

FORECASTING DATA RUNTUN WAKTU MUSIMAN MENGGUNAKAN METODE SINGULAR SPECTRUM ANALYSIS (SSA)

Dadang Ruhiat¹, Dini Andiani², Wulan Nurul Kamilah³

^{1,2,3} Universitas Bale Bandung, Jl. Raden AA Wiranatakusumah No.7, Bandung, Indonesia
Email: dadangwiraruhiat@gmail.com

ABSTRACT

Recently, modeling and forecasting of time series data continue to grow and are widely used in various fields including in the field of hydrology. A very important hydrological parameter is the river rate flows which in Indonesia as in the other tropical countries, the magnitude and fluctuation are influenced by two seasonal factors, the rainy and the dry seasons. Modeling and forecasting based on statistical time series can be done through two approaches, parametric statistics and non-parametric statistics. However, the facts show that modeling and forecasting time series through parametric statistical approaches are commonly used. In this study, a non-parametric statistical approach using the Singular Spectrum Analysis (SSA) method is used to model and forecast river flow rates time series data. The purpose of this study is to determine the results of the SSA forecasting method and to find out the results of the comparison with the results of parametric statistical forecasting that have been done previously through the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The research process method is a theoretical study which is then continued with the computational process. The result of the analysis shows that the SSA method provides forecasting with a Mean Absolute Percentage Error (MAPE) value is smaller than the SARIMA model. Thus, it is concluded that the forecasting of the river flow rates which seasonally patterned through the SSA method is relatively better than the SARIMA model.

Keywords: SSA, SARIMA, Box-Jenkins, MAPE

ABSTRAK

Pemodelan dan *forecasting* data runtun waktu akhir-akhir ini terus berkembang dan digunakan di berbagai bidang termasuk di bidang hidrologi. Parameter hidrologi yang sangat penting adalah debit sungai di Indonesia sebagaimana halnya di negara tropis lainnya, besaran dan fluktuasinya dipengaruhi oleh dua faktor musiman, yaitu musim hujan dan kemarau. Pemodelan dan *forecasting* runtun waktu berbasis statistik pada dasarnya dapat dilakukan melalui dua pendekatan, yaitu statistik parametrik dan statistik non-parametrik. Namun fakta menunjukkan pemodelan dan *forecasting* runtun waktu melalui pendekatan statistik parametrik lebih banyak dilakukan. Pada penelitian ini dilakukan pemodelan dan *forecasting* data runtun waktu debit sungai melalui pendekatan statistik non-parametrik dengan menggunakan metode *Singular Spectrum Analysis* (SSA). Tujuan dari penelitian ini adalah untuk mengetahui hasil *forecasting* metode SSA dan mengetahui hasil komparasinya dengan hasil *forecasting* statistik parametrik yang telah dilakukan sebelumnya melalui model *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Metode proses penelitian ini adalah berupa kajian teori yang kemudian dilanjutkan dengan proses komputasi. Hasil analisis menunjukkan metode SSA memberikan hasil *forecasting* dengan nilai *Mean Absolute Percentage Error* (MAPE) yang lebih kecil dari model SARIMA. Dengan demikian disimpulkan *forecasting* runtun waktu debit sungai yang berpola musiman melalui metode SSA relatif lebih baik dari hasil *forecasting* model SARIMA.

Kata kunci: SSA, SARIMA, Box-Jenkins, MAPE

Dikirim: 10 Februari 2020; Diterima: 15 Februari 2020; Dipublikasikan: 30 Maret 2020

Cara sitasi: Ruhiat, D., Andiani, D., & Kamilah, W. N. (2020). Forecasting data runtun waktu musiman menggunakan metode singular spectrum analysis (ssa). *Teorema: Teori dan Riset Matematika*, 5(1), 47-60.

PENDAHULUAN

Debit sungai merupakan parameter hidrologi yang sangat penting dalam rekayasa dan pengelolaan sumber daya air (SDA). Karakteristik debit sungai di Indonesia sebagaimana halnya di negara tropis lainnya, besaran dan fluktuasinya dipengaruhi oleh dua faktor musiman, yaitu musim hujan dan kemarau. Informasi terkait pola dan besaran debit sungai untuk beberapa waktu ke depan sangat penting dalam rangka pengelolaan dan rekayasa sumber daya air (SDA). Informasi tersebut dapat diperoleh melalui peramalan (*forecasting*) runtun waktu debit sungai melalui pendekatan metode yang berbasis statistik, baik pendekatan statistik parametrik maupun statistik non parameterik. Metode parametrik mensyaratkan bahwa historis dari data runtun waktu yang digunakan berdistribusi normal, sedangkan metode non parametrik sebaliknya distribusi data runtun waktu tidak harus berdistribusi normal.

Pemodelan dan *forecasting* debit sungai melalui pendekatan berbasis statistik parametrik, sebelumnya telah dilakukan oleh beberapa peneliti. Metode yang digunakan umumnya adalah metode Box-Jenkins, baik yang melibatkan maupun tidak melibatkan faktor musiman pada proses pemodelannya. Beberapa model yang termasuk metode Box-Jenkins tersebut diantaranya adalah model *Autoregressive Integrated Moving Average* (ARIMA), *Seasonal Autoregressive Integrated Moving Average* (SARIMA), *Autoregressive Fractionally Integrated Moving Average* (ARFIMA), dan *Seasonal Autoregressive Fractionally Integrated Moving Average* (SARFIMA).

Fakta saat ini menunjukkan bahwa perkembangan/penggunaan metode pemodelan dan *forecasting* untuk data runtun waktu debit sungai menggunakan metode parametrik lebih banyak digunakan daripada metode non parametrik. Juwono (2010), melakukan kajian mengenai pengaruh perbedaan rerata data debit pada pemodelan deret berkala untuk peramalan debit sungai dengan metode *Autoregressive Fractionally Integrated Moving Average* (ARFIMA). Suprayogi *et.al*, (2015) melakukan pengembangan model hidrologi runtun waktu untuk peramalan debit sungai menggunakan *Daubechies Wavelet-Adaptive Neuro Fuzzy Inference System* pada Sub DAS Siak Bagian Hulu. Namun demikian metode-metode tersebut tidak melibatkan faktor musiman pada proses pemodelannya.

