

ISSN: 2087-2590

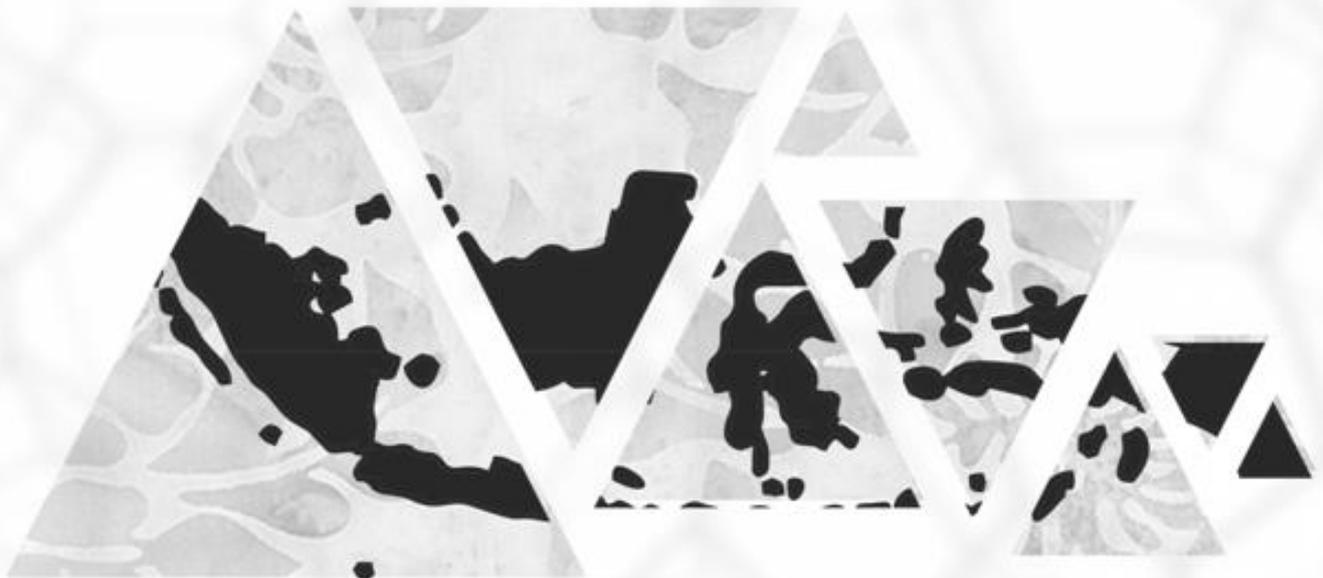


# PROSIDING

## SEMINAR NASIONAL STATISTIKA V

*"Peranan Statistika dan Sistem Informasi  
dalam Pemetaan Kualitas Pendidikan"*

Sabtu, 17 Oktober 2015  
Bale Sawala, Universitas Padjadjaran



DEPARTEMEN STATISTIKA FMIPA  
UNIVERSITAS PADJADJARAN





**PROSIDING SEMINAR NASIONAL STATISTIKA V**

*"Peranan Statistika dan Sistem Informasi dalam Pemataan  
Kualitas Pendidikan"*

**ISSN : 2087-2590**

**VOL 5, NOVEMBER 2015**

**REDAKSI:**

Trysni Malldina

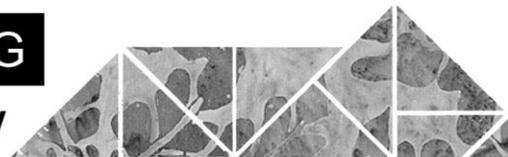
Elsa Febriani

Ristanti Putri Febriana

Riestiawati Dahlan

**SABTU, 17 OKTOBER 2015**

**BALE SANTIKA, UNIVERSITAS PADJADJARAN**



## KATA PENGANTAR

Sege nap puji dan syukur kami panjatkan kehadirat illahi robbi, yang telah memberikan kekuatan dan limpahan karunia yang sangat besar. Sholawat serta salam semoga senantiasa tercurah bagi Nabi Muhammad SAW, panutan dan romhat bagi semesta alam.

Panitia seminar nasional merasa sangat berbahagia dan bersyukur, karena dapat mengadakan acara akbar Seminar Nasional Statistika yang diadakan setiap tahun dan pada tahun ini merupakan tahun kelima.

Prosiding ini merupakan bagian yang tidak terpisahkan dari Seminar Nasional Statistika 2015, karena berisi kumpulan makalah yang dipresentasikan pada sesi paralel dari para pemakalah yang diperoleh dari hasil penelitian, pemikiran, ataupun kajian yang mendalam mengenai suatu topik tertentu dalam ruang lingkup statistika dan aplikasinya.

Ucapan terima kasih kami sampaikan kepada seluruh peserta pemakalah yang telah berkontribusi pada acara seminar ini, selain itu juga tim editor dan redaksi yang telah bekerja keras menyunting dan menyusun prosiding ini sehingga siap disajikan untuk pembicara.

