Perbandingan Model Peramalan *Singular Spectrum Analysis* (SSA) dan *Fourier Series Analysis* (FSA) pada Data Suhu Udara di Surabaya

Comparative Analysis of Singular Spectrum Analysis (SSA) and Fourier Series Analysis (FSA) on Air Temperature Data in Surabaya

Stefanie Intan Christienova *, Evi Wahyu Pratiwi, Gumgum Darmawan

Program Studi Magister Statistika, Fakultas MIPA, Universitas Padjadjaran, Bandung
e-mail: *si.christienova@gmail.com, evi.wepe@gmail.com, gumstat@gmail.com

Abstrak

Pada dekade terakhir terjadi peningkatan suhu di kota besar tidak terkecuali di Surabaya. Kenaikan suhu juga sangat berkaitan dengan kelembaban udara di suatu wilayah yang juga akan mempengaruhi cuaca. Peramalan yang tepat akan suhu udara sangat dibutuhkan. Dalam penelitian ini akan dibandingkan hasil peramalan dengan menggunakan model Singular Spectrum Analysis (SSA) dengan model Fourier Series Analysis (FSA). Kedua metode ini tidak memerlukan pemenuhan asumsi parametrik dan baik diterapkan pada data musiman. Dari hasil pengujian, data suhu udara yang digunakan dalam penelitian ini memiliki pola musiman. Berdasarkan analisis dengan menggunakan kedua model tersebut, yang memberikan nilai MAPE terkecil adalah FSA sebesar 1,8897 dibandingkan model SSA sebesar 2,00932. Namun, jika dilihat plot data asli dengan hasil rekonstruksi, penghitungan dengan SSA mempunyai plot yang hampir mirip dibandingkan dengan FSA.

***Kata kunci**: Peramalan, Suhu udara, Analisis deret waktu, Singular Spectrum Analysis, Fourier Series Analysis*

Abstract

In the last decade, there was an increase in temperature in big cities, include Surabaya. The increase in temperature is strongly related to the air humidity in an area that will affect the weather. The appropriate forecasting of air temperture is needed therefore in this study we will compare the results of forecasting methods by using Singular Spectrum Analysis (SSA) and Fourier Series Analysis (FSA). Parametric assumptions are not required in this method. Beside that, these methods are good to apply to seasonal data. From the test result, the air temperature data used in this study has a seasonal pattern. Based on the analysis by using both models, FSA provides the smallest MAPE worth 1,8897 compared to the SSA which is 2,00932. However, when looking at the plot of the original data and the reconstruction result, SSA has a similar plot compared to the FSA.

***Keywords**: Forecasting, Air temperature, Time-series anaylisis, Singular Spectrum Analysis, Fourier Series Analysis*

1. Pendahuluan

Kenaikan suhu permukaan bumi yang dikenal dengan *global warming* menyebabkan perubahan pola iklim. Perubahan pola iklim ini menyebabkan tidak menentunya kondisi iklim, dampak perubahan iklim adalah perubahan distribusi curah hujan baik secara spasial maupun temporal serta memicu peningkatan peluang kejadian cuaca dan iklim ekstrem (Trenberth et al, 2004).

Indonesia termasuk negara beriklim tropis sehingga di wilayah ini tidak ditemukan musim dingin tetapi memiliki tingkat curah hujan yang cukup tinggi. Hal ini dikarenakan wilayah perairan Indonesia mendapatkan sinar matahari yang kuat sepanjang tahun serta posisi matahari yang tepat

melintasi khatulistiwa dua kali dalam setahun. Indonesia hanya memiliki dua musim, yaitu kemarau dan hujan. Tingkat curah hujan di Indonesia cukup tinggi walaupun suhu udara rata-rata di Indonesia tidak memiliki perbedaan yang mencolok antar musim, namun demikian suhu udara merupakan salah satu unsur iklim yang perlu diamati. Hal ini berkaitan dengan aplikasi suhu udara untuk berbagai keperluan, antara lain mendeteksi daerah rawan banjir, prakiraan cuaca maupun iklim, mengetahui kondisi pemasan global dan lain sebagainya.

Jawa Timur merupakan salah satu provinsi besar di Indonesia dengan jumlah kabupaten/kota terbanyak dan jumlah penduduk terbesar kedua setelah Jawa Barat. Surabaya sebagai ibukota Jawa Timur tentunya menjadi salah satu kota terpadat di Indonesia. Pada dekade terakhir terjadi peningkatan temperatur di kota besar tidak terkecuali di Surabaya. Tingkat pemanasan rata-rata selama lima puluh tahun terakhir hampir dua kali lipat dari rata-rata seratus tahun terakhir, dimana pemanasan lebih dirasakan pada daerah daratan daripada lautan (Ali dan Brodjol. 2012). Pada sebelas tahun terakhir merupakan tahun-tahun terhangat dalam temperatur permukaan global sejak 1850. Hal tersebut juga berpengaruh pada cuaca di Surabaya. Menurut BMKG suhu panas di Kota Surabaya diperkirakan mencapai puncaknya sejak awal Oktober 2011.

