Retrieving Pressure-Temperature and Water Vapour Profiles in Earth's Atmosphere from INSAT 3DR data using Machine Learning

Fida Salim
Soumik Bhattacharya

Under Supervision of Dr.Subhankar Mishra and Dr.Jayesh Ghoyal

RECAP

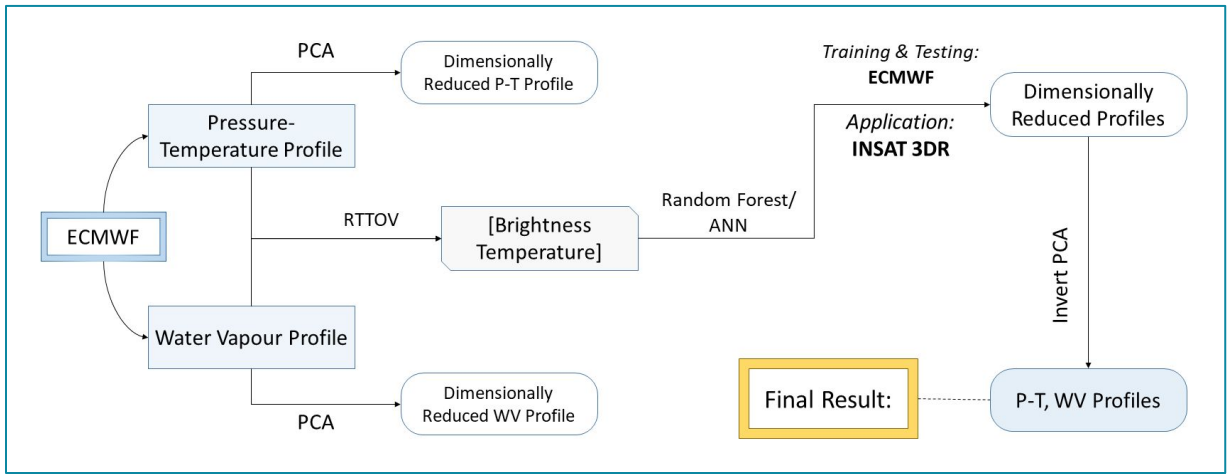
Owing to lack of in-situ measurements of sea surface temperature, brightness temperature (BT) derived from satellite data is used in regression based models to retrieve actual SST.

In this project, we will use in-situ measured pressure-temperature (P-T) and water vapor (w-v) profiles from ECMWF, and use a radiative transfer model (RTTOV) to derive BT for each corresponding profile.

The main part of the project involves developing a machine learning algorithm to model the P-T and w-v profiles from BT and use the model on real time data of INSAT-3DR satellite images.

PROBLEM STATEMENT!

Whether the Water-Vapour profile or Temperature profile can be obtained from Brightness Temperature?



Literature Survey

The relationship between water vapor and sea surface temperature has long been a topic of interest.

- *Stephens (1990)* provides an useful literature review in this field. He also presents his own analysis of satellite observations to illustrate the spatial and temporal patterns of this relationship.
- ML is adopted in the field of atmospheric science mostly for applications related to forecasting.
- Some of the previous similar works include *Tripathy et al., 2006* (ANN to predict SST anomalies in Indian Ocean); *Xiao et al., 2019* (Long short-term memory (LSTM) and AdaBoost to predict SST) etc.
- In the next slide, we'll look into a more recent paper, *Sarkar et al., 2020*, in detail.

Literature Survey (cont.)

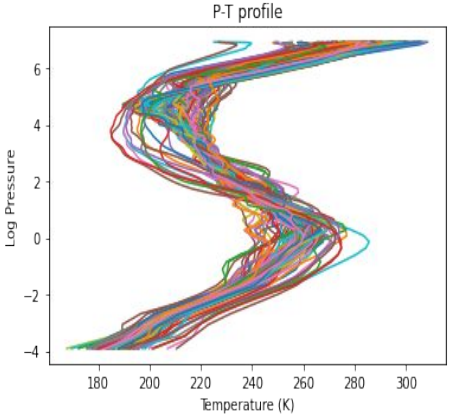
The paper by *Sarkar et al., 2020* proposes a deep learning approach to forecast sea surface temperatures (SST) of five different locations from Indian Ocean.

