BAB II

TINJAUAN PUSTAKA

2.1. Wisatawan

Sektor pariwisata sangat berkembang cepat seiring dengan pertumbuhan negara-negara di dunia. Sektor pariwisata merupakan salah satu sumber devisa negara khususnya di bidang ekonomi. Salah satu faktor penentu perkembangan pariwisata adalah banyaknya kunjungan wisatawan baik wisatawan nusantara maupun wisatawan mancanegara. Indonesia sendiri merupakan salah satu tujuan utama kunjungan wisatawan mancanegara. Pada tahun 2014, data kunjungan terbanyak wisatawan mancanegara mengunjungi Bali dengan total 39,80% terhadap wisatawan nusantara. Pulau Bali yang menjadi destinasi wisata terpopuler di Indonesia dengan total kunjungan wisatawan mancanegara ke Indonesia melalui pintu masuk Bandara Ngurah Rai Bali mencapai 39% terhitung sampai bulan Oktober 2016, dengan nilai penerimaan devisa Bali untuk Indonesia dari sektor pariwisata sebesar 70 Triliun Rupiah.

2.2. Peramalan

Peramalan merupakan kata yang berasal dari kata ramal yang artinya adalah suatu kejadian yang diduga atau diperkirakan akan terjadi dimasa yang akan datang. Sedangkan untuk peramalan merupakan bentuk kegiatannya. Peramalan menurut J. Supranto merupakan dugaan atau perkiraan mengenai terjadinya suatu kejadian diwaktu yang akan datang (*Ischak*, 2018).

Suatu ramalan yang layak dari deret waktu dapat dilakukan hanya jika kondisi berikut terpenuhi :

1. Deret memiliki struktur
2. Sebuah mekanisme (metode, algoritma) yang mengidentifikasi struktur ditemukan.
3. Sebuah metode kelanjutan deret waktu (*time series continuation*) berdasarkan pada struktur yang diidentifikasi tersedia.
4. Struktur dari deret waktu dipertahankan untuk periode waktu mendatang, yang mana kita akan meramalkan (melanjutkan) deret. (Golyandina, 2001)

2.3. Deret Waktu

Deret waktu adalah sekumpulan data yang berupa pengamatan yang diukur selama kurun waktu tertentu, berdasarkan waktu dengan interval yang sama. Dengan adanya deret waktu maka menjadikan adanya analisis deret waktu yang merupakan metode untuk melalukan peramalan. Peramalan deret waktu adalah penggunaan model untuk memprediksi nilai pada kejadian dimasa yang akan datang. Oleh karena itu data yang digunakan harus memiliki hubungan antara kejadian mendatang dengan kejadian di waktu sebelumnya (Ischak, 2018).

Dalam pemilihan metode peramalan terdapat faktor penting yaitu mengidentifikasi dan memahami pola data. Pengidentifikasian pola data dapat dilakukan dengan melihat plot data deret waktu yang disajikan sesuai dengan data.

Pola data pada deret waktu terbagi menjadi 4 jenis, diantaranya (Makridakis S, *et al*, 1997) :

1. Pola Horizontal (H) yaitu pola data yang terjadi jika data berfluktuasi di sekitar nilai rata-rata konstan.
2. Pola Musiman (S) yaitu pola dataterjadi jika data dipengaruhi faktor musiman.
3. Pola Siklis (C) yaitu pola data yang terjadi jika data dipengaruhi fluktuasi ekonomi jangka panjang.
4. Pola Data Trend (T) yaitu pola data yang terjadi jika terjadi kenaikan ataupun penurunan jangka panjang pada data.

2.4. *Singular Spectrum Analysis* (SSA)

Singular Spectrum Analysis (SSA) adalah teknik baru dari analisis deret waktu. Metode ini tidak membuat asumsi statistik mengenai sinyal atau *noise* saat melakukan analisis dan menyelidiki sifat-sifat algoritma (Golyandina, 2001). Dasar dari *Singular Spectrum Analysis* (SSA) adalah dekomposisi dengan menguraikan data deret waktu menjadi komponen-komponennya yaitu tren, musiman, siklis, dan *noise* (Ischak, 2018). Tahapan dalam melakukan analisis menggunakan *Singular Spectrum Analysis* (SSA) terbagi menjadi dua yaitu dekomposisi dan rekonstruksi (Golyandina, 2001).

