

Peramalan jumlah kedatangan wisatawan mancanegara ke bali menggunakan metode hibrida SSA-WFTS

Nadia Uli Clarissa^{a,1}, Winita Sulandari^{a,2,*}, Respatiwan^{a,3}

^a Program Studi Statistika, Universitas Sebelas Maret, Jl. Ir. Sutami No.36, Ketingan, Kec. Jebres, Kota Surakarta, Jawa Tengah, 57126, Indonesia

*² winita@mipa.uns.ac.id

*Correspondent Author

KATAKUNCI

Singular Spectrum Analysis
Weighted Fuzzy Time Series
Peramalan

ABSTRAK

Salah satu provinsi di Indonesia dengan jumlah kunjungan wisatawan mancanegara terbanyak yaitu provinsi Bali. Perlunya peramalan kunjungan wisatawan mancanegara ke Bali yang dapat dijadikan acuan oleh Pemerintah untuk menetapkan strategi dalam memperbaiki kualitas pariwisata di Bali. Metode yang digunakan untuk meramalkan yaitu metode hibrida Singular Spectrum Analysis (SSA) – Weighted Fuzzy Time Series (WFTS). Pemodelan SSA dilakukan untuk menganalisis komponen linear, lalu nilai residu dari model SSA dimodelkan dengan WFTS. Peramalan dilakukan dengan 4 metode yaitu, SSA dengan R-forecasting, SSA-WFTS dengan metode Chen, Yu, Cheng ($\alpha = 0,9$), dan Lee ($c = 1,1$). Keempat metode ini akan dibandingkan untuk memperoleh model terbaik. Hasil peramalan diperoleh nilai MAPE sebesar 14,515% untuk model SSA R-forecasting, 9,029% untuk model SSA-WFTS metode Chen, 9,067% untuk model SSA-WFTS metode Yu, 9,125% untuk model SSA-WFTS metode Cheng ($\alpha = 0,9$), dan 9,028% untuk model SSA-WFTS metode Lee ($c = 1,1$). Model terbaik diperoleh dengan pemodelan hibrida SSA-WFTS metode Chen dengan nilai MAPE terkecil dibanding model lainnya.

Forecasting the number of international tourist arrivals to Bali using the SSA-WFTS hybrid method

KEYWORDS

Singular Spectrum Analysis
Weighted Fuzzy Time Series
Forecasting

Bali is one of a province in Indonesia with the most number of foreign tourist arrivals. The result of forecasting the number of tourist arrivals in Bali can be used to improve the quality of tourism in Bali by the government. In this research, hybrid Singular Spectrum Analysis (SSA) – Weighted Fuzzy Time Series (WFTS) method is used for forecasting. The linear component will be modeled by SSA and the residual from the SSA model will be forecasted with WFTS. There are 4 methods for analysis and will be compared to obtain the best model. There are, SSA R-forecasting, SSA-WFTS Chen, SSA-WFTS Yu, SSA-WFTS Cheng ($\alpha = 0.9$), and SSA-WFTS Lee ($c = 1.1$). The MAPE value obtained from the results of forecasting using the SSA R-forecasting is 14.515%, using the SSA-WFTS Chen is 9.029%, using the SSA-WFTS Yu is 9.067%, using the SSA-WFTS Cheng ($\alpha = 0.9$) is 9.125% , and using the SSA-WFTS Lee ($c = 1.1$) is 9.028%. It shows that forecasting the number of foreign arrivals in Bali using the SSA-WFTS Chen method provides the best forecasting results.

This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Pendahuluan

Sektor pariwisata memiliki peran penting terhadap devisa negara, pendapatan daerah, pengembangan wilayah, dan penciptaan lapangan tenaga kerja [1]. Apabila berbicara tentang pariwisata, maka tidak akan jauh kaitannya dengan istilah wisatawan. Wisatawan berdasarkan tempat asalnya dibagi menjadi dua, yaitu wisatawan nusantara dan wisatawan mancanegara (Wisman). Wisatawan mancanegara adalah setiap orang yang melakukan perjalanan ke suatu negara (di luar negara tempat tinggalnya), kurang dari satu tahun, didorong oleh suatu tujuan utama (bisnis, berlibur, atau tujuan pribadi lainnya), selain untuk bekerja pada penduduk negara yang dikunjungi [2]. Indonesia menduduki peringkat keempat dalam kunjungan wisatawan mancanegara ke negara-negara ASEAN dengan jumlah kedatangan yang meningkat tiap tahunnya [3]. Tahun 2019, daya saing pariwisata Indonesia berada pada peringkat 40 dari 140 negara di dunia dengan hasil bahwa Indonesia masih relatif rendah dalam segi lingkungan berkelanjutan, kesehatan dan kebersihan, dan infrastruktur jasa pariwisata [4]. Salah satu provinsi di Indonesia yang menjadi destinasi wisatawan mancanegara adalah Provinsi Bali. Menurut BPS Provinsi Bali pada tahun 2018, 38,8% wisatawan mancanegara yang mengunjungi Indonesia juga mengunjungi Bali. Hingga tahun 2018, jumlah kunjungan wisatawan mancanegara ke Bali terus meningkat tiap tahunnya [5]. Perlunya peramalan kunjungan wisatawan mancanegara ke Bali yang dapat dijadikan acuan oleh Pemerintah untuk menetapkan strategi dalam memperbaiki kualitas Pariwisata di Bali khususnya dalam segi lingkungan berkelanjutan, kesehatan dan kebersihan, dan infrastruktur jasa pariwisata, seiring dengan bertambahnya jumlah kunjungan wisatawan mancanegara ke Bali tiap tahunnya.

Singular Spectrum Analysis (SSA) merupakan salah satu teknik analisis runtun waktu dan peramalan. *SSA* bertujuan mendekomposisi deret awal menjadi sejumlah dari bilangan kecil yang dapat diinterpretasikan seperti *trend* yang bervariasi lambat, komponen *oscillatory* dan *noise* [6]. Beberapa hasil penelitian menunjukkan *SSA* memberikan hasil yang baik dalam meramalkan data yang memiliki komponen musiman dan tren [7][8][9]. *SSA* memiliki kekurangan yaitu tidak dapat menangani komponen nonlinear stokastik pada data, sehingga diperlukan suatu cara untuk menangani *noise* yang mana menunjukkan hubungan nonlinear. Zhang (2001) memperkenalkan metode hibrida model *Autoregressive Integrated Moving Average (ARIMA)* dan model *Artificial Neural Network (ANN)* [10]. *ARIMA* memodelkan hubungan linear pada data dan hubungan nonlinear dimodelkan dengan model *ANN*. Metode hibrida *ARIMA-ANN* terdiri dari 2 langkah, yaitu pertama dilakukan pemodelan *ARIMA* untuk menganalisis komponen linear, lalu nilai residu dari model *ARIMA* akan dimodelkan dengan *ANN*. Metode hibrida ini meningkatkan performa peramalan pada data yang memiliki struktur korelasi linear dan nonlinear. Penanganan komponen nonlinear juga dapat dilakukan dengan model nonlinear lainnya yaitu *Weighted Fuzzy Time Series (WFTS)*.