Jatinangor, 11 November 2015

Panitia Seminar Nasional Statistika 2015

## DAFTAR ISI

KATA PENGANTAR..... vii

DAFTAR ISI..... ix

## AKTUARIA/INVESTASI/ANTRIAN/INVENTORI/MATEMATIKA

A.1 EKSPLORASI <i>ADAPTIVE NEURO FUZZY INFERENCE SYSTEM (ANFIS)</i> PADA PERAMALAN INFLASI INDONESIA.....	1
A.2 TIPE HIMPUNAN PADA RUANG PSEUDO METRIK .....	9
A.3 MENENTUKAN <i>BUFFERSTOCK</i> BERDASARKAN RAMALAN KEBUTUHAN DENGAN ADANYA <i>LEAD TIME</i> .....	15
A.4 Matriks Simetrik Secara Eksak Satu Nilai Eigen Positif .....	25
A.5 PERBANDINGAN BANYAKNYA <i>IDLE FACILITY</i> DAN PENGUNJUNG TERLAYANI BERDASARKAN KARAKTERISTIK KEDATANGAN PADA JARINGAN JACKSON DENGAN ENAM WORKSTATION .....	31
A.6 PENDUGAAN PERSENTASE PENDUDUK MISKIN DI KOTA BANDAR LAMPUNG DENGAN METODE PREDIKSI TAK BIAS LINEAR TERBAIK EMPIRIK PADA MODEL PENDUGAAN AREA KECIL .....	41
A.7 OPTIMISASI PORTOFOLIO <i>MEAN- VaR</i> DENGAN KOEFISIEN <i>RISK AVERSION</i> DI BAWAH MODEL DERET WAKTU .....	55

## DATA MINING/DESAIN EKSPERIMEN/SURVIVAL

D.1 APLIKASI <i>LATENT DIRICHLET ALLOCATION (LDA)</i> DALAM SOSIAL MEDIA ...	70
D.2 OPTIMASI MULTI RESPON MENGGUNAKAN METODE <i>OVERLAID RESPONSE SURFACE</i> .....	78
D.3 OPTIMASI MULTI RESPON PADA <i>CENTRAL COMPOSITE DESIGN</i> .....	88
D.4 EVALUASI HASIL PEMBELAJARAN DIKLAT MENGGUNAKAN DISAIN <i>PRE TEST POST TEST NONEQUIVALENT CONTROL GROUP</i> .....	101
D.5 PENENTUAN <i>FITNESS FUNCTION</i> BERDASARKAN PEMBOBOTAN PADA FUNGSI RELIABILITAS DAN BIAYA.....	118

## GENERALIZED LINEAR MODEL/NONPARAMETRIK

G.1 PENENTUAN FUNGSI PEMBOBOT PADA <i>GEOGRAPHICALLY WEIGHTED POISSON REGRESSION SEMIPARAMETRIC</i> .....	131
G.2 ANALISIS KEMATIAN BAYI DENGAN MENGGUNAKAN <i>RANDOM EFFECT DYNAMIC PROBIT MODEL</i> DENGAN MENGGUNAKAN <i>HECKMAN ESTIMATOR</i> .....	143
G.3 PENGELOMPOKAN KECAMATAN BERBASIS SUMBER DAYA MANUSIA DI PROVINSI JAWA BARAT MENGGUNAKAN METODE <i>LATENT CLASS CLUSTER</i> .....	151
G.4 ESTIMASI TOTAL POPULASI MENGGUNAKAN <i>PENALIZED SPLINES MODEL- ASSISTED</i> .....	161

**REGRESI**

R.1 MENENTUKAN METODE ESTIMASI TERBAIK UNTUK PENERAPAN MODEL GRAVITASI MIGRASI SEDERHANA DAN PENGARUH PENGAMBILAN JUMLAH SAMPEL SERTA PROPORSI NILAI NOL DALAM MODEL REGRESI.....169

R.2 PENERAPAN MODEL SPASIAL DATA PANEL PADA DATA PRODUK DOMESTIK REGIONAL BRUTO (PDRB)..... 183

R.3 APAKAH JENJANG PENDIDIKAN DI INDONESIA BERPERAN DALAM MENGURANGI KEMISKINAN? .....194

R.4 PENGARUH MODAL SOSIAL TERHADAP KEMISKINAN RUMAH TANGGA DI INDONESIA TAHUN 2012 (PENERAPAN METODE *TWO STAGE HUBER*) .....210

R.5 PEMODELAN KEJADIAN DEMAM BERDARAH DENGUE MENGGUNAKAN *GEOGRAPHICALLY WEIGHTED REGRESSION* .....222

R.6 PENDUGAAN FAKTOR-FAKTOR YANG MEMPENGARUHI KEMISKINAN DI JAWA TENGAH MENGGUNAKAN ANALISIS REGRESI SPASIAL .....231

**SEM/MULTIVARIATE**

S.1 PENDEKATAN *PARTIAL LEAST SQUARE PATH MODELING* UNTUK DATA NON METRIK (*NON METRIC PLS PM*).....248

S.2 PENGKLASIFIKASIAN WILAYAH BERDASARKAN TIPOLOGI KLASSEN DAN ANALISIS KETIMPANGAN .....258

S.3 ANALISIS SEM DAN PENENTUAN JUMLAH SEGMENT PADA DATA *UNOBSERVED HETEROGENEITY* DENGAN *FINITE MIXTURE PARTIAL LEAST SQUARE*.....269

S.4 *MULTILEVEL STRUCTURAL EQUATION MODELING* (MSEM) PADA PENGGUNAAN FASILITAS KESEHATAN IBU DI PULAU JAWA TAHUN 2012..288

S.5 MULTI DIMENSIONAL SCALING UNTUK MEMETAKAN KABUPATEN/KOTA DI JAWA BARAT BERDASARKAN SEKTOR LAPANGAN USAHA.....306

S.6 EFEK UKURAN SAMPEL DAN *MISSING DATA* DALAM ESTIMASI PARAMETER MODEL SEM DENGAN MAXIMUM LIKELIHOOD, PLS DAN GSCA .....316

