

Peramalan Data *Time Series Seasonal* Menggunakan Metode Analisis Spektral

Anis Mahfud Al'afi¹, Widiarti*, Dian Kurniasari¹, dan Mustofa Usman¹

Jurusan Matematika, Fakultas MIPA, Universitas Lampung
Jl. Prof. Dr. Sumantri Brojonegoro No. 1 Bandar Lampung

*Email korespondensi: widiarti08@gmail.com

Dikirim: 27-02-2020, Diterima: 14-03-2020, Diterbitkan: 31-03-2020

Abstrak

Transportasi udara kini menjadi mode transportasi yang menjadi pilihan utama. Walaupun biaya transportasi ini tergolong mahal namun dapat menghemat banyak waktu untuk sampai ke tujuan. Oleh karena itu, meramalkan jumlah penumpang pesawat menjadi hal yang menarik untuk diteliti. Pada paper ini dilakukan peramalan jumlah penumpang pesawat di Bandar Udara Raden Intan II menggunakan metode analisis spektral. Metode ini digunakan untuk mendapatkan informasi yang lebih lengkap mengenai karakteristik data runtun waktu yang diperlukan untuk menelaah periodesitas yang terjadi. Setelah mendapatkan periodesitas maka data dimodelkan dengan metode Seasonal ARIMA. Berdasarkan data yang tersedia diperoleh model terbaik untuk peramalan penumpang pesawat di Bandar Udara Raden Intan II adalah Seasonal ARIMA (0,1,1) (0,1,1)³.

Kata kunci : Analisis Spektral, Deret Waktu, Penumpang Pesawat, SARIMA

Abstract

Air transportation is now a mode of transportation that is often the first choice. Although the transportation costs are relatively expensive, it can save a lot of time to get to the destination. Therefore, predicting the number of aircraft passengers is an interesting thing to study. In this paper we forecast the number of aircraft passengers at Raden Intan II Airport using spectral analysis methods. This method is used to obtain more complete information about the time series data characteristics to examine the periodicity. After getting the periodicity, the data are modeled using the ARIMA Seasonal Method. Based on available data. The results show that the best model for forecasting aircraft passengers at Raden Intan II Airport is Seasonal ARIMA (0,1,1) (0,1,1)³.

Keywords: Airplane Passengers, SARIMA, Spectral Analysis, Time Series

1. Pendahuluan

Data *time series* merupakan serangkaian pengamatan yang terurut berdasarkan waktu dengan jarak yang sama [1]. Jenis data ini sering ditemui dalam keseharian karena data tersebut dikumpulkan melalui waktu interval yaitu harian, mingguan atau bulanan. Dari data yang terkumpul dapat dilihat ada suatu pola didalamnya. Dalam *time series* pola tersebut dibagi menjadi tiga yaitu pola trend, siklis dan musiman (*seasonal*). Pola musiman (*seasonal*) adalah pola yang mengalami pengulangan yang sama berkali-kali pada interval tertentu.

Berdasarkan pembagian daerah (*domain*), data *time series* terbagi menjadi dua daerah (*domain*) yaitu daerah waktu (*time domain*) dan daerah frekuensi (*frekuensi domain*) [2]. Daerah waktu menelaah tentang signifikansi autokorelasi, kestasioneran data, penaksiran parameter model regresi deret waktu dan peramalan (*forecasting*). Sedangkan daerah frekuensi (*frequency domain*) menelaah frekuensi tersembunyi pada data musiman yang sulit diperoleh dalam daerah waktu. Tujuannya adalah untuk mengetahui hal-hal istimewa atau kondisi tertentu pada data. Salah satu cara yang dapat digunakan dalam analisis daerah frekuensi adalah analisis spektral.

Analisis spektral atau juga disebut analisis spektrum diperkenalkan oleh Schuster pada abad ke-19 yang digunakan pada bidang kelautan, meteorologi, dan astronomi [3]. Analisis spektral merupakan suatu metode untuk penaksiran dalam domain frekuensi untuk menelaah periodisitas tersembunyi, yaitu periodesitas yang sulit ditemukan dalam domain waktu [4]. Analisis spektral dapat digunakan untuk melihat pola periodik yang terjadi dalam data tersebut.

Setelah periode atau *seasonal* data diketahui selanjutnya periode tersebut dapat digunakan untuk melakukan pemodelan. Pemodelan data yang sesuai dengan dugaan data *seasonal* adalah menggunakan analisis *Seasonal*

Autoregressive Integrated Moving Average (SARIMA). Metode SARIMA merupakan metode ARIMA yang mengandung unsur musiman [5]. Metode ini bisa memodelkan masalah data yang mempunyai pola musiman.

