



DOI: 10.5281/zenodo.18986555

ESTIMATIVA DOS PARÂMETROS DO MODELO ARIMA PARA A VELOCIDADE DO VENTO EM PIRASSUNUNGA

Gabriel Figueiredo Chades de Carvalho¹

João Paulo Martins dos Santos²

Antonio Luiz Ferrari³

Resumo

O comportamento extremo do vento pode ser prejudicial a muitas atividades humanas, inclusive para a aviação. Em particular, a instrução aérea da Academia da Força Aérea (AFA) situada na cidade de Pirassununga, São Paulo, velocidades do vento acima dos limites permitidos para instrução geram cancelamento de voos, o que pode atrasar o cronograma anual das missões. Nesse contexto surge a motivação para uma análise da série temporal de velocidades do vento em Pirassununga. A análise foi baseada no modelo ARIMA suportado pela linguagem R com utilização de pacotes específicos de séries temporais. A análise dos resíduos sugere que o modelo ARIMA(2,1,2) é adequado à representação, pois o teste de Ljung-Box fornece $Q^*=1.7892$, $df=6$, valor $p=0,938$. Dentre as consequências estão a obtenção de intervalos de previsão para intervalos subsequentes e a análise dos impactos dos valores médios estimados. De forma menos intuitiva, o estudo reforça a importância e necessidade de domínio de linguagens de programação de alto nível, como R, para conduzir análises e implementar modelos relacionados à tomada de decisões em diferentes cenários operacionais. **Palavras-chave:** ARIMA, Velocidade do vento, Séries temporais, Previsão.

Abstract

Extreme wind behavior can be detrimental to many human activities, including aviation. In particular, in the flight training at the Air Force Academy (AFA) located in Pirassununga, São Paulo, wind speeds exceeding the allowed limits for training result in flight cancellations, which may delay the annual mission schedule. In this context, the motivation arises for an analysis of the wind speed time series in Pirassununga. The analysis was based on the ARIMA model, supported by the R programming language with the use of specific time series packages. Residual analysis suggests that the ARIMA(2,1,2) model is appropriate for representation, as the

¹Academia da Força Aérea, gfcc33@gmail.com

²Academia da Força Aérea, jpmdossantos@yahoo.com.br

³Academia da Força Aérea, antonioluizalf@fab.mil.br

Ljung-Box test yields $Q^* = 1.7892$, $df = 6$, $p\text{-value} = 0.938$. Among the outcomes is the possibility of obtaining prediction intervals for subsequent periods and analyzing the impact of estimated average values. Less intuitively, the study reinforces the importance and necessity of mastering high-level programming languages, such as R, to conduct analyses and implement models related to decision-making in various operational scenarios. **Keywords:** ARIMA, Wind speed, Time series, Forecast, Aviation instruction.

Introdução

Em geral, situações em que as condições do vento são extremas geram perigo para a operação aérea, especialmente durante o pouso e decolagem, momentos mais críticos do voo, segundo estudo da Boeing (2022, p. 14). Por esse motivo, cada modelo de aeronave tem seus próprios limites de velocidade de vento para operação, de acordo com suas características. Tais valores devem ser respeitados pois, caso contrário, a segurança de voo seria comprometida, sendo recomendado que se proceda uma arremetida ou que se cancele a decolagem, por exemplo, em casos de vento acima do limite.

Nesse contexto surge a motivação para o desenvolvimento de um estudo envolvendo a variável velocidade do vento. Por meio de ajustes de modelos auto-regressivos integrados de médias móveis (ARIMA) a um conjunto de dados, é possível realizar previsões de séries temporais (Morettin; Toloi, 2006, p. 105), como é o caso da variação da velocidade do vento ao longo do tempo. Também conhecido como metodologia de Box-Jenkins, esse é um processo clássico para a estimativa de valores futuros por sua versatilidade, pois pode lidar com séries temporais estacionárias ou não, com ou sem sazonalidade (Teixeira *et al.*, 2019, p. 2).

Objetivo

O objetivo foi analisar os dados de velocidade de vento entre 2013 a 2023 da estação meteorológica da Academia da Força Aérea, no município de Pirassununga, e obter as estimativas dos parâmetros do modelo ARIMA para as velocidades médias diárias.