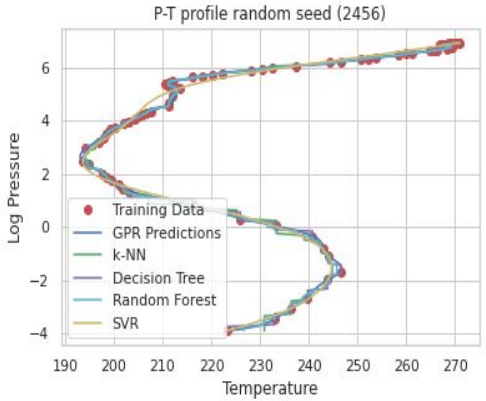
- The ML models they used are Long short-term memory (LSTM), a special type of Recurrent Neural Network and Autoregressive integrated moving average with exogenous input (ARIMAX).
- Additionally, the authors perform various regularization techniques and sensitivity analysis respectively to prevent overfitting of the model and to explore the impact of different input variables on the model's performance.
- Although their model outperforms other traditional ML models in use, they acknowledge that the model is not well-suited to predict SST for a long-term forecast.
- A reason for inaccuracy in long-term prediction is the propagated errors from previous iterations in these time series regression models.

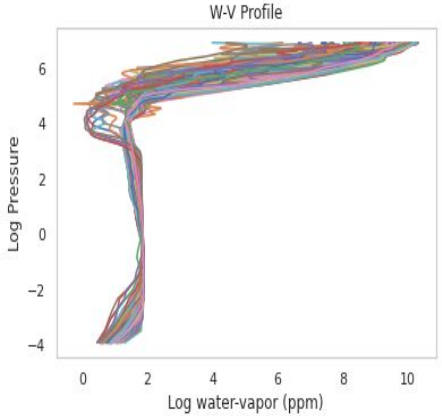
BASELINE IMPLEMENTATION

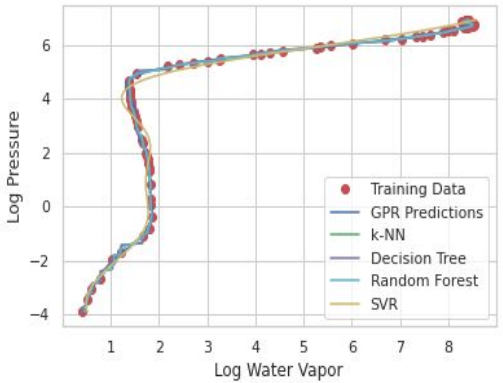
Regression Models

- Gaussian Process Regression
- k-NN regression
- Support Vector Regression
- Decision Tree Regression
- Random Forest Regression





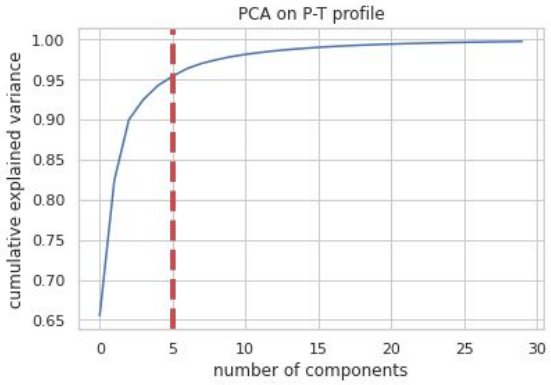




BASELINE IMPLEMENTATION (cont.)

