Comparative analysis of South Indian climate and forecast of certain key features

Jeffrey S Varghese
Department of CSE
PES University
Bangalore, India
jeff0716042@gmail.com

Thejaswi A
Department of CSE
PES University
Bangalore, India
thejaswia03@gmail.com

Rachitha Suresh
Department of CSE
PES University
Bangalore, India
rachithas.08@gmail.com

Abstract**—** In this paper, we aim to compare the climates of South Indian cities- Chennai and Mangalore. Libraries in python were used to perform exploratory data analysis (EDA) on the weather data obtained from the NASA POWER Project. We plan to implement machine learning methods to forecast features like temperature and rainfall of these cities.

Keywords—climate, EDA, South India

I. INTRODUCTION

Climate in South India is unique compared to other regions of the country characterized by its high humidity and temperature values throughout the year. The climate is known to be quite stable and extreme climatic conditions are not usually observed in any of the South Indian cities. But lately, incidents like heatwaves, persistent floods, unusual number of cyclones and irregular rainfalls have proven that South India is one of the most vulnerable regions to climate change. How does this affect the livelihoods of people in South India? Cities like Bangalore where the main driving sector of the economy is IT, traffic caused by unprecedented amounts of rain holds back the employees to the extent where the loss observed by the companies is immense. Hyderabad observed heatwaves on 10 different occasions this year alone showing that the effect of climate change and global warming on cities in South India is worth notice. These are just a few among tens of such incidents that made it to the news.

In this project we aim to make a comparative analysis of climates of different South Indian cities and forecast some of the features of weather using appropriate models.

II. RELATED WORK

There are many published papers related to weather forecasting that consider models to perform classification and predict weather state. Concepts of physics, differential equations, Artificial intelligence and Machine learning are used to build models to forecast the weather. Deep learning models perform better than the traditional machine learning models as they can extract features of the past data to predict future weather conditions. Thus, many papers utilized neural networks (deep learning) and other works use other machine learning techniques.

Abrahamsen et al. [2] investigated the effects of an artificial neural network on the prediction of weather conditions using large data sets. The authors used multiple sources of meteorological data online. They use the Python API to read meteorological data and ANN models have been developed using TensorFlow framework. Kumar Abhishek et al. [6], develop a nonlinear predictive model using ANN to analyse the weather. They used different transfer functions to forecast the maximum temperature for 365 days of the year. Their work can be used to study trends and weather patterns over a long period of time in a particular area.

Guerra et al. [3] suggest a weather classification model based on convolutional neural networks, which were trained using new multi-class dataset. Sawaitul et al. [4] propose a weather classification and forecasting models using back propagation feedforward neural network. Lu et al. [5], developed a binary classification model for weather prediction based on SVM classifier even though the accuracy of the model was lower than the neural network approaches. Ibrahim Gade et al. [1] performed a comparative study of different machine learning methods including linear regression, support vector machine (SVM), decision tree ,linear discriminant analysis (LDA), Gaussian NB, random forest k-nearest neighbors (KNN), AdaBoost, Extreme Gradient Boosting (XGBoost), artificial neural networks (multi-layer perceptron MLP), and deep learning to predict or classify weather states using the NCDC weather data. They concluded that decision tree , XGB, and AdaBoost methods outperform other methods in terms of classification accuracy and KNN model, random forest and XGBoost models works best on the training data for weather forecasting , but these models showed a decreasing performance with the test data indicating overfitting.

Mohsen Hayati et al. [7] have used Multi-Layer Perceptron in short term temperature forecasting. They trained and tested their model with 10 years of data observed in Kermanshah city, west of Iran. In their model they use a three-layer network with sigmoid transfer function for hidden layer and linear transfer function for output layer which can represent any functional relationship between inputs and outputs, if the sigmoid layer has enough neurons. From an optimization point of view learning in a neural network is equivalent to minimizing a global error function, which is a multivariate function that depends on the weights in the network. The result of MLP network model used for one day ahead temperature forecast in the Kermanshah, Iran, showed that MLP network has a good performance and reasonable prediction accuracy. It's forecasting reliabilities were evaluated by computing the mean absolute error between the exact and predicted values. The results suggest that this neural network could be an important tool for temperature forecasting.

III. PROBLEM STATEMENT

Comparative analysis of South Indian climate and forecast of certain key features.

A. Dataset

We obtained our dataset from the NASA POWER Project. The POWER project by NASA targets three user communities. Agroclimatology, Renewable energy, and Sustainable buildings. The parameters vary based on the targeted user communities. We chose the Renewable energy community. We had options for hourly, daily, monthly, yearly, and climatology weather/climate data. We chose hourly data because higher granularity in data is essential for better analysis and forecasting of weather. With regard to our problem statement, we have taken the data of 2 south Indian cities, Chennai and Mangalore. Our dataset has weather details of both the cities from 01-01-2012 to 02-01-2022 and consists of 87697 rows and 9 attributes which are represented in Figure 1.