Material e Método

A série temporal de dados utilizada é resultante do agrupamento das séries parciais provenientes da estação meteorológica da AFA, no município de Pirassununga, originada no site da Divisão de Pesquisa do Instituto de Controle do Espaço Aéreo (ICEA), por meio do “CLIMAER Superfície”, e separada por anos entre 2013 a 2023.

Após o pré-processamento, os dados foram analisados de acordo com o ciclo iterativo de Box-Jenkins, que, segundo Morettin e Toloi (2006, p. 106), é um método que se destaca dos demais pelo pequeno número de parâmetros e precisão nas previsões. O processo

consiste em ajustar modelos $ARIMA(p, d, q)$ a um conjunto de dados em quatro etapas: especificação, identificação, estimação (estimativa dos parâmetros p , d e q) e diagnóstico. No processo de identificação, são

realizadas estimativas preliminares de p e q do modelo identificado, como valores iniciais para o processo de estimação de máxima verossimilhança dos parâmetros. Para realizar essas estimativas, são usadas as autocorrelações amostrais (Box *et al.*, 2016, p. 11; Morettin; Toloi, 2006, p. 172).

Quanto ao parâmetro d , é o número de vezes a se derivar a série temporal para torná-la estacionária, podendo-se, dessa forma, trabalhar com modelos $ARMA(p, q)$. Na prática, d costuma assumir os valores 0, 1 ou no máximo 2, em que 0 define um comportamento estacionário (Box *et al.*, 2016, p. 11; Morettin; Toloi, 2006, p. 152).

Para otimizar o processo matemático, a linguagem R (R Core Team, 2023) por meio da interface gráfica *RStudio Posit* (R Studio Team, 2022) foi utilizada. Os gráficos de autocorrelação e autocorrelação parcial foram gerados, respectivamente, por meio das funções *acf* e *pacf* do pacote *forecasts*. A estimativa dos parâmetros foi realizada com a função *auto.arima*, que realiza o ciclo iterativo de Box-Jenkins e avalia a adequação dos parâmetros à série temporal. Por fim, a análise dos resíduos com a função *checkresiduals*, a fim de verificar a adequação do modelo por meio da análise dos gráficos gerados como o resultado do teste de Ljung-Box. Ver (Hyndman; Athanasopoulos, 2018), (Hyndman; Khandakar, 2008) para mais informações.

Resultados e Discussão

A análise dos resultados identificou a presença de valores nulos recorrentes em faixas específicas de horários, resultantes, provavelmente, da inoperância da estação de medições naqueles horários (Colabone, 2011). A **Figura 1** ilustra os valores das velocidades horárias da série de dados, enquanto que a **Figura 2** ilustra a série de médias diárias.