Beberapa metode parametrik untuk pemodelan dan *forecasting* data runtun waktu yang berpola musiman sebelumnya sudah digunakan di berbagai bidang, antara lain: Sakhabakhsh & Yarmohammadi (2012), melakukan studi tentang penggunaan model *Seasonal Autoregressive Integrated Moving Average* (SARIMA) dalam ilmu energi; Bako *et.al* (2013), menggunakan model Seasonal ARIMA untuk memprediksi hasil tangkapan ikan; Ruhiat (2016), menggunakan metode SARFIMA untuk peramalan debit sungai Cimanuk yang mengandung pola musiman; Ruhiat & Effendi (2018), melakukan kajian mengenai pengaruh faktor musiman terhadap hasil *forecasting* debit sungai; Ruhiat & Suwanda (2019), melakukan *forecasting* debit sungai dengan menggunakan metode regresi Spektral.

Ruhiat & Effendi (2018), dalam penelitiannya menunjukkan hasil bahwa model SARIMA yang melibatkan faktor musiman dalam proses pemodelannya memberikan hasil *forecasting* yang lebih baik dari model ARIMA yang tidak melibatkan faktor musiman dalam proses pemodelannya, baik berdasarkan besarnya nilai MAPE maupun berdasarkan perilaku dari hasil *forecasting* dalam menirukan perilaku dari data historisnya. Namun demikian, hasil *forecasting* melalui model SARIMA memiliki nilai *Mean Absolute Percentage Error* (MAPE) yang relatif masih cukup besar, dimana model terbaiknya memiliki nilai MAPE sebesar 36,94%.

Berdasarkan hal-hal tersebut di atas peneliti terdorong untuk mencoba kembali melakukan pemodelan dan *forecasting* dengan menggunakan alternatif metode non parametrik, yaitu metode *Singular Spectrum Analysis* (SSA). SSA tidak mensyaratkan perlunya asumsi kenormalan untuk data deret waktu. Data yang digunakan adalah sama dengan yang digunakan untuk pemodelan dan *forecasting* melalui metode SARIMA yang telah dilakukan sebelumnya. Penelitian ini dilakukan dengan tujuan untuk mengetahui hasil *forecasting* melalui metode SSA yang merupakan metode non

parametrik dan mengetahui bagaimana komparasi hasil *forecasting* metode SSA tersebut dengan hasil *forecasting* model SARIMA yang merupakan statistik parametrik yang sebelumnya dilakukan oleh Ruhiat & Effendi (2018).

METODE PENELITIAN

Lokasi dan Data Penelitian

Data studi kasus yang digunakan pada penelitian ini adalah data sekunder berupa debit sungai hasil pencatatan pos duga air (PDA) Sungai Citarum yang secara geografis terletak pada koordinat 6° 57' 00" LS dan 107° 32' 00" BT, tepatnya berada di Kecamatan Batujajar Kabupaten Bandung Provinsi Jawa Barat. Data diperoleh dari Pusat Penelitian dan Pengembangan Sumber Daya Air (PUSAIR), Kementerian Pekerjaan umum, Jl. Ir. H. Juanda No. 193 Bandung. Panjang data yang digunakan adalah hasil pencatatan selama 23 (dua puluh tiga) tahun terakhir atau meliputi 276 data debit sungai bulanan.

Identifikasi data

Identifikasi data merupakan tahap awal dari proses pemodelan dan *forecasting* runtun waktu. Beberapa jenis identifikasi yang perlu dilakukan adalah identifikasi gejala autokorelasi, kestasioneran dan identifikasi ada tidaknya gejala musiman.

Metode Analisis Data

Kajian Teori

Kajian teori dilakukan dengan mengkaji teori *Basic Singular Spectrum Analysis* (SSA) dan mengkaji teori-teori yang terkait dan mendukung SSA berdasarkan referensi-referensi yang ada. Algoritma *basic* SSA pada dasarnya terdiri atas dua tahapan, yaitu dekomposisi dan rekonstruksi yang masing-masing memiliki dua tahapan. Tahap dekomposisi terdiri atas *embedding* dan *Singular Value Decomposition* (SVD), dan pada tahap rekonstruksi terdiri atas *grouping* dan *diagonal*.

1) Dekomposisi

Embedding

Pada langkah *embedding*, runtun waktu *unidimensional* $Y_N = \{y_1, \dots, y_N\}$ dipetakan menjadi seri *multidimensional* X_1, \dots, X_K dengan vectors $X_i = (y_i, \dots, y_{i+L-1})^T \in R^L$, dimana $K = N-L+1$. Vektor X_i disebut vektor lag L . Pada tahapan ini diperlukan penentuan parameter *Window Length* (L), dengan ketentuan $2 \leq L \leq N/2$. Hasil dari langkah *embedding* adalah matrix X *trajectory* berukuran $L \times K$, yang juga merupakan matrix *Hankel* dimana semua elemen anti diagonalnya mempunyai nilai yang sama, dan dinyatakan dalam bentuk sebagai berikut:

$$X = [X_1 : \dots : X_K] = (x_{ij})_{i,j=1}^{L,K} = \begin{pmatrix} y_1 & y_2 & \dots & y_K \\ y_2 & y_3 & \dots & y_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \dots & y_N \end{pmatrix} \tag{1}$$

Singular Value Decomposition

Langkah kedua adalah membuat *Singular Value Dekomposition* (SVD) dari matriks *trajectory* X dan mewakilinya sebagai jumlah dari matriks-matriks *biorthogonal*. SVD dari matriks *trajectory* dapat dinyatakan sebagai berikut:

$$X = \sum_{i=1}^d U_i \sqrt{\lambda_i} V_i^T \tag{2}$$

Matriks X terbentuk dari *eigenvector* U_i , *Singular Value* $\sqrt{\lambda_i}$ dan *pricipal component* V_i^T . Ketiga elemen tersebut merupakan pembentuk SVD dan disebut *eigentriple*.

2) Rekonstruksi

Grouping

Tahapan *grouping* memecah matriks X_i menjadi beberapa kelompok dan menjumlahkan matriks-matriks tersebut ke dalam masing-masing kelompok. Misalkan $I = \{i_1, \dots, i_p\}$ adalah kelompok dari indeks i_1, \dots, i_p . Kemudian matriks X_I disesuaikan dengan kelompok I didefinisikan sebagai $X_I = X_{i_1} + \dots + X_{i_p}$. Pemisahan set indeks $J = \{1, \dots, d\}$ menjadi subset disjoin I_1, \dots, I_m sesuai pernyataan berikut:

$$X = X_{I_1} + \dots + X_{I_m} \quad (3)$$

Diagonal Averaging

Diagonal *averaging* adalah suatu proses yang menstransformasikan setiap matriks X_{I_j} dari grup dekomposisi (6) ke dalam bentuk matriks Hankel yang kemudian dapat dikonversi menjadi runtun waktu, yang merupakan komponen tambahan dari seri awal Y_N .

$$y_t = \sum_{k=1}^m y_t^{(k)} \quad , t = 1, \dots, m \quad (4)$$

Dimana $\tilde{y}_N^{(k)} = \tilde{y}_1^{(k)} + \dots + \tilde{y}_m^{(k)}$ sesuai dengan matriks \tilde{X}_{I_k}

Kajian teori-teori yang terkait dengan SSA berdasarkan referensi-referensi yang mendukung, antara lain: *A new approach for selecting the number of eigenvalues in SSA* oleh Alharby & Hassani (2016); *Pengujian pola musiman pada data deret waktu dengan menggunakan regresi spektral* oleh Darmawan et,al (2012); *Literatur review singular spectral analysis* oleh Darmawan (2017); *Singular spectrum analysis based on the minimum variance estimator* oleh Hassani (2010); dan *An improved ssa forecasting result base on filtered reccurent forecasting algorithm* oleh Hassani et, al (2017).