Akibat kenaikan suhu, temperatur udara juga sangat berkaitan dengan kelembaban udara di suatu wilayah yang juga akan mempengaruhi cuaca. Peramalan yang tepat akan temperatur udara sangat dibutuhkan. Dalam prakteknya berbagai macam metode peramalan dapat digunakan untuk melakukan prediksi terdapat nilai sebuah data runtun waktu. Namun, pemilihan metode bergantung pada berbagai aspek yang mempengaruhi, yaitu aspek waktu, pola data, tipe model sistem yang diamati, hingga tingkat keakuratan peramalan yang diinginkan. Disamping itu juga menerapkan suatu metode data juga harus memenuhi asumsi-asumsi yang digunakan. *Singular Spectrum Analysis* merupakan metode peramalan dengan pendekatan non-parametrik, yang artinya metode ini fleksibel karena terbebas dari asumsi parametriknya. Selain peramalan dengan model *Singular Spectrum Analysis*, peramalan musiman yang serupa dapat juga dilakukan dengan menggunakan model *Fourier Series Analysis*. Tujuan dari penelitian ini adalah membuat dan membandingkan sistem peramalan suhu udara dengan metode *Singular Spectrum Analysis* dan *Fourier Series Analysis*.

2. Metode Penelitian

Dalam penelitian ini model yang digunakan adalah model persamaan *Singular Spectrum Analysis* dan *Fourier Series Analysis*. Data yang digunakan merupakan data sekunder yang diperoleh dari BMKG Kota Surabaya. Obyek dalam penelitian ini adalah suhu udara di Stasiun Pengukuran Juanda, Surabaya. Data yang digunakan dalam peramalan ini berdasarkan data suhu udara bulanan Januari 2013-Desember 2015.

2.1 Singular Spectrum Analysis

(1) Dekomposisi

Pada dekomposisi terdapat dua tahap yaitu *Embedding* dan *Singular Value Decomposition* (SVD). Parameter yang memiliki peran penting dalam dekomposisi adalah *Window Length* (L).

Embedding

Misal terdapat data deret waktu $F = (f_0, f_1, \dots, f_{N-1})$ dengan panjang N dan tidak terdapat data hilang. Langkah pertama dalam SSA adalah *embedding* dimana F ditransformasi ke dalam matriks lintasan berukuran L x K. Pada tahap ini diperlukan penentuan parameter *window length* (L) dengan ketentuan 2<L<N/2. *Embedding* dapat dikatakan sebagai pemetaan yang mentransfer data deret waktu F *unidimensional* ke dalam *multidimensional* $X_1, X_2, \dots, X_k$ dengan *lag vector* X_i tersebut kemudian dibentuk matriks lintasan berukuran L x K. Matriks lintasan ini merupakan matriks dimana semua elemen pada anti diagonalnya bernilai sama.

$$X_{i,j} = \begin{pmatrix} f_0 & f_1 & \cdots & f_{k-1} \\ f_1 & f_2 & \cdots & f_k \\ \vdots & \cdots & \ddots & \vdots \\ f_{L-1} & f_L & \cdots & f_{N-1} \end{pmatrix} \quad (1)$$

Konsep dasar pada tahap *embedding* ini adalah melakukan pemetaan yang mentransfer data deret waktu F unidimensional ke dalam multidimensional $X_1, X_2, \ldots, X_k$ sehingga didapatkan *output* sebuah matriks, yaitu matriks *Hankel* dimana semua elemen pada anti diagonalnya bernilai sama.

Singular Value Decomposition

Langkah kedua dalam dekomposisi adalah membuat *Singular Value Decomposition* (SVD) dari matriks lintasan. Secara formal, SVD dari matriks *M* riil atau kompleks berukuran *mxn* adalah faktorisasi dari bentuk *UΣV**, dengan *U* adalah matriks *unitary* berukuran *mxn*, yaitu mempunyai sifat *U*U = UU* = I, Σ* adalah matriks diagonal *rectangular* berukuran *mxn* non negatif dan *V* adalah matriks *unitary* berukuran *nxn*.

Diagonal entri dari *Σii* dari *Σ* di kenal sebagai *singular value* dari *M*. Kolom matriks *U* dan kolom matriks *V* disebut sebagai *left-singular vectors* dan *right-singular vectors* dari *M*. SVD standar dapat di hitung dengan langkah–langkah sebagai berikut :

- Vektor *singular* kiri dari *M* di buat orthonormal dari *MM**
- Vektor *singular* kanan dari *M* di buat orthonormal dari *M*M*
- Akar kan nilai dari *singular value matriks M* yang positif dari *M*M* dan *MM**

Untuk penentuan *singular value* dalam analisis *singular spectrum* adalah sebagai berikut, misalkan $\lambda_1, \ldots, \lambda_L$adalah *eigenvalue* dari matriks ***S*** (dimana $S = XX^T$) dengan urutan yang menurun $\lambda_1 \geq \ \ldots \ \geq \lambda_L \geq 0$ dan $U_1, \ldots, U_L$adalah *eigenvector* dari masing- masing *eigenvalue.* Rank dari matriks ***X*** dapat ditunjukkan dengan $d = max\{i, \lambda_i > 0\}$. Jika dinotasikan$V_i = \frac{X^T U_i}{\sqrt{\lambda_i}}$ untuk *i*=1,...,*d* maka SVD dari matriks lintasan adalah sebagai berikut.

$$\begin{aligned} X &= X_1 + X_2 + \cdots + X_d \\ &= U_1\sqrt{\lambda_1 V_1^T} + U_2\sqrt{\lambda_2 V_2^T} + \cdots + U_d\sqrt{\lambda_d V_d^T} \\ &= \textstyle\sum_{i=1}^{d} U_i\sqrt{\lambda_i V_i^T} \end{aligned} \quad (2)$$

Matriks **X** adalah terbentuk dari eigenvector U_i, singular value $\sqrt{\lambda_i}$dan principal component V_i^T. Ketiga elemen pembentuk SVD ini disebut dengan eigentriple.

Konsep dasar pada tahap ini adalah mendapatkan barisan matriks dari matriks ***S*** dimana pada masing- masing matriks dalam barisan tersebut mengandung *eigenvector* U_i, *singular value* $\sqrt{\lambda_i}$ dan *principal component* V_i^Tyang menggambarkan karakteristik pada masing-masing matriks dalam barisan tersebut.

