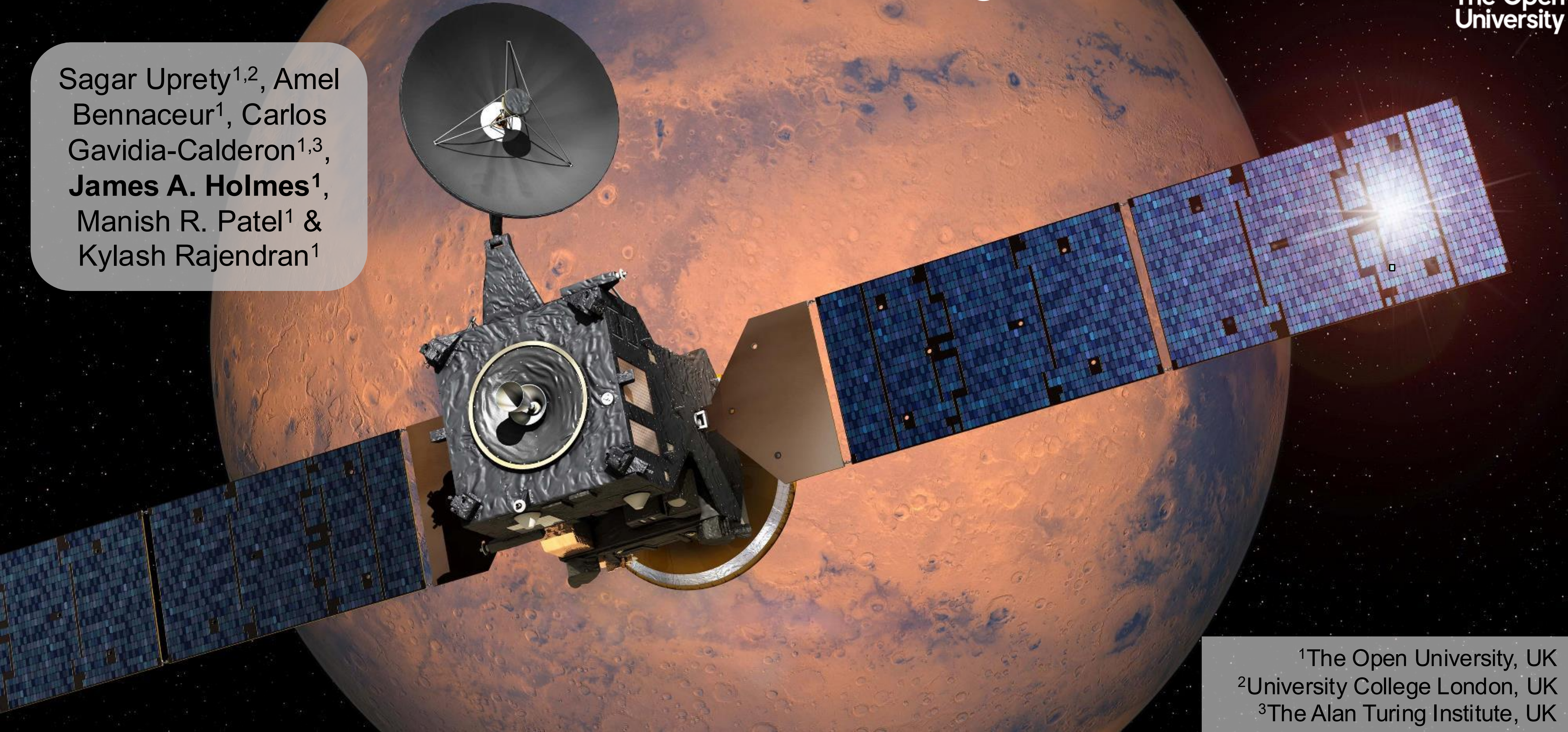
Weather Prediction on Mars as a Multivariate Time Series Forecasting Problem



Sagar Uprety[1,2], Amel Bennaceur[1], Carlos Gavidia-Calderon[1,3], **James A. Holmes[1]**, Manish R. Patel[1] & Kylash Rajendran[1]

