



INTISARI

***SINGULAR SPECTRUM ANALYSIS* UNTUK DATA HILANG**

Oleh

NOVIA NILAM NURLAZUARDINI

16/403764/PPA/05281

Data mengenai cuaca merupakan salah satu data yang banyak mengandung data hilang. Hal ini terjadi karena beberapa faktor, salah satunya adalah kerusakan alat ukur. Peramalan yang tepat dalam sektor tersebut memberikan dampak yang besar pada sektor kehidupan yang lain. *Recurrent Imputation Method* (RIM) dapat digunakan untuk mengatasi masalah tersebut dengan mengaplikasikan *Singular Spectrum Analysis* (SSA) terlebih dahulu terhadap data. RIM membagi data menjadi tiga bagian, yaitu data sebelum data hilang, data hilang, dan data setelah data hilang, yang kemudian digunakan untuk imputasi data hilang. SSA merupakan teknik nonparametrik yang merupakan teknik model bebas, hal ini berarti SSA terbebas dari semua asumsi statistika. Pada penelitian ini, digunakan data curah hujan di California dari periode 1895 sampai dengan 2017 sebagai *insample*, dengan tiga data hilang pada periode Mei sampai dengan Juli 1947. Kemudian, data periode Januari sampai dengan Agustus 2018 digunakan sebagai data *outsample*. Diramalkan empat data, yaitu data bulan September sampai dengan Desember 2018. Sehingga diperoleh nilai RMSE peramalan *outsample* data adalah 2,117, nilai MAPE sebesar 2,64 %, dan nilai MAE sebesar 1,503.

Keywords: SSA, Data Hilang, Recurrent Imputation Method, Weather Dataset



ABSTRACT

SINGULAR SPECTRUM ANALYSIS FOR MISSING DATA

By

NOVIA NILAM NURLAZUARDINI

16/403764/PPA/05281

Weather data set is one of data that contains a lot of missing values. This happens because of several factors, one of them is the breakdown of gauge. Proper forecasting in the sector has a large impact on other sectors of life. Recurrent Imputation Method (RIM) can be used to overcome this problem by applying Singular Spectrum Analysis (SSA) to the data first. RIM divides the data into three parts, namely data before missing value, data missing value, and data after missing value, which used for imputation of data missing value. SSA is a nonparametric technique which is a free model technique, this means that SSA is free from all statistical assumptions. In this study, rainfall data in California were used from the period 1895 to 2017 as an insample data, with three missing values in May to July 1947 periode. Then, data from January to August 2018 were used as outsample data. The prediction data starts from September to December 2018. So that the value of RMSE obtained from forecasting outsample data is 2.117, MAPE is 2.64 %, and MAE is 1.503.

Keywords: SSA, Missing Data, RIM, Weather Dataset