(2) Rekonstruksi

Grouping

Pada langkah ini, matriks lintasan berukuran L x K diuraikan menjadi beberapa sub-kelompok, yaitu pola *trend*, musiman, periodik, dan *noise*. Pengelompokan berhubungan erat dengan pemecahan matriks X_i menjadi beberapa kelompok dan menjumlahkan matriks dalam masing-masing kelompok. Matriks X_i akan dipartisi ke dalam m subset disjoin $I = \{I_1, I_2, \ldots, I_m\}$.

Misalkan $I = \{i_1, i_2, \ldots, i_p\}$ adalah matriks X_I dengan indeks $i_1, i_2, \ldots, i_p$ sesuai dengan kelompok I yang dapat didefinisikan $X_I = X_{i1} + \cdots + X_{ip}$. Kemudian X_i disesuaikan dengan kelompok $I = \{I_1, I_2, \ldots, I_m\}$. Maka $X = X_1 + X_2 + \cdots + X_d$ dapat diekspansi menjadi $X = X_{I1} + X_{I2} + \cdots + X_{Im}$.

Pengujian Pola Musiman

Pengujian pola musiman pada penulisan ini digunakan analisis spektral. Analisis spektral adalah analisis deret waktu yang dapat menguraikan data ke dalam himpunan gelombang sinus dan atau kosinus pada berbagai frekuensi yang dapat digunakan untuk mencari periodisitas tersembunyi. Analisis spektral dapat mengidentifikasi apakah sebuah data memiliki pola musiman atau tidak kemudian mendeteksi besarnya periode musiman pada data. Jika masing- masing *eigenvector* diklaim memiliki pola musiman kemudian akan ditentukan perioditas musimannya, dimana kelompok *eigen vector* yang memiliki periode yang sama akan dikelompokkan menjadi satu kelompok. Berikut adalah persamaan spektral.

$$Z_t = \sum_{k=0}^{n/2}(a_k \cos \omega_k t + b_k \sin \omega_k t) + e_t \quad (3)$$

Dengan Z_t : series data deret waktu pada periode ke-t
a_k dan b_k : koefisien fourier
$\cos \omega t \; dan \; \sin \omega t$: fungsi kontinu yang tidak berkorelasi
e_t : error pada periode waktu ke-t
ωt : frekuensi fourier
t : periode waktu

Berikut adalah tahapan untuk melakukan pengujian musiman dengan menggunakan analisis spektral:

1. Hitung a_k dan $b_k$pada persamaan (3) dengan rumusan sebagai berikut.

$$a_k = \begin{cases} \frac{1}{n}\sum_{t=1}^{n} Z_t \cos \omega_k t \; ; untuk \; k = 0 \; dan \; k = \frac{n}{2} jika \; n \; genap \\ \frac{2}{n}\sum_{t=1}^{n} Z_t \cos \omega_k t \; ; untuk \; k = 1,2,\dots,\frac{n-1}{2} \end{cases} \quad (4)$$

dan

$$b_k = \frac{2}{n}\sum_{t=1}^{n} Z_t \sin \omega_k t \; ; untuk \; k = 1,2,\dots,\frac{n-1}{2} \quad (5)$$

2. Uji keberartian terhadap masing-masing frekuensi fourier yang telah dihitung pada langkah pertama.
   Hipotesis statistik:

   H_0: $a_k = b_k = 0$ (koefisien fourier tidak berarti)
   H_1: $a_k \neq 0 \; atau \; b_k \neq 0$ (koefisien fourier berarti)
   Statistik uji:

$$F = \frac{(n-3)(a_k^2+b_k^2)}{2\sum_{\substack{j=1 \\ j\neq k}}^{n/2}(a_j^2+b_j^2)} \quad (6)$$

   Mengikuti dstribusi F(2,n-3).
   Jika H_0 signifikan atau koefisien fourier berarti, maka hal ini mengindikasikan bahwa terdapat pola musiman pada data yang akan diujikan.

3. Hitung nilai ordinat $I(\omega_k)$ dengan rumusan sebagai berikut.

$$I(\omega_k) = \begin{cases} na_0^2 \; ; untuk \; k = 0 \\ \frac{n}{2}(a_k^2 + b_k^2) \; ; untuk \; k = 1,2,\dots,\frac{n-1}{2} \\ na_k^2 \; ; untuk \; k = \frac{n}{2} \end{cases} \quad (7)$$

4. Melakukan pengujian untuk melihat dimana letak pola musiman, menggunakan statistik uji T sebagai berikut.

$$T = \frac{I^{(1)}(\omega_{(1)})}{\frac{n}{\sum_{k=1}^{2} I(\omega_k)}} \quad (8)$$

dengan, $I^{(1)}(\omega_{(1)})$: ordinat maksimum dari periodogram pada frekuensi fourier

$I(\omega_k)$: nilai ordinat periodogram pada frekuensi fourier ke-k

Kriteria uji : tolak hipotesis nol jika $T_{hitung} > g_{\alpha}$.