Weighted Fuzzy Time Series (WFTS) merupakan pengembangan dari *Fuzzy Time Series (FTS)*. *Fuzzy Time Series* pertama kali diperkenalkan oleh Song dan Chissom (1993) [11], lalu dikembangkan oleh Chen (1996) dengan cara yang lebih sederhana [12], namun metode Chen tidak memperhatikan pengulangan relasi *fuzzy* dan menganggap semua relasi *fuzzy* memiliki bobot yang sama, yang mana nilai ramalan yang dihasilkan tidak memuaskan. Yu (2004) mengembangkan metode Chen untuk mengatasi dua permasalahan dalam peramalan *Fuzzy Time Series*, yaitu pengulangan dan pembobotan dalam relasi *fuzzy* [13]. Selanjutnya, Cheng (2008) mengembangkan model adaptif dengan nilai ramalan saat t diperoleh dari nilai observasi pada saat $t - 1$ dijumlahkan dengan bobot kesalahan peramalan terakhir. Lalu, Lee dan Suhartono (2010), mengembangkan metode Yu dengan mengatur bobot dari relasi *fuzzy* sebagai fungsi *exponential growth* [14].

Beberapa penelitian menggabungkan metode lain dengan metode *WFTS*, diantaranya dengan metode *SSA-FLR* [15], [16], dan dengan metode *Winter* [17]. Metode hibrida *SSA-FLR-WFTS* memberikan hasil yang lebih baik dibandingkan metode *SSA-FLR* [15]. Selain itu, peramalan jumlah kedatangan wisatawan mancanegara ke Bali menggunakan metode hibrida

Winter-WFTS memberikan hasil peramalan yang lebih baik dibandingkan metode dekomposisi, Winter, dan ARIMA [17].

Berdasarkan penjelasan di atas, penelitian ini bertujuan memperoleh model terbaik dan mendapatkan hasil ramalan jumlah kedatangan wisatawan mancanegara ke Bali dengan menggunakan metode hibrida SSA-WFTS.

Metode

Data yang digunakan dalam penelitian ini adalah jumlah kedatangan wisatawan mancanegara ke Bali tahun 2009-2019. Data merupakan data sekunder yang diperoleh dari laman resmi Badan Pusat Statistik (BPS) Provinsi Bali yaitu <https://bali.bps.go.id/> dan merupakan data bulanan. Data akan dibagi menjadi dua, yaitu data *in-sample* dan data *out-sample*. Data *in-sample* yaitu data pada tahun 2009-2018 sedangkan data *out-sample* yaitu data pada tahun 2019. Analisis dilakukan menggunakan metode hibrida SSA-WFTS. Pada penelitian ini, metode hibrida SSA-WFTS dibagi menjadi 2 langkah, langkah pertama yaitu SSA dengan *R-forecasting* untuk memodelkan komponen deterministik dan langkah kedua yaitu menggabungkan SSA dengan WFTS, yang mana algoritma WFTS dijalankan untuk memodelkan komponen nonlinear stokastik dari nilai residu model SSA. Adapun langkah-langkah hibrida SSA-WFTS adalah sebagai berikut :

1) Pemodelan dan peramalan menggunakan metode SSA dengan *R-forecasting* [18]:

a) *Decomposition*

i) *Embedding*

Tahap *embedding* mengubah deret waktu menjadi matriks lintasan (*Trajectory Matrix*) dengan mentransformasi data yang berdimensi satu (vektor) menjadi data yang multidimensional (matriks). Didefinisikan L merupakan panjang jendela (*windows length*) dengan $2 < L < N/2$ dan $K = N - L + 1$, maka akan membentuk matriks sebagai berikut :

$$\mathbf{X} = (x_{ij})_{i,j=1}^{L,K} = \begin{bmatrix} f_0 & f_1 & \cdots & f_{K-1} \\ f_1 & f_2 & \cdots & f_K \\ \vdots & \vdots & \ddots & \vdots \\ f_{L-1} & f_L & \cdots & f_{N-1} \end{bmatrix} \quad (1)$$

Matriks \mathbf{X} disebut matriks *Hankel*, yang mana semua elemen anti-diagonalnya bernilai sama.

ii) *Singular Value Decomposition (SVD)*

Pada tahap ini, dilakukan *singular value decomposition* pada matriks (1). Definisikan $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ dan *eigenvalues* dari \mathbf{S} adalah $\lambda_1, \dots, \lambda_L$. *Eigenvectors* dari tiap *eigenvalues* yaitu U_1, \dots, U_L . Selanjutnya, $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$, maka SVD dari matriks trajektori \mathbf{X} dapat ditulis dalam persamaan (2).

$$\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d, \quad (2)$$

dengan $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$ untuk $i = 1, \dots, d$, $d = \max\{i: \lambda_i > 0\}$. Kumpulan $(\sqrt{\lambda_i}, U_i, V_i^T)$ dapat disebut dengan *eigentruple* ke- i dari SVD dan \mathbf{X}_i disebut *principal component (PC)* [19].

b) *Reconstruction*

i) *Grouping*

Pada tahap *grouping*, matriks dikelompokkan ke dalam subkelompok berdasarkan pola pembentuk data deret waktu yakni *trend*, *seasonal*, dan *noise*. Tahap *grouping* pada langkah rekonstruksi merupakan partisi himpunan dari indeks $\{1, \dots, d\}$ menjadi m disjoint subset I_1, \dots, I_m . Jika $\mathbf{I} = \{i_1, \dots, i_p\}$, maka *resultant matrix* \mathbf{X}_I yang dapat didefinisikan sebagai $\mathbf{X}_I = \mathbf{X}_{i_1} + \mathbf{X}_{i_2} \dots + \mathbf{X}_{i_p}$, berkorespondensi dengan $\mathbf{I} = \{I_1, \dots, I_m\}$. Maka, $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_d$, dapat ditulis menjadi persamaan (3).

$$\mathbf{X} = \mathbf{X}_{I_1} + \mathbf{X}_{I_2} \dots + \mathbf{X}_{I_m} \tag{3}$$

Prosedur untuk memilih set I_1, \dots, I_m disebut *eigentriple grouping*.