Pada era modern ini, mode transportasi sangatlah penting dikarenakan telah menjadi kebutuhan pokok bagi kebanyakan orang. Salah satu model transportasi yang banyak peminatnya adalah transportasi udara. Meskipun biaya transportasi ini tergolong mahal namun dapat menghemat banyak waktu untuk sampai ke tujuan. Lonjakan penumpang pesawat umumnya terjadi saat musim libur panjang. Berdasarkan hal tersebut, paper ini mengkaji peramalan jumlah penumpang pesawat perbulannya dengan menggunakan analisis SARIMA dengan melihat periode yang terjadi menggunakan analisis spektral.

2. Metode Penelitian

Data yang digunakan dalam paper ini adalah data penumpang pesawat Bandar Udara Raden Intan II yang disajikan dalam Tabel 1.

Tabel 1. Data penumpang pesawat terbang periode Januari 2012 – Agustus 2018

Bulan	Jumlah Penumpang per Tahun						
	2012	2013	2014	2015	2016	2017	2108
Januari	47599	50692	50567	47987	67549	88164	110466
Februari	47014	42754	45170	47589	66136	76767	106672
Maret	54061	47492	50136	50027	71366	94570	117545
April	50457	46048	45838	46301	72420	86557	115736
Mei	55302	48990	50970	51889	79222	105516	96929
Juni	53958	51061	55936	54817	76631	91007	112658
Juli	50503	42369	51219	42544	87899	133685	117779
Agustus	38258	54672	51250	58789	75710	107517	101826
September	48882	53701	49620	54855	73220	113221	
Oktober	52107	50561	51450	52518	74813	116735	
November	53400	49787	52615	52297	76767	111915	
Desember	54906	54232	57424	57984	95956	128098	

Analisis data dalam paper ini menggunakan *software* R Studio 3.4.2. Langkah - langkah dalam membuat peramalan penumpang pesawat terbang adalah sebagai berikut :

1. Menguji kestasioneran data dengan uji *Augmented Dickey Fuller* (ADF).
2. Melakukan transformasi data dengan metode *Box-Cox* dan pembedaan (*differencing*) apabila data tidak stasioner terhadap *varian* dan rata-rata.
3. Melakukan analisis spektral untuk menentukan periode menggunakan periodogramnya.
4. Menguji komponen periodik tersembunyi dari periodogramnya menggunakan uji Fisher.
5. Melakukan pendugaan model sementara melalui plot ACF dan plot PACF.
6. Melakukan pemeriksaan diagnostik model
7. Melakukan pemilihan model terbaik dengan menggunakan kriteria AIC dengan pengambilan keputusan didasarkan pada nilai AIC minimum.
8. Meramalkan jumlah penumpang pesawat Bandar Udara Raden Intan II.

3. Hasil dan Pembahasan

Pemodelan SARIMA dapat dilakukan dengan terlebih dahulu mencari periodesitas tersembunyi dari data penumpang penerbangan Bandara Raden Intan II menggunakan analisis spektral. Sebelum melakukan analisis spektral, data diuji kestasionerannya melalui uji ADF. Uji hipotesis yang digunakan adalah

H_0 : data penumpang penerbangan tidak stasioner

H_1 : data penumpang penerbangan stasioner

Tabel 2 Hasil Uji kestasioneran data

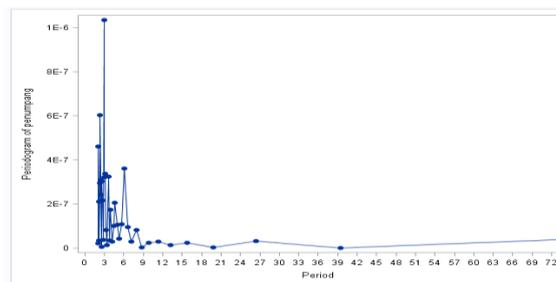
Augemented Dickey-Fuller Unit Root Tests			
Type	Lags	F	Pr > F
Zero Mean	0		
Single Means	0	1.671	0.6489
Trend	0	5.977	0.075

Berdasarkan Tabel 2, uji ADF memberikan nilai $Pr > F$ lebih besar dari $\alpha = 0,05$. Dengan demikian tidak tolak H_0 pada taraf nyata atau dengan kata lain data tidak stasioner. Oleh karena itu data perlu ditransformasi dan *differencing* agar menjadi stasioner. Setelah data ditransformasi dan *differencing* maka data diuji kembali dengan uji ADF yang hasilnya disajikan dalam Tabel 3. Berdasarkan Tabel 3 nilai $Pr > F$ yang diperoleh kurang dari 0.0010. Karena nilai tersebut lebih kecil dari taraf signifikan 5%, maka tolak H_0 yang berarti data penumpang penerbangan telah stasioner.