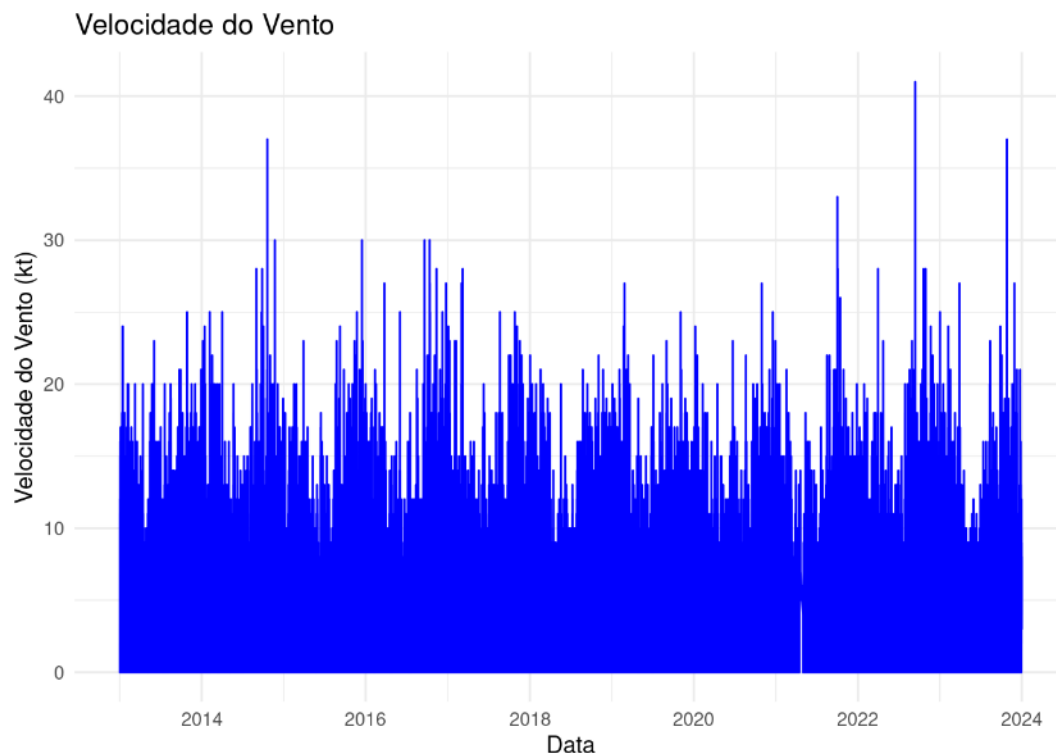


Figura 1: Série das velocidades para os anos entre 2013 e 2023.

Fonte: Autores, 2024

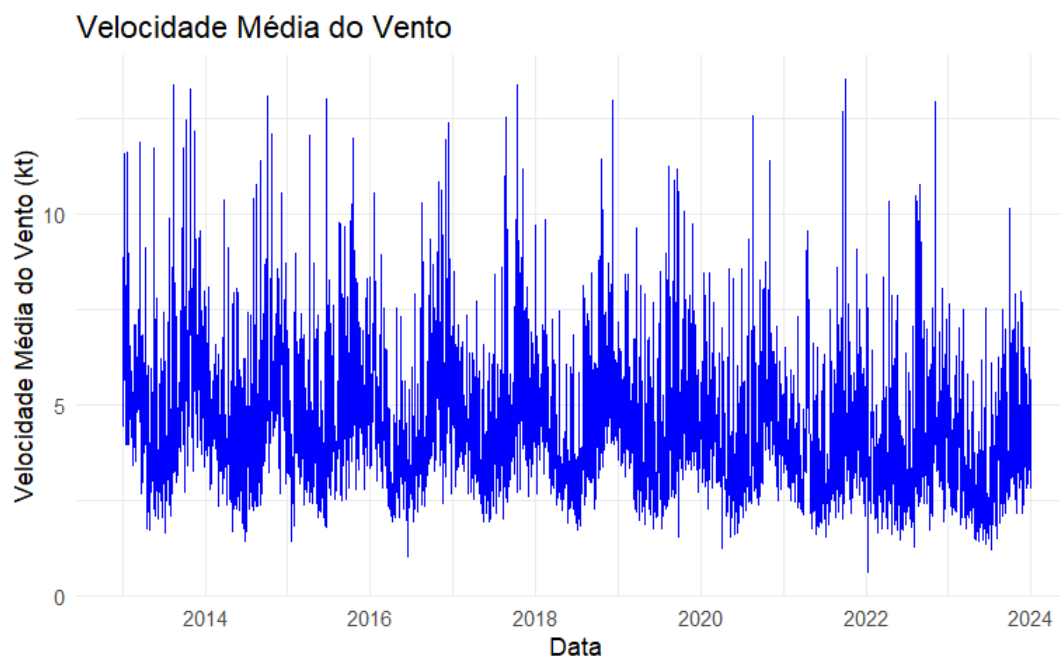


Figura 2: Série das velocidades médias diárias para os anos entre 2013 e 2023.

Fonte: Autores, 2024

A **Figura 3** mostra os resultados da utilização da função *auto.arima* aos dados de velocidades médias diárias.