**TIME SERIES**

T.1 PENGGUNAAN SOFTWARE R UNTUK MENGESTIMASI PARAMETER *GENERALIZED SPACE-TIME AUTOREGRESSIVE* (GSTAR) DAN MEMBANDINGKAN MATRIKS BOBOT TERBAIK.....338

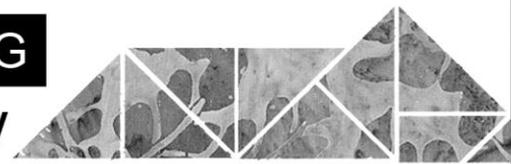
T.2 MENENTUKAN UKURAN OBSERVASI UNTUK KAJIAN PERAMALAN *TIME SERIES* .....352

T.3 PERAMALAN CURAH HUJAN DENGAN *SINGULAR SPECTRUM ANALYSIS WITH MISSING DATA (SSAM)*.....358

T.4 PENGGUNAAN METODE *WEIGHT FUZZY TIME SERIES* PADA DATA INDEKS HARGA SAHAM GABUNGAN.....371

T.5 ANALISIS DERET WAKTU MENGGUNAKAN METODA DEKOMPOSISI MULTIPIKATIF .....388

T.6 ANALISIS DATA EKSPOR *CRUDE PALM OIL (CPO)* DENGAN PENDEKATAN MODEL *SPACE TIME* .....397



T.7 PENDEKATAN MODEL GSTARI DENGAN MELIBATKAN VARIABEL EKSOGEN  
UNTUK MERAMALKAN PESERTA KB AKTIF .....410

T.8 METODE WAVELET UNTUK PERAMALAN NILAI TUKAR RUPIAH .....424

## T.3

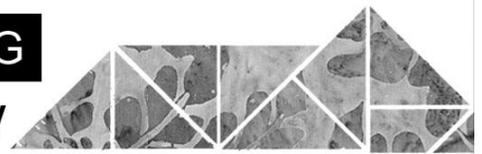
**PERAMALAN CURAH HUJAN DENGAN *SINGULAR SPECTRUM ANALYSIS WITH MISSING DATA (SSAM)*****Sisti Nadia Amalia <sup>1)</sup>, Toni Toharudin <sup>2)</sup>, Gumgum Darmawan <sup>3)</sup>**<sup>1)</sup> Mahasiswa Magister Statistika Terapan FMIPA - UNPAD<sup>2,3)</sup> Staf Pengajar Departemen Statistika FMIPA - UNPADEmail : [neng.sistinadiamalia@gmail.com](mailto:neng.sistinadiamalia@gmail.com); [gumstat@gmail.com](mailto:gumstat@gmail.com)**ABSTRAK**

Pola curah hujan di suatu wilayah memiliki karakteristik yang berbeda-beda. Hal ini dikarenakan adanya perbedaan lintang, gerak semu matahari, letak geografis, topografi serta interaksi berbagai macam sirkulasi udara. Data curah hujan sangat penting untuk perencanaan teknik khususnya untuk bangunan air misalnya irigasi, bendungan, drainase perkotaan, pelabuhan, dermaga, dan lain-lain. Namun sering terjadi data curah hujan di periode tertentu hilang karena terganggunya proses pencatatan di stasiun hujan baik kesalahan alat maupun petugas pencatat. Untuk melakukan peramalan curah hujan cenderung diperlukan data lengkap, sehingga yang sering dilakukan pada data curah hujan yang hilang terlebih dulu di estimasi kemudian dilanjutkan dengan teknik peramalannya. Penelitian ini berbeda dari kebanyakan penelitian yang telah dilakukan sebelumnya, dengan dasar teknik Singular Spectrum Analysis (SSA) yang dimodifikasi dengan data hilang menjadi Singular Spectrum Analysis With Missing Data (SSAM). Sehingga dapat digunakan langsung pada data yang tidak lengkap tanpa melakukan estimasi untuk data yang hilang terlebih dulu.

**Kata Kunci:** *Singular Spectrum Analysis (SSA), Missing Data, Curah Hujan.*

**1. PENDAHULUAN**

Tarogong merupakan salah satu kecamatan di Kabupaten Garut. Kabupaten Garut terletak di Provinsi Jawa Barat bagian Selatan pada koordinat  $6^{\circ}56'49''$  -  $7^{\circ}45'00''$  Lintang Selatan dan  $107^{\circ}25'8''$  -  $108^{\circ}7'30''$  Bujur Timur. Kabupaten Garut memiliki luas wilayah administratif sebesar 306.519 Ha ( $3.065,19 \text{ km}^2$ ). Secara umum iklim wilayah Kabupaten Garut dapat dikategorikan sebagai daerah beriklim tropis basah. Kabupaten Garut yang secara geografis berdekatan dengan Kota Bandung sebagai ibukota provinsi Jawa Barat, merupakan daerah penyangga dan hinterland bagi pengembangan wilayah Bandung Raya. Kabupaten Garut mempunyai kedudukan strategis dalam memasok kebutuhan warga Kota dan Kabupaten Bandung sekaligus pula berperan di dalam mengendalikan keseimbangan lingkungan. Untuk itu penting dilakukan peramalan curah hujan di Kabupaten Garut.