Tabel 3 Hasil uji kestasioneran data hasil transformasi dan *differencing*

Augemented Dickey-Fuller Unit Root Tests			
Type	Lags	F	Pr>F
Zero Mean	0		
Single Means	0	96.303	<.0010
Trend	0	95.238	<.0010

Untuk mengetahui periodisitas tersembunyi dari data penumpang penerbangan, maka dilakukan analisis spektral dengan fungsi periodogram sebagai berikut:



Gambar 1. Periodogram

Berdasarkan Gambar 1, titik tertinggi pada periodogramnya yaitu 0.000001035 dengan nilai spektral sebesar 0,00823668. Nilai tertinggi periodogram tersebut berada pada periode $2.9259 \approx 3$. Periode dengan nilai periodogram dan spektral tertinggi ini menunjukkan periode penumpang pesawat yang paling sering berulang yaitu 3 bulan. Selanjutnya dilakukan uji Fisher untuk melihat adanya komponen periodik dari periodogram dengan uji hipotesis

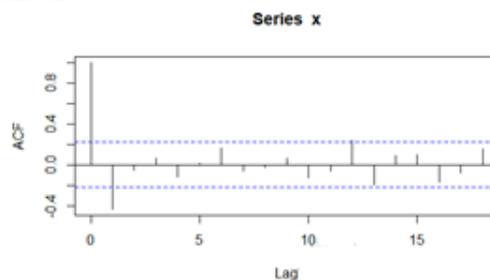
- H_0 : periodogram tidak terdapat komponen periodik
- H_1 : periodogram terdapat komponen periodik

Tabel 4 menunjukkan bahwa dengan taraf nyata 5% dan ukuran sampel 80, nilai Fisher lebih besar dari nilai tabel untuk pengujian hipotesis tersebut yaitu 0,15738. Dengan demikian diperoleh kesimpulan tolak H_0 pada taraf nyata 5% yang berarti data penumpang penerbangan mengandung komponen periodik.

Tabel 4 Hasil Uji komponen periodik data

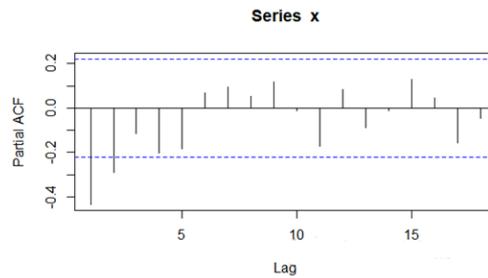
Fisher's Kappa	
Nilai	0.16189

Setelah diperoleh periode dari data, selanjutnya data dimodelkan menggunakan metode SARIMA. Identifikasi model dilihat dari plot ACF dan PACF baik yang non musiman maupun musimannya yang ditunjukkan oleh Gambar 2 – Gambar 5.



Gambar 2. Plot ACF Nonmusiman

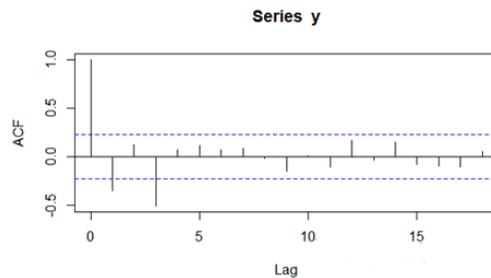
Dari Gambar 2 terlihat bahwa nilai autokorelasi *cut off* pada lag pertama. Sehingga dapat diduga nilai q adalah 1.



Gambar 3. Plot PACF Nonmusiman

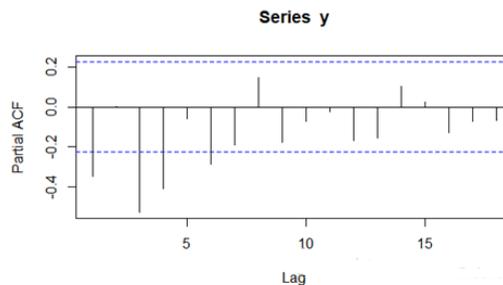
Dari hasil Gambar 3. terlihat bahwa nilai PACF *dyng down* sehingga dapat diduga nilai q adalah 2 karena nilai lag yang keluar adalah dua. Maka model proses nonmusiman yang mungkin cocok adalah (1,1,2), (1,1,1), (0,1,2), (0,1,1), dan (1,1,0).