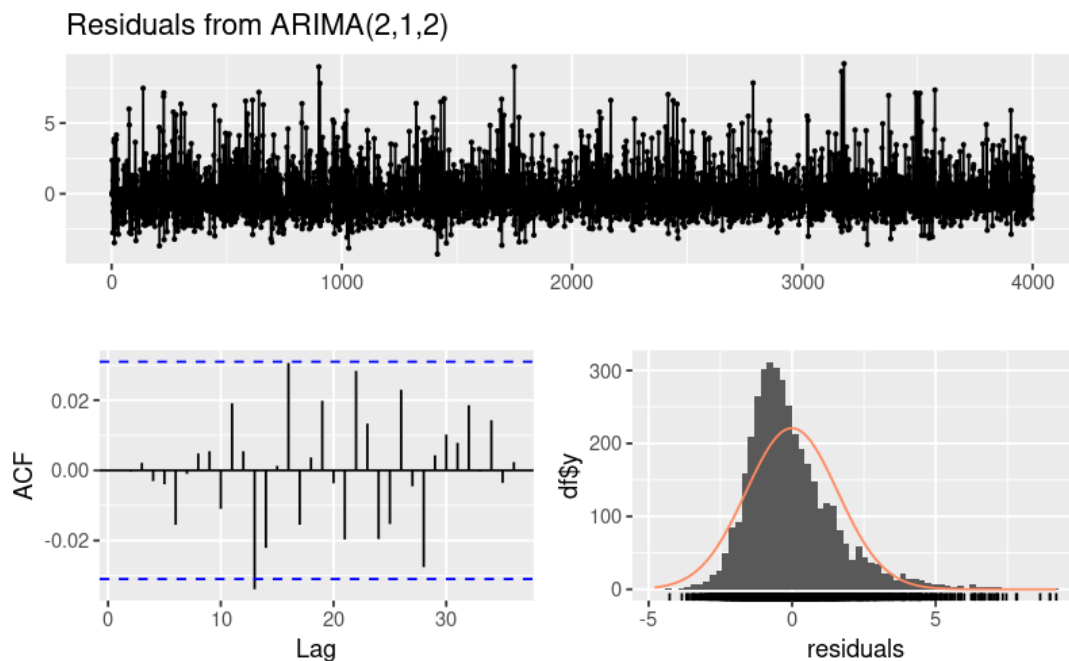


Figura 3: Resultados da função *auto.arima* para as médias diárias.

Fonte: Autores, 2024

O melhor modelo identificado foi o ARIMA(2,1,2) e o teste de Ljung-Box resultou em $Q^*=1.7892$, $df=6$, valor $p=0,938$; logo, não há evidências suficientes para rejeitar a hipótese nula de que os resíduos independentes e não correlacionados, o que sugere que o modelo ARIMA(2,1,2) captura o comportamento da série temporal de forma adequada.

Conclusão

A análise foi motivada pelo contexto da necessidade de estudos da variável velocidade do vento, a qual influencia diretamente nas atividades de instrução aérea. Esta, por sua vez, é influenciada por uma variedade de elementos além do clima, tais como disponibilidade de aeronaves, instrutores, entre outros.

Devido às restrições de disponibilidade dos dados, foi necessário considerar a média diária. Ao obter um modelo ARIMA para a série temporal, por meio da função *auto.arima* do pacote *forecasts*, torna-se possível estimativas de velocidades de vento para períodos de tempo subsequentes e, por consequência, possibilitar uma análise do impacto nas atividades. Ao obter um modelo ARIMA para as médias diárias, é possível obter os intervalos de predição para intervalos subsequentes, mas a estabilidade da série de médias no sentido de não incorporar os valores extremos horários observados sugere certa estabilidade, no mesmo sentido, dos intervalos de previsões do modelo ARIMA. Portanto, estudos posteriores devem considerar máximos diários ou horários, levando em conta que há intervalos faltantes.

Outras possibilidades a serem consideradas são a inclusão de dados provenientes de

outras fontes, estimativas das velocidades nos intervalos de inoperância da estação. Por fim, essa análise põe em evidência a importância de conhecimentos de uma linguagem computacional de alto nível.