2.4.1. Dekomposisi

Pada tahap dekomposisi, parameter yang digunakan adalah *Window Length* (L) (Ischak, 2018). Parameter ini berfungsi untuk menentukan banyaknya dimensi matriks lintasan. Nilai L ini merupakan dimensi dari matriks lintasan yang merupakan matriks dari perkalian *Hankel*. Penentuan dari nilai L dilakukan dengan proses pengecekan melalui *trial and eror*.

2.4.1.1. *Embedding*

Pada proses *Embedding* mengubah data deret waktu menjadi urutan Lag vektor dengan ukuran L dengan membentuk K=N-L+1 lag vektor.

$$X_i = (x_i,\ldots, x_i+L-1)^T \qquad (1 \leq i \leq K)$$

dengan ukuran L. Jika ditekankan ukuran (dimensi) vektor X_i, maka akan disebut sebagai L-lag vektor.

Matriks Lintasan dari **X** adalah (Zhigljavsky, 2011)

$$\mathbf{X} = [X_1 : \ldots : X_K] = (x_{ij})_{i,j}^{L,K} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_k \\ x_2 & x_3 & x_4 & \cdots & x_{K+1} \\ x_3 & x_4 & x_5 & \cdots & x_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & x_{L+2} & \cdots & x_N \end{bmatrix} \qquad (2.1)$$

Lag vektor X_i adalah kolom dari matriks lintasan **X**. Baris dan kolom dari **X** bagian dari data deret waktu.

Elemen (i,j) dari matriks **X** adalah $x_{ij} = x_{i+j-1}$ yang menghasilkan X bernilai sama dengan elemen “anti diagonals” i+j = konstan. (Oleh karena itu lintasan matriks yang terbentuk dinamakan matriks *Hankel*). Matriks yang terbentuk mendefinisikan kemiripan satu dengan yang lain antara matriks lintasan L x K dan deret waktu.

2.4.1.2. Singular Value Decomposition (SVD)

Pada tahap ini, akan menghasilkan *singular value decomposition* (SVD) dari matriks **X**. Didefinisikan **S** = $\mathbf{XX^T}$ dan ditunjukkan dengan $\lambda_1, \dots, \lambda_L$ *eigenvalue* dari S sedangkan *singular value* dinotasikan dengan $U_{1,\dots,}$ U_L merupakan *eigenvector* yang sesuai dengan *eigenvalues* pada matriks S.

Selanjutnya didefinisikan d = rank **X** = max {i, dimana $\lambda_i > 0$} atau biasanya dengan d = L* = min {K,L}. Sedangkan *principal component* dinotasikan dengan $V_i = X^T U_i / \sqrt{\lambda_i}$ (i = 1,..., d). Sehingga didapatkan SVD dari matriks lintasan **X** sebagai berikut (Zhigljavsky, 2011)

$$\mathbf{X} = X_1 + \dots + X_d \qquad (2.2)$$

dimana $X_i = \sqrt{\lambda_i}\, U_i V_i^T$. Matriks X mempunyai rank 1, karena merupakan matriks elementer. Kumpulan dari $(\sqrt{\lambda_1}, U_i, V_i)$ disebut *eigentriple* ke i dari SVD (Zhigljavsky, 2011). SVD dari matriks lintasan dapat ditulis dengan persamaan berikut :

$$X = X_1 + X_2 + \dots + X_d$$

$$X = \sqrt{\lambda_1} U_1 V_1^T + \sqrt{\lambda_2} U_2 V_2^T + \dots + \sqrt{\lambda_d} U_d V_d^T$$

$$X = \sum_{i=1}^{d} \sqrt{\lambda_i}\, U_i V_i^T \qquad (2.3)$$

2.4.2. Rekonstruksi

Parameter yang digunakan dalam tahap rekonstruksi adalah *groupping eeffect* (r). Fungsi dari parameter ini adalah menentukan pola pada plot data. Sebelumnya pada tahap dekomposisi dengan penggunaan parameter L, dan menyajikan serangkaian seri awal yang telah dipisahkan dengan baik pada SVD maka *eigentriples* yang terbentuk akan membantu dalam penentuan parameter *grouping effect*. Hasil dari tahap rekonstruksi akan mendekatkan hasil peramalan dengan data asli. Oleh karena itu pengelompokkan yang tepat dilakukan akan mendukung hasil peramalan dengan baik dengan menunjukkan nilai MAPE dari nilai ramal dengan data asli.