ii) *Diagonal Averaging*

Pada tahap ini, tiap matriks dari tahap *grouping* menjadi deret baru dengan panjang N . Misal \mathbf{Y} adalah matriks berukuran $L \times K$ dengan elemen y_{ij} . Dengan $1 \leq i \leq L$ dan $1 \leq j \leq K$. Tentukan $L^* = \min(L, K)$, $K^* = \max(L, K)$, $N = L + K - 1$, dan $y_{ij}^* = y_{ij}$ jika $L > K$ dan $y_{ij}^* = y_{ij}$ lainnya. Matriks \mathbf{Y} ditransformasikan ke dalam bentuk deret kembali kembali, f_0, \dots, f_{N-1} menggunakan metode *diagonal averaging* dengan menggunakan persamaan berikut :

$$\tilde{f}_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} y_{m, k-m+2}^* & \text{untuk } 1 \leq k < L^* - 1, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m, k-m+2}^* & \text{untuk } L^* - 1 \leq k < K^*, \\ \frac{1}{N-k} \sum_{m=k-K^*+2}^{N-K^*+1} y_{m, k-m+2}^* & \text{untuk } K^* \leq k < N. \end{cases} \tag{4}$$

Diagonal averaging (4) diterapkan pada *resultant matrix* \mathbf{X}_{I_k} pada persamaan (3), sehingga \mathbf{X}_{I_k} akan membentuk deret rekonstruksi $\tilde{F}^{(k)} = (\tilde{f}_0^{(k)}, \dots, \tilde{f}_{N-1}^{(k)})$. Oleh karena itu, deret asli f_0, f_1, \dots, f_{N-1} akan menjadi jumlah dari m deret sebagai berikut:

$$f_n = \sum_{k=1}^m \tilde{f}_n^{(k)}, (n = 0, 1, \dots, N - 1)$$

2) Peramalan menggunakan algoritma *R-forecasting*

a) Deret waktu $F_{N+M} = (f_0, \dots, f_{N+M-1})$ didefinisikan dengan,

$$\hat{f}_i = \begin{cases} \tilde{f}_i & \text{untuk } i = 0, \dots, N - 1, \\ \sum_{j=1}^{L-1} a_j f_{i-j} & \text{untuk } i = N, \dots, N + M - 1, \end{cases}$$

dengan vektor dari koefisien $\mathcal{R} = (a_{L-1}, \dots, a_1)^T$ yang dapat diekspresikan sebagai,

$$\mathcal{R} = \frac{1}{1-v^2} \sum_{i=1}^r \pi_i U_i^\nabla.$$

Definisikan U_i^∇ adalah vektor komponen $L-1$ pertama dari *eigenvector* U_i dan $v^2 = \pi_1^2 + \dots + \pi_r^2$. Dengan π_i adalah komponen terakhir dari vektor U_i , untuk $i = 1, \dots, r$.

b) Angka-angka $\hat{f}_N, \dots, \hat{f}_{N+M-1}$ membentuk M langkah kedepan dari peramalan *SSA*.

3) Meramalkan nilai residu dari model *SSA* dengan *WFTS*

a) Metode Chen

Algoritma metode Chen adalah sebagai berikut [12]:

i) Mendefinisikan semesta pembicaraan (*universe of discourse*) dari data historis,

$$U = [D_{\min} - D1, D_{\max} - D2],$$

dengan D_{\min} dan D_{\max} adalah data minimum dan maksimum dan $D1$ dan $D2$ adalah bilangan positif.

ii) Membagi semesta pembicaraan ke dalam interval yang sama, $U = \{u_1, u_2, \dots, u_m\}$.

- iii) Menentukan himpunan *fuzzy* pada U . Melakukan fuzzifikasi pada data historis yang diamati. Misal A_1, A_2, \dots, A_i adalah himpunan *fuzzy* yang mempunyai nilai linguistik dari suatu variabel linguistik. Pendefinisian himpunan *fuzzy* A_1, A_2, \dots, A_i pada himpunan semesta U sebagai berikut:

$$A_1 = \frac{a_{11}}{u_1} + \frac{a_{12}}{u_2} + \dots + \frac{a_{1m}}{u_m},$$

$$A_2 = \frac{a_{21}}{u_1} + \frac{a_{22}}{u_2} + \dots + \frac{a_{2m}}{u_m},$$

$$A_k = \frac{a_{k1}}{u_1} + \frac{a_{k2}}{u_2} + \dots + \frac{a_{km}}{u_m},$$

dengan nilai a_{ij} mempunyai *range* $[0,1]$, $1 \leq i \leq k$ dan $1 \leq j \leq m$. Nilai dari a_{ij} menandakan derajat keanggotaan dari u_j dalam himpunan *fuzzy* A_i . Apabila suatu data mempunyai nilai derajat keanggotaan maksimum pada himpunan *fuzzy* A_k , maka data tersebut masuk ke dalam himpunan *fuzzy* A_k atau variabel linguistik A_k .

- iv) Menentukan *Fuzzy Logical Relationship (FLR)*. Dua himpunan *fuzzy* berturut-turut $A_i(t-p)$ dan $A_j(t)$ dapat dibentuk menjadi *FLR* ($A_i \rightarrow A_j$). A_i disebut sebagai *Left Hand Sides (LHS)* dan A_j disebut dengan *Right Hand Sides (RHS)*.
- v) Menentukan *Fuzzy Logical Relationship Group (FLRG)* untuk semua *FLR* berdasarkan *LHS* yang sama. Contoh, jika *FLR* berbentuk $A_1 \rightarrow A_2, A_1 \rightarrow A_1, A_1 \rightarrow A_2, A_1 \rightarrow A_3$, maka *FLRG* yang terbentuk ialah $A_1 \rightarrow A_1, A_2, A_3$.
- vi) Melakukan peramalan. Jika $F(t-1) = A_1$, maka terdapat beberapa aturan sebagai berikut yang meliputi :
- (1) Jika tidak terdapat *FLR* dari A_i , ($A_1 \rightarrow \#$), maka $F(t) = A_i$.
 - (2) Jika hanya terdapat satu *FLR*, ($A_1 \rightarrow A_j$), maka $F(t) = A_j$.
 - (3) Jika terdapat banyak *FLR* ($A_{j1}, A_{j2}, \dots, A_{jk}$), maka $F(t) = A_{j1}, A_{j2}, \dots, A_{jk}$.
- vii) Defuzzifikasi. Misalkan $F(t) = A_{j1}, A_{j2}, \dots, A_{jk}$, maka $\hat{y} = \frac{\sum_{p=1}^k m_{jp}}{k}$, dengan $\hat{y}(t)$ merupakan defuzzifikasi dan m_{jp} adalah nilai tengah dari A_{jp} .