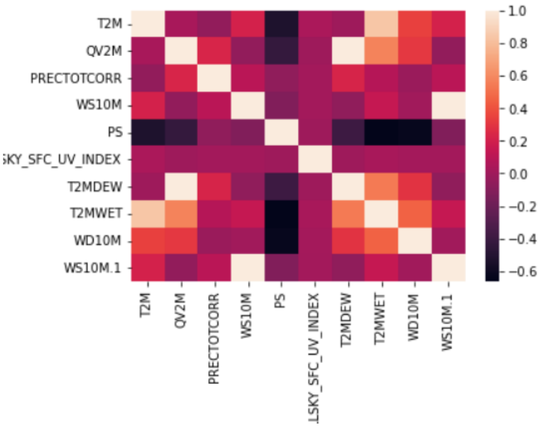
Sl.No	Parameter	Abbreviation	Definition
1	Temperature at 2 Meters	T2M	The average air (dry bulb) temperature at 2 meters above the surface of the earth.
2	Specific Humidity at 2 Meters	QV2M	The ratio of the mass of water vapor to the total mass of air at 2 meters (g water/kg total air).
3	Precipitation Corrected	PRECTOTCORR	The bias corrected average of total precipitation at the surface of the earth in water mass (includes water content in snow).
4	Wind Speed at 10 Meters	WS10M	The average of wind speed at 10 meters above the surface of the earth.
5	Surface Pressure	PS	The average of surface pressure at the surface of the earth.
6	All Sky Surface UV Index	ALLSKY_SFC_UV_INDEX	The ultraviolet radiation exposure index.
7	Dew/Frost Point at 2 Meters	T2MDEW	The dew/frost point temperature at 2 meters above the surface of the earth.
8	Wet Bulb Temperature at 2 Meters	T2MWET	The adiabatic saturation temperature which can be measured by a thermometer covered in a water-soaked cloth over which air is passed at 2 meters above the surface of the earth.
9	Wind Direction at 10 Meters	WD10M	The average of the wind direction at 10 meters above the surface of the earth.

Fig 1. Description of attributes of the dataset.

B. *Exploratory Data Analysis*

It was found that all 9 of the attributes were of type – float64 and one column representing the date has datetime64 as its datatype. Null values were found only in the ultraviolet exposure radiation index. 44746 and 44818 null values were found in Chennai and Mangalore datasets.

It is observed from Figure 2 that Specific humidity (QV2M) and Dew/Frost point (T2MDEW) have a strong positive correlation, so we can drop one of them. We chose to drop Dew/Frost point. The correlation can be observed in Figure 6 as well.



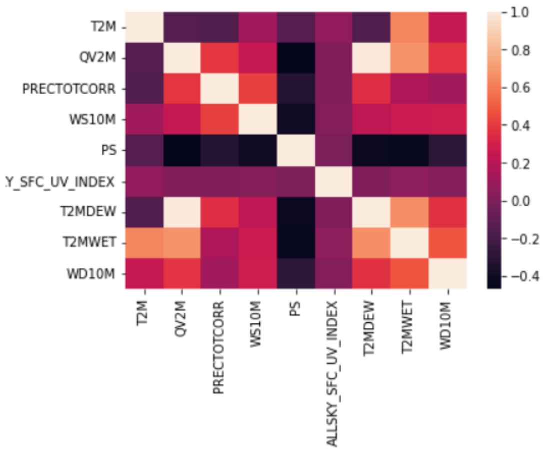


Fig. 2 Heatmaps of Chennai and Mangalore datasets.

It can also be inferred from Figure 2 the surface pressure of Chennai has a strong negative correlation with Temperature, specific humidity, and wind direction. The surface pressure of Mangalore has a strong negative correlation with wind speed and precipiation as well.

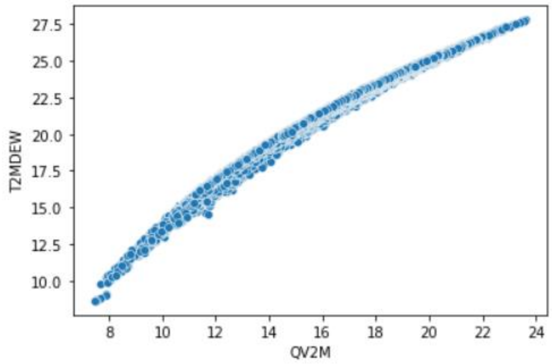


Fig 6. Scatter plot between Dewfrost point and Specific humidity.