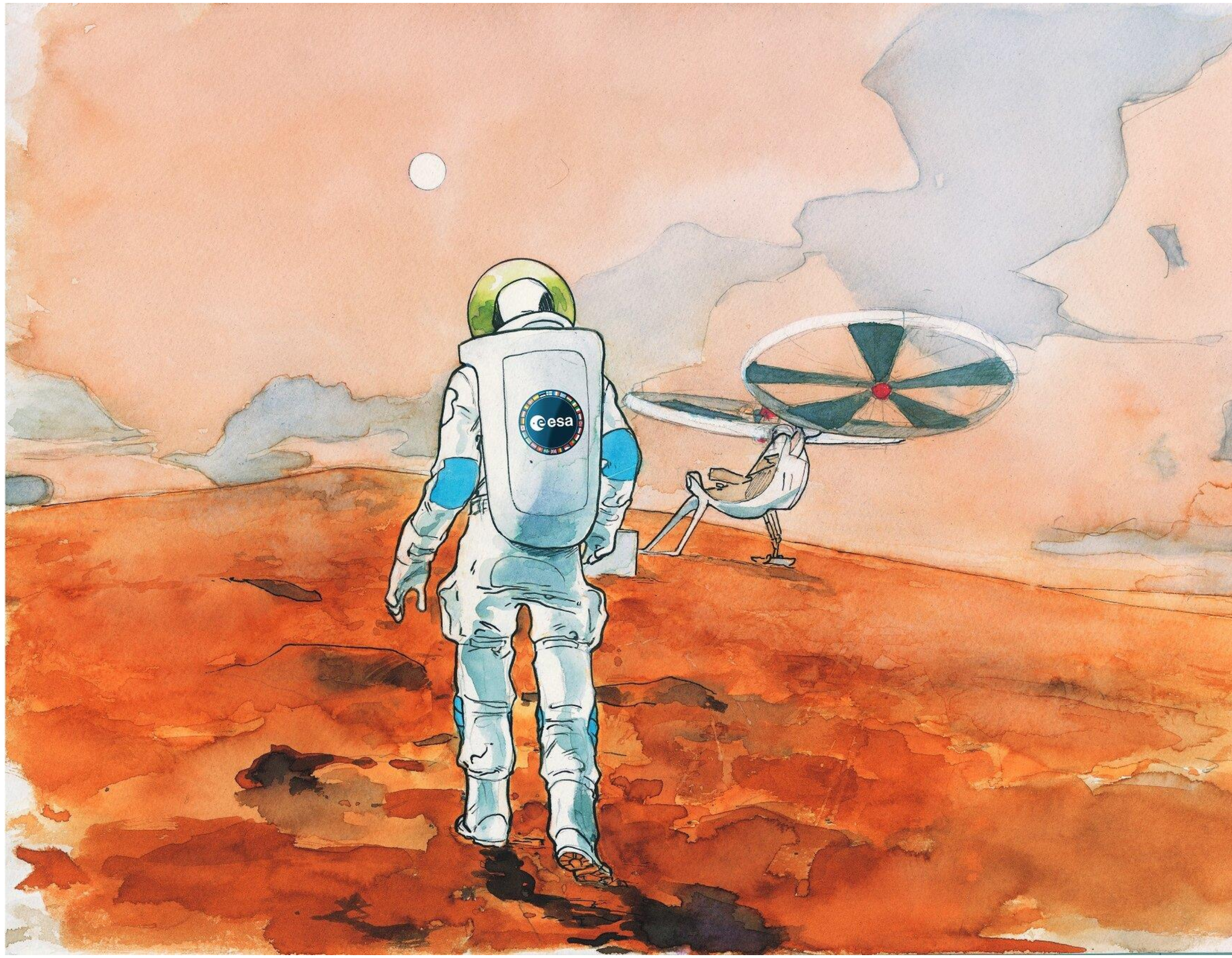
[1]The Open University, UK
[2]University College London, UK
[3]The Alan Turing Institute, UK

Motivation

- Accurate weather prediction on Mars is imperative for the safety of future human explorers

- Conventional physics-based numerical weather prediction face challenges due to sparse observational data and intricacies of the martian atmosphere

- We propose to use Machine Learning (ML) to forecast martian weather using the OpenMARS reanalysis dataset



Research questions

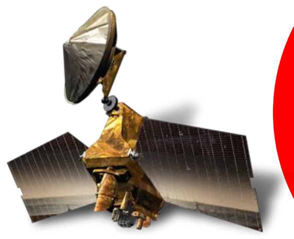
- Can an ML model provide a rapid forecast of the expected local weather on a par with complex physics-based numerical weather prediction models?

- Can an ML model predict an impending global dust storm?