Proses Komputasi

Proses komputasi merupakan proses perhitungan untuk seluruh tahapan analisis dan *forecasting basic SSA*, dilakukan dengan menggunakan bantuan *Software R*, dalam hal ini menggunakan *Package Rssa*.

1. Ukuran Kebaikan Model

Ukuran kebaikan model yang digunakan adalah *Mean Absolute Percentage Error (MAPE)* dengan rumus:

$$MAPE = \frac{\sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{y_t}}{T} \times 100 \% \quad (5)$$

dimana :

T = Banyak periode ramalan/dugaan

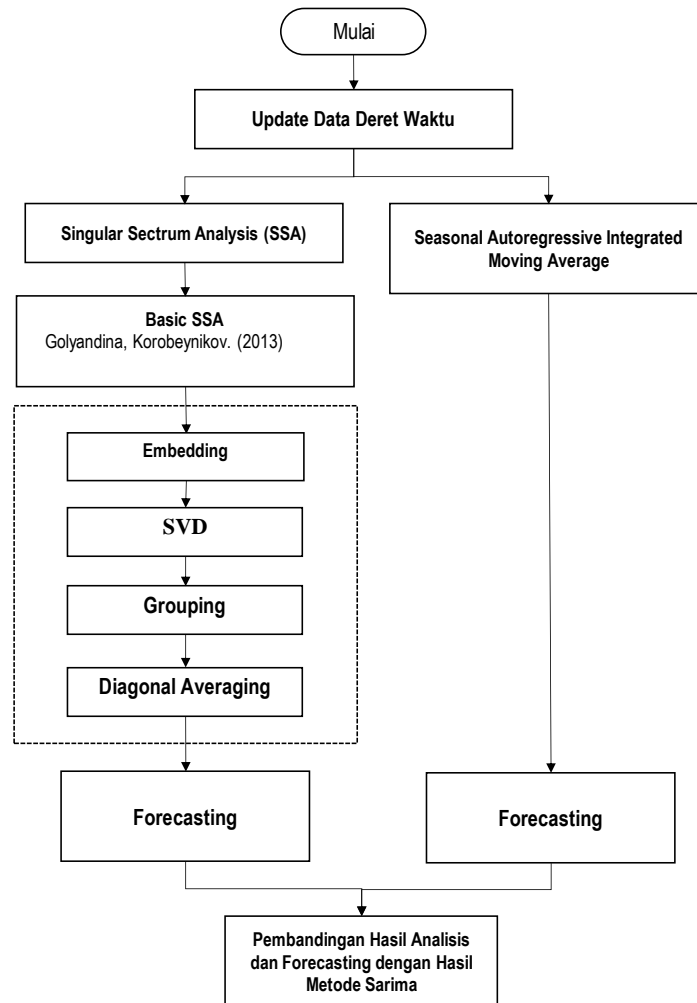
2. Perbandingan Hasil *Forecasting*

Hasil *forecasting* metode non parametrik SSA selanjutnya dibandingkan dengan hasil *forecasting* metode parametrik SARIMA dengan indikator pembandingnya adalah nilai MAPE dan perilaku hasil *forecasting* dari kedua metode tersebut.

3. Kalibrasi Model

Tahap akhir proses *forecasting* adalah kalibrasi atau verifikasi hasil *forecasting* terhadap data runtun waktu historis.

Penggambaran penyelesaian masalah dalam bentuk langkah-langkah pelaksanaan penelitian disajikan dalam Gambar 1 berikut.



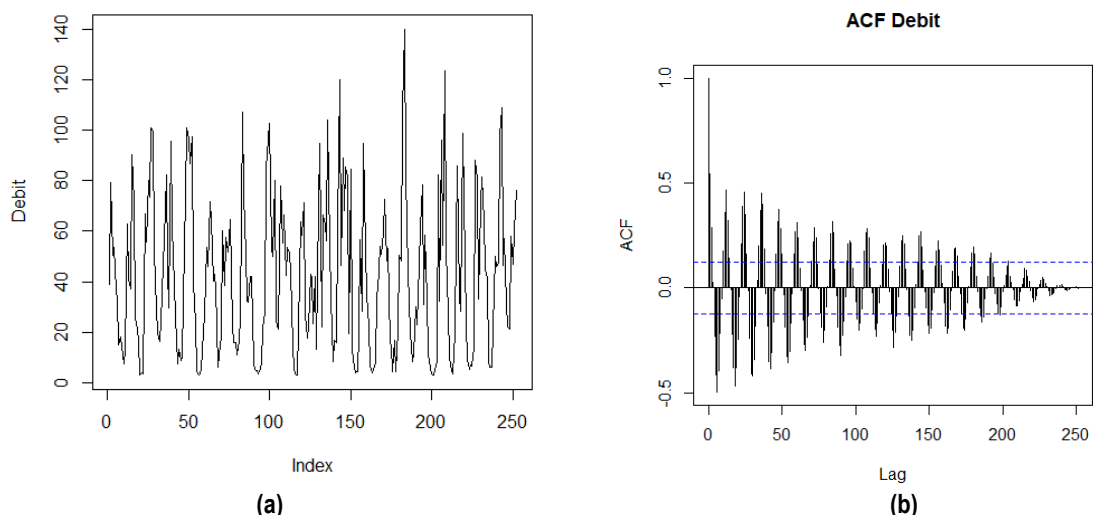
Gambar 1. Alur basic SSA pada proses forecasting

HASIL DAN PEMBAHASAN

Identifikasi data deret waktu

Identifikasi karakteristik data perlu dilakukan guna mengetahui karakteristik data runtun waktu sebelum dilanjutkan tahap selanjutnya. Identifikasi bisa dilakukan melalui *ploting* data dan atau melalui beberapa pengujian. Karakteristik data yang perlu diidentifikasi adalah autokorelasi, stasioneritas dan pola musiman. Identifikasi autokorelasi dilakukan untuk mendeteksi apakah terjadi korelasi antara series data runtun waktu pada suatu periode waktu dengan data di periode waktu lainnya. Apabila autokorelasi terpenuhi maka analisis data runtun waktu bisa dilanjutkan.