Salah satu unsur iklim yang paling banyak digunakan dalam penelitian adalah curah hujan. Curah hujan merupakan jumlah air hujan yang turun pada suatu daerah dalam kurun waktu tertentu. Satuan yang digunakan dalam mengukur curah hujan adalah millimeter (mm). Curah hujan satu millimeter artinya dalam luasan satu meter persegi pada tempat yang datar tertampung air setinggi satu millimeter atau sebanyak satu liter. Namun data curah hujan yang diamati pada periode tertentu seringkali tidak tersedia, atau dengan kata lain terdapat data missing. Ketidakterersediaan data tersebut menjadi masalah penting untuk proses pengolahan data selanjutnya. Adanya data missing akan mengurangi informasi yang terkandung pada data. Pengumpulan data curah hujan yang tidak lengkap dapat disebabkan oleh beberapa hal, antara lain : perubahan tipe sensor atau instrumen, perubahan penentuan waktu observasi, kelalaian petugas pencatat, dan lain sebagainya.

Secara konvensional masalah data tidak lengkap (terdapat *missing data*) seringkali diselesaikan dengan cara mengestimasi terlebih dahulu data yang hilang. Namun pada penelitian ini ditawarkan teknik peramalan tanpa memerlukan tahap estimasi tersebut. Dengan menggunakan metode *Singular Spectrum Analysis* yang dimodifikasi untuk data data tidak lengkap yaitu *Singular Spectrum Analysis with Missing Data (SSAM)*. Secara ringkas *Singular Spectrum Analysis* merupakan teknik analisis time series yang powerful. Teknik ini termasuk ke dalam pendekatan *non-parametrik*, dimana teknik ini fleksibel terhadap data dan terbebas dari asumsi parametriknya. Sehingga bila digunakan dalam peramalan memungkinkan memiliki ketepatan peramalan yang baik.

## 2. METODOLOGI

### 2.1 *Singular Spectrum Analysis*

Algoritma dasar SSA ada dua tahap: dekomposisi dan rekonstruksi. Tahap awal dari SSA yaitu dekomposisi, dengan melakukan embedding dan singular value decomposition (SVD). Embedding adalah proses pembentukan seri asli ke dalam matriks lintasan; SVD mendekomposisi matriks lintasan dan memecah data menjadi tren, musiman, komponen bulanan, maupun *noise* sesuai dengan *singular values* mereka. Kemudian tahap selanjutnya yaitu rekonstruksi yang melibatkan grouping untuk membuat kelompok dari dekomposisi matriks lintasan dan diagonal averaging untuk merekonstruksi time series baru dari subkelompok.

- **Step 1**

(*Dekomposisi*) Tahap dasar dari dekomposisi pada SSA yaitu embedding. Konsep dari embedding adalah matriks lintasan, yang dibentuk dari data deret waktu awal yang berdimensi satu diubah menjadi deret multidimensi. Pertama, menentukan *window length* ( $L$ ), dan  $K = N - L + 1$ . Kemudian membangun matriks lintasan  $\mathbf{X}$  berdimensi  $L \times K$  dengan  $x_i$  adalah elemen baris dari  $\mathbf{X}$  untuk  $i = 1, \dots, K$ . Sehingga diperoleh matriks lintasan atau trajectory matriks sebagai berikut:

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{bmatrix} y_1 & y_2 & y_3 & \dots & y_K \\ y_2 & y_3 & y_4 & \dots & y_{K+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \dots & y_N \end{bmatrix} \quad (1)$$

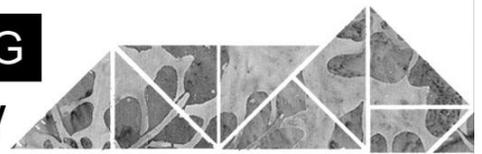
Tahap selanjutnya dalam dekomposisi adalah *Singular Value Decomposition* (SVD). Tahap ini merupakan tahap yang paling penting dalam SSA yaitu mendekomposisi matriks lintasan untuk memperoleh singular value. Menurut Golyandina, N., et.al(2001) dalam proses dekomposisi matriks lintasan untuk data yang panjangnya  $N$  tidak cukup besar dan diasumsikan seri stasioner, maka direkomendasikan untuk mengganti matriks  $S = \mathbf{X}\mathbf{X}^T$  menjadi matriks lain yang memperhitungkan stasioneritas dari deret waktu. Sehingga dalam penelitian ini menggunakan matriks *Toeplitz*, adapun konsep dari matriks *Toeplitz* yaitu menempatkan nilai-nilai yang sama dari  $c_{ij}$  pada diagonal matriksnya. Keutamaannya yaitu menggunakan estimasi standar fungsi kovarians dari seri dan mengubahnya menjadi  $L \times L$  matriks. Berikut ilustrasi dari Matriks *Toeplitz* ( $\mathbf{C}$ ):

$$\mathbf{C} = \begin{bmatrix} c(0) & c(1) & \dots & c(L-1) \\ c(1) & c(0) & \ddots & \vdots \\ \vdots & \vdots & \ddots & c(1) \\ c(L-1) & \dots & \dots & c(0) \end{bmatrix} \quad (2)$$

Adapun perhitungan untuk elemen matriks *Toeplitz*, yaitu

$$c_j = \frac{1}{N-j} \sum_{i=1}^{N-j} x_i x_{i+j}, \quad 0 \leq j \leq L-1 \quad (3)$$