b) Metode Yu

Algoritma Yu adalah sebagai berikut [13]:

- i) Mendefinisikan semesta pembicaraan (*universe of discourse*) dari data historis.
- ii) Membagi semesta pembicaraan ke dalam interval yang sama, $U = \{u_1, u_2, \dots, u_m\}$.
- iii) Menentukan himpunan *fuzzy* pada U dan melakukan fuzzifikasi pada data historis yang diamati.
- iv) Menentukan *Fuzzy Logical Relationship (FLR)*.
- v) Menentukan *Fuzzy Logical Relationship Group (FLRG)* untuk semua *FLR* berdasarkan *LHS* yang sama sesuai urutan waktu. Contoh, jika *FLR* berbentuk $A_1 \rightarrow A_2, A_1 \rightarrow A_1, A_1 \rightarrow A_2, A_1 \rightarrow A_3$, maka *FLRG* yang terbentuk ialah $A_1 \rightarrow A_2, A_1, A_2, A_3$.
- vi) Melakukan peramalan sama dengan metode Chen.
- vii) Defuzzifikasi. Misalkan $F(t) = A_{j1}, A_{j2}, \dots, A_{jk}$, maka matriks yang didefuzzifikasi yaitu matriks nilai titik tengah dari $A_{j1}, A_{j2}, \dots, A_{jk}$ yang dapat ditulis dengan,

$$M(t) = [m_{j1}, m_{j2}, \dots, m_{jk}],$$

dengan $M(t)$ merupakan nilai defuzzifikasi peramalan dari $F(t)$.

- viii) Menghitung bobot. Pembobot untuk $F(t) = A_{j1}, A_{j2}, \dots, A_{jk}$ ialah dengan w'_1, w'_2, \dots, w'_k dengan $w'_i = \frac{w_i}{\sum_{h=1}^k w_h}$, dengan $w_1 = 1, w_i = w_{i-1} + 1$ untuk $2 \leq i \leq k$, dan w_h adalah bobot untuk A_{jh} . Sehingga matriks pembobot dapat ditulis sebagai,

$$W(t) = [w'_1, w'_2, \dots, w'_k] = \left[\frac{1}{\sum_{h=1}^k w_h}, \frac{2}{\sum_{h=1}^k w_h}, \dots, \frac{k}{\sum_{h=1}^k w_h} \right].$$

- ix) Menghitung nilai ramalan akhir. Nilai ramalan akhir diperoleh dari perkalian matriks yang didefuzzifikasi dan tranpos dari matriks pembobot.

$$\hat{F}(t) = M(t) \times w(t)^T$$

$$\hat{F}(t) = [m_{j_1}, m_{j_2}, \dots, m_{j_k}] \times \left[\frac{1}{\sum_{h=1}^k w_h}, \frac{2}{\sum_{h=1}^k w_h}, \dots, \frac{k}{\sum_{h=1}^k w_h} \right]^T,$$

dengan \times merupakan operator perkalian matriks.

c) Metode Cheng

Algoritma metode Cheng adalah sebagai berikut [20]:

- i) Mendefinisikan semesta pembicaraan (*universe of discourse*) dari data historis.
- ii) Membagi semesta pembicaraan ke dalam interval yang sama, $U = \{u_1, u_2, \dots, u_m\}$.
- iii) Melakukan fuzzifikasi.
- iv) Menentukan *Fuzzy Logical Relationship (FLR)*.
- v) Menentukan *Fuzzy Logical Relationship Group (FLRG)* untuk semua *FLR* berdasarkan *LHS* yang sama sesuai urutan waktu. Contoh, apabila *FLR* berbentuk $A_1 \rightarrow A_2, A_1 \rightarrow A_1, A_1 \rightarrow A_2, A_1 \rightarrow A_3$, maka *FLRG* yang terbentuk ialah $A_1 \rightarrow A_2, A_1, A_2, A_3$ dengan matriks pembobot $w(t) = [w_1, \dots, w_4] = [1, 1, 2, 1]$.
- vi) Menghitung nilai ramalan dengan $F(t) = L_{df}(t-1) \cdot W_n(t-1)$, dengan $L_{df}(t-1)$ adalah matriks defuzzifikasi dan $W_n(t-1)$ adalah matriks pembobot. Sebagai contoh, jika $F(t)$ adalah A_2, A_3, A_1, A_2 , maka nilai ramalan ialah,

$$F(t) = [m_2, m_3, m_1, m_2] \times \left[\frac{1}{1+1+2+1}, \frac{1}{5}, \frac{1}{5}, \frac{2}{5} \right]^T.$$

- vii) Menghitung nilai ramalan akhir dengan $\hat{y}(t) = y(t-1) + (\alpha \times [F(t) - y(t-1)])$, dengan $y(t-1)$ adalah nilai pengamatan pada saat $t-1$ dan α adalah parameter pembobot.

d) Metode Lee

Algoritma metode Lee adalah sebagai berikut [14]:

- i) Mendefinisikan semesta pembicaraan (*universe of discourse*) dari data historis.
- ii) Membagi semesta pembicaraan ke dalam interval yang sama, $U = \{u_1, u_2, \dots, u_m\}$.
- iii) Melakukan fuzzifikasi.
- iv) Menentukan *Fuzzy Logical Relationship (FLR)*
- v) Menentukan *Fuzzy Logical Relationship Group (FLRG)* untuk semua *FLR* berdasarkan *LHS* yang sama. Contoh, jika *FLR* berbentuk $A_1 \rightarrow A_2, A_1 \rightarrow A_1, A_1 \rightarrow A_2, A_1 \rightarrow A_3$, maka *FLRG* yang terbentuk ialah $A_1 \rightarrow A_2, A_1, A_2, A_3$.
- vi) Melakukan peramalan sama dengan metode Chen.
- vii) Defuzzifikasi. Misalkan $F(t) = A_{j_1}, A_{j_2}, \dots, A_{j_k}$, maka matriks yang didefuzzifikasi yaitu matriks nilai titik tengah dari $A_{j_1}, A_{j_2}, \dots, A_{j_k}$ yang dapat ditulis dengan,

$$M(t) = [m_{j_1}, m_{j_2}, \dots, m_{j_k}],$$

dengan $M(t)$ merupakan nilai defuzzifikasi peramalan dari $F(t)$.