Kestasioneran data runtun waktu terjadi dalam rata-rata dan dalam varians. Identifikasi kestasioneran dalam rata-rata dapat diketahui melalui uji Augmented Dickey Fuller (ADF). Jika hasil uji menunjukkan data runtun waktu tidak stasioner maka dapat distasionerkan melalui proses pembedaan (*differencing*). Kestasioneran dalam varians dapat diketahui melalui uji *Lavene* atau uji Bartlet. Data runtun waktu yang tidak stasioner dalam varians dapat diatasi dengan melakukan transformasi Box-Cox. Kemudian identifikasi pola musiman pada data runtun waktu, selain dapat dilakukan secara visual melalui plot ACF (*Auto-Correlations Function*) dan plot data per periodenya (*Buys Ballot*), juga dapat dideteksi dengan menggunakan metode regresi spektral yang dapat mendeteksi periodesitas tersembunyi. Hasil *ploting* data deret waktu terkait dengan identifikasi data runtun waktu disajikan pada Gambar 2 berikut.



Sumber: Hasil olah data, 2019
Gambar 2. (a) Plot data per periodenya (Buys Ballot); (b) Plot ACF

Grafik pada Gambar 2 (a), secara visual menunjukkan bahwa data runtun waktu debit sungai Citarum PDA Nanjung teridentifikasi berpola fluktuatif yang tersebar di sekitar rata-rata sehingga hal ini mengindikasikan adanya gejala stasioner dalam rata-rata dan menunjukkan adanya pola musiman pada data runtun waktu tersebut. Kemudian secara seksama pada plot data tersebut bisa dilihat cukup jelas bahwa variasi lebar pola plot data runtun waktu tidak seragam dan hal ini mengindikasikan adanya gejala ketidakstasioneran dalam varians. Sedangkan grafik pada Gambar 2 (b) secara visual menunjukkan adanya pola sinusoidal yang mengindikasikan bahwa data runtun waktu berpola musiman.

Berdasarkan hasil beberapa pengujian diperoleh informasi terkait dengan karakteristik data runtun waktu sebagai berikut:

- 1) Hasil pengujian autokorelasi dengan menggunakan *Software* Minitab menunjukkan adanya gejala autokorelasi, dimana hasil pengujian menunjukkan terdapat beberapa nilai T hitung yang lebih besar dari nilai $Z_{0,05} = 1,645$ ($T > 1.645$) atau lebih kecil dari nilai $-Z_{0,05} = -1,645$ ($T < -1.645$).
- 2) Pengujian kestasioneran melalui uji Dickey-Fuller dengan nilai $\alpha = 5\%$, dan perhitungannya menggunakan *software R*, menunjukkan bahwa data debit sungai citarum PDA Nanjung adalah stasioner dalam rata-rata.
 - **Sintaks software R :**
`# Uji stasionaritas (Dickey-Fuller test)`
`adf.test(x,alternative = c("stationary"))`
 - **Output software R :**
 Augmented Dickey-Fuller Test
 data: x
 Dickey-Fuller = -6.4937, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary
- 3) Identifikasi pola musiman dilakukan melalui metode regresi spektral, yang dalam proses perhitungannya menggunakan bantuan *software R*. Hasil memberi informasi bahwa nilai $T = 0,442893 < g_{\alpha} = 0.13135$, dengan demikian disimpulkan bahwa data debit Sungai Citarum PDA Nanjung adalah data runtun waktu yang dipengaruhi oleh faktor musiman dan memiliki periodesitas 12.

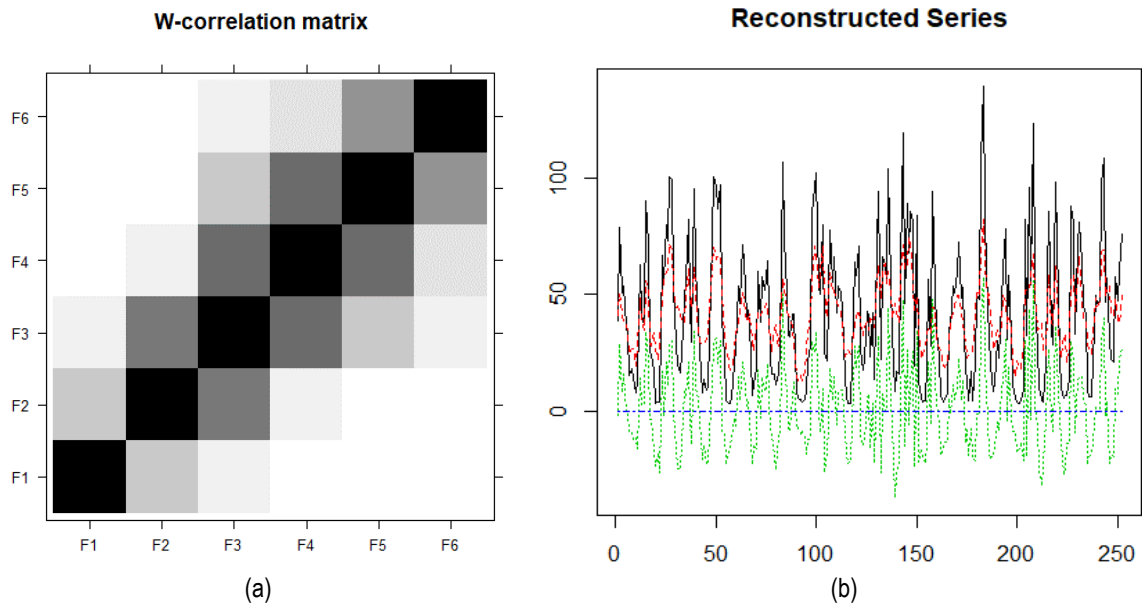
Hasil Forecasting

Algoritma *basic SSA* pada dasarnya terdiri atas dua tahapan, yaitu dekomposisi dan rekonstruksi yang masing-masing memiliki dua tahapan. Tahap dekomposisi terdiri atas *embedding*

dan *Singular Value Decomposition* (SVD), dan pada tahap rekonstruksi terdiri atas *grouping* dan *diagonal averaging*. Setelah melalui beberapa tahapan tersebut selanjutnya dilakukan proses *forecasting*.