(3) Diagonal Averaging

Pada tahap ini akan dilakukan transformasi dari hasil pengelompokkan matriks X_{Ii} ke dalam seri baru dengan panjang N. Tujuan dari tahap ini adalah mendapatkan *singular value* dari komponen- komponen yang telah dipisahkan, kemudian akan digunakan dalam peramalan. Hasil pada tahap ini merupakan matriks **F** sebagai berikut.

$$F = \begin{pmatrix} f_{11} & f_{21} & \cdots & f_k \\ f_{21} & f_{22} & \cdots & f_{k+1} \\ \vdots & \cdots & \ddots & \vdots \\ f_L & f_{L+1} & \ldots & f_N \end{pmatrix} \quad (9)$$

Untuk mencari rata- rata diagonal matriks dapat digunakan persamaan sebagai berikut.

$$g_k = \begin{cases} \frac{1}{k}\sum_{m=1}^{k} f^*_{m,k-m+1} \; ; untuk \; 1 \le k \le L^* \\ \frac{1}{L^*}\sum_{m=1}^{L^*-1} f^*_{m,k-m+1} \; ; untuk \; L^* \le k \le K^* \\ \frac{1}{N-k+1}\sum_{m=k-K^*+1}^{N-K^*+1} f^*_{m,k-m+1} \; ; untuk \; K^* \le k \le N \end{cases} \quad (10)$$

dimana $L^* = \min(L, K)$ dan $K^* = \max(L, K)$. Persamaan (10) jika diaplikasikan kedalam matriks resultan X_{im} akan membentuk deret $\tilde{Y}^{(k)} = (\tilde{Y}_1^{(k)}, \ldots, \tilde{Y}_N^{(k)})$. Oleh karena itu, deret asli akan didekomposisi menjadi jumlah dari m deret:

$$\boldsymbol{y_n = \sum_{k=1}^{m} \tilde{Y}_N^{(k)}} \quad (11)$$

(4) Evaluasi Peramalan

Evaluasi peramalan dilakukan untuk melihat kecocokan metode terhadap data. Dengan menggunakan ukuran ketepatan dan pengujian keandalan peramalan.

Ukuran Ketepatan Peramalan

Setelah dilakukan tahapan demi tahapan, selanjutnya menghitung seberapa besar ketepatan peramalan tersebut. Dalam Makridakis (1999) dijelaskan bahwa ukuran ketepatan peramalan dipandang sebagai kriteria penolakan untuk memilih suatu metode peramalan sehingga dapat digunakan untuk menentukan kemungkinan yang lebih baik. Metode yang digunakan adalah *Mean Absolute Percentage Error* (MAPE). Berikut adalah bentuk perhitungannya:

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{Y_T - \hat{Y}_t}{Y_t}\right| \times 100\%$$

Dalam Lewis (1982) di dalam Tsai (2012) dijelaskan bahwa kriteria MAPE sebagai berikut:

Tabel 1. The Standard Level of MAPE (%) model evaluation

< 10%	highly accurate forecasting
10-20%	good forecasting

20-50%	reasonable forecasting
>50%	weak and inaccurate predictability

Pengujian Keandalan Peramalan

Di dalam metode peramalan yang mengasumsikan kesinambungan beberapa pola historis di masa yang akan datang berdasarkan dari masa sebelumnya, *tracking signal* merupakan ukuran toleransi yang dapat digunakan untuk menentukan kemungkinan digunakannya hasil peramalan tersebut yang memperkirakan apabila pola dasar berubah. (Bovas dan Ledolter, 1983) menyatakan bahwa jika nilai-nilai *tracking signal* berada di luar batas yang dapat diterima, yaitu $\pm$ 5 maka model peramalan harus ditinjau kembali dan akan dipertimbangkan model baru. Dengan perhitungan sebagai berikut:

Tracking signal = $\frac{\sum_1^n e_n}{\sum_1^n \frac{|e_n|}{n}}$

(5) Peramalan

Peramalan yang digunakan dalam penelitian ini adalah SSA *recurent*. Dengan bantuan *Linear Recurrent Formula* (LRF) untuk membangun modelnya. Metode Forecasting SSA awalnya di usulkan oleh [5] dan [6] yang biasa disebut singkatan LRF.

$$X_{i+d} = \sum_{k=1}^{d} r_k X_{i+d-k}$$

dimana $1 \leq i \leq N-d$. Untuk menaksir koefisien LRF, yaitu $(r_1, r_2, \dots, r_d)$ digunakan eigenvector yang diperoleh dari langkah SVD. Dengan $P = (p_1, p_2, \dots, p_{L-1}, p_L)^T$, $P^{\overline{V}} = (p_1, p_2, \dots, p_{L-1})^T$, π_i komponen terakhir dari vektor $(p_1, p_2, \dots, p_{L-1,} p_{L,})$, dan $v^2 = \sum_{i=1}^{L-1} \pi_i^2$ maka koefisien LRF (vektor R) dapat dihitung dengan persamaan:

$$(r_{L-1}, \dots, r_1) = \frac{1}{1-v^2} \sum_{i=1}^{L-1} \pi_i P_i^{\overline{V}}$$

Dalam peramalan SSA *forecasting* ini, deret waktu yang digunakan adalah deret hasil rekonstruksi yang berupa kombinasi linear komponen pertama dan vektor R. Kemudian akan ditentukan M buah titik data baru yang akan diramalkan.

$$g_i \begin{cases} \tilde{X}_i & , 0 \leq i \leq N \\ \sum_{j=1}^{L-1} r_j g_{i-j} & , N+1 \leq i \leq N+M \end{cases}$$

Maka terbentuk deret hasil peramalan, yaitu $G_{N+M} = (g_1, \dots, g_{N+M})$ dimana $g_{N+1}, \dots, g_{N+M}$ adalah hasil ramalan dari SSA *forecasting*.

2.2 Fourier Series Analysis

Model *Fourier Series Analysis* atau yang dikenal dengan model regresi spektral merupakan model peramalan yang memungkinkan untuk meramalkan pola suhu udara. Dalam penelitian ini, dilakukan analisis untuk data suhu udara yang mempunyai pola musiman. Pola musiman di plot lalu orde spektralnya ditentukan kemudian dilakukan peramalan dengan menggunakan bantuan software R. MAPE digunakan untuk mengukur kebaikan peramalan dari hasil yang didapat.