Untuk matriks  $\mathbf{C}$ , dapat dihitung *eigenvalue* ( $\lambda_k$ ) dan *eigenvector* ( $\mathbf{v}_k$ ) ditentukan dan diurutkan secara menurun dari  $\lambda_k$  ( $1 \leq k \leq L$ ). Kemudian menghitung elemen ke- $i$  dari  $k$  *principal component* yang akan terbentuk, yaitu



$$a_{k,i} = \sum_{j=1}^L x_{i+j} v_{j,k}, \quad 1 \leq i \leq N - L \tag{4}$$

• **Step 2**

(*Rekonstruksi*) Tahap awal dari rekonstruksi pada SSA yaitu grouping. Konsep daei grouping yaitu mengelompokkan hasil dekomposisi matriks lintasan, dan menyatukan yang dianggap mirip sehingga membentuk beberapa kelompok. Selanjutnya pada tahap rekonstruksi yaitu diagonal averaging, langkah ini akan mengubah setiap matriks hasil dekomposisi yang telah dikelompokkan dengan perhitungan menggunakan principal component ke-k, sebagai berikut

$$x_i^k = \begin{cases} \frac{1}{i} \sum_{j=1}^i a_{k,i-j+1} v_{j,k}, & 1 \leq i \leq L - 1 \\ \frac{1}{L} \sum_{j=1}^L a_{k,i-j+1} v_{j,k}, & L \leq i \leq N - L + 1 \\ \frac{1}{N-i+1} \sum_{j=i-N+L}^L a_{k,i-j+1} v_{j,k}, & N - L + 2 \leq i \leq N \end{cases} \tag{5}$$

**2.2 Singular Spectrum Analysis with Missing Data**

Modifikasi SSAM terletak pada bagian SVD (*Singular Value Decomposition*). Dalam tahap SVD tetap menggunakan matriks *Toeplitz*. Namun didalam perhitungannya mengabaikan pasangan data yang terdapat data hilang, adapun perhitungannya sebagai berikut :

$$c_j = \frac{1}{N_l} \sum_{l \leq N-j} x_l x_{l+j}, \quad 0 \leq i \leq L - 1 \tag{6}$$

Untuk  $N_l$  pasangan yang tidak terdapat data hilang. Kemudian dari matriks *Toeplitz* akan dihitung *eigenvalue* dan *eigenvector*. *Eigenvalue* maupun *eigenvector* akan tetap berisi tanpa data hilang. Begitu juga *principal component*, berikut perhitungan untuk *principal component* ke -k dengan mengabaikan data hilang, sebagai berikut :

$$a_{k,i} = \frac{L}{N_l} \sum_{l \leq L} x_{i+l} v_{l,k}, \quad 1 \leq i \leq N - L \tag{7}$$

Tahap rekonstruksi dilakukan seperti didalam tahap SSA biasa, namun jika ada nilai *principal component* yang dijumlahkan dan hasilnya tidak ada (*missing*), maka nilai rekonstruksinya juga tidak ada (*missing*).

**2.3 Peramalan dengan SSA Reccurent Forecasting**

Peramalan dengan menggunakan SSA konsepnya adalah dengan bantuan dari *Linear Recurrent Formula* (LRF) untuk membangun model.

$$y_{i+d} = \sum_{k=1}^d a_k y_{i+d-k}, 1 \leq i \leq N - d \quad (8)$$

Untuk menaksir koefisien LRF, yaitu  $a_1, \dots, a_d$  dapat digunakan *eigenvector* yang diperoleh pada tahap SVD.

$$(a_{L-1}, a_{L-2}, \dots, a_2, a_1)^T = \frac{1}{1-v^2} \sum_{i=1}^v \pi_i U_i^v, v^2 = \sum_{i=1}^v \pi_i^2 \quad (9)$$

Dalam peramalan menggunakan SSA *Recurrent Forecasting* ini, deret waktu yang digunakan adalah deret hasil rekonstruksi yang diperoleh dari hasil *diagonal averaging*. Kemudian akan ditentukan  $M$  buah titik data baru yang akan diramalkan. Sehingga akan terbentuk deret hasil peramalan yaitu :

$$Z_{N+M} = (z_1, z_2, \dots, z_{N+M}) \quad (10)$$

Dari perhitungan :

$$z_i \begin{cases} \tilde{y}_i, & i = 0, 1, \dots, N \\ \sum_{j=1}^{L-1} a_j z_{i-j}, & i = N + 1, \dots, N + M \end{cases} \quad (11)$$

#### 2.4 Evaluasi Peramalan

Evaluasi peramalan dilakukan untuk melihat kecocokan metode terhadap data. Hal ini dapat dilihat dari nilai residual yang kecil (selisih antara nilai prediksi dengan nilai aktual). Metode evaluasi peramalan yang digunakan adalah *Mean Absolute Percentage Error (MAPE)*. *MAPE* digunakan untuk mengukur ketepatan hasil prediksi dengan data aktual dalam bentuk rata-rata persentase absolute kesalahan, dengan perhitungan sebagai berikut :

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right| \times 100\% \quad (12)$$