- Do some particular ML models outperform others at the above tasks?

What is OpenMARS?

Credit: ESA/ATG medialab

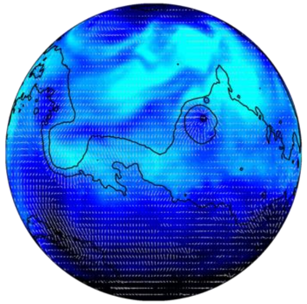


Credit: mars.nasa.gov/mro

OpenMARS global reanalysis dataset

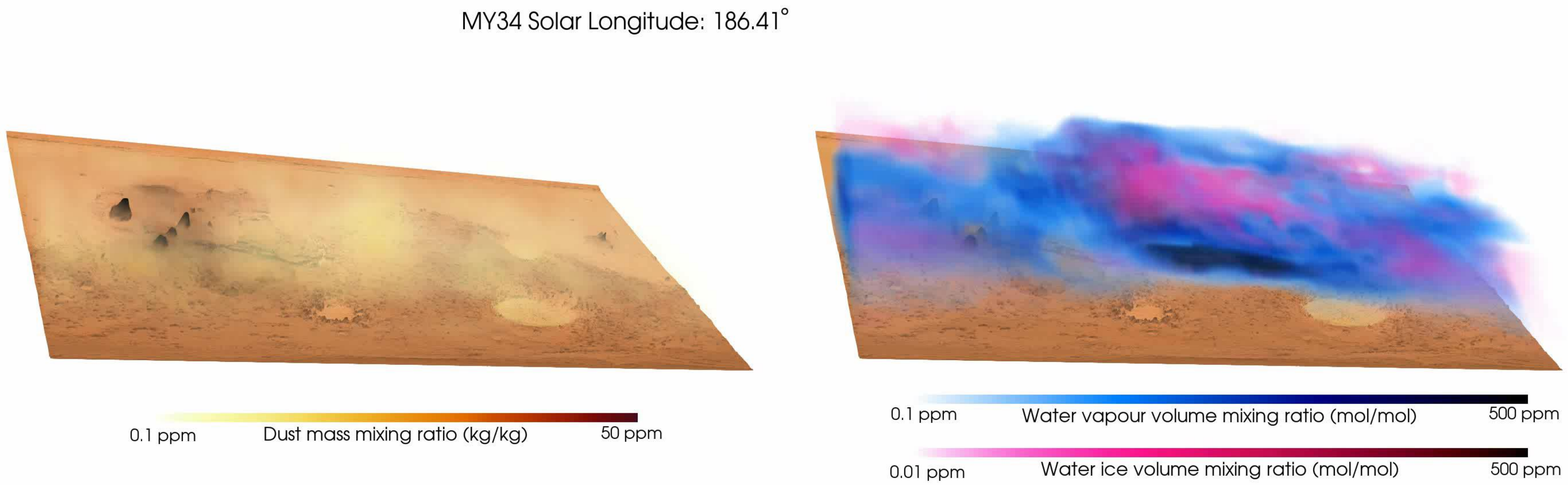
Observations
- Trace Gas Orbiter
- Mars Climate Sounder

Global Circulation Model



- Open access to Mars Assimilated Remote Soundings (OpenMARS) dataset.
- Reanalysis product combining past spacecraft observations with a Mars Global Circulation Model
- Global 4-D surface/atmosphere reference database of surface and atmospheric properties from July 1998 to April 2019 (equivalent to around eleven Mars years).

Dust/water during a Mars Global Dust Storm



What data are we using from OpenMARS?

- Focus on prediction of:
  - surface pressure
  - surface temperature
  - near-surface zonal wind speed
  - dust column

Variable	Description	Unit
Tsurf	Surface temperature	Kelvin
Psurf	Surface pressure	Pascals
Cloud	Water ice optical depth at infrared wavelength	No unit
Vapour	Water vapour column	kg/m^2
u_wind	Near-surface (~4 m) zonal wind	m/s
v_wind	Near-surface (~4 m) meridional wind	m/s
Dust	Dust column optical depth at visible wavelength	No unit
Temp	Atmospheric temperature at ~20 km altitude	Kelvin

First few rows of the cleaned dataset

Time	Tsurf	Psurf	Cloud	Vapour	u_wind	v_wind	Dust	Temp
1998-07-15 21:23:39	264.042	721.113	0.092	0.027	-7.451	8.604	0.428	179.686
1998-07-15 23:26:53	274.736	705.090	0.145	0.026	-7.053	4.934	0.427	174.502
1998-07-16 01:30:07	265.939	700.691	0.105	0.026	-6.825	-0.063	0.427	173.429
1998-07-16 03:33:21	238.624	697.252	0.134	0.025	-5.373	-4.048	0.426	173.556
1998-07-16 05:36:35	213.634	717.146	0.139	0.026	-3.899	-3.133	0.426	174.789

What coverage does OpenMARS contain?