Persamaan umum Model *Fourier Series Analysis* (FSA) atau yang dikenal juga dengan persamaan regresi spektral mempunyai persamaan sebagai berikut:

$$\widehat{Y}_t = a_0 + b_0 t + a_1 \cos(\omega t) + b_1 \sin(\omega t) + \cdots + a_k \cos(\omega t) + b_k \sin(k\omega t) \qquad (12)$$

dengan:

$\hat{Y}_t$ = nilai fitted atau ramalan pada waktu ke-t
a_0 = konstanta yang digunakan untuk menentukan tingkat dari data deret waktu
b_0 = taksiran trend dari data deret waktu
$a_1, b_1, a_2, b_2, \ldots$ = koefisien yang mendefinisikan amplitude dan phase
ω = $2\pi f/n$ (omega)
k = harmonik dari ω

Walaupun persamaan matematiknya sampai dengan orde k, akan tetapi biasanya hanya sampai orde ke-5. Ini dikarenakan jika ordenya lebih dari 5, maka persamaan tersebut tidak sederhana (*parsimony*) lagi.

Metode pemodelan dalam regresi spektral meliputi dua tahap, yang pertama adalah penentuan orde dari persamaan, lalu langkah kedua melakukan peramalan dengan menggunakan persamaan di atas. Setelah dilakukan peramalan ditentukan nilai MSE dan MAPE dari data *outsample* yang telah ditentukan. Penentuan data *outsample* disesuaikan dengan banyaknya pengamatan yang akan diramalkan (*forecast*), adalah seperti berikut:

1. Dalam *R-Language*, tentukan dahulu rata-rata pembedaan musiman data deret waktu. Ini merupakan taksiran dari rata-rata *trend* tahunan dengan persamaan:

$$Mean\ (Y_t - Y_{t-12})$$

Tentukan trend bulanan dengan membagi rata rata tahunan dengan panjang musiman (S).

$$\frac{Mean\ (Y_t - Y_{t-12})}{S}$$

Pusatkan trend dari rata rata Yt dan rata rata dari waktu-t, kemudian tentukan persamaan garis trend-nya.

Tentukan nilai trend untuk data awal dan data terakhir, t =1 dan t= N.

Tentukan nilai deviasi dari trend dengan mengurangi proyeksi trend dari setiap data aktual untuk menghasilkan data runtun baru dengan data terpusat di nol (0).

Input runtun detrended dan nilai trigonometri terhadap nilai regresi multipel nilai-nilai aktual detrended-nya adalah $\cos(\omega_t), \sin(\omega_t), \cos(2\omega_t), \sin(2\omega_t), \ldots$

Cocokan koefisien model FSA (Fourier Series Analysis) terhadap nilai detrended-nya dengan menggunakan model regresi multipel untuk mendapatkan koefisien-koefisien $a_1, b_1, a_2, b_2, \ldots$ yang akan meminimumkan jumlah kuadrat dari error.

Hitung amplitudo untuk setiap frekuensi. Nilai amplitude dapat digunakan persamaan

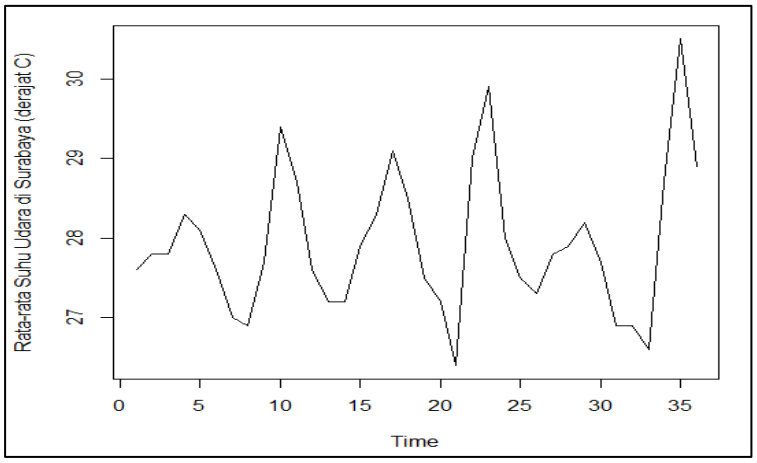
$$A_i = \sqrt{\left(a_i{}^2 + b_i{}^2\right)}$$

Buang frekuensi-frekuensi yang Nilai $A_i > 0{,}5$. Jika frekuensi pada orde 1 dan 2 mempunyai nilai Amplitudo yang lebih besar dari 0,5 dan frekuensi pada orde 3 mempunyai amplitudo lebih kecil dari 0,5, maka data tersebut mempunyai persamaan regresi spektral dengan orde 2.

Ramalkan nilai out-sample sesuai dengan orde yang telah ditentukan pada langkah 8, dengan memproyeksikan komponen-kompenon trend dan musiman-nya.

3. Hasil dan Pembahasan

Langkah pertama dalam analisis data deret waktu adalah memetakan data deret waktu. Berikut pola data rata-rata suhu udara di Kota Surabaya selama periode Januari 2013 sampai Desember 2015. Berdasarkan gambar di bawah, data cenderung mengalami perubahan pola berulang dalam periode sekitar 6 bulanan. Oleh karena itu dapat disimpulkan bahwa data memiliki pola musiman.