#### 2.5 Validasi Peramalan

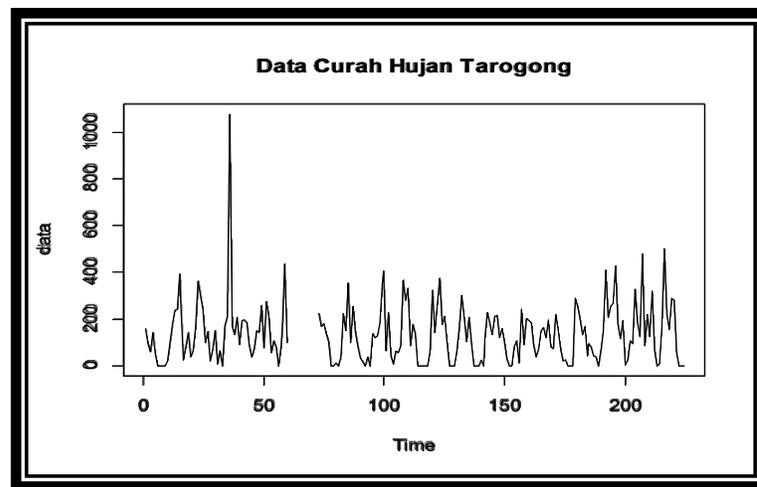
Didalam metode peramalan yang mengasumsikan kesinambungan beberapa pola historis di masa yang akan datang berdasarkan dari masa sebelumnya, *tracking signal* merupakan ukuran toleransi yang dapat digunakan untuk menentukan kemungkinan digunakannya hasil peramalan tersebut yang memperkirakan apabila pola 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 :

$$\text{Tracking Signal} = \frac{\sum_1^n e_n}{\sum_1^n \frac{|e_n|}{n}} \quad (13)$$

### 3. HASIL DAN PEMBAHASAN

Data yang digunakan dalam penelitian ini adalah data curah hujan bulanan di stasiun pengamatan Tarogong dengan periode pengamatan Januari 1997 – April 2015. Jumlah data sebanyak 220. Namun data curah hujan pada periode Januari 2002 – Desember 2002 hilang, sehingga jumlah data yang terisi lengkap sebanyak 208. Berikut plot data curah hujan Tarogong



Gambar 3.1 Plot Data Curah Hujan Tarogong (dalam mm)

Sumber : Dinas Pengelolaan Sumber Daya Air Provinsi Jawa Barat

Syarat untuk dapat melakukan teknik *Singular Spectrum Analysis* adalah data harus berstruktur baik trend, musiman maupun siklik. Gambar 3.1 menunjukkan bahwa curah hujan di stasiun pengamatan Tarogong memiliki pola musiman.

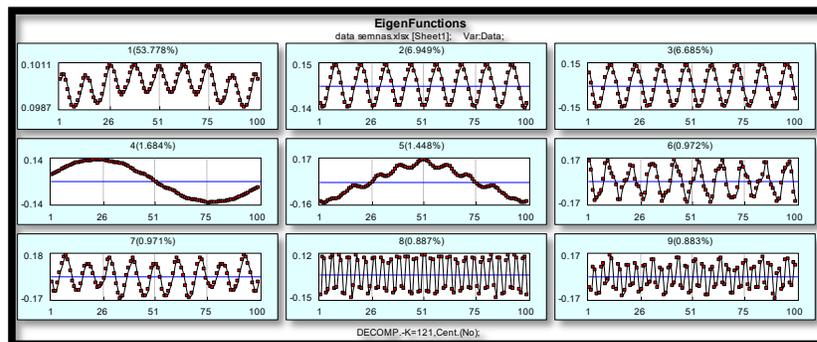
#### 3.1 Algoritma *Singular Spectrum Analysis With Missing Data*

##### • Step 1 (*Dekomposisi*).

Pada langkah awal ini, terlebih dulu yang dilakukan adalah membentuk matriks lintasan  $X$ . Dengan menentukan *window length* ( $L$ ), pemilihan *window length* harus mempertimbangkan komponen yang terkandung dar data sebab ini adalah dasar dari tahapan *singular spectrum analysis*. Hasil dari matriks lintasan  $X$  akan

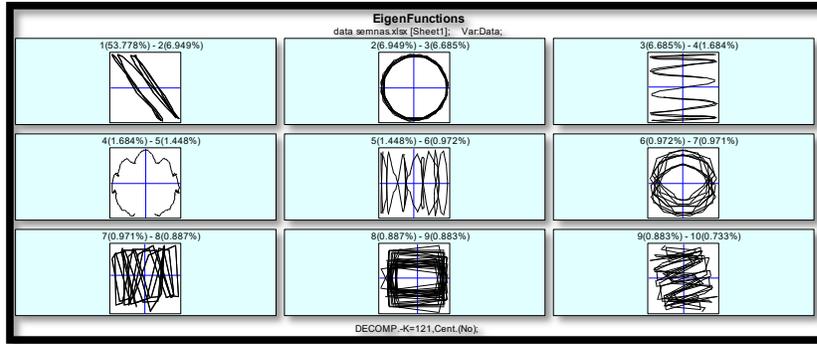
mempengaruhi hasil dari tahapan-tahapan berikutnya. Dalam penelitian itu dipilih  $L = 100$ , sehingga didapatkan nilai  $K = 121$ . Dengan demikian akan diperoleh matriks lintasan  $X$  berukuran  $100 \times 121$ .

Selanjutnya adalah tahap *singular value decomposition* (SVD). SVD adalah tahapan yang paling penting dalam *singular spectrum analysis*, hasil dari tahapan ini akan diperoleh *eigenvalue*, *eigenvector* dan *principal component* yang dibutuhkan dalam proses rekonstruksi pada tahap *grouping*. Untuk memperoleh *eigenvalue*, *eigenvector* dan *principal component* menggunakan matriks *Toeplitz C*. Dengan catatan data pada penelitian ini stasioner dan terdapat data hilang, namun tetap akan terbentuk 100 nilai *eigenvalue* ( $\sqrt{\lambda_i}, U_i, V_i$ ). Kemudian ke 100 nilai *eigenvalue* ini akan mendekomposisikan matriks lintasan menjadi perjumlahan dari beberapa matriks sesuai dengan nilai singularnya yang disebut dengan *singular value decomposition*. Berikut adalah plot 1 dimensi dan plot 2 dimensi *eigenvector* yang berpasangan dari matriks *Toeplitz C*.



