MISSING VALUES IMPUTATION FOR WIND SPEED

NUR ARINA BAZILAH KAMISAN, SITI MARIAM NORRULASHIKIN, SITI FATIMAH HASSAN

NUR ARINA BAZILAH KAMISAN

UNIVERSITI TEKNOLOGI MALAYSIA, UNIVERSITI MALAYA

4 - 5 AUGUST 2021

1

CONTENTS

2

Introduction

- Missing data or missing values is a common problem in real data. Thus, it becomes a challenge for researchers to find a solution for the missing values.
- Different series may require different strategies to estimate these missing values.
- In this study, we will impute the missing values with a temporary values that follows the data seasonality. When data is now complete, AR model will be used to forecast the data and the position where the missing values are in the original data will be replaced with the forecast from the AR model.

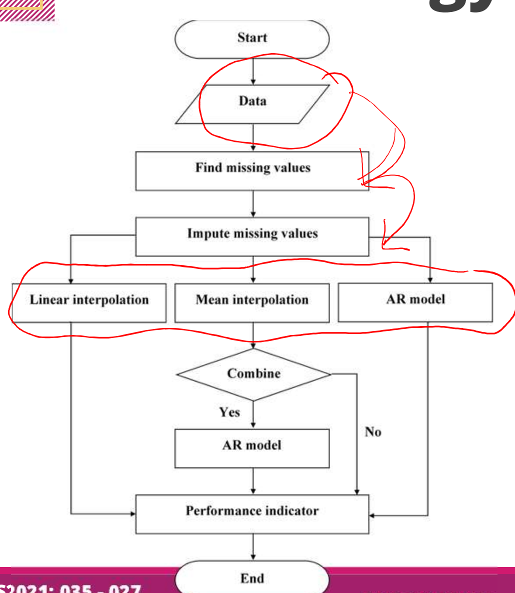
Objectives

- To develop a technique for imputation of missing values in wind speed data.
- To compare few techniques with the technique that we have developed.

Literature Review

- Missing values in the data collection will influence the level of results on data classification, according to Okutan et al. (2018) so ignoring the missing values will lead to the incorrect inference.
- Mean interpolation is the most straightforward way to deal with missing values is to calculate the mean of the measured data depending on the precise timing and it does not require any values or calculations (Strike et al., 2001).
- Moon et al. (2019) explain that linear interpolation is a statistical method which is always being used to estimate with a function based on the known data.
- Autoregressive model is a very random mechanism that is used to characterize time-varying processes (Yuan et al., 2016). Traditionally, according to Liu et al. (2019), AR was used to understand how time series operate, but it was also used to look for missing values due to computer malfunction or machine record failure.

Methodology



1. Missing data will first being imputed with three selected methods.
2. Mean interpolation will be also used as the temporary imputation before being forecast with AR model.
3. Calculate the MAPE to compare the performance of all the selected techniques.

Result and Discussion

Technique	MAPE
Mean interpolation	2.556
Linear interpolation	3.301
AR model	2.323
Mean interpolation + AR model	2.286

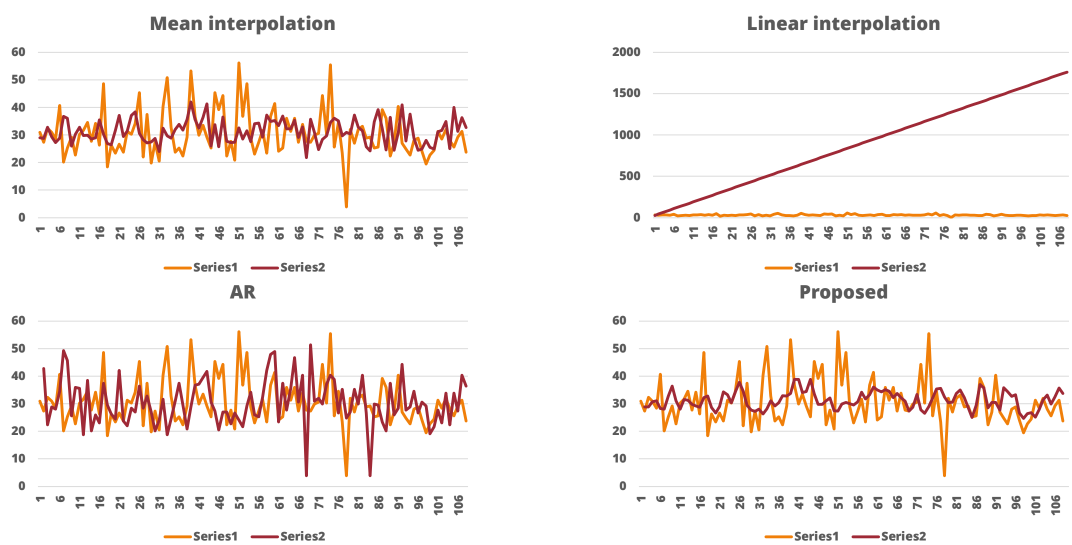
iCMS2021: 035 - 027

MISSING VALUES IMPUTATION FOR WIND SPEED

7

7

Result and Discussion



iCMS2021: 035 - 027

MISSING VALUES IMPUTATION FOR WIND SPEED

8

8

Universiti Teknologi MARA | Cawangan Kedah Kampus Sungai Petani

2021 ICMS INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND STATISTICS

THANK YOU

INTERNATIONAL CONFERENCE ON COMPUTING, MATHEMATICS AND STATISTICS

MISSING VALUES IMPUTATION FOR WIND SPEED

9

9