Setelah memperoleh model dugaan proses nonmusiman, maka selanjutnya dilakukan pendugaan model pada proses musiman. Data penumpang pesawat memiliki kompenen periodik sebesar 3. Dengan demikian menduga model proses musiman, hal yang harus dilakukan adalah melakukan pembeda musiman dengan s sebesar 3. Selanjutnya mengidentifikasi model musimannya dengan melihat plot ACF dan PACF musiman sebagai berikut:



Gambar 4. Plot ACF Musiman

Dari hasil Gambar 4 terlihat bahwa nilai autokorelasi *cut off* pada lag pertama. Sehingga dapat diduga nilai P adalah 1.



Gambar 5. Plot PACF Musiman

Dari hasil Gambar 5 terlihat bahwa nilai PACF *cut off*. Sehingga dapat diduga nilai Q adalah 1 karena nilai musiman juga sudah stasioner maka nilai D adalah 1. Berdasarkan ordo autokorelasi dan autokorelasi parsial yang diperoleh, maka model proses musiman yang mungkin cocok adalah (1.1.1) , (1.1.0) dan (0.1.1).

Dari model nonmusiman dan musiman yang dihasilkan maka terdapat model SARIMA sebanyak 15 model yang akan diuji signifikan model. Dengan teknik *try and error*, diperoleh model berikut uji signifikannya seperti yang disajikan pada Tabel 5. Dari Tabel 5 tersebut model SARIMA yang signifikan sebanyak 5 model yaitu (0,1,2) (1,1,1)³, (0,1,1) (1,1,0)³, (1,1,0) (1,1,0)³, (0,1,1) (0,1,1)³ dan (1,1,0) (0,1,1)³. Kemudian akan dilakukan pemeriksaan diagnostik masing-masing model yang signifikan.

Model-model yang nilai dugaan parameternya signifikan selanjutnya akan dilakukan uji diagnostik untuk mengetahui apakah galat dari model sudah bersifat *white noise* dan berdistribusi normal. Untuk mengetahui galat bersifat *white noise* maka dilakukan uji *Ljung Box*. Galat pada lag dari model dikatakan tidak berkorelasi apabila *p-value* untuk masing-masing galat pada setiap lag lebih besar dari tingkat kesalahan. Hasil dari uji *Ljung Box* disajikan dalam Tabel 6.

Tabel 5. Hasil Uji Signifikan Model

Hasil Uji Signifikan	
Model	Hasil
$(1,1,2)(1,1,1)^3$	Tidak
$(1,1,1)(1,1,1)^3$	Tidak
$(0,1,2)(1,1,1)^3$	Signifikan
$(1,1,0)(1,1,1)^3$	Tidak
$(0,1,1)(1,1,1)^3$	Tidak
$(1,1,2)(1,1,0)^3$	Tidak
$(1,1,1)(1,1,0)^3$	Tidak
$(0,1,2)(1,1,0)^3$	Tidak
$(1,1,0)(1,1,0)^3$	Signifikan
$(0,1,1)(1,1,0)^3$	Signifikan
$(1,1,2)(0,1,1)^3$	Tidak
$(1,1,1)(0,1,1)^3$	Tidak
$(0,1,2)(0,1,1)^3$	Tidak
$(1,1,0)(0,1,1)^3$	Signifikan
$(0,1,1)(0,1,1)^3$	Signifikan

Tabel 6 Uji Ljung Box

Model	Lag			
	12	24	36	48
SARIMA $(0,1,2)(1,1,1)^3$	0,95	0,94	0,97	0,80
SARIMA $(0,1,1)(1,1,0)^3$	0,52	0,66	0,79	0,66
SARIMA $(1,1,0)(1,1,0)^3$	0,03	0,03	0,17	0,02
SARIMA $(0,1,1)(0,1,1)^3$	0,65	0,57	0,66	0,50
SARIMA $(1,1,0)(0,1,1)^3$	0,29	0,12	0,27	0,14

Berdasarkan Tabel 6 terdapat empat model yang memenuhi asumsi galat *white noise*. Model tersebut adalah SARIMA $(0,1,2)(1,1,1)^3$, SARIMA $(0,1,1)(1,1,0)^3$, SARIMA $(0,1,1)(0,1,1)^3$ dan SARIMA $(1,1,0)(0,1,1)^3$.

Langkah selanjutnya adalah menguji kenormalan galat dari keempat model tersebut dengan menggunakan uji Shapiro Wilk dengan hipotesis yang digunakan adalah :

H_0 = galat berdistribusi normal,

H_1 = galat tidak berdistribusi normal.