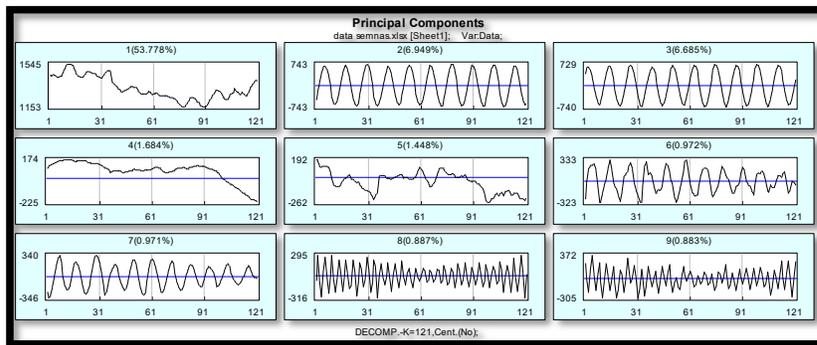
Gambar 3.2 Plot 12 *Eigenvector* Pertama

Berdasarkan Gambar 3.2, dapat diidentifikasi beberapa komponen, misalnya komponen musiman pada *eigenvector* kedua dan ketiga. Untuk komponen musiman, tidak dapat ditentukan hanya melalui plot satu dimensinya saja, namun harus diperiksa melalui plot dua dimensi *eigenvector* yang berpasangan.



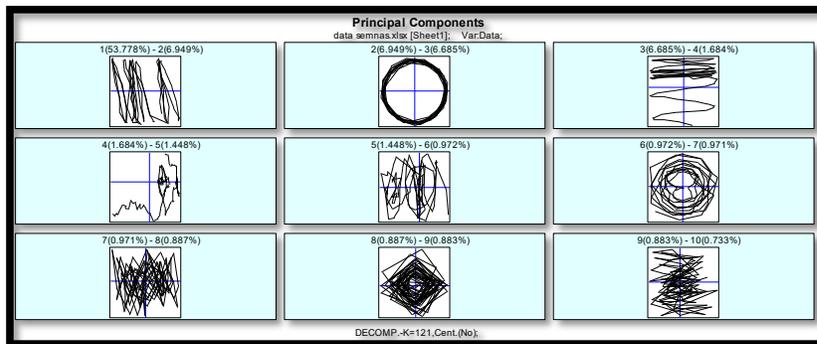
Gambar 3.3 Plot 2 Dimensi 12 Pasangan *Eigenvector* Pertama

Berdasarkan Gambar 3.3, terlihat bahwa pasangan dari *eigenvector* 2-3 membentuk pola musiman. Hal ini sejalan dengan identifikasi pada plot satu dimensi sebelumnya. Dari gambar diatas juga dapat dicari komponen musiman dengan cara mencari plot yang membentuk “lingkaran” dan memeriksa periodenya. Kemudian didapatkan bahwa plot pasangan *eigenvector* 6-7 dan 8-9 memiliki bentuk yang menyerupai lingkaran. Kemudian adalah proses identifikasi plot *principal component*.



Gambar 3.4 Plot 12 *Principal Component* Pertama

Plot *principal component* (PC) ini juga digunakan untuk mengidentifikasi komponen yang terdapat dalam deret waktu. Berdasarkan Gambar 3.4 terlihat bahwa PC 2 dan 3 mengandung pola musiman.

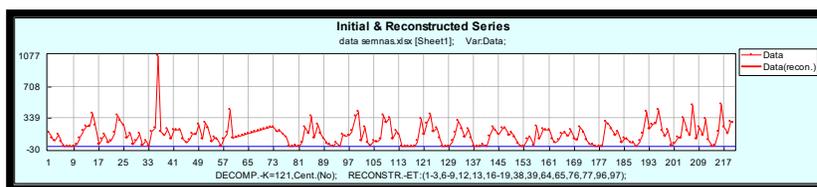


Gambar 3.5 Plot 2 Dimensi 12 Pasangan *Principal Component* Pertama

Seperti plot dua dimensi sebelumnya, plot dua dimensi dari pasangan *principal component* pun menghasilkan identifikasi yang sama. Yaitu komponen musiman yang terdapat pada pasangan *principal component* 2-3, 6-7 dan 8-9. Karena hasil dari identifikasi plot *eigenvector* dan *principal component* adalah sama, maka tidak masalah hasil mana yang akan digunakan sebagai acuan. Namun, pada *principal component* terkandung semua unsur *eigen triple*, sehingga *principal component* digunakan sebagai acuan.

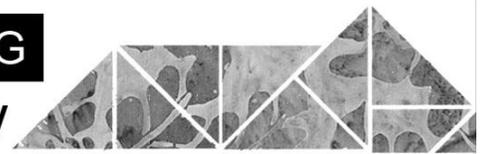
- **Step 2. (Rekonstruksi).**

Tahap awal pada rekonstruksi adalah grouping, berdasarkan plot dari *Principal Component* (PC) dapat disimpulkan bahwa terdapat 2 group ( $m=2$ ) yang terdiri dari komponen musiman dan *noise*. Dalam tahap ini komponen-komponen tersebut akan dikelompokkan sesuai dengan *principal component*nya. setelah melakukan tahap *grouping*, tahap selanjutnya yaitu *diagonal averaging*. Konsep dari *diagonal averaging* adakah merekonstruksi group yang dipilih untuk menentukan deret yang baru. Berikut adalah gambar yang menunjukkan hasil rekonstruksi.