Referências

- [1] Boeing. Statistical summary of commercial jet airplane accidents: Worldwide operations 1959 – 2021. Technical report, 2022. Available at: https://www.boeing.com/resources/boeingdotcom/company/about_bca/pdf/statsum.pdf.
- [2] G. E. P. Box et al. Time series analysis: forecasting and control. Wiley, Hoboken, 6th edition, 2016.
- [3] R. J. Hyndman and G. Athanasopoulos. Forecasting: principles and practice. OTexts, Melbourne, 2nd edition, 2018.
- [4] R. J. Hyndman and Y. Khandakar. Automatic time series forecasting: the forecast package for r. Journal of Statistical Software, 27(3):1–22, 2008.
- [5] P. A. Morettin and C. M. C. Toloi. Análise de séries temporais. Blucher, São Paulo, 2nd edition, 2006.
- [6] R. Team. Rstudio: Integrated development environment for r. Technical report, RStudio, PBC, Boston, MA, 2022. Available at: <http://www.rstudio.com/>.
- [7] R. C. Team. The r project for statistical computing. Technical report, 2023. Available at: <https://www.r-project.org/>.
- [8] R. Teixeira et al. Previsão de séries temporais de velocidade do vento utilizando redes neurais artificiais e métodos estatísticos na região de arraial do cabo - rj. In Anais do Congresso Brasileiro de Inteligência Computacional, 2019. Available at: <https://sbic.org.br/wp-content/uploads/2019/12/CBIC2019-101.pdf>.

Anexo A

```
if(!require(ggplot2)){install.packages("ggplot2")}

library(ggplot2)

if(!require(tidyverse)){install.packages("tidyverse")}

library(tidyverse)

library(kableExtra)

#-----

folder=getwd()

pasta =file=paste(folder,"Dados/ICEA/",sep = "/")

library(readxl); library(dplyr)

files <- list.files(pasta, pattern = "\\\\.xlsx$", full.names = TRUE)

dfs <- list()

for (file in files) {

# Leitura do arquivo Excel e armazenamento no dataframe

df <- read_excel(file,skip = 1)

df=df[,1:5]

names(df)=c("Data","Cabeceira","Velocidade","Direcao","Rajada")

df$Cabeceira <- as.numeric(ifelse(df$Cabeceira == "-", NA, df$Cabeceira))

df$Velocidade <- as.numeric(ifelse(df$Velocidade == "#VALUE!", NA,

df$Velocidade))

df$Velocidade <- as.numeric(ifelse(df$Velocidade == "-", NA, df$Velocidade))

df$Direcao <- as.numeric(ifelse(df$Direcao == "-", NA, df$Direcao))

df$Rajada <- as.numeric(ifelse(df$Rajada == "-", NA, df$Rajada))

df$Data=as.POSIXct(df$Data, format = "%d/%m/%Y - %H:%M:%S")

dfs <- c(dfs, list(df))

}

#-----

merged_df <- bind_rows(dfs)

# Criar gráfico de série temporal com ggplot2

grafico <- ggplot(merged_df, aes(x = Data, y = Velocidade)) +

geom_line(color = "blue") + labs(title = "Velocidade do Vento",
```

```
x = "Data",
y = "Velocidade do Vento (kt)" +
theme_minimal() # Estilo do tema
grafico

#-----
# Carregar pacotes necessários

library(dplyr)
library(lubridate)
# Supondo que seu data frame seja chamado 'dados'
# Criar uma coluna de horas para facilitar a filtragem
dados <- merged_df %>% mutate(Hora = hour(Data))
# Filtrar os dados para excluir os horários entre 23h e 5h
dados_filtered <- dados %>%
filter(!(Hora >= 23 | Hora < 6))
media_diaria <- dados_filtered %>%
group_by(Dia = as.Date(Data)) %>%
summarise(Media_Velocidade = mean(Velocidade, na.rm = TRUE))

#-----

names(media_diaria)=c("Dia", "Velocidade")
x=na.omit(media_diaria$Velocidade)
library(forecast)

#-----

# Arima no R: autoarima <https://otexts.com/fpp2/arima-r.html>
fit.arima <- auto.arima(x,trace=T)

#-----

checkresiduals(fit.arima)
```