- viii) Menghitung bobot. Pembobot untuk $F(t) = A_{j_1}, A_{j_2}, \dots, A_{j_k}$ ialah dengan w'_1, w'_2, \dots, w'_k dengan $w'_i = \frac{w_i}{\sum_{h=1}^k w_h}$, $w_1 = 1$, $w_i = c^{i-1}$ untuk $2 \leq i \leq k$, $c \geq 1$ dan w_h adalah bobot untuk A_{j_h} . Sehingga matriks pembobot dapat ditulis sebagai,

$$W(t) = [w'_1, w'_2, \dots, w'_k] = \left[\frac{1}{\sum_{h=1}^k w_h}, \frac{c}{\sum_{h=1}^k w_h}, \dots, \frac{c^{k-1}}{\sum_{h=1}^k w_h} \right].$$

- ix) Menghitung nilai ramalan akhir. Nilai ramalan akhir diperoleh dari perkalian matriks yang didefuzzifikasi dan tranpos dari matriks pembobot.

$$\hat{F}(t) = M(t) \times w(t)^T$$

$$\hat{F}(t) = [m_{j1}, m_{j2}, \dots, m_{jk}] \times \left[\frac{1}{\sum_{h=1}^k w_h}, \frac{c}{\sum_{h=1}^k w_h}, \dots, \frac{c^{k-1}}{\sum_{h=1}^k w_h} \right]^T,$$

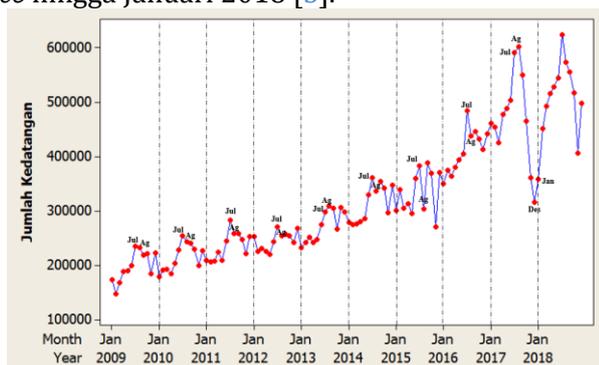
dengan \times merupakan operator perkalian matriks.

- 4) Menghitung nilai ramalan akhir dengan menjumlahkan nilai ramalan SSA dengan *R-forecasting* dan nilai ramalan residu dari peramalan menggunakan *WFTS*.
- 5) Menghitung *MAPE* dan menentukan model terbaik.

Hasil dan pembahasan

Identifikasi Data

Plot data jumlah kedatangan wisatawan mancanegara ke Bali pada tahun 2009-2018 ditunjukkan pada [Gambar 1](#). Plot menunjukkan bahwa data memiliki pola tren dan musiman. Pola kunjungan menunjukkan pola yang hampir sama tiap tahunnya. Pada awal tahun, jumlah kedatangan cenderung rendah, lalu puncak kedatangan tertinggi terjadi pada bulan-bulan triwulan ketiga, seperti bulan Juli dan Agustus. Bulan selanjutnya mengalami penurunan, lalu meningkat kembali di akhir tahun dikarenakan libur akhir tahun. Pada bulan Agustus sampai Desember 2017, jumlah kedatangan wisman mengalami penurunan terus menerus dikarenakan bencana erupsi Gunung Agung yang dimulai sejak bulan Agustus 2017. Jumlah kedatangan wisman yang berangsur-angsur mengalami penurunan hingga mencapai titik terendah di Bulan Desember, cukup berimbas ke bulan Januari 2018 terkait dengan pemberlakuan *travel advice* hingga Januari 2018 [5].



Gambar 1. Plot Jumlah Kedatangan Wisman ke Bali

Peramalan dengan Metode *SSA-WFTS*

1) Pemodelan SSA

a) *Decomposition*

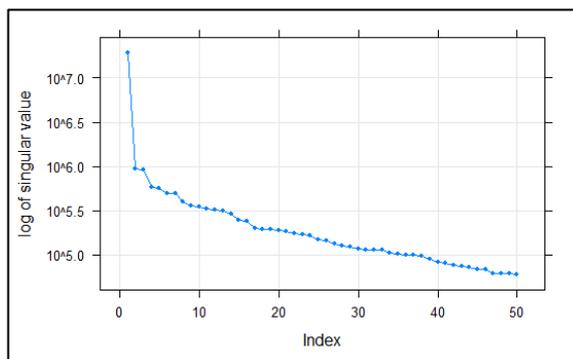
i) *Embedding*

Pada penelitian ini, ditentukan nilai *windows length* sebesar 59 ($L = 59$), maka $K = 120 - 59 + 1 = 60$. Matriks lintasan $L \times K$ yang terbentuk sebagai berikut:

$$\mathbf{X} = (x_{ij})_{i,j=1}^{59,60} = \begin{bmatrix} 174541 & 147704 & \dots & 299013 \\ 147704 & 168205 & \dots & 279257 \\ \vdots & \vdots & \ddots & \vdots \\ 307276 & 299013 & \dots & 498819 \end{bmatrix}$$

ii) *Singular Value Decomposition (SVD)*

Plot yang terbentuk dari 50 *singular values* jumlah kedatangan wisatawan mancanegara ke Bali dapat dilihat pada [Gambar 2](#).



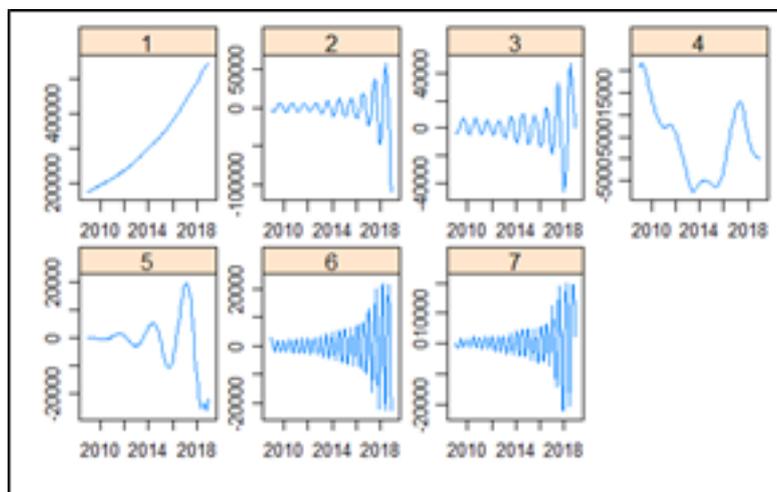
Gambar 2. Plot dari 50 Singular Values Jumlah Kedatangan Wisatawan Mancanegara ke Bali

Effect grouping (r) digunakan untuk membantasi nilai *eigen triples* yang akan digunakan untuk mengidentifikasi komponen tren dan musiman. Nilai r ditentukan berdasarkan jumlah *eigen triples* yang tidak menunjukkan *noise* pada plot *singular value*. Rentetan *singular values* yang menurun secara perlahan pada plot *singular value* menunjukkan komponen *noise* [21]. Gambar 2 menunjukkan bahwa pergerakan antar *singular values* mulai melambat setelah *eigen triple* ke-7, maka *eigen triple* ke-1 sampai ke-7 merupakan komponen tren dan musiman. *Eigen triple* ke-8 sampai ke-50 akan digabung lalu dikategorikan sebagai komponen *noise*.