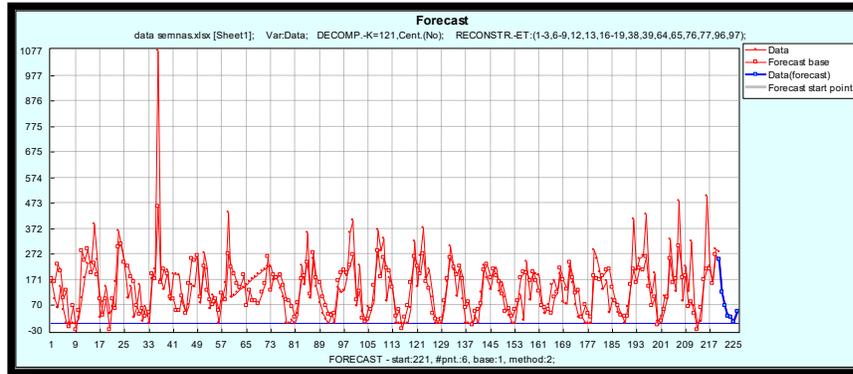
Gambar 3.6 Perbandingan Deret Asli dan Deret Rekonstruksi

Berdasarkan Gambar 3.6 terlihat bahwa hasil rekonstruksi yang melibatkan hampir mendekati data aslinya.



3.2 Peramalan dengan *SSA Recurrent Forecasting*

Berikut merupakan plot hasil ramalan dengan menggunakan *SSA recurrent forecasting* disertai dengan deret asali dan deret rekonstruksinya.



Gambar 3.7 Plot Deret Asli, Deret Rekonstruksi, dan Hasil *Forecasting*

Untuk melihat nilai hasil ramalan, disajikan Tabel 3.1 yang merupakan hasil ramalan untuk tingkat curah hujan di stasiun Pengamatan untuk 6 bulan ke depan dengan menggunakan metode *SSA Recurrent Forecasting* dengan melibatkan komponen musiman. Sebagai berikut :

Tabel 3.1 Hasil Ramalan

Bulan	Hasil Ramalan
Mei	118.69
Juni	67
Juli	25.38
Agustus	20.61
September	2.72
Oktober	42.89

**3.3 Evaluasi Peramalan**

Untuk melihat ketepatan akan dilihat dari nilai MAPE. Berikut Tabel 3.2 disajikan untuk memperlihatkan nilai MAPE.

Tabel 3.2 Ketepatan Peramalan dengan MAPE

Nilai Aktual	Nilai Prediksi	Residual	MAPE
172	168.45	3.55	13,96%
501.5	212.45	289.05	
224	212.85	11.15	
153	153.61	-0.61	
290.5	267.34	23.16	
281.5	251.35	30.15	

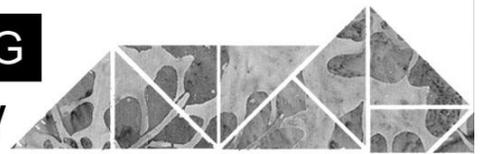
Berdasarkan Tabel 3.2, diperoleh bahwa hasil ramalan memiliki nilai MAPE 13,96%. Sehingga dapat disimpulkan bahwa hasil ramalan dapat digunakan.

**3.4 Validasi Peramalan**

Untuk melakukan validasi peramalan digunakan *tracking signal*. Berikut Tabel 3.3 disajikan untuk menunjukkan *tracking signal* hasil evaluasi data curah hujan di Tarogong 6 bulan terakhir.

Tabel 3.3 Validasi Peramalan dengan *Tracking Signal*

Nilai Aktual	Nilai Prediksi	<i>Tracking Signal</i>
172	168.45	1.00
501.5	212.45	1.98
224	212.85	2.00
153	153.61	2.00



290.5	267.34	2.08
281.5	251.35	2.21

Berdasarkan table 3.3 diketahui bahwa besarnya nilai-nilai *tracking signal* dari 6 periode waktu yang diramalkan berada dalam batas toleransi. menunjukkan bahwa model peramalan bisa digunakan untuk meramalkan  $M$  periode waktu ke depan.

#### 4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dilakukan, maka diperoleh:

Peramalan curah hujan di stasiun pengamatan Tarogong pada periode Januari 1997 – April 2015 dengan *Singular Spectrum Analysis with Missing data* memiliki nilai MAPE sebesar 13,96%.

#### 5. DAFTAR PUSTAKA

- Abraham, B., and Ledolter, J. (1983). *Statistical Methods for Forecasting*, John Wiley and Sons, New York.
- Ardaya, K.N.D., Setiawan., Suhartono. (2013). Pemodelan VAR-NN dan GSTAR-NN untuk Peramalan Curah Hujan di Kabupaten Malang. *Jurnal Sains Dan Seni Pomits*. Volume 2, No.1, 2337-3520.
- Golyandina, N., Netkrutkin, V., and Zhigljavsky, A. A. (2001). *Analysis of Time Series Structure : SSA and related techniques*, Chapman & Hall/ CRC, New York – London.
- Schoellhamer, D. H. (2001). Singular Spectrum Analysis for Time Series with Missing Data. *Geophys. Res. Lett*, Volume 28, 3187-3190.
- Shen, Y., Peng, F., and Li, B. (2015). Improved Singular Spectrum Analysis for Time Series with Missing Data. *Nonlin. Processes Geophys*, Volume 22, 371-376.