Tahap Dekomposisi

Pada langkah *embedding*, runtun waktu unidimensional $Y_N = \{y_1, \dots, y_N\}$ dipetakan menjadi seri multidimensional X_1, \dots, X_K dengan *vectors* $X_i = (y_i, \dots, y_{i+L-1})^T \in R^L$, dimana $K = N-L+1$. Vektor X_i disebut vektor lag L. Pada tahapan ini diperlukan penentuan parameter *Window Length* (L), dengan ketentuan $2 \leq L \leq N/2$. Hasil dari langkah *embedding* adalah matrik *X trajectory* berukuran L x K,



Sumber: Hasil olah data, 2019
Gambar 3. (a) Matrik korelasi; (b) Tahapan rekonstruksi

Perhitungan algoritma *Basic SSA* untuk *forecasting* runtun waktu debit sungai pada penelitian ini menggunakan bantuan *software R*. Untuk mendapatkan hasil *forecasting* terbaik dengan nilai MAPE terkecil, pada tahap dekomposisi harus ditentukan nilai parameter *Window Length* (L), dengan ketentuan $2 \leq L \leq N/2$. Untuk nilai $L = 6$ diperoleh matrik korelasi yang secara visual disajikan pada Gambar 3.(a).

Dengan jumlah data runtun waktu sebanyak 252, kemungkinan nilai L berada antara $2 \leq L \leq 126$. Setelah dilakukan perhitungan diperoleh hasil bahwa *forecasting* data runtun waktu debit sungai Citarum PDA Nanjung terbaik terjadi pada nilai $L = 6$, dengan nilai MAPE sebesar 0,00000632% untuk panjang *forecasting* 24 bulan dan sebesar 0,00000457% untuk panjang *forecasting* 12 bulan.

Beberapa gambar terkait tahapan proses *forecasting* melalui algoritma *Basic SSA*, diantaranya adalah matrik korelasi pada tahapan dekomposisi disajikan pada Gambar 3 (a); Gambar 3 (b) menggambarkan tahapan rekonstruksi; Gambar 4 menggambarkan *output* perhitungan berupa matrik korelasi dan Gambar 5 menggambarkan hasil rekonstruksi.

```
RGui (32-bit)
File Edit View Misc Packages Windows Help

R Console

$`3`
[1] 4 5 6

attr(,"hclust")

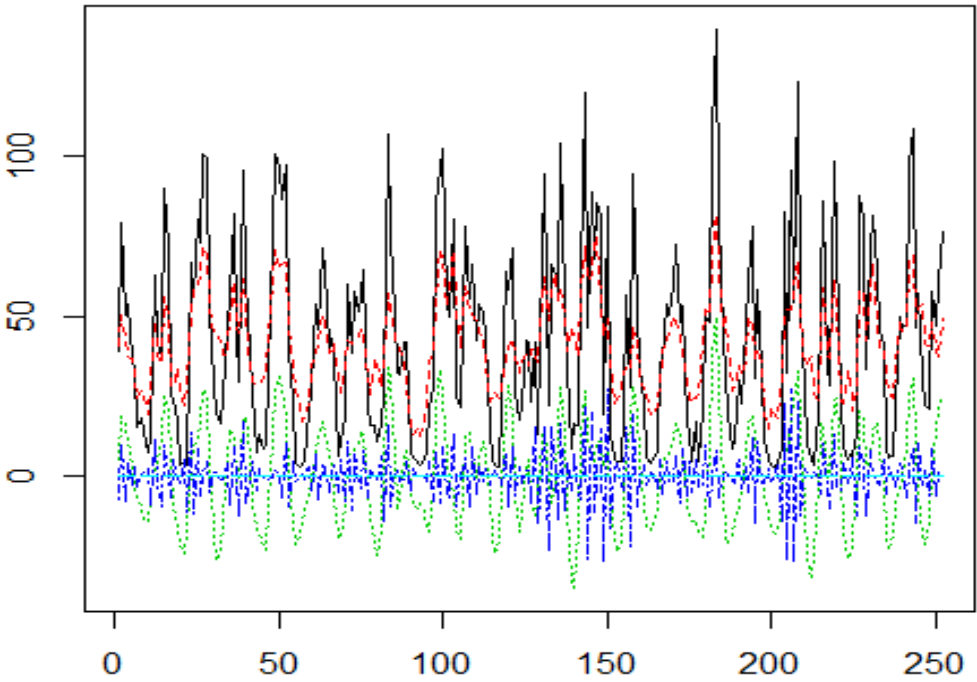
Call:
hclust(d = as.dist((1 - w)/2))

Cluster method : complete
Number of objects: 6

attr(,"wcor")
      F1      F2      F3      F4      F5      F6
F1 1.00000000 0.028345257 0.008736487 0.0322457 0.003837558 0.001698703
F2 0.028345257 1.000000000 0.960029361 0.1345671 0.024597655 0.004906242
F3 0.008736487 0.960029361 1.000000000 0.1781180 0.059613338 0.020308641
F4 0.032245705 0.134567080 0.178118001 1.0000000 0.695885132 0.397349072
F5 0.003837558 0.024597655 0.059613338 0.6958851 1.000000000 0.753391122
F6 0.001698703 0.004906242 0.020308641 0.3973491 0.753391122 1.000000000
attr(,"class")
[1] "wcor.matrix"
attr(,"class")
[1] "grouping.auto.wcor"
> |
```

Sumber: Hasil olah data, 2019
Gambar 4. Matriks korelasi

Data hasil *forecasting* runtun waktu debit Sungai Citarum PDA Nanjung melalui Algoritma *Basic Singular Spectrum Analysis (Basic SSA)* yang dalam proses perhitungannya menggunakan bantuan program *R* disajikan pada Gambar 6 dan secara visual diilustrasikan pada grafik Gambar 7.