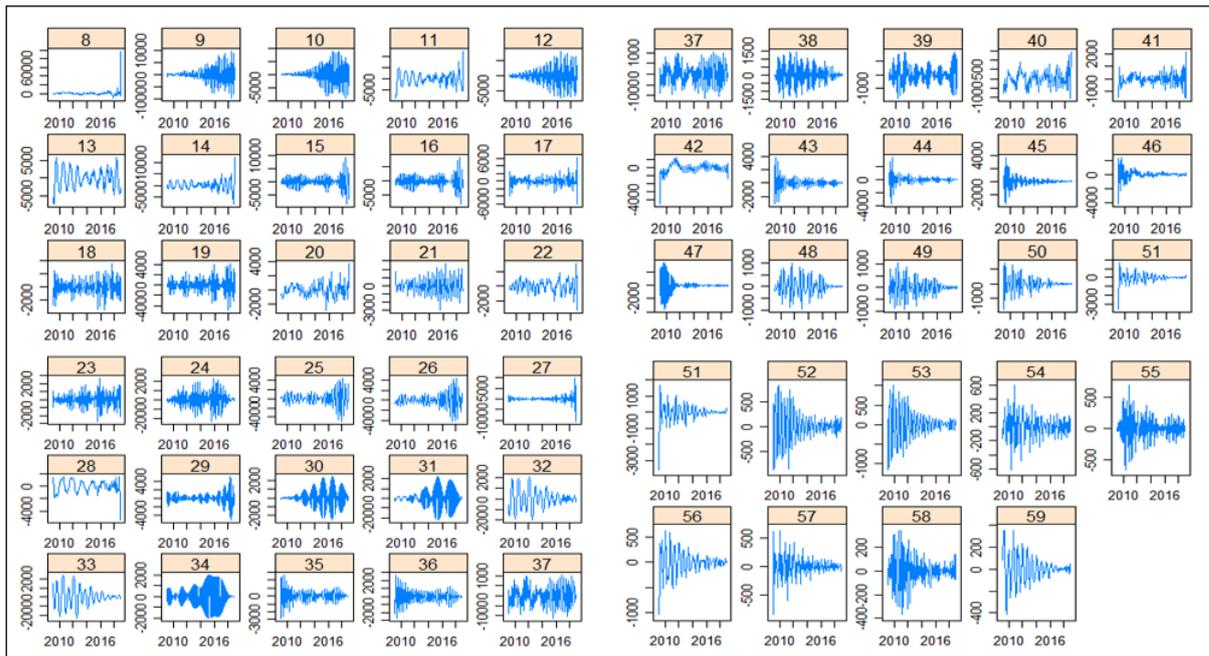
b) *Reconstruction*

i) *Grouping*

Pengelompokan *eigen triple* akan dibagi menjadi komponen tren dan komponen musiman. Plot dari seri yang sudah direkonstruksi dapat digunakan untuk mengidentifikasi *eigen triple* yang memiliki komponen tren dan komponen musiman. Pengelompokan *eigen triple* dapat dilihat dari kemiripan *singular value* dari masing-masing *eigen triple*. Berdasarkan Gambar 3, pengelompokan dibagi menjadi 4 bagian yaitu tren, musiman 1, musiman 2, dan musiman 3. Tren terdiri dari *eigen triple* ke-1. Musiman 1 terdiri dari *eigen triple* ke-2 dan ke-3. *Eigen triple* ke-4 dan ke-5 dikelompokkan menjadi musiman 2. Lalu, musiman 3 terdiri dari *eigen triple* ke-6 dan ke-7. Terakhir, *noise* terdiri dari *eigen triple* ke-8 sampai ke-59 yang dapat dilihat pada Gambar 4.



Gambar 3. Plot dari deret jumlah kedatangan wisatawan ke bali yang sudah direkonstruksi



Gambar 4. Plot dari yang sudah direkonstruksi dan dikelompokkan ke dalam Noise

ii) Diagonal Averaging

Deret yang sudah direkonstruksi yang berasal dari hasil *diagonal averaging* ditampilkan pada [Tabel 1](#).

Tabel 1. Hasil *Diagonal Averaging* deret Jumlah Kedatangan Wisatawan Mancanegara ke Bali yang sudah dikelompokkan

Tahun	Bulan	Data Asli	Rekonstruksi			Diagonal Averaging	Residu	
			Tren	Musiman 1	Musiman 2			Musiman 3
2009	Jan	17454	174756,8	-	21264,4	2975,90	192417,3	-
2009	Feb	14770	176111,4	-	21191,3	231.917	189247,4	-
2009	Mar	16820	177528,5	-	21534,8	-2847,4	188211,6	-
...
2018	Des	49881	542010,2	-108090	22136,5	20363,7	391419,7	107399,3

2) Peramalan dengan R-forecasting

Hasil peramalan data *out-sample* dengan metode SSA dengan *R-forecasting* ditunjukkan pada [Tabel 2](#). Peramalan jumlah kunjungan wisatawan mancanegara ke Bali menggunakan metode SSA dengan *R-forecasting* menghasilkan nilai MAPE sebesar 14,515%.

Tabel 2. Hasil Peramalan dengan metode SSA dengan R-forecasting

Tahun	Bulan	Data Asli	Hasil Ramalan
2019	Jan	455570	396720
2019	Feb	437456	462990
...
2019	Des	552403	462683
MAPE			14,515%

Meramalkan nilai residu dari model SSA dengan WFTS

a) Menentukan *Universe of Discourse*

Rentang nilai *universe of discourse* ditentukan dari rentang -190000 sampai 110000, maka diperoleh,

$$U = [-190000, 110000]$$

b) Membagi Semesta Pembicaraan ke dalam Interval yang Sama dan Menentukan Himpunan Fuzzy pada U

Interval ditentukan sebesar 5000, sehingga semesta pembicaraan terbagi menjadi 60 interval ($u_1, u_2, u_3, \dots, u_{60}$). Selanjutnya, himpunan fuzzy ditentukan dengan menambahkan variabel linguistik A_1, A_2, \dots, A_{60} pada setiap $u_1, u_2, u_3, \dots, u_{60}$, yang ditampilkan pada Tabel 3.

