

Spatio-temporal visualization and forecasting of PM10 in the Brazilian state of Minas Gerais

Kim Leone Souza da Silva¹, Paulo Jorge Canas Rodrigues¹

¹ Federal University of Bahia, Brazil.

Abbreviated abstract: Air pollution is a matter of great concern these days. An important type of air pollution is particulate matter (PM10). It is of extreme interest to monitor them so that they do not exceed the established critical levels. In this work, we will analyse the behavior of particulate materials in 29 locations in the Brazilian state of Minas Gerais. Besides a visualization analysis, we consider six time series models that allow us to forecast PM10 concentrations, with application to three stations.



k.leone_silva@hotmail.com - 1



3rd Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
October 28-30, 2021

Problem, Data, Previous Works

PM10 are particulate materials less than 10 micrometers in diameter. One of the determining factors that affect human health is the size of particles in the atmosphere, due to the degree of permanence and penetration they have in the respiratory system.

Data:

- State Environment Foundation (FEAM), Brazil. (<http://www.feam.br/qualidade-do-ar/dados>)
- Hourly observations between 01/01/2015 and 31/12/2019
- Extraction
- Cleaning

Initially we had access to 58 stations. However, in some there were many missing values. Therefore, we decided to only analyze stations that contained less than 30% of missing values.

- 29 air quality monitoring stations with 43824 hourly observations each.



k.leone_silva@hotmail.com - 2



3rd Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
October 28-30, 2021

Methods

Besides a descriptive and visualization analysis, we consider six time series models that allow us to fit and forecast PM10 concentrations, with application to three stations, one in Belo Horizonte the state capital, the station with the lowest concentration levels and the station with the highest concentration levels.

Modelos de previsão

- SNAIVE - The SNAIVE model is an extension of the NAIVE model that considers a seasonal component of period T in the time series.
- NAIVE + Decomposition
- Exp Smoothing + Decomposition - Forecasts are the weighted average of past observations.
- SARIMA - The SARIMA model extends ARIMA by adding a linear combination of seasonal values and forecast errors.
- NNetar - A neural network is formed by the input, output and hidden layers.
- NNetar + Decomposition

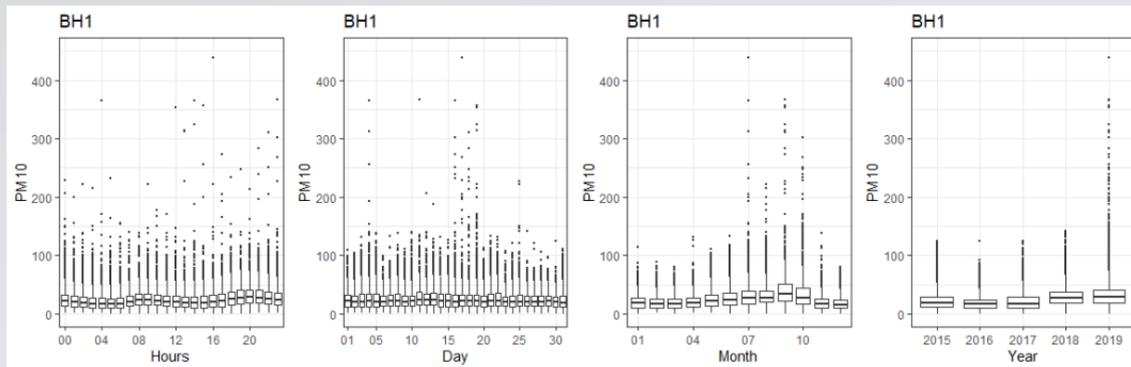


k.leone_silva@hotmail.com - 3



3rd Conference on
**Statistics and
Data Science**
Salvador, Brazil (online)
October 28-30, 2021

Results and Conclusions



Here we can visualize the hourly, daily, monthly and yearly behavior of PM10 concentrations In Belo Horizonte.

	BH1		Itabira 4		S.J.daLapa2	
	RMSE	SMAPE	RMSE	SMAPE	RMSE	SMAPE
Seasonal Naive	19,88	0,56	13,13	0,53	110,77	0,98
Naive+Decomposition	17,73	0,74	13,92	0,81	99,54	0,64
Exp Smoothing + Decomposition	17,57	0,46	11,85	0,64	79,01	0,55
SARIMA	21,09	0,60	12,32	0,70	80,64	0,55
NNETAR	19,77	0,55	13,05	0,47	73,55	0,51
NNETAR + Decomposition	19,44	0,53	8,81	0,31	84,60	0,60

Analyzing the forecasts for the last 15 days of 2019 for the municipality of BH1 we have that the best model was the Exp Smoothing + Decomposition. For the station in Itabira4 the best model was the NNETAR + Decomposition and for the station in São José da Lapa 2 the best model was the NNETAR.

Finally, we reiterate the importance of collecting and analyzing particulate materials as they have a major impact on human and environmental health.