- Temporal resolution of the dataset is every two hours and in total there are 88,560 data points for each variable

- Choose one atmospheric column from the complete OpenMARS spatial grid (5º lat-lon) corresponding to InSight landing site location for local forecasting

- The dataset covers the timespan from the start of Mars Year 24 to just after the end of Mars Year 34, equivalently from 15th July 1998 to 23rd April 2019.

- Split ratio of 70, 20 ,10 for training, validation and testing of ML models

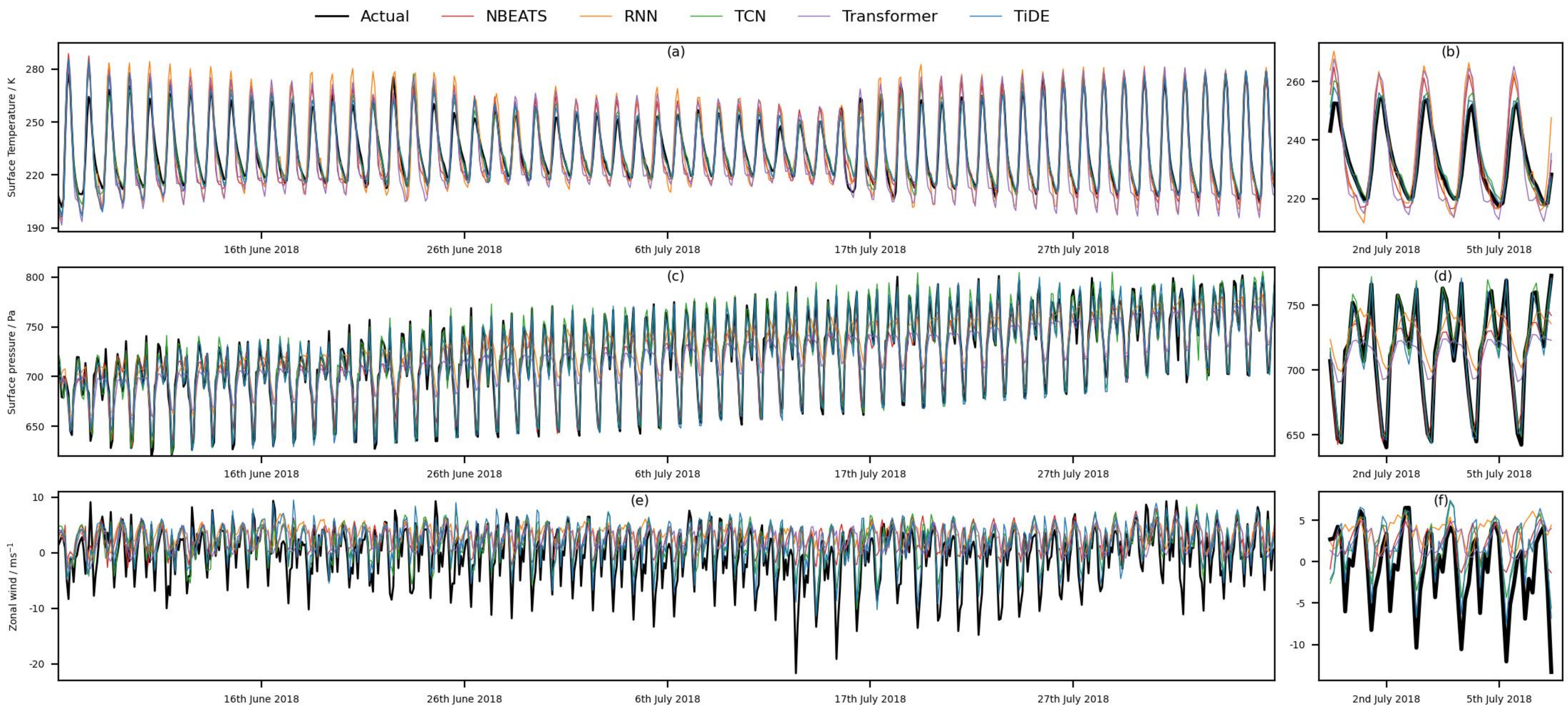
What ML models are we using?