2.4.2.1. Grouping

Pada persamaan yang telah terbentuk (2.3), matrisk **Xi** akan dipartisi ke m *disjoint subset* $I_{1, \dots ,} I_{m.}$ Jika I = {i_1 , ... , i_p} maka yang dihasilkan dari matriks $\mathbf{X_i}$ sesuai dengan kelompok I yang didefinisikan sebagai $\mathbf{X_1}$ = X_{i1} + ... + X_{ip} . Matriks yang dihasilkan dihitung untuk pengelompokkan I = I_1 , ... , I_m maka persamaan yang terbentuk adalah (Zhigljavsky, 2011)

$$\mathbf{X} = X_{I1} + \dots + X_{Im} \qquad (2.4)$$

Prosedur pemilihan set I_1 , ... , I_m dinamakan *eigentriple* *grouping*. Jika m=d dan $I_j = \{j\}$, j= 1,..., d maka pengelompokkan yang terbentuk disebut sebagai *elementary*.

2.4.2.2. Diagonal Averaging

Pada tahap ini mengubah setiap matriks $\mathbf{X}_{I_j}$ dari persamaan (2.4) menjadi data deret waktu baru dengan panjang N yang dimisalkan denga **Y** berikut

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_k \\ y_{21} & y_{22} & \cdots & y_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & \cdots & y_N \end{bmatrix}$$

Jika **Y** adalah matriks LxK dengan elemen y_{ij}, $1 \le i \le L$, $1 \le j \le K$. Definisikan $L^* = \min(L,K)$, $K^* = \max(L,K)$ dan $N = L+K-1$. Maka $y^*_{ij} = y_{ij}$ jika $L < K$ dan $y^*_{ij} = y_{ji}$ untuk yang lain. Dengan membuat *diagonal averaging* mengubah matriks **Y** menjadi deret waktu $y_1, \ldots, y_N$ dengan rumus (Zhigljavsky, 2011)

$$\mathbf{Y_k} = \begin{cases} \frac{1}{k}\sum_{m=1}^{k} y^*_{m,k-m+1} & untuk\ 1 \le k < L^* \\ \frac{1}{L^*}\sum_{m=1}^{L^*} y^*_{m,k-m+1} & untuk\ L^* \le k \le K^* \\ \frac{1}{N-k+1}\sum_{m=k-K^*+1}^{N-K^*+1} y^*_{m,k-m+1} & untuk\ K^* < k \le N \end{cases} \quad (2.5)$$

Berdasarkan persamaan (2.5) misal pada matriks **Y**, dipilih k=1 maka $y_1 = y_{11}$, untuk k=2 maka $y_2 = \left(\frac{y_{12} + y_{21}}{2}\right)$, untuk k=3 maka $y_3 = \left(\frac{y_{13} + y_{22} + y_{31}}{3}\right)$, dan begitu seterusnya. Untuk catatan jika matriks **Y** adalah matriks lintasan dari beberapa deret $(z_1, \ldots, z_N)$, maka $y_i = z_i$

untuk semua i.

Persamaan (2.5) diterapkan pada hasil matriks $\mathbf{X}_{I_k}$ membuat deret baru $\widetilde{X^{(k)}} = \left(\widetilde{x_1^{(k)}}, \dots, \widetilde{x_N^{(k)}}\right)$. Oleh karena itu deret $x_1, \dots, x_N$ adalah didekomposisi menjadi jumlah m deret yaitu

$$x_n = \sum_{k=1}^{m} \widetilde{x_n^{(k)}} \qquad (n = 1,2, \dots, N) \qquad (2.6)$$