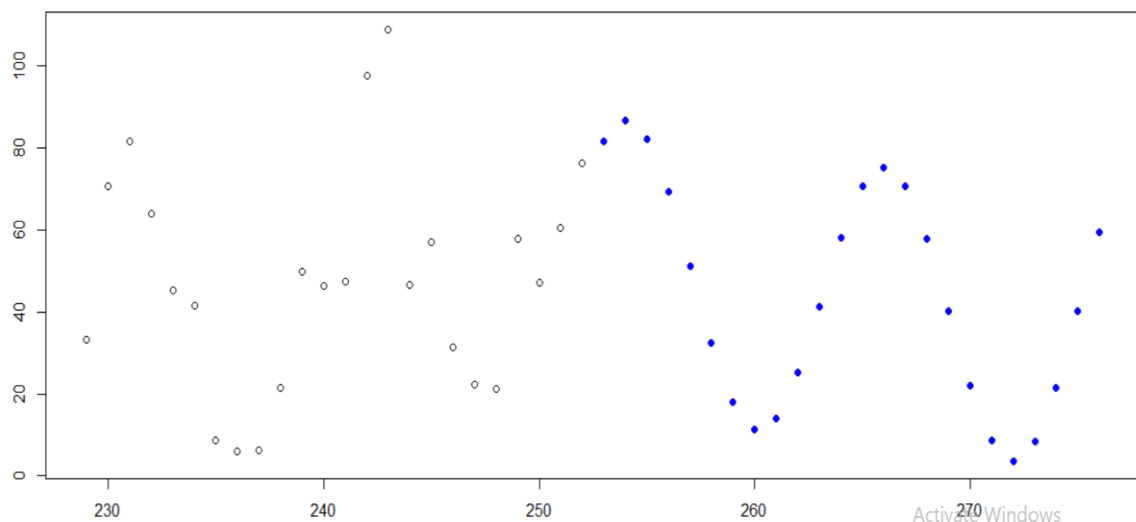
Sumber: Hasil olah data, 2019
Gambar 5. Hasil rekonstruksi


```

File Edit View Misc Packages Windows Help
[Icons]
R Console
7      17.917208
8      11.274480
9      13.997905
10     25.095186
11     41.325577
12     58.072579
13     70.588277
14     75.275846
15     70.665647
16     57.821640
17     40.067172
18     22.101765
19     8.744312
20     3.639202
21     8.272808
22     21.565448
23     40.150141
24     59.265338
> ##Plot the result including the last 24 points of the series##
> plot(f, include = 24, shadecols = "green", type = "l")
>
> ###MAPE###
> MAPE=( (abs((res[229:252])/Debit[229:252]))/length(Debit[229:252]))*100
> sum(MAPE)
[1] 6.316045e-14
    
```

Sumber : Hasil olah data, 2019

Gambar 6. Data dan nilai MAPE hasil *forecasting* untuk 24 bulan



Sumber : Hasil olah data, 2019

Gambar 7. Grafik data hasil *forecasting* untuk 24 bulan

Perbandingan data debit sungai hasil pengamatan dan hasil *forecasting* untuk 2 tahun (24 bulan) disajikan pada Tabel 1 dan digambarkan pada Gambar 8.

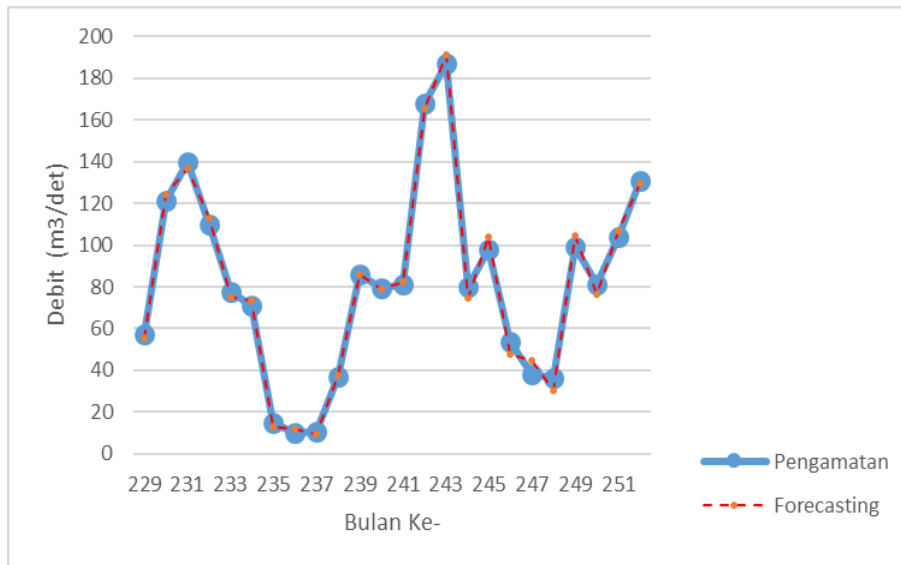
Tabel 1.
Data debit sungai citarum PDA Nanjung hasil pengamatan dan hasil *forecasting*

Bulan Ke	Aliran /Km ² (lt/det)		
	Pengamatan	Residu	<i>Forecasting</i>
1	57,04	-1,79	55,25
2	121,46	2,63	124,09
3	140,02	-2,98	137,04
4	109,95	3,05	113,00
5	77,83	-2,88	74,95
6	71,13	2,57	73,69
7	14,95	-2,13	12,82
8	10,26	1,55	11,81
9	10,70	-0,98	9,73
10	36,94	0,46	37,39
11	85,73	0,03	85,76
12	79,37	-0,53	78,84
13	81,36	1,31	82,67
14	167,87	-2,64	165,23
15	186,95	4,22	191,17
16	80,01	-5,44	74,57
17	97,87	6,07	103,94
18	53,78	-6,14	47,63
19	38,13	6,45	44,58
20	36,27	-6,06	30,21
21	99,29	5,28	104,57
22	80,96	-4,37	76,59
23	103,69	3,16	106,85
24	130,95	-1,61	129,34

Sumber : Hasil analisis, 2019

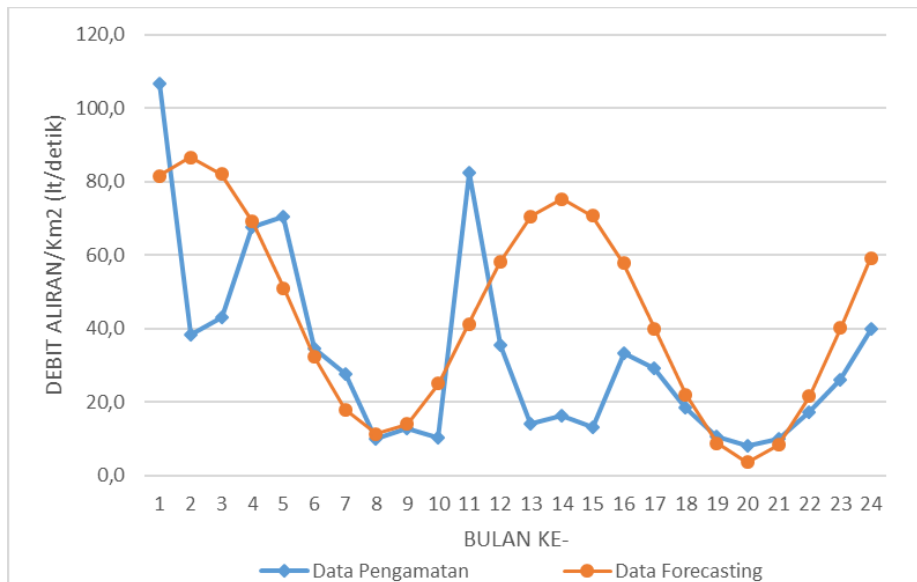
Grafik pada Gambar 8 memperlihatkan bahwa dengan MAPE *in sample* yang mendekati nol (0,00000632), *ploting* data hasil *forecasting* dengan data hasil pengamatan (data historis) untuk 24 bulan terakhir nampak berimpit, atau terjadi dispersi yang sangat kecil. Hal ini menunjukkan bahwa hasil *forecasting* data runtun waktu dengan menggunakan Basic SSA dengan nilai parameter *window Leng* sangat baik. Namun demikian hasil *forecasting* Basic SSA pada nilai parameter L (*Window Length*) = 6 tersebut akan dikalibrasikan dengan data hasil pengamatan. Panjang data untuk kalibrasi adalah data runtun waktu debit sungai untuk 24 bulan terakhir. Hasil kalibrasi disajikan pada Tabel 2 dan Gambar 9.

