Day-ahead Forecasts of Air Temperature

Hewei Wang*, Muhammad Salman Pathan†‡, Yee Hui Lee§, and Soumyabrata Dev†‡

* Beijing University of Technology, Beijing, China
† ADAPT SFI Research Centre, Dublin, Ireland
‡ School of Computer Science, University College Dublin, Ireland
§ School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore
1

Problem statement

- Problem statement
- Related works
- Contribution of this paper
Air temperature is an essential factor that directly impacts the weather. Temperature can be counted as an important sign of climatic change, that profoundly impacts our health, development, and urban planning. Hence, it is vital to design a framework that can accurately predict the temperature values for considerable lead times.
Related work

Astsatryan et al. implemented several neural network architectures to predict the hourly air temperature for up to 24 hours in the Ararat valley of Armenia.

Chen et al. used SARIMA (Seasonal Autoregressive Integrated Moving Average) techniques to predict the monthly mean air temperature in Nanjing city of China from 1951–2017.

The daily temperature between 1980–2010 for four different European cities was forecasted. Murat et al. used Box-Jenkins and Holt Winters seasonal autoregressive integrated moving-average to forecast the future temperature values.

Manandhar et al. use satellite data to analyze other atmospheric events. They analyze and compare the cloud cover obtained from both satellite-and ground-based-images.
Contributions of this paper

1. We present a robust framework to forecast ground-based air temperature values using historical data.

2. In the spirit of reproducible research, we share the sourcecode of our approach to the community for further benchmarking. The proposed model code, dataset, and experimental results are available at: https://github.com/Soumyabrata/temperature-forecasting
Proposed Method
We represent the air temperature values up to time $t$ as $a_1, a_2, \ldots, a_t$. We use triple exponential smoothing technique to model the seasonality of the temperature values. We model the future air temperature values $a_{t+m}$ as:

$$a_{t+m} = s_t + mb_t + c_{t-L+1+(m-1)} \mod L,$$  \hspace{1cm} (1)

Here, $L$ is the season length, $s_t$ is the smooth version, $b_t$ is the linear trend estimate, and $c_t$ is seasonal corrections. In this paper, we benchmark our proposed method with persistence model and average model. The persistence model assumes that the forecasted temperature value is same as latest value, indicated by $a_{t+m} = a_t$. The average model forecasts the future temperature value as the average of the historical values, indicated by $a_{t+m} = \frac{1}{t} \sum_t a_t$. 
Experiment and Discussion

- Dataset
- Qualitative evaluation
- Quantitative evaluation
We obtain the air temperature data from National Oceanic and Atmospheric Administration (NOAA) Climate Data Online service (CDO2). We choose the weather station that is situated at Alpena Regional Airport, based in Michigan, United States. This data is the daily averaged value of the air temperature measured from the ground-based weather station. We use 6 years worth of data for the period 2015-2020.

| STATION | NAME                        | DATE | AWND | PGTM | PRCP | SNOW | SNWD | TAVG | TMAX | TMIN | WDF2 | WDF5 | WSF2 | WSF5 | WT01 | WT02 | WT03 | WT04 | WT05 | WT06 | WT07 | WT08 | V |
|---------|-----------------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 1       |                             |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 2       | USW0094849                  | 1/1/15 | 12.3 | 0    | 0    | 0    | 21   | 28   | 17   | 260  | 260  | 21.9 | 33.1 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 3       | USW0094849                  | 2/1/15 | 4.03 | 0    | 0    | 0    | 19   | 25   | 8    | 240  | 260  | 15   | 19.9 |      |      |      |      |      |      |      |      |      |      |      |      |
| 4       | USW0094849                  | 3/1/15 | 5.82 | 0.2  | 3    | 0    | 21   | 32   | 10   | 130  | 140  | 14.1 | 19   | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 5       | USW0094849                  | 4/1/15 | 10.07 | 0.18 | 2    | 3.1  | 28   | 32   | 8    | 320  | 310  | 21.9 | 30   | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 6       | USW0094849                  | 5/1/15 | 9.62 | 0.01 | 0    | 5.1  | 7    | 9    | 1    | 270  | 270  | 23   | 31.1 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 7       | USW0094849                  | 6/1/15 | 10.29 | 0    | 0    | 3.9  | 7    | 16   | 2    | 240  | 230  | 18.1 | 25.9 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 8       | USW0094849                  | 7/1/15 | 13.65 | 0.1  | 1    | 3.1  | 5    | 10   | -2   | 320  | 330  | 28   | 40   | 1    | 1    |      |      |      |      |      |      |      |      |      |      |      |
| 9       | USW0094849                  | 8/1/15 | 13.65 | 0.1  | 1    | 3.1  | 9    | 16   | 5    | 180  | 260  | 21   | 27.1 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 10      | USW0094849                  | 9/1/15 | 9.4   | 0.01 | 0    | 3.9  | 8    | 12   | 2    | 260  | 300  | 18.1 | 25.1 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 11      | USW0094849                  | 10/1/15 | 9.62 | 0    | 0    | 3.9  | 6    | 18   | 0    | 250  | 240  | 21   | 30   |      |      |      |      |      |      |      |      |      |      |      |      |      |
| 12      | USW0094849                  | 11/1/15 | 5.82 | 0    | 0    | 4    | 18   | 26   | 15   | 210  | 220  | 13   | 18.1 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
| 13      | USW0094849                  | 12/1/15 | 10.07 | 0.02 | 1.5  | 3.9  | 19   | 23   | 2    | 330  | 320  | 25.1 | 31.1 | 1    |      |      |      |      |      |      |      |      |      |      |      |      |
Qualitative evaluation

Our proposed method can effectively capture the temperature values and provide a basis for short-term and long-term prediction. Fig. 1 provides a subjective evaluation of our proposed method. We observe that our proposed method can capture the peaks and troughs of the variation of the ground-based air temperature. We use historical data of 5 years to forecast the future temperature values. We observe that our proposed technique can accurately capture both the rising and falling trends of temperature values in the two subplots of Fig. 1.

Fig. 1: We demonstrate sample illustrations of prediction of ground-based temperature. We observe that our proposed technique can clearly capture the fluctuations of the air temperature.
**Method:** We compute the root mean square error (RMSE) value between the measured data and the forecasted temperature value in order to provide an objective evaluation of our proposed method. The performance of the temperature prediction is determined by two primary factors – the amount of historical data for training and the length of the lead time.

**Result:** Table I shows the RMSE value (measured in K) averaged across 50 experiments for the benchmarking methods. We observed that the average model performs the worst. Our proposed method shows a consistent improvement over the persistence model for the varying lead times.

| Lead Time | Proposed | Persistence | Average  |
|-----------|----------|-------------|----------|
| 1 day     | 3.141    | 5.320       | 15.979   |
| 2 days    | 4.098    | 5.669       | 14.721   |
| 3 days    | 4.618    | 6.819       | 15.123   |
| 4 days    | 5.322    | 7.835       | 16.055   |

**TABLE I:** We compute the RMSE values (measured in K) of the air temperature for varying lead times.
Conclusion
Conclusion

Summary of content

- We use triple exponential smoothing method for predicting future temperature values using past temperature data.
- Our proposed method shows better performance as compared to the other models.

Future work

- We intend to evaluate the impact of the length of historical data on the forecasts estimates.
- We plan to further improve the forecasting performance by incorporating other sensor data in addition to temperature data.