Tabel 3. Hasil Pembagian Semesta ke dalam Interval yang Sama

Semesta	Min	Mean	Max	Himpunan Fuzzy
u_1	-190000	-187500	-185000	A_1
u_2	-185000	-182500	-180000	A_2
u_3	-180000	-177500	-175000	A_3
...
u_{60}	105000	107500	110000	A_{60}

c) Melakukan Fuzzifikasi

Fuzzifikasi merupakan proses penentuan data aktual ke dalam himpunan fuzzy. Hasil fuzzifikasi ditampilkan pada Tabel 4.

Tabel 4. Hasil Fuzzifikasi

No.	Residu dari Model SSA	Himpunan Fuzzy
1	-17876,3	A_{35}
2	-41543,4	A_{30}
3	-20006,6	A_{34}
...
120	107399,3	A_{60}

d) Menentukan Fuzzy Logical Relationship (FLR) dan Fuzzy Logical Relationship Group (FLRG)

Hasil dari Fuzzy Logical Relationship (FLR) ditampilkan pada Tabel 5.

Tabel 5. Fuzzy Logical Relationsip (FLR)

$A_{35} \rightarrow A_{30}$
$A_{30} \rightarrow A_{34}$
$A_{34} \rightarrow A_{38}$
$A_{38} \rightarrow A_{37}$
...
$A_{36} \rightarrow A_{60}$
$A_{60} \rightarrow \#$

FLRG akan dibagi menjadi 2 macam. Pertama, *FLRG* dengan tidak memperhatikan pengulangan relasi *fuzzy* yang akan digunakan untuk metode Chen. Kedua, yaitu *FLRG* dengan memperhatikan pengulangan relasi *fuzzy* untuk metode Yu, Cheng, dan Lee. [Tabel 6](#) menunjukkan hasil *FLRG* dengan tidak memperhatikan pengulangan dan peramalannya. [Tabel 7](#) menunjukkan hasil *FLRG* dengan memperhatikan pengulangan dan peramalannya.

Tabel 6. *FLRG* dengan tidak Memperhatikan Pengulangan dan Peramalannya

<i>FLRG</i>	<i>F (t)</i>
$A_{24} \rightarrow A_{44}$	A_{44}
$A_{26} \rightarrow A_{45}$	A_{45}
$A_{30} \rightarrow A_{34}$	A_{34}
$A_{32} \rightarrow A_{37}, A_{34}$	A_{37}, A_{34}
$A_{33} \rightarrow A_{41}, A_{37}, A_{42}, A_{35}$	$A_{41}, A_{37}, A_{42}, A_{35}$
...	...
$A_{60} \rightarrow \#$	A_{60}

Tabel 7. *FLRG* dengan Memperhatikan Pengulangan dan Peramalannya

<i>FLRG</i>	<i>F (t)</i>
$A_{24} \rightarrow A_{44}$	A_{44}
$A_{26} \rightarrow A_{45}$	A_{45}
$A_{30} \rightarrow A_{34}$	A_{34}
$A_{32} \rightarrow A_{37}, A_{34}$	A_{37}, A_{34}
$A_{33} \rightarrow A_{41}, A_{37}, A_{41}, A_{42}, A_{35}$	$A_{41}, A_{37}, A_{42}, A_{35}$
...	...
$A_{60} \rightarrow \#$	A_{60}

e) Defuzzifikasi

Proses defuzzifikasi adalah proses untuk menentukan nilai ramalan berdasarkan pada *FLR* pada Tabel 5 dan *FLRG* pada Tabel 6 dan Tabel 7. Sebagai contoh, data Januari 2009 memiliki *FLR* yaitu $A_{35} \rightarrow A_{30}$ dan *FLRG* yaitu $A_{35} \rightarrow A_{30}$, maka nilai ramalan merupakan nilai tengah u_{30} yaitu -42500. Defuzzifikasi dilakukan menggunakan algoritma Chen, Yu, Cheng dengan $\alpha = 0,9$, dan Lee dengan nilai $c = 1,1$. Hasil defuzzifikasi untuk data *in-sample* dapat dilihat pada [Tabel 8](#) dan untuk data *out-sample* ditampilkan pada [Tabel 9](#).

Tabel 8. Hasil Defuzzifikasi untuk Data *In-Sample*

Tahun	Bulan	Residu dari Model SSA	Nilai ramalan			
			Chen	Yu	Cheng ($\alpha = 0,9$)	Lee ($c = 1,1$)
2009	Jan	-17876.31	-	-	-	-
2009	Feb	-41543.39	1875	6944.44	-1754.5	4529.02
2009	Mar	-20006.57	-22500	-22500	-24404	-22500
...
2018	Nov	-13095.3	-4500	-7733.9	-2230.8	-6882.7
2018	Des	107399.3	25500	24722.2	11440.5	19630.7

Tabel 9. Hasil Defuzzifikasi untuk Data Out-Sample

Tahun	Bulan	Residu dari Model SSA	Nilai ramalan			
			Chen	Yu	Cheng ($\alpha = 0,9$)	Lee ($c = 1,1$)
2019	Jan	58849,99	-	-	-	-
2019	Feb	-25534,4	57500	57500	57635	57500
2019	Mar	-100736	1250	1166.67	596.559	2792.05
...
2018	Nov	-2700,14	-4166.67	-4587.91	-3019.79	-3982.5
2018	Des	89719,65	9375	9871.79	3854.99	8194.88

3) Menghitung Nilai Ramalan Akhir

Nilai ramalan akhir diperoleh dengan menjumlahkan nilai ramalan SSA dengan *R-forecasting* dan nilai ramalan residu yang diperoleh menggunakan metode *WFTS*. Nilai ramalan akhir untuk data *out-sample* ditampilkan pada [Tabel 10](#) dan untuk data *in-sample* ditampilkan pada [Tabel 11](#).

Tabel 10. Nilai Ramalan Akhir untuk Data *In-sample*

Tahun	Bulan	Nilai Ramalan Akhir			
		SSA-WFTS Chen	SSA-WFTS Yu	SSA-WFTS Cheng ($\alpha = 0,9$)	SSA-WFTS Lee ($c = 1,1$)
2009	Jan	-	-	-	-
2009	Feb	191122	196192	187493	193776
2009	Mar	165712	165712	163807	165712
...
2018	Nov	415320	412086	417589	412938
2018	Des	416920	416142	402860	411050

Tabel 11. Nilai Ramalan Akhir untuk Data *Out-sample*

Tahun	Bulan	Nilai Ramalan Akhir			
		SSA - WFTS Chen	SSA - WFTS Yu	SSA - WFTS Cheng ($\alpha = 0,9$)	SSA - WFTS Lee ($c = 1,1$)
2019	Jan	-	-	-	-
2019	Feb	520490	520490	520625	520490
2019	Mar	551555	551472	550902	553097
...
2019	Nov	496458	496037	497605	496643
2019	Des	472058	472555	466538	470878

4) Menghitung MAPE dan Menentukan Model Terbaik

Berdasarkan Tabel 12, metode SSA- WFTS dengan algoritma Chen memiliki nilai MAPE terkecil dibanding metode lainnya untuk data *out-sample*, yaitu sebesar 9,029%. Maka dari itu, pemodelan untuk meramalkan jumlah kunjungan wisatawan mancanegara ke Bali menggunakan SSA-WFTS algoritma Chen merupakan model terbaik dibandingkan model SSA R-forecasting, SSA-WFTS algoritma Yu, SSA-WFTS algoritma Cheng ($\alpha = 0,9$), dan SSA-WFTS algoritma Lee ($c = 1,1$).