South Indian cities are known to have a Tropical climate and this is evident from the temperature plots of our data set. The range of temperature in Chennai is 20 º C which is typical of a tropical climate which is characterised by short range of temperatures.

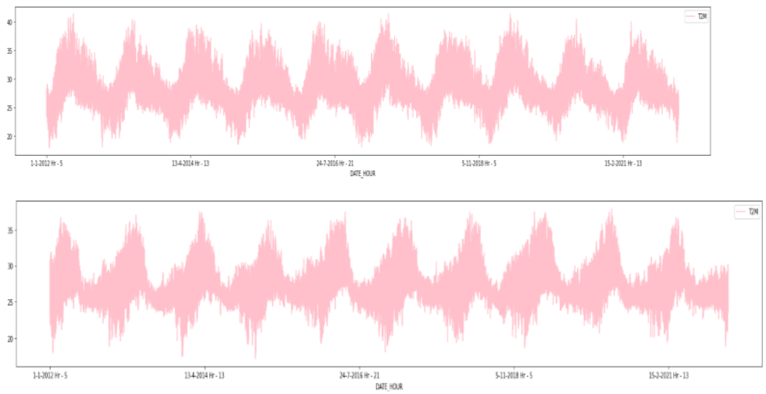


Fig 3. Temperature Variations over the years in Chennai and Mangalore.

Chennai and Mangalore have similar elevations and lie on the same latitude, both of them are coastal cities and have similar geographical features which should imply that they must have similar weather observations. Figure 4 is an indication of the same.

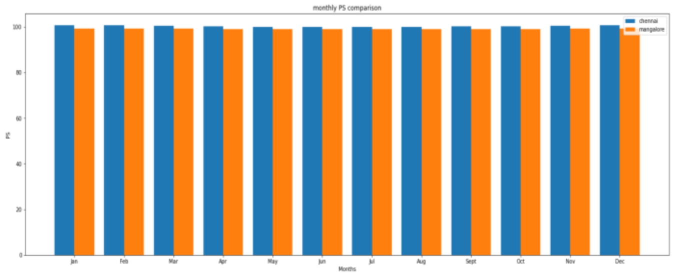


Fig 4. Comparison of surface pressure between Chennai and Mangalore.

Figure 5 compares the monthly total rainfall between Chennai and Mangalore., but the pattern of precipitation differs, this is because Mangalore receives rainfall from the southwest monsoon winds(June – Sept) as opposed to Chennai which lies in the rain-shadow region of the Western Ghats. Chennai gets rainfall from the northeast winds in winter because Chennai faces the northeast monsoon(Oct-Dec).

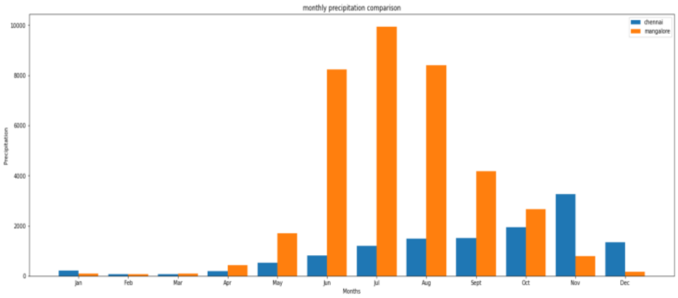


Fig 5. Comparison of rainfall between Chennai and Mangalore.

References

[1] Gad, I. and Hosahalli, D., 2022. "*A comparative study of prediction and classification models on NCDC weather data*"

[2] Abrahamsen, Erik et al. "Machine Learning In Python For Weather Forecast Based On Freely Available Weather Data". *Linköping Electronic Conference Proceedings*, 2018. *Linköping University Electronic Press*, https://doi.org/10.3384/ecp18153169.

[3] Villarreal Guerra, Jose Carlos et al. "Weather Classification: A New Multi-Class Dataset, Data Augmentation Approach And Comprehensive Evaluations Of Convolutional Neural Networks". *2018 NASA/ESA Conference On Adaptive Hardware And Systems (AHS)*, 2018. *IEEE*,

[4] Sawaitul SD, Wagh K, Chatur P. Classification and prediction of future weather by using back propagation algorithm-an approach. Int J Emerg Technol Adv Eng. 2012

[5] Lu C, Lin D, Jia J, et al. Two-class weather classification. IEEE Trans Pattern Anal Mach Intell. 2017 Dec;39(12):2510–2524.

[6] Abhishek, Kumar et al. "Weather Forecasting Model Using Artificial Neural Network". *Procedia Technology*, vol 4, 2012, pp. 311-318.

[7] Hayati, Mohsen & Mohebi, Zahra. (2007). "Application of Artificial Neural Networks for Temperature Forecasting". World Academy of Science, Engineering and Technology. 28. 275-279.