# AQI scale prediction based on Prophet-RF optimization model

Tianjun Chang[1,2], Zhongyang Guo[1,2]

[1]Key Laboratory of Geographical Information Science,East China Normal University, Shanghai, China
[2]School of Geographic Sciences,Shanghai, China

## Introduction

In recent years, the issue of air pollution has attracted wide attention. In many cities in China, there have been different levels of haze. In order to give the public a more intuitive understanding of the pollution situation, the air quality index is calculated according to the above pollutant concentration to illustrate air pollution and air quality. In China, the prediction methods of air quality time series in China mainly focus on time series method, gray system and neural network. The traditional time series analysis theory considers that the general time series can be decomposed into deterministic trends and random changes. The Prophet model adheres to this theory and adds the holiday effect to the deterministic trend of time series, which makes the uncertain holiday effect become a deterministic trend.
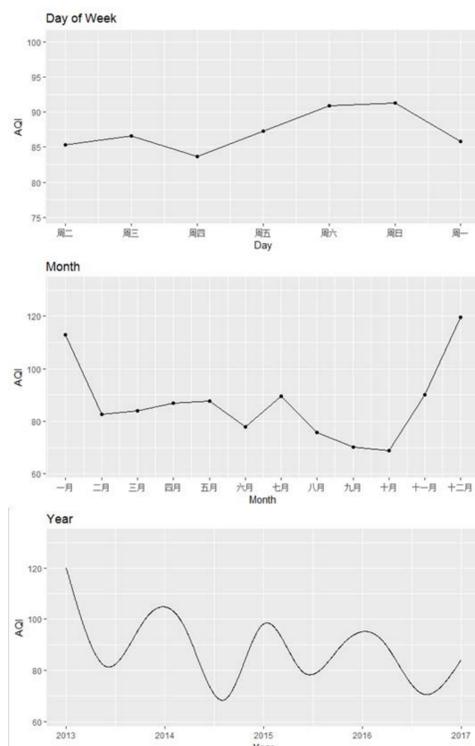
## DATA CHARACTERISTICS



**Figure 1:** *The weekly, monthly and annual trend changes of AQI time series.*

- On a weekly cycle, AQI peaks on Saturdays and Sundays, and is low around Thursday.
- On a monthly cycle, AQI rises sharply in November, to reach the highest level in December and January, the rest in a relatively small fluctuations, AQI usually achieve the lowest level in October.
- On a yearly cycle, every year there are some fluctuations, the overall trend is declining, but speed is slow, and the distance between peak and valley value reduced year by year, every year the AQI curve shows strong cyclical, it also corresponding changes of the weeks and months.
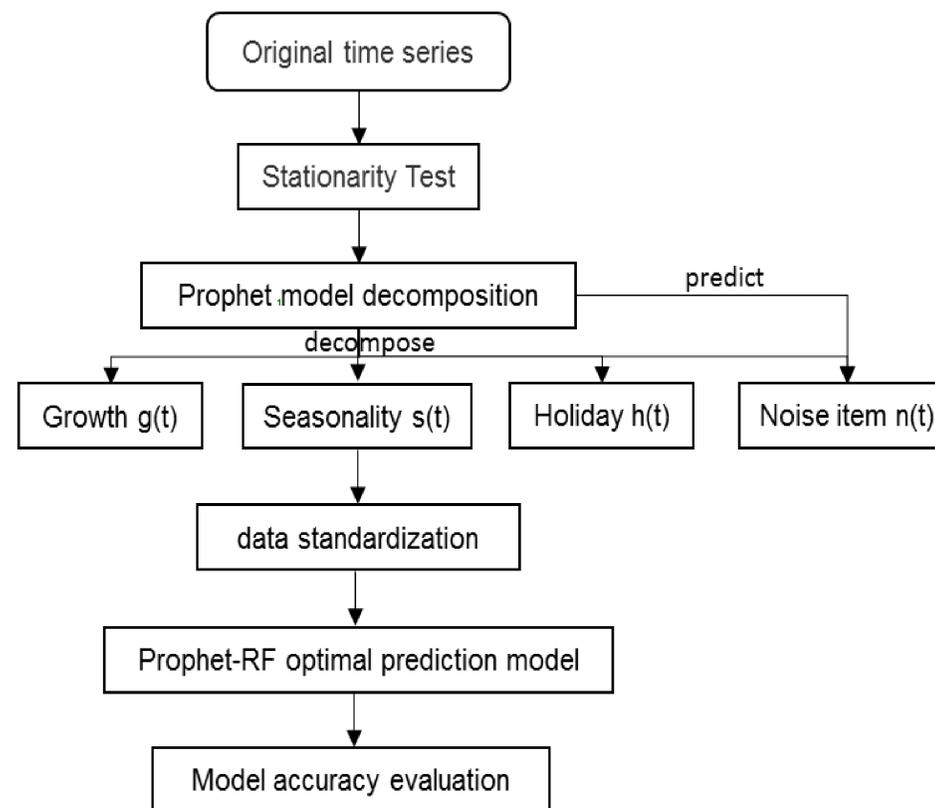
## FORCASTING PROCESS of COMPOSITE MODEL



**Figure 2:** *The weekly, monthly and annual trend changes of AQI time series.*

The Prophet model divides the time series into three parts:

$$y(t) = g(t) + s(t) + h(t) + n(t)$$

Among them,g(t) is the growth term; s(t) is the Seasonal items; h(t)is a holiday term; n(t) is the noise term.

## Prediction results of the composite model

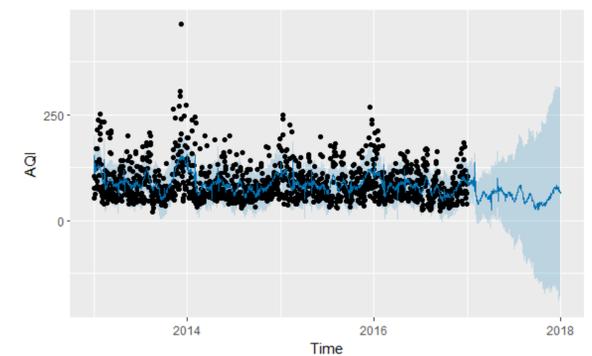The following figure shows the performance composite model in fitting AQI historical data sets:



**Figure 3:** *The historical fitting curves of three models*

According to the statistic and fitting curves, prediction results of composite model is better than single model. Prophet-RF composite prediction model is suitable for long time scale prediction with high precision, the figure below shows the predicted daily AQI values in 2017,it provides a new perspective for related research and has good application prospect in reality.



**Figure 4:** *AQI calendar heat map of Shanghai in 2017*

## References

[1] Sean J.Taylor,Benjamin Letham, "Forecast at Scale," PeerJ Preprints, 2017.
[2] Jing Chai,Cao peng, "Study on the comparison of air quality index between holidays and non-holidays in Changsha in 2016," Health, vol. 11(27), 2017, pp.33-34.
[3] Han J W, Micheline K,Pei J,"Data Mining Concepts and Techniques," Beijing: China Machine Press, 2015.