- Darts python library for training the time series forecasting ML models
- Input_chunk_length = a week of Mars data
- Output_chunk_length = one day of Mars data

Best hyperparameter values

Model	Hyperparameters
RNNModel	n_rnn_layers=2, hidden_dim=30, batch_size=96, dropout=0.25, learning_rate=0.0005
TCNModel	kernel_size=2, num_filters=6, dilation_base=2, batch_size=96, dropout=0.05, learning_rate=0.0005
Transformers	d_model=12, n_head=6, num_encoder_layers=2, num_decoder_layers=4, dim_feedforward=64, batch_size=96, dropout=0.05, learning_rate=0.0005
NBEATS	num_blocks=3, num_layers=4, layer_widths=512, batch_size=96, dropout=0.05, learning_rate=0.000953
TiDe	num_encoder_layers=2, num_decoder_layers=2, decoder_output_dim=1, temporal_decoder_hidden=1, batch_size=96, dropout=0.05, learning_rate=0.0005, hidden_size=30

ML forecasting results



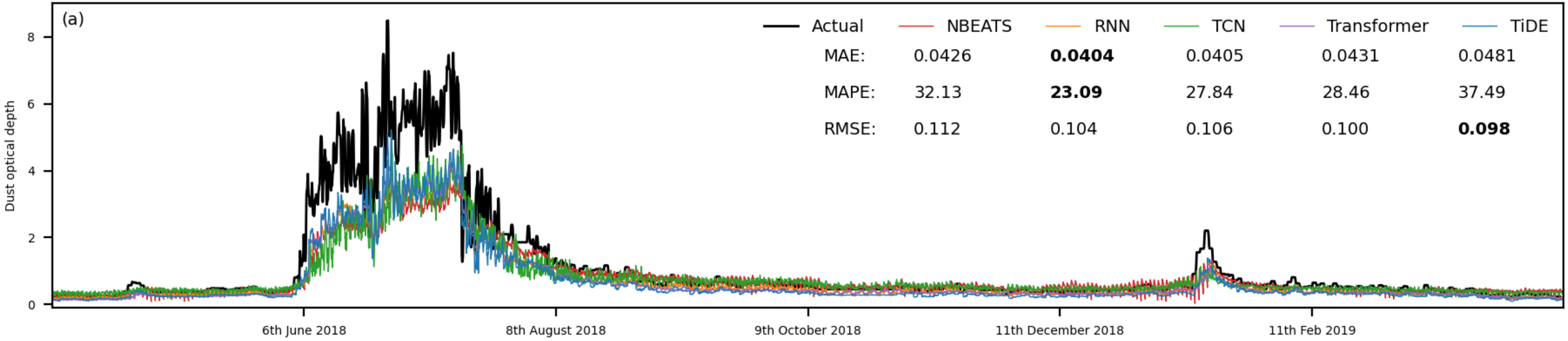
ML forecasting metrics

- Three different ML models perform best for the three different variables predicted
  - Surface temperature: **TCN**
  - Surface pressure: **TiDE**
  - Zonal wind: **NBEATS**

- Overall **TCN** model is perhaps the best model at predicting across all three variables and showed potentially interesting results for dust column optical depth too….

