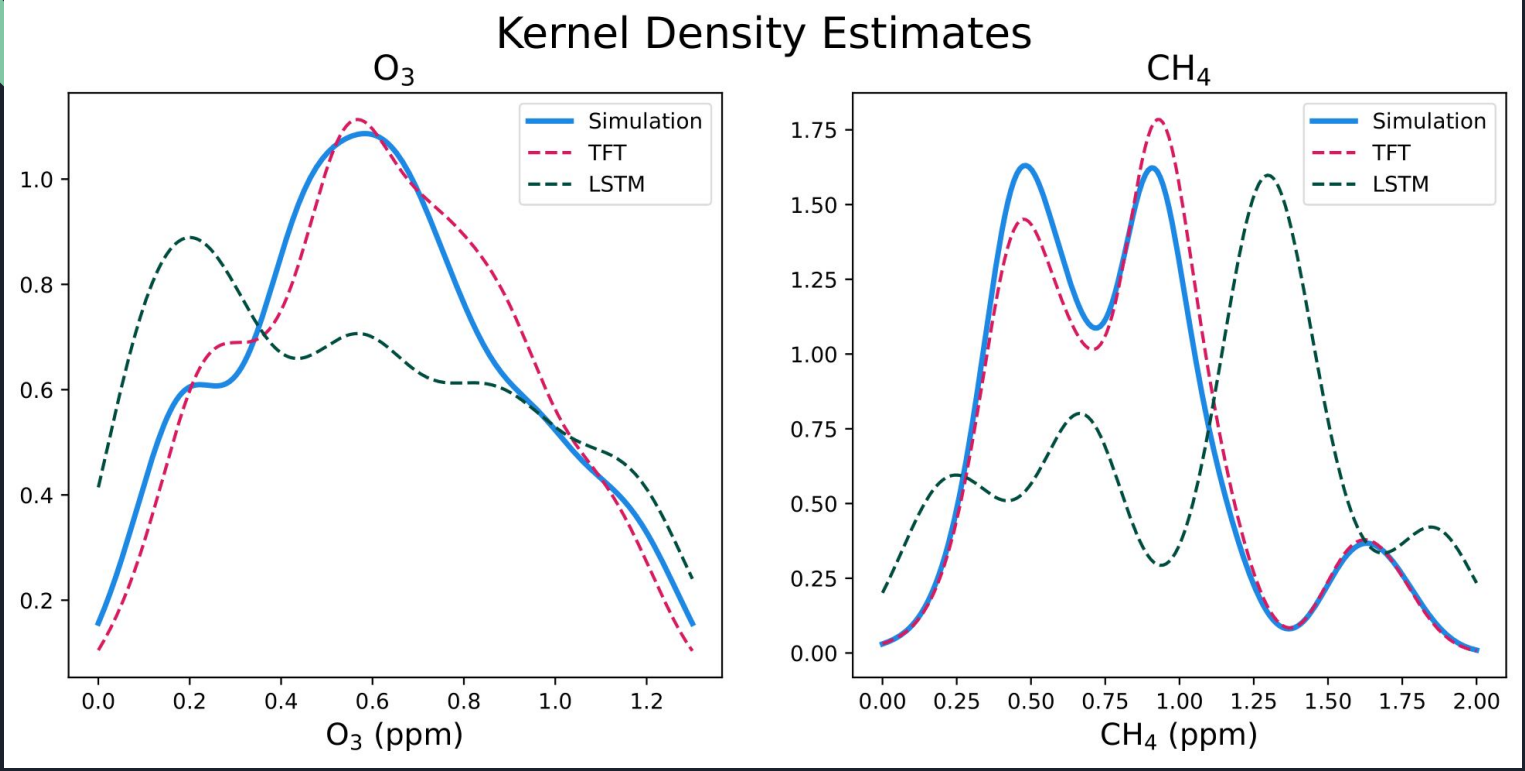


Results

Model	MSE	MAE	R^2	O_3 error (ppm)
TFT	**0.0274**	**0.1672**	0.965	**0.1238**
Autoencoder [10]	0.0306	0.1980	0.958	0.1429
LSTM	0.0289	0.2884	**0.971**	0.2087
Random Forest	0.0297	0.2721	0.916	0.2925

- **Accuracy:** We find that the TFT outperforms comparable time series forecasting baselines on the MOZART-4 dataset by **~8%,** as well as providing the extra quantile forecasts.
- **Computational cost:** creating the MOZART-4 dataset (5000 runs of 3 days each) took ~30 minutes to simulate on a single 32-core CPU, while predictions for the same data using the TFT model took ~20 seconds on a V100 GPU.
- This demonstrates that on a computing cluster consisting of mostly CPUs with a small number of GPUs, as is typical in climate modelling, this emulator setup can potentially achieve orders of magnitude speedups for global atmospheric chemistry modelling.

Results - 2



Un-normalized kernel density estimates for Ozone and Methane (in parts per million), using the TFT and LSTM baseline for emulation predictions.

Conclusions & Future Work

- Our results show that a transformer specialised for time-series modeling can supply significantly improved predictions for a chemical box model emulation task over baselines, and can greatly speed up the simulation at the cost of some accuracy.
- Directions for future work:
  - Physics-constrained NNs—incorporating information about the chemical system into the loss function may enable more chemically plausible outputs
  - Partial emulation—emulate some parts of the chemical system and numerically solve others in order to increase emulator accuracy while still providing significant speedups.
  - Global model emulation—incorporate our box model emulator into a global atmospheric chemistry model.