Gambar 1. Plot Rata-rata Suhu Udara di Kota Surabaya, 2013-2015

Gambar di atas menunjukkan rata-rata suhu udara bulanan di Kota Surabaya. Jika dilihat secara rata-rata, suhu udara bulanan di Surabaya masih berkisar di angka 28^0c. Angka ini tidak menunjukkan tingkat suhu udara yang tinggi di Surabaya jika dibandingkan dengan wilayah lain di Indonesia. Hal ini mungkin terjadi karena data yang diperoleh adalah data rata-rata bulanan yang pasti akan berbeda jika menggunakan data suhu udara harian.

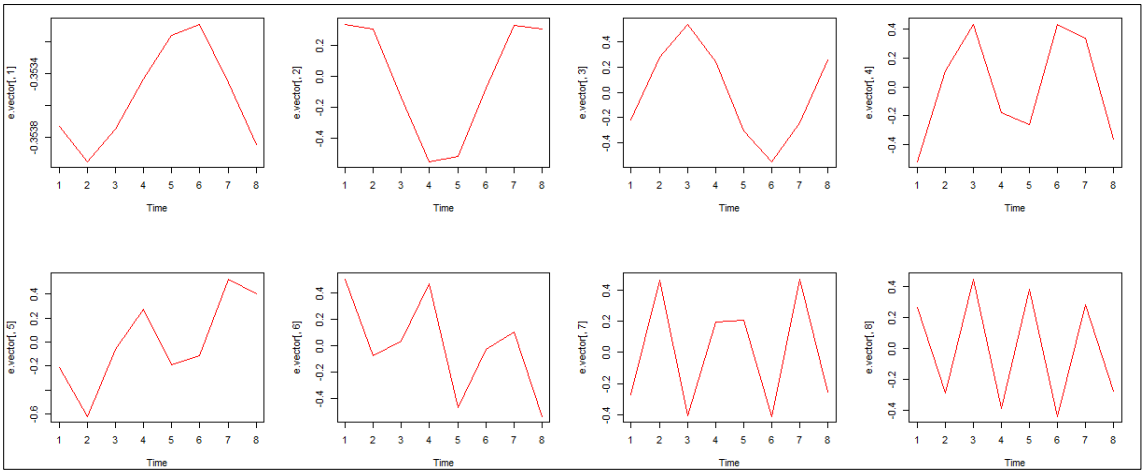
Singular Spectrum Analysis

Langkah pertama dalam SSA adalah *embedding.* Awalnya harus ditentukan parameter *window length* (L) dengan ketentuan 2<L<N/2. *Window length* dihitung dengan membandingkan *forecasting outsample* dengan data *outsample*. Penentuan *window length* (L) ini dilakukan dengan tanpa *grouping*. Dari hasil penghitungan *trial and error* diperoleh L=8 dengan nilai MAPE yang terkecil adalah 3,86 (Tabel 1).

Tabel 2. Perbandingan MAPE dengan Berbagai *Window Length*

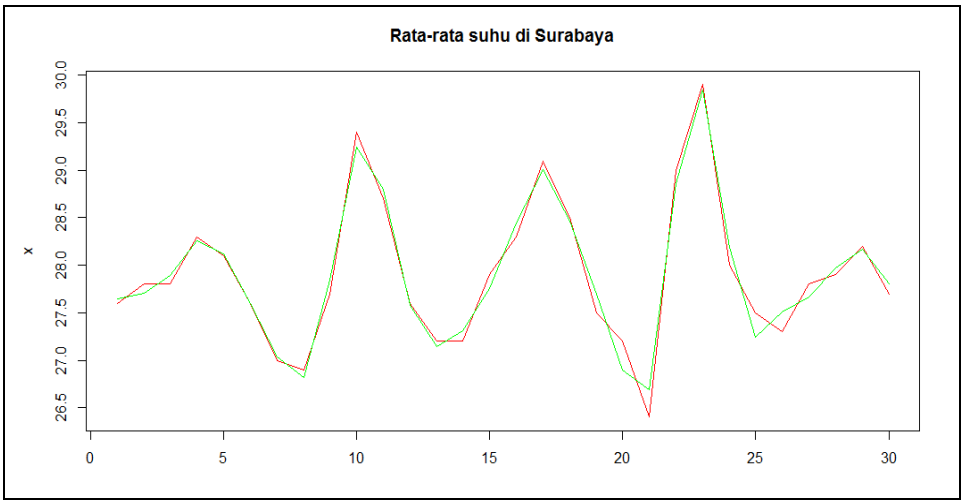
Window Length (L)	MAPE
7	4.13
8	3.86
9	4.21
10	6.40
11	5.20
12	5.58
13	5.39
14	4.83
15	4.58
16	4.55
17	4.46

Selanjutnya mendapat nilai K=30-8+1=23 sehingga pada proses SVD akan membuat matriks dengan L x K. Konsep dasar pada tahap ini adalah mendapatkan barisan matriks dari matriks $\boldsymbol{S}$ dimana pada masing- masing matriks dalam barisan tersebut mengandung *eigenvector* U_i, *singular value* $\sqrt{\lambda_i}$ dan *principal component* V_i^T yang menggambarkan karakteristik pada masing- masing matriks dalam barisan tersebut. Untuk mempermudah melihat pola dari tiap *eigenvector*, berikut ditampilkan plot *eigenvector* yang mengikuti beberapa komponen pola terkecuali komponen *noise*:



Gambar 2. Plot *Eigenvector*

Tahapan selanjutnya adalah menentukan *grouping* dari pola-pola yang hampir sejenis pada *eigentriple*. Dari penghitungan R untuk periode masing-masing *vector* dari matriks **S** diperoleh informasi bahwa terdapat 2 pola, yaitu musiman 1 dengan *eigentriple* 1,2,3 dan musiman 2 dengan *eigentriple* 4,5,6. Setelah diketahui banyaknya *grouping*, langkah selanjutnya dilakukan verifikasi. Langkah ini merupakan suatu langkah untuk meyakinkan banyakya pengelompokkan yang harus dilakukan agar memperoleh hasil terbaik. Verifikasi dilakukan dengan *window length* (L) = 8 dan *grouping* = 2 kelompok (musiman 1 [*eigentriple* 1,2,3] dan musiman 2 [*eigentriple* 4,5,6]) menggunakan metode SSA.