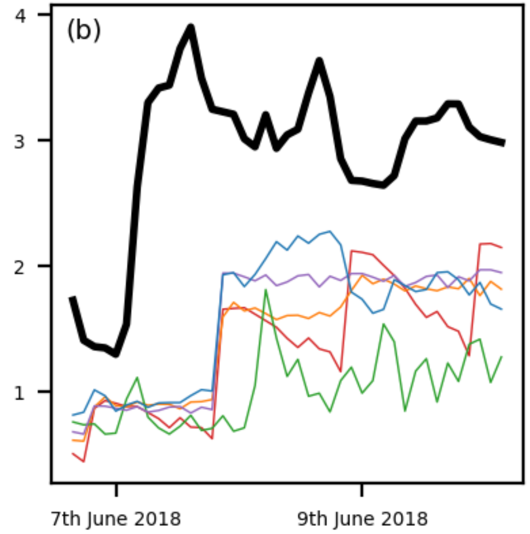
Model	Variable	MAE	MAPE	RMSE
RNN	Tsurf	0.0247	11.07	0.0341
	Psurf	0.0275	6.02	0.0407
	u_wind	0.0646	16.67	0.0901
TCN	Tsurf	**0.00602**	**2.54**	**0.0109**
	Psurf	0.0130	2.67	0.0172
	u_wind	0.0612	15.25	0.0764
Transformer	Tsurf	0.0281	14.25	0.0362
	Psurf	0.0279	6.01	0.0408
	u_wind	0.0605	15.42	0.0861
NBEATS	Tsurf	0.0188	8.71	0.0258
	Psurf	0.0321	6.16	0.0404
	u_wind	**0.0534**	**13.76**	**0.0751**
TiDE	Tsurf	0.00699	2.99	0.0126
	Psurf	**0.0118**	**2.48**	**0.0160**
	u_wind	0.0665	16.14	0.0781

Case study: ML forecasting of a dust storm

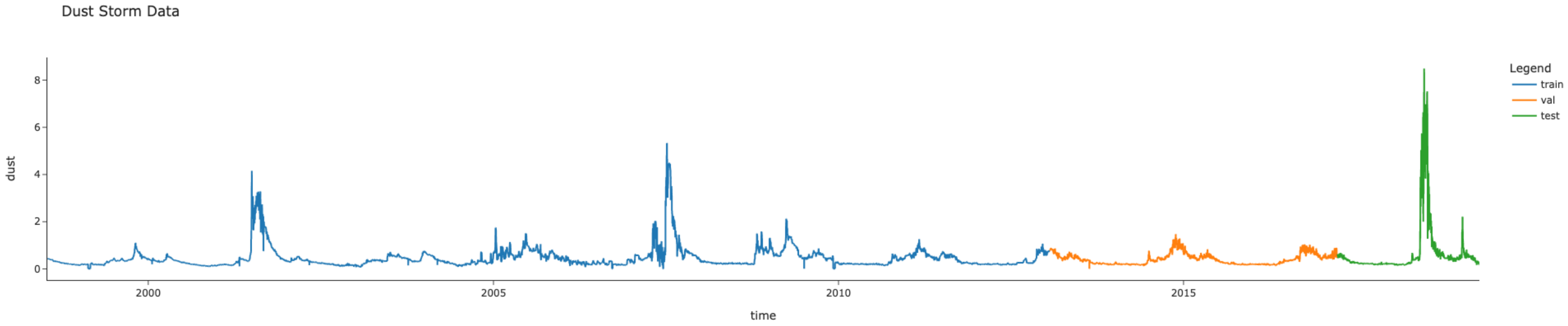


- Not a huge amount of intrinsic variability between the different ML model forecasts

- TCN model shows abrupt increase just as the global dust storm is initiating in reality (within the noise?)



Why is it difficult to forecast a global dust storm?



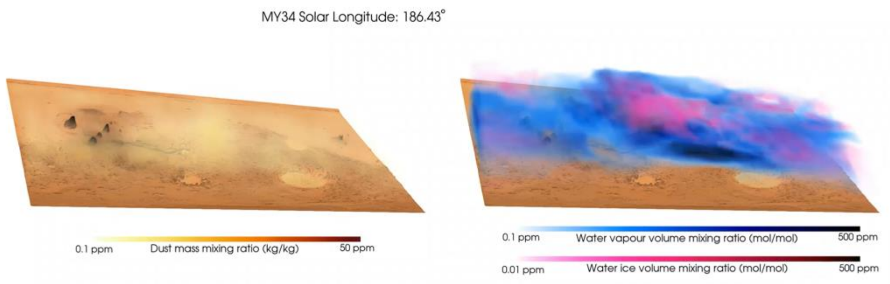
- There were only two instances of global dust storms in training dataset over the nine Mars years included

- The three global dust storms in the entire dataset occur at different times of the year

Summary

- TCN and TiDE ML models most efficient in reproducing realistic tidal structures in surface pressure

- Dust storms are difficult to predict on Mars!



- Future work:
  - Testing ML models on actual data alone (e.g. InSight lander)
  - Extend the training data through including more recent reanalysis data
  - Open question: are there other known methods to tackle dust forecasting?

- Replication package and benchmarking results are publicly accessible here: https://github.com/amelBennaceur/OpenMarsML

OpenMARS database version 5 →