Grafik pada Gambar 9 memberi gambaran bahwa gejala data hasil *forecasting* untuk tahun pertama mengikuti perilaku data hasil pengamatan (historis), namun pada periode 6 bulan (semester) ke tiga tidak mengikuti perilaku data hasil pengamatan, kemudian memasuki periode 6 bulan (semester) ke-4 data hasil *forecasting* kembali mengikuti perilaku dari data hasil pengamatan.



Sumber: Hasil olah data, 2019

Gambar 8. Data debit sungai PDA Nanjung hasil pengamatan dan hasil forecasting



Sumber: Hasil olah data, 2019

Gambar 9. Ploting kalibrasi data debit sungai PDA Nanjung hasil forecasting dengan hasil pengamatan

Tabel 2.
Kalibrasi data hasil forecasting dengan data historis

Bulan Ke	Aliran /Km ² (lt/det)	
	Pengamatan	Forecasting
1	106,9	81,56
2	38,3	86,61
3	43,2	82,15
4	67,6	69,22
5	70,4	51,22
6	34,5	32,40
7	27,8	17,92
8	10,1	11,27

Bulan Ke	Aliran /Km ² (lt/det)	
	Pengamatan	Forecasting
9	12,7	14,00
10	10,4	25,10
11	82,5	41,33
12	35,5	58,07
13	14,2	70,09
14	16,2	75,28
15	13,2	60,67
16	33,4	57,82
17	29,3	40,07
18	18,6	22,10
19	10,6	8,74
20	8,0	3,64
21	9,9	8,27
22	17,3	21,57
23	26,2	40,15
24	39,8	59,27

Sumber: Hasil analisis, 2019

Komparasi Nilai MAPE in Sample Hasil Forecasting SSA dan SARIMA

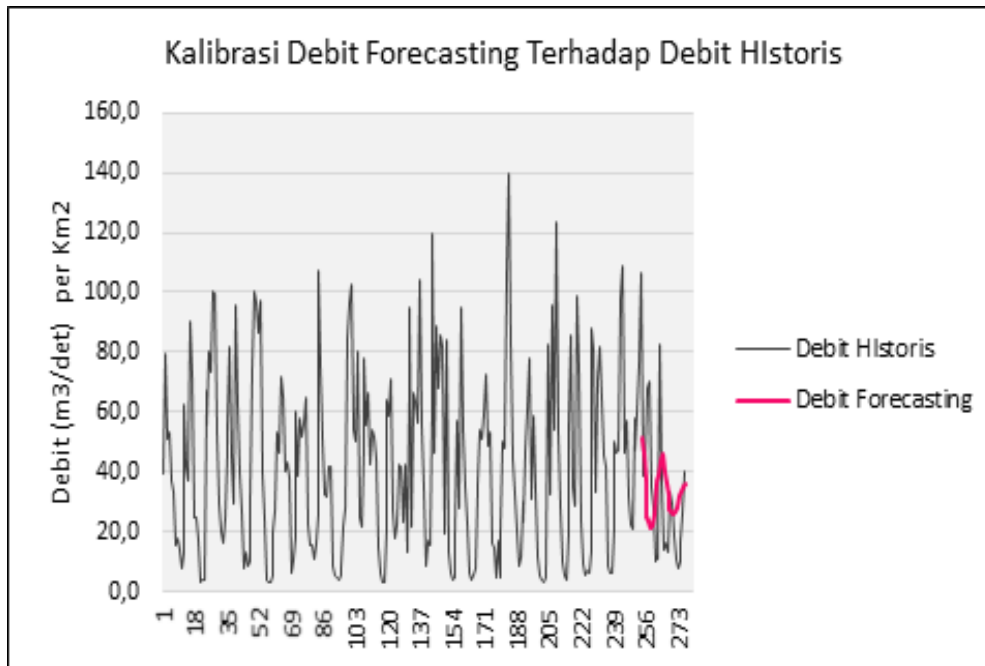
Pemodelan dan *forecasting* data runtun waktu terhadap debit sungai citarum dilakukan oleh Ruhiat & Effendi (2018). Adapun hasil penelitiannya tersebut dapat dilihat pada tabel berikut.

Tabel 3.
Nilai MAPE In Sample Model SARIMA (p,d,q)(P,D,Q)^s

MODEL	MAPE	
	12 Bulan	24 Bulan
SARIMA(2,0,2)(1,0,0) ¹²	36,94	50,42
SARIMA(2,0,2)(1,0,0) ¹²	38,44	56,36
SARIMA(2,0,2)(1,0,0) ¹²	38,81	53,81
SARIMA(2,0,2)(1,0,0) ¹²	40,10	50,92
SARIMA(2,0,2)(1,0,0) ¹²	40,94	48,91

Sumber: Ruhiat & Effendi (2018)

Berdasarkan hasil penelitian sebelumnya diketahui bahwa beberapa model SARIMA terbaik berikut dengan nilai MAPE masing-masing untuk panjang peramalan 12 bulan dan 24 bulan disajikan pada Tabel 3. Data pada tabel tersebut menunjukkan bahwa hasil peramalan (*forecasting*) model SARIMA terbaik yaitu SARIMA (2,0,2)(1,0,0)¹² memiliki nilai MAPE 36,94 untuk panjang peramalan 12 bulan dan 50,42 untuk panjang peramalan 24 bulan. Sedangkan perilaku data *forecasting* terhadap data historis disajikan pada Gambar 10.