Tabel 12. Nilai MAPE

<i>Data</i>	<i>Metode</i>	<i>MAPE</i>
In-sample	SSA R-forecasting	5,156%
	SSA-WFTS metode Chen	5,154%
	SSA-WFTS metode Yu	5,139%
	SSA-WFTS metode Cheng ($\alpha = 0,9$)	4,762%
	SSA-WFTS metode Lee ($c = 1,1$)	5,057%
Out-sample	SSA R-forecasting	14,515%
	SSA-WFTS metode Chen	9,029%
	SSA-WFTS metode Yu	9,067%
	SSA-WFTS metode Cheng ($\alpha = 0,9$)	9,125%
	SSA-WFTS metode Lee ($c = 1,1$)	9,082%

Simpulan

Berdasarkan analisis yang telah dilakukan, peramalan jumlah kunjungan wisatawan mancanegara ke Bali menggunakan metode hibrida SSA-WFTS diperoleh model terbaik yaitu model SSA-WFTS dengan algoritma Chen. Model SSA-WFTS dengan metode Chen menghasilkan nilai MAPE terkecil yaitu sebesar 9,029%, dibandingkan dengan model SSA R-forecasting, SSA-WFTS metode Yu, SSA-WFTS metode Cheng ($\alpha = 0,9$), dan SSA-WFTS metode Lee ($c = 1,1$).

Daftar Pustaka

- [1] A. Hutauruk and S. Harto, "Pengaruh Kunjungan Wisatawan Asing terhadap Ekonomi Pariwisata Indonesia.," Jurnal Online Mahasiswa Fakultas Ilmu Sosial dan Ilmu Politik Universitas Riau, vol. 4, no. 2, pp. 1-14, 2017.
- [2] U. Nations, International Recommendations for Tourism Statistics 2008, New York: United Nations Publication, 2010.
- [3] T. A. Secretariat, ASEAN STATISTICAL YEARBOOK 2020, Jakarta: Association of Southeast Asian Nations (ASEAN) 2020, 2020.
- [4] L. U. Calderwood and M. Soshkin, "The Travel & Tourism Competitiveness Report 2019," World Economic Forum's Platform, Geneva, 2019.
- [5] B. S. Distribusi, "STATISTIK WISATAWAN MANCANEGARA KE BALI 2018," BPS Provinsi Bali, Denpasar, 2019.
- [6] N. Golyandina and A. Zhigljavsky, Singular Spectrum Analysis for Time Series, vol. 120, Berlin: Springer, 2013.
- [7] H. Khaeri, E. Yulian and G. Darmawan, "Penerapan Metode Singular Spectrum Analysis (SSA) pada Peramalan Jumlah Penumpang Kereta Api di Indonesia Tahun 2017," Jurnal Euclid, vol. 5, no. 1, pp. 8-19, 2017.

-
- [8] R. S. K. Siregar, D. Prariesa and G. Darmawan, "Aplikasi Metode Singular Spectral Analysis (SSA) dalam Peramalan Pertumbuhan Ekonomi Indonesia Tahun 2017," *Jurnal Matematika "MANTIK"*, vol. 3, no. 1, pp. 5-12, 2017.
- [9] M. Sari, I. W. Sumarjaya and M. Susilawati, "Peramalan Jumlah Kunjungan Wisatawan Mancanegara ke Bali Menggunakan Metode Singular Spectrum Analysis," *E-Jurnal Matematika*, vol. 8, no. 4, pp. 303-308, 2019.
- [10] G. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159-175, 2001.
- [11] Q. Song and B. Chissom, "Fuzzy time series and its model," *Fuzzy Sets and System*, vol. 54, no. 3, pp. 269-277, 1993.
- [12] S. Chen, "Forecasting enrollments based on fuzzy time series," *Fuzzy Sets and System*, vol. 81, pp. 311-319, 1996.
- [13] H. Yu, "Weighted fuzzy time series models for TAIEX forecasting," *Physics A*, vol. 349, no. 2, pp. 609-624, 2005.
- [14] M. Lee and Suhartono, "A novel weighted fuzzy time series model for forecasting seasonal data," *Proceeding the 2nd International Conference on Mathematical Sciences*, pp. 332-340, 2010.
- [15] W. Sulandari, Subanar, M. Lee and P. Rodrigues, "Indonesia electricity load forecasting using singular spectrum analysis, fuzzy system, and neural networks," *Energy*, vol. 190, pp. 1-11, 2020.
- [16] Subanar and W. Sulandari, "A comparison forecasting methods for trend and seasonal Indonesia tourist arrivals time series," *AIP Conference Proceedings*, vol. 2329, p. 060012, 2021.
- [17] Suhartono and M. H. Lee, "A Hybrid Approach based on Winter's Model and Weighted Fuzzy Time Series for Forecasting Trend and Seasonal Data," *Journal of Mathematics and Statistics*, pp. 177-183, 2011.
- [18] N. Golyandina, V. Nekrutkin and A. Zhigljavsky, *Analysis of Time Series Structure: SSA and Related Techniques*, Florida: Chapman & Hall, 2001.
- [19] Q. Zhang, B. Wang, B. He, Y. Peng and M. Ren, "Singular Spectrum Analysis and ARIMA Hybrid Model for Annual Runoff Forecasting," *Water resources management*, vol. 25, no. 11, pp. 2683-2703, 2011.
- [20] C. Cheng, T. Chen, H. J. Teoh and C. Chiang, "Fuzzy time-series based on adaptive expectation model for TAIEC forecasting," *Expert System with Application*, vol. 34, pp. 1126-1132, 2008.
- [21] A. A. Ete, M. Fitriyanawati and M. T. Arifin, "Forecasting The Number Of Tourist Arrivals To Batam Using The Singular Spectrum Analysis And Arima Methods," *Advances in Social Science, Education and Humanities Research*, vol. 317, pp. 119-126, 2019.