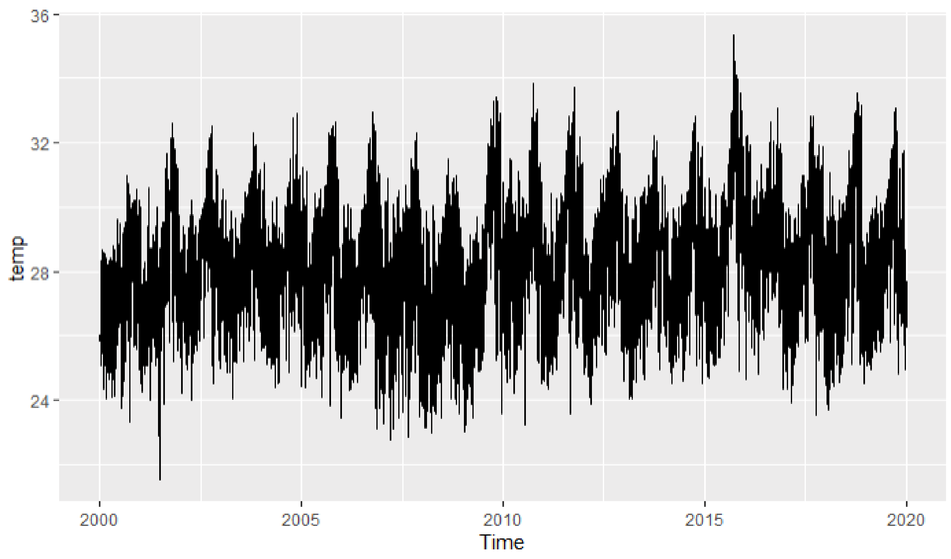
Basic forecasting of daily average temperature in Brazil cities

Hieu Ng. Minh, Long Ng. Hoang
Applied Forecasting course, VNUK Institute for Research and Executive Education
January 2023

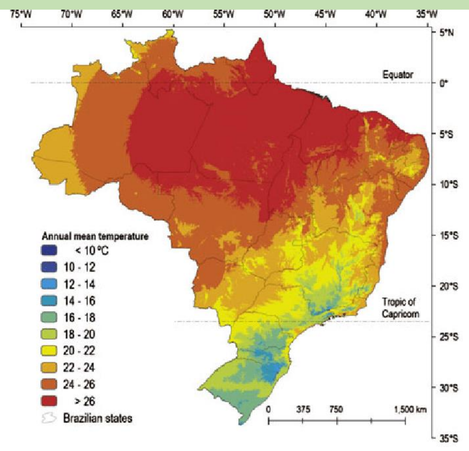


Project Introduction

- Climate change has dramatic impacts on temperature.
- Build basic models to forecast daily temperature based on historical data and other weather elements.
- Daily data from the National Institute of Meteorology - INMET distributed in the Brazilian territory, from 1961-2019.
- Daily temperature of Brazil's cities doesn't depend much on other weather elements.
- Other weather elements depend partially on daily temperature.

Executive Summary

- Global warming has negative influences on temperature-related practices.
- Normal daily temperatures in certain periods are not the same (after 2000 – before 2000)



- Agricultural practices and normal-life practices are affected by temperatures, humidity, atmospheric pressure, wind speed, and cloudiness.
- Measures from different stations have different problems with missing data to deal with.
- Forecasting model built from 1 city in Brazil might be used to forecast for nearby cities only.
- Long-term forecasting is better with deep learning models (RNN or LSTM).
- Short-term forecasting is better with basic forecasting models.
- Among 4 elements – Humidity, Atmospheric Pressure, Wind speed, Cloudiness - Humidity has the largest impact on daily temperature, wind speed has the smallest impact on daily temperature.
- Daily temperature increases slightly through time/
- Temperature doesn't depend strongly on other weather elements.
- More research is needed to find impact factors on daily temperature.

About the data

- The data obtained from INMET is very large and "messy". Data is measured from 265 different stations in different cities and regions in Brazil.
- Every day, data on temperature and weather factors are measured 3 times at 3 time frames.
- Measurements started on January 1, 1961 and ended on December 31, 2019. With 3 measurements per day, 265 stations, the amount of data from the INMET source is really huge.
- The team performed data cleaning by:

-take data every day by the average of 3 measurements per day

-deal with missing data by retrieving data of the same time frame on the previous or subsequent day

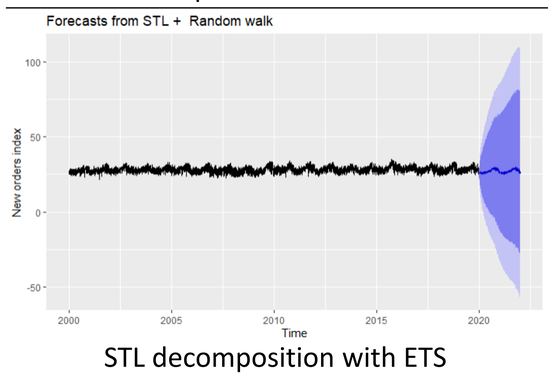
-select data from the most complete measurement station of the day

-handle leap years

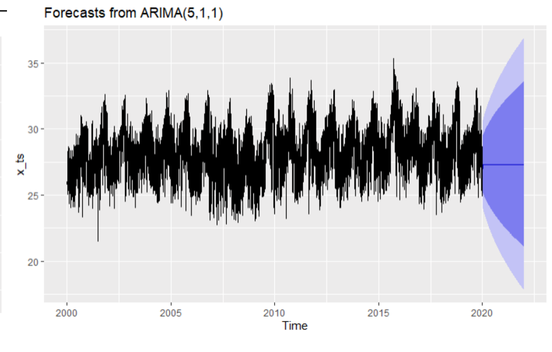
-choose a reasonable period to include in training the predictive model

Forecasting results

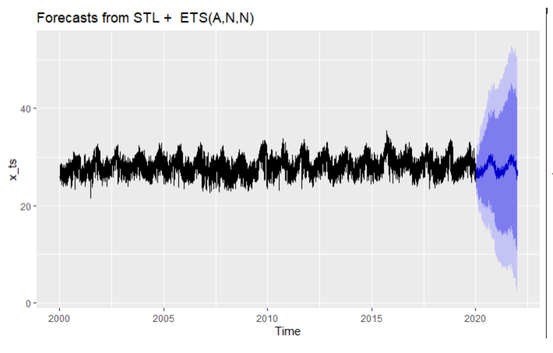
STL decomposition with random walk



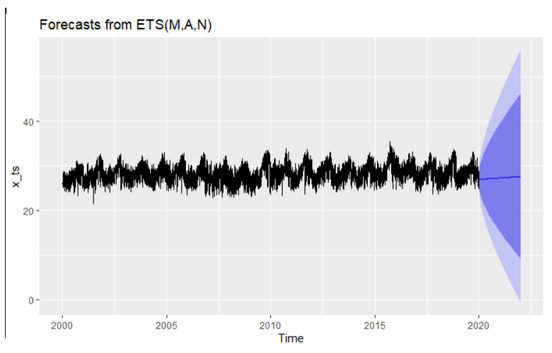
ARIMA



STL decomposition with ETS



ETS



References

1. Forecast of weather parameters using time series data, A. K. ShuklaYogesh GardeYogesh GardeIna Jain, 2014.
2. Transductive LSTM for time-series prediction: An application to weather forecasting, Zahra Karevan, Johan A.K. Suykens, 2020.
3. Time Series Prediction Based on Facebook Prophet: A Case Study, Temperature Forecasting in Myintkyina, Zar Zar OO Sabai PHYU, 2020.