Sumber: Ruhiat & Effendi (2018)

Gambar 10. Grafik model SARIMA (2,0,2)(1,0,0)¹²

Komparasi data hasil *forecasting* metode SSA dan metode SARIMA terhadap data runtun waktu debit sungai Citarum hasil pencatatan selama 23 tahun terakhir. Nilai kebaikan model MAPE untuk panjang peramalan 12 bulan dan 24 bulan disajikan pada Tabel 4 berikut.

Tabel 4.
Nilai MAPE *in Sample* Model SARIMA (2,0,2)(1,0,0)² dan SSA

METODE	MAPE	
	12 Bulan	24 Bulan
SSA	0,00000457	0,0000063
SARIMA (2,0,2)(1,0,0) ²	36,94	50,42

Sumber: Hasil analisis, 2019

Berdasarkan data nilai MAPE seperti yang disajikan pada Tabel 4 di atas, dapat disimpulkan bahwa metode SSA yang merupakan metode non parametrik memberikan hasil *forecasting* yang lebih baik dari hasil *forecasting* metode SARIMA yang merupakan metode parametrik.

KESIMPULAN

Berdasarkan kajian teori dan proses komputasi, diperoleh hasil penelitian yaitu: (1) Nilai MAPE hasil *forecasting* untuk panjang peramalan 12 bulan melalui metode *Singular Spectrum Analysis* (SSA) yang proses perhitungannya menggunakan *Software R* diketahui mendekati nilai nol yaitu sebesar 0,00000457% atau mendekati nilai nol terjadi pada nilai parameter L (*Window Length*) = 6. Nilai MAPE terkecil kedua terjadi pada nilai $L = 7$ dengan nilai MAPE sebesar 9,7596%. Sedangkan untuk panjang peramalan 24 bulan adalah sebesar 0,0000063%; dan (2) Nilai MAPE hasil *forecasting* debit sungai Citarum dengan menggunakan metode SSA (*Basic SSA*) ternyata jauh lebih kecil jika dibandingkan dengan nilai MAPE hasil *forecasting* dengan menggunakan metode SARIMA yang dilakukan peneliti pada penelitian sebelumnya terhadap data yang sama. Dari hasil penelitian sebelumnya diketahui besar nilai MAPE model terbaik yaitu model SARIMA (2,0,2)(1,0,0)¹² memiliki nilai MAPE *in sample* sebesar 36,94% untuk panjang peramalan 12 bulan dan nilai MAPE sebesar 50,42% untuk panjang peramalan 24 bulan. Hal ini menunjukkan bahwa metode SSA cukup handal digunakan untuk *forecasting* data yang mengandung pola musiman, seperti data debit sungai.

REKOMENDASI

Forecasting data deret waktu untuk debit Sungai Citarum PDA Nanjung dengan menggunakan metode *Basic Singular Spectrum Analysis* (SSA) atau algoritma *Basic SSA* dan memberikan hasil yang relatif baik dimana diperoleh nilai MAPE yang kecil, baik untuk panjang *forecasting* 12 bulan maupun 24 bulan. Seiring dengan semakin berkembangnya pemakaian metode SSA dan *improvement SSA*, penulis merekomendasikan agar pada *forecasting* data deret waktu pola musiman, khususnya debit sungai dapat dicoba menggunakan metode SSA hasil *improvement*.

UCAPAN TERIMA KASIH

Paper ini merupakan salah satu luaran dari penelitian pada skema Penelitian Dosen Pemula. Untuk itu penulis mengucapkan terima kasih yang sebesar-besarnya kepada Direktorat Riset dan Pengabdian kepada Masyarakat (DRPM), Kementerian Riset, Teknologi dan Perguruan Tinggi (Kemristekdikti) yang telah mendanai penelitian ini.

DAFTAR PUSTAKA

- Alharby, N., & Hassani, H. (2016). A new approach for selecting the number of eigenvalues in singular spectrum analysis. *Journal of Franklin Institute*, 353, 1-16.
- Bako, H. Y., Matias, & Monica. H. (2013). Predictive modeling of pelagic fish catch in Malaysia using seasonal arima models. *Science Group Publishing*, 2(3), 136-140, Technology and Human Development, Universiti Tun Hussein Onn Malaysia.
- Darmawan, G., Mulyani, S., & Sudartianto. (2012). *Pengujian pola musiman pada data deret waktu dengan menggunakan regresi spektral*. Universitas Padjdjaran Djatinangor.
- Darmawan, D. (2017). *Literatur review singular spectral analysis*. Universitas Padjdjaran Djatinangor.
- Hassani, H. (2010). *Singular spectrum analysis based on the minimum variance estimator*. *Real Word Applications*, 11, 2065-207.
- Hassani, H., Kalantari, M., & Yarmohammadi, M. (2017). *An improved ssa forecasting result base on filtered reccurent forecasting algorithm*. *C.R. Acad. Sci. Paris, Ser. I* 355, 1026-1036.
- Juwono, P. T. (2010). Pengaruh perbedaan rerata data debit pada pemodelan deret berkala untuk peramalan debit sungai dengan metode arfima. *Jurnal Pengairan*, 1(2), Fakultas Teknik UNIBRAW.
- Ruhat, D. (2016). Penerapan model seasonal autoregressive fractionally integrated moving average (sarfima) untuk peramalan debit air sungai cimanuk. *Tesis Statistika Terapan*. Universitas Padjdjaran.
- Ruhat, D., & Effendi, A. (2018). Pengaruh faktor musiman pada pemodelan deret waktu untuk peramalan debit sungai dengan metode sarima. *Teorema: Teori dan Riset matematika*, 2(2), 117-128, p-ISSN 2541-0660, e-ISSN 2597-7237.
- Ruhat, D., & Suwanda, C. (2019). Peramalan data deret waktu berpola musiman menggunakan metode regresi spektral. *Teorema: Teori dan Riset matematika*, 4(1), 1-12, p-ISSN 2541-0660, e-ISSN 2597-7237.
- Sakhabakhsh, L., & Yarmohammadi. (2012). An empirical study of the usefulness of sarfima in energi science, *Internastional Journal of Energy Science*, 2(2), 59-63.
- Suprayogi, I., Fauzi, Manyuk, Efrizal, & Eki. (2015). Pengembangan model hidrologi runtun waktu untuk peramalan debit sungai menggunakan daubechies wavelet-adaptive neuro fuzzy inference system (study kasus sub das siak bagian hulu). *Annual Civil Engineering Conference*, ISBN :978-979-792-636-6, Fakultas Teknik, Universitas Riau, Pekanbaru.