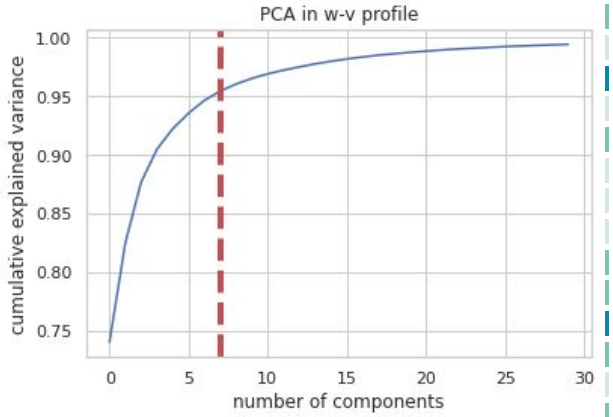
Dimensionality Reduction

- PCA



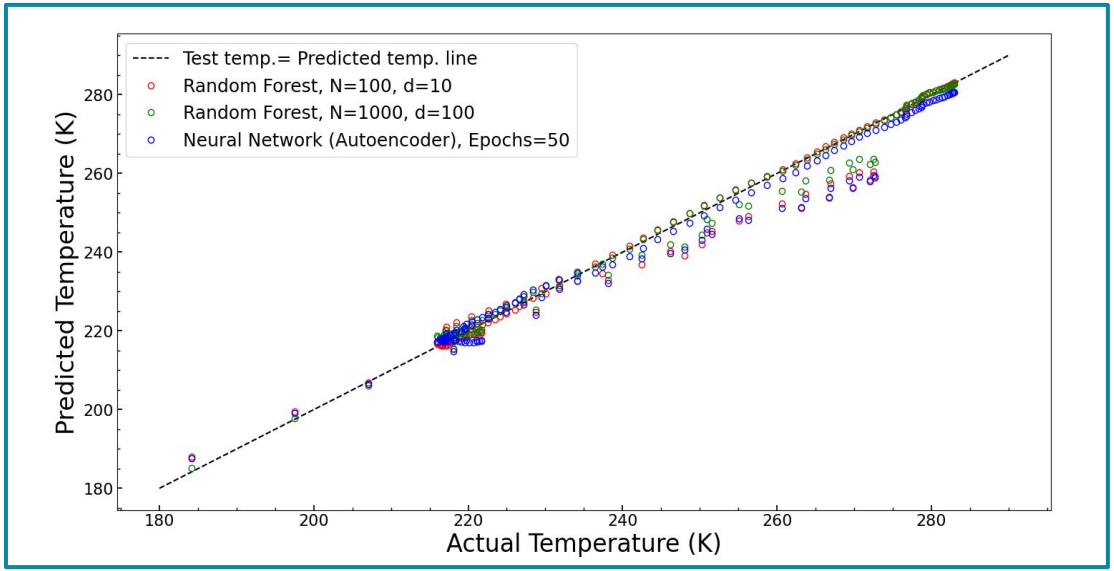
The first 5 and above principle components provide exceeds 95% cumulative explained variance in P-T profile.

The first 7 and above principle components provide exceeds 95% cumulative explained variance in W-V profile.



Modelling Temperature Profile from BRIGHTNESS Temperature

- Random Forest
- Artificial Neural Network

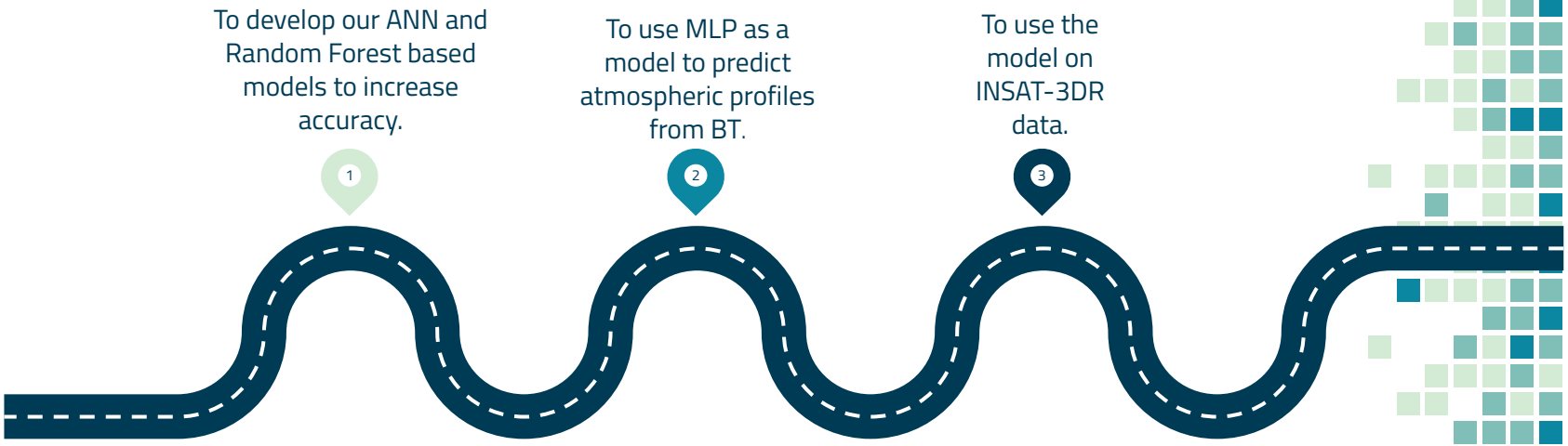


No of Trees (N)	Nodes per tree (d)	Model Score (%)	Error in Temp (%)
100	10	61.58	1.30
1000	100	70.55	1.05

Summary

- We downloaded Pressure-Temperature (PT) and Water Vapour (WV) profiles from ECMWF and applied RTTOV model to get the corresponding Brightness Temperatures
- Different Regression Models were used to parameterize the PT and WV profiles
- PCA was used to reduce dimension of the profiles. The minimum number of components required to achieve 95% cumulative explained variance is 5 and 7, respectively for PT and WV profiles
- Random Forest Regressor and Autoencoder were used as initial approaches to model temperature from brightness temperature. Preliminary results show that the Random Forest based model outperforms the ANN based model.

Future Plans



THANKS!

Any questions?