2.5. Peramalan dalam *Singular Spectrum Analysis* (SSA)

Prinsip-prinsip peramalan dalam *Singular Spectrum Analysis* (SSA) memiliki sifat penting dari dekomposisi SSA jika seri asli f_n memenuhi rumus *linear recurrent formula* (LRF)

$$f_n = a_1 f_{n-1} + \dots + a_d f_{n-d} \qquad (2.7)$$

dimana d ($d \leq L$) adalah angka dari *nonzero singular values* dari matriks **X**. Dari beberapa dimensi *d* dengan beberapa koefisien $a_1, \dots, a_d$ kemudian untuk setiap *N* dan *L* terdapat banyak nilai singular *d nonzero* dalam SVD dari matriks lintasan **X**. Oleh karena itu jika *window length* L dan K = N-L+1 lebih besar dari *d*, maka hanya dibutuhkan lebih banyak *d* pada matrisk $\mathbf{X}_i$ untuk merekonstruksi data (Golyandina, 2001).

Asumsikan tujuan yang ingin dicapai dengan menggunakan SSA adalah suatu komponen aditif tertentu $F_N^{(1)}$ dapat diekstrak dari suatu deret F_N. Dalam algoritma ini, suatu *window length* L yang sesuai, SVD matriks lintasan diperoleh dari deret F_N dan *eigentriples* $(\sqrt{\lambda}, U, V)$ dipilih yang sesuai dengan $F_N^{(1)}$. Pada langkah *diagonal averaging*, deret yang direkonstruksi $\tilde{F}_N^{(1)}$ yang mengestimasi $F_N^{(1)}$ akan diperoleh (Ete, 2017).

2.5.1. *Reccurent Forecasting*

Algoritma *Recurrent Forecasting* dapat dirumuskan sebagai berikut:

1. Deret waktu $\mathbf{Y_{N+M}} = (y_1, \ldots, y_{N+M})$ didefinisikan dengan

$$y_i = \begin{cases} \tilde{x}_i & \text{for } i=1,\ldots,N \\ \sum_{j=1}^{L-1} a_j y_{i-j} & \text{for } i=N+1,\ldots,N+M \end{cases} \quad (2.8)$$

2. Angka-angka $\mathbf{Y_{N+1}}, \ldots, \mathbf{y_{N+M}}$ membentuk istilah *M* dari *Recurrent Forecasting*.

   Jadi, peramalan *Recurrent Forecasting* dilakukan dengan penggunaan langsung LRR dengan koefisien $\{a_j, j = 1, \ldots, L-1\}$.

   Definisikan operator linear $\wp_{Rec} : R^L \mapsto R^L$ dengan rumus

$$\wp_{Rec} Y = \begin{pmatrix} \overline{Y} \\ R^T \overline{Y} \end{pmatrix} \quad (2.9)$$

dengan,

$$Z_i = \begin{cases} \hat{X}_i & \text{for } i = 1, \ldots, K \\ \wp_{Rec} Z_{i-1} & \text{for } i = K+1, \ldots, K+M \end{cases} \quad (3.0)$$

Dapat dilihat bahwa matriks $\mathbf{Z} = [Z_1 : \ldots : Z_{K+M}]$ adalah matriks lintasan dari $\mathbf{Y_{N+M}}$. Karena itu (3.0) dapat dianggap bentuk vector dari (2.8) .

2.6. Ketepatan Pencapaian Model Peramalan

Dalam melakukan suatu peramalan yang merupakan kegiatan memprediksi masa depan dengan menggunakan data di masa lampau, hasil yang akan didapatkan tidaklah sama dengan data sesungguhnya (Ishcak, 2018). Maka dari itu usaha untuk membuat nilai *error* seminimal mungkin dibutuhkan pada proses peramalan.

Salah satu tingkat akurasi peramalan dapat diukur dari nilai *Mean Absolute Persentage Error* (MAPE) yaitu rata-rata persentase kesalahan pertama dari beberapa periode. Tingkat keakuratan dapat dijelaskan dengan membandingkan nilai yang diproyeksikan dengan nilai aktual. Untuk melakukan peramalan dan untuk mengetahui akuratnya sebuah model maka nilai akurasinya harus semakin kecil.

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

Untuk ketentuan MAPE adalah sebagai berikut :

MAPE	Keterangan
<10 %	Sangat Baik
<20 %	Baik
<30 %	Cukup Baik
>30%	Tidak Akurat

Sumber : (Lewis, 1997).