Gambar 3. Perbandingan Deret Asli dengan Deret Rekonstruksi SSA

Gambar tersebut menunjukkan bahwa hasil rekonstruksi (digambarkan dengan garis berwarna hijau) hampir mendekati data aslinya (digambarkan dengan garis berwarna merah) kecuali pada kondisi suhu udara minimum yang menunjukkan sedikit adanya perbedaan. Dengan demikian dapat dikatakan bahwa rekonstruksi menggunakan SSA *Linear Recurrent Formula* dengan *window length* (L) = 8 dan *grouping* = 2 kelompok cukup baik.

Untuk menentukan dapat dilakukannya peramalan maka terlebih dulu dilakukan evaluasi hasil ramalan, yang akan dilihat dari nilai MAPE ukuran ketepatan peramalannya. Hasilnya sebagai berikut:

Tabel 3. Hasil Penghitungan MAPE

Aktual	Prediksi	MSE	MAPE
26,9	27,37967	0,44687	2,00932
26,9	26,70771		
26,6	27,41306		
28,7	29,01980		
30,5	29,27897		
28,9	28,50012		

Tabel di atas menunjukkan bahwa hasil ketepatan ramalan dari deret rekonstruksinya memiliki nilai MAPE 2,00932. Ini artinya hasil peramalan dapat dikatakan sangat baik. Dengan demikian, peramalan suhu udara dengan metode *Singular Spectrum Analysis* memadai.

Selanjutnya diperlukan pengujian keandalan peramalan. Tabel berikut ini menunjukkan *tracking signal* hasil evaluasi suhu udara 6 bulan terakhir. Nilai-nilai *tracking signal* dari 6 periode waktu yang diramalkan menunjukkan besaran yang beragam. Dari hasil penghitungan *Tracking Signal* dapat disimpulkan bahwa peramalan masih bisa digunakan selama 6 periode waktu ke depan, karena batas toleransi yang bisa diterima, yaitu $\pm$ 5 (Bovas dan Ledolter, 1983).

Tabel 4. Perhitungan *Tracking Signal*

Data Ke-	Aktual	Prediksi	Tracking Signal
31	26,9	27,30802	-1,00
32	26,6	26,77778	-0,85
33	28,7	27,41215	-2,22
34	30,5	28,99488	-3,15
35	28,9	29,28656	-0,33
36	26,9	28,46410	0,35

Tahapan terakhir adalah dilakukannya peramalan. Karena peramalan menggunakan *Recurrent Forecasting* maka terlebih dulu dihitung nilai Koefisien Linear formula r_j. Tabel di bawah menyajikan hasil Koefisien Linear Formula r_j, yang akan digunakan dalam perhitungan peramalan.

Tabel 5. Koefisien Linear Recurrent Formula

No	r_j
1	-0,1599357
2	-0,1823291
3	-0,1708013
4	-0,1522614
5	0,4825232
6	-0,4949938
7	0,3079442

Dengan sistem peramalan sebagai berikut:

$$g_{37} = \sum_{j=1}^{12} r_j g_{i-j} = -0{,}1599357(28.46) - 0{,}1823291(29.29) \ldots + 0{,}3079442(27.87)$$

$\vdots$

dst

Dari sistem peramalan yang diperoleh, maka hasil ramalan bulanan rata-rata suhu udara di Kota Surabaya tahun 2016 dengan menggunakan metode *Singular Spectrum Analysis* untuk 6 bulan kedepan sebagai berikut:

Tabel 6. Hasil Ramalan Suhu Udara Bulanan

Bulan	Januari	Februari	Maret	April	Mei	Juni
Hasil Ramalan	28,135	27,695	27,230	27,717	28,230	27,566

Fourier Series Analysis

Kemudian dengan data yang sama dilakukan proses peramalan dengan menggunakan metode FSA. Karena data curah hujan bersifat stasioner maka data di*invers differencing* terlebih dahulu agar pola *trend* terlihat sehingga model regresi spektralnya sesuai. Dengan menggunakan langkah-langkah pada BAB III (Metode Penelitian), diperoleh FSA berorde 1 dan 5. Nilai MAPE untuk 6 data *outsample* adalah 1,8897. Sedangkan nilai MSE 6 data *outsample* adalah 0,3412.

$$\hat{Y}_t = a_0 + b_0 t + a_1 \cos(\omega t) + b_1 \sin(\omega t) + a_5 \cos(5\omega t) + b_5 \sin(5\omega t)$$

Dengan Nilai:

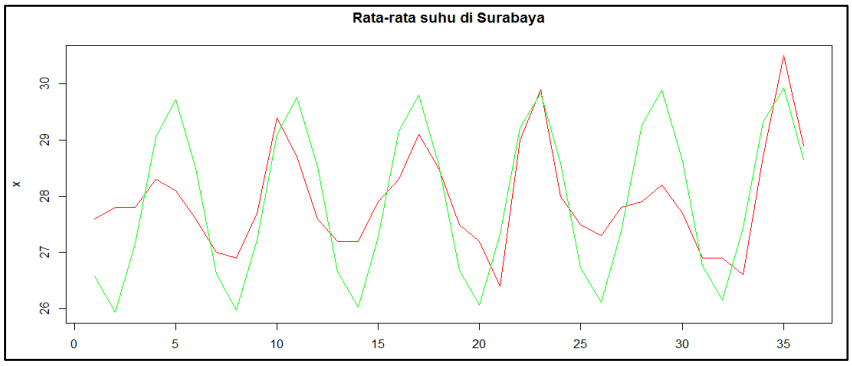
$a_0 = 27.79972$ $a_1 = 0.3023562$ $a_5 = 0.3603278$