Pengambilan keputusan yaitu apabila $p - value \geq \gamma$ maka terima H_0 sedangkan jika $p - value < \gamma$ maka tolak H_0 dengan $\gamma = 0,05$.

Tabel 7 Uji Shapiro Wilk

Model	<i>p-value</i>
SARIMA $(0,1,2)(1,1,1)^3$	0.2935
SARIMA $(0,1,1)(1,1,0)^3$	0.5877
SARIMA $(0,1,1)(0,1,1)^3$	0.2973
SARIMA $(1,1,0)(0,1,1)^3$	0.2506

Hasil uji Shapiro Wilk pada Tabel 7 menunjukkan bahwa $p - value > \gamma$. Dengan demikian tidak tolak H_0 pada taraf nyata $\gamma = 5\%$ dan dapat disimpulkan bahwa keempat model yang diuji memenuhi asumsi galat berdistribusi normal.

Keempat kandidat model yang cocok telah memenuhi asumsi galat *white noise* dan uji kenormalan galat. Selanjutnya akan dipilih model terbaik berdasarkan kriteria AIC yaitu model yang memiliki nilai AIC terkecil. Nilai AIC dari keempat model tersebut disajikan dalam Tabel 8.

Tabel 8 Nilai AIC

Model	AIC
SARIMA (0,1,2)(1,1,1) ³	1.583,04
SARIMA (0,1,1)(1,1,0) ³	-1026,63
SARIMA (0,1,1)(0,1,1) ³	-1.038,69
SARIMA (1,1,0)(0,1,1) ³	-1.031,95

Berdasarkan Tabel 8, model yang dipilih adalah SARIMA (0,1,1)(0,1,1)³ karena model ini memiliki nilai AIC terkecil yaitu sebesar -1.038,688. Dari persamaan SARIMA (0,1,1)(0,1,1)³ diperoleh nilai $\phi = 0,6457$ dan $\theta = 0,9114$. Dengan demikian model peramalannya dapat dinyatakan sebagai berikut:

$$Y_t = 1,6457Y_{t-1} - 0,6457Y_{t-2} + 0,3543Y_{t-3} - Y_{t-4} + 0,6457Y_{t-5} + e_t - 0,9114e_{t-3} \quad (1)$$

Dengan menggunakan model Persamaan (1), maka diperoleh prediksi jumlah penumpang pesawat untuk 8 bulan kedepan yaitu September 2018 sampai April 2019 seperti yang disajikan pada Tabel 9.

Tabel 9 Peramalan Jumlah Penumpang Penerbangan Bandara Raden Intan II

Bulan	Jumlah Penumpang
Sep-18	114456
Oct-18	117363
Nov-18	116956
Dec-18	121686
Jan-19	124691
Feb-19	124382
Mar-19	129210
Apr-19	132314

4. Kesimpulan

Berdasarkan hasil dan pembahasan dapat diambil kesimpulan sebagai berikut:

1. Hasil kajian periodesitas dari analisis spektral diperoleh periode terkuat pada data penumpang penumpang adalah 3 bulan.
2. Model terbaik dari data deret waktu untuk meramalkan data penumpang pesawat yaitu model SARIMA (0,1,1)(0,1,1)³ dengan persamaan modelnya adalah sebagai berikut:

$$Y_t = 1,6457Y_{t-1} - 0,6457Y_{t-2} + 0,3543Y_{t-3} - Y_{t-4} + 0,6457Y_{t-5} + e_t - 0,9114e_{t-3}$$

Daftar Pustaka:

- [1] Wei, W.W.S. 2006. *Time Series Analysis Univariate and Multivariate Methods*. Pearson Education, Inc., New York.
- [2] Chatfield, C. 1996, *The Analysis of Time Series: An Introduction*, Fifth Edition, Chapman and Hall, New York.
- [3] Bloomfield, P. 2000. *Fourier Analysis of Time Series. Second Edition*. A Wiley Interscience Publication. New York
- [4] Mulyana. 2004. *Analisis Spektral Untuk Menelaah Periodesitas Tersembunyi Dari Data Deret Waktu*. Fakultas Matematika dan Ilmu Pengetahuan Alam. UNPAD: Bandung.
- [5] Wahyuni, R. 2017. *Peramalan Kunjungan Wisatawan Mancanegara Melalui Bandara Internasional Hang Nadim Kota Batam Dengan Model Seasonal Autoregressive Integrated Moving Average (SARIMA)*. Fakultas Matematika dan Ilmu Pengetahuan Alam. UNAND: Padang