$b_0 = 0.007222222$ $b_1 = -0.9111571$ $b_5 = 0.8805421$

Sehingga persamaannya sebagai berikut:

$$\hat{Y}_t = 27{,}80 + 0{,}0072t + 0{,}3024 \cos(\omega t) - 0{,}9112 \sin(\omega t) + 0{,}3603 \cos(5\omega t) + 0{,}8805 \sin(5\omega t)$$

Dari persamaan di atas kemudian dibuat *forecast* data untuk membandingkan dengan data asli. Perbandingan data deret asli dengan *forecast* data dapat dilihat pada gambar di bawah. Gambar di dibawah menunjukkan bahwa hasil f*orecast* data (digambarkan dengan garis berwarna hijau) memiliki periode yang sama dengan data aslinya (digambarkan dengan garis berwarna merah). Namun, *forecast* data kurang begitu mirip dengan data asli di periode awal.



Gambar 4. Perbandingan Deret Asli dengan Forecast Data FSA

Selanjutnya diperlukan pengujian keandalan peramalan. Tabel berikut ini menunjukkan *tracking signal* hasil evaluasi suhu udara 6 bulan terakhir. Nilai-nilai *tracking signal* dari 6 periode waktu yang diramalkan menunjukkan besaran yang beragam. Dari hasil penghitungan *Tracking*

Signal dapat disimpulkan bahwa peramalan masih bisa digunakan selama 6 periode waktu ke depan, karena batas toleransi yang bisa diterima, yaitu $\pm$ 5 (Bovas dan Ledolter, 1983).

Tabel 7. Perhitungan *Tracking Signal*

Data Ke-	Aktual	Prediksi	Tracking Signal
31	26,9	26,75532	1.00
32	26,6	26,15896	2.00
33	28,7	27,43787	0.83
34	30,5	29,31798	-0.97
35	28,9	29,92298	0.01
36	26,9	28,65410	0.48

Dari sistem peramalan yang diperoleh, maka hasil ramalan bulanan rata-rata suhu udara di Kota Surabaya tahun 2016 dengan menggunakan metode *Fourier Series Analysis* untuk 6 bulan kedepan sebagai berikut:

Tabel 8. Hasil Ramalan Suhu Udara Bulanan

Bulan	Hasil ramalan
Januari	26,790
Februari	26,205
Maret	27,493
April	29,370
Mei	29,963
Juni	28,686

4. Kesimpulan

Nilai MAPE menunjukkan bahwa metode analisis dengan menggunakan model FSA menghasilkan nilai yang lebih kecil dibandingkan dengan model SSA. Dapat dikatakan bahwa untuk melakukan peramalan terhadap data suhu udara, analisis dengan FSA lebih akurat daripada SSA. Tetapi jika dilihat berdasarkan plot data asli dengan hasil rekostruksi, penghitungan dengan SSA mempunyai plot yang hampir mirip dengan penghitungan FSA.Hasil ramalan untuk periode 6 bulan ke depan dengan metode SSA dan FSA menunjukkan hasil yang cukup berbeda, berkisar antara 0,5oC – 1,5oC.

Daftar Pustaka

Abraham, B., Ledolter, J., 2005. Statistical Methods for Forecasting. Wiley Interscience.

Ahadiansyah, A., 2009. Perbandingan Model Autoregressive dan Model Analisis Path untuk Data Suhu Minimum Pondok Betung Tangerang Tahun 2007. Universitas Islam Negeri.

Caraka, R.E., 2016. Long Memory Models to Forecasting Temperature. Seminar Nasional Meteorologi dan Klimatologi 2016.

Darmawan, G. 2016. "Identifikasi Pola Data Curah Hujan pada Proses Grouping dalam Metode Singular Spectrum Analysis". Seminar Nasional Pendidikan Matematika 2016.

Darmawan, G., Hendrawati, T., Arisanti, R. 2015. "Model Auto Singular Spectrum untuk Meramalkan Kejadian Banjir di Bandung dan Sekitarnya". Seminar Nasional Matematika dan Pendidikan Matematika UNY 2015.

Darmawan, G., Toharudin, T., Handoko, B. 2016. "Model Regresi Spektral untuk Memodelkan Data Musiman". Seminar Nasional Matematika dan Pendidikan Matematika UNY 2016.

Machmudin, A. Dan Ulama, B.S.S., 2012. Peramalan Temperatur Udara di Kota Surabaya dengan Menggunakan ARIMA dan Artificial Neural Network. Jurnal Sains dan Seni ITS Vol. 1, No. 1.

Pankratz, A. 1983. Forecasting with Univariate Box-Jenkins Models:Concepts and Cases. Wiley Online Library.

Pratopo, A.K.F., 2012. Program Studi Meteorologi, Fakultas Ilmu dan Teknologi Kebumian, Institut Teknologi Bandung.

Ramdani, A.L., 2011. Penggunaan Model Arima dalam Peramalan Suhu Udara di Sekitar Palangkaraya. Institut Pertanian Bogor.

Trenberth, K., Overpeck, J., Solomon, S. 2004. Exploring Drought and Its Implications For the Future. Eos, Transactions American Geophysical Union 85:3, 27.