

Forecasting the Existence of Chocolate with Variation and Seasonal Calendar Effects Using the Classic Time Series Approach

Peramalan Eksistensi Cokelat dengan Efek *Calendar Variation* dan *Seasonal* Menggunakan Pendekatan Time Series Klasik

*I Gusti Bagus Ngurah Diksa

Abstract

Chocolate is the raw material for making cakes, so consumption of chocolate also increases on Eid al-Fitr. However, this is different in the United States where the tradition of sharing chocolate cake is carried out on Christmas. To monitor the existence of this chocolate can be through the movement of data on Google Trends. This study aims to predict the existence of chocolate from the Google trend where the use of chocolate by the community fluctuates according to the calendar variance and seasonal rhythm. The method used is classic time series, namely naive, double exponential smoothing, multiplicative decomposition, additive decomposition, holt winter multiplicative, holt winter additive, time series regression, hybrid time series, ARIMA, and ARIMAX. Based on MAPE in sample, the best time series model to model the existence of chocolate in Indonesia is ARIMAX (1,0,0) while for the United States it is Hybrid Time Series Regression-ARIMA(2,1,[10]). For forecasting the existence of chocolate in Indonesia, the best models in forecasting are ARIMA (([11],[12]),1,1) and Naïve Seasonal. In contrast to the best forecasting model for the existence of chocolate in the United States, namely Hybrid Naïve Seasonal-SARIMA (2,1,0)(0,0,1)¹² Hybrid Time Series Regression- ARIMA(2,1,[10]), Time Series Regression, Winter Multiplicative, ARIMAX([3],0,0).

Keywords : Chocolate, Calendar Varians, Seasonal, MAPE

Abstrak

Cokelat menjadi bahan baku pembuatan kue, maka konsumsi cokelat meningkat juga di hari raya Idul Fitri Namun hal ini berbeda yang terjadi di United States yang mana tradisi berbagi kue cokelat dilaksanakan di hari raya Natal. Untuk memantau eksistensi cokelat ini dapat melalui pergerakan data di google trend. penelitian ini bertujuan untuk meramalkan eksistensi cokelat dari google trend yang mana penggunaan cokelat oleh masyarakat ini mengalami fluktuasi mengikuti ritme *calendar varians* dan musiman. Metode yang digunakan adalah time series klasik yaitu *naïve*, *dekomposisi multiplicative*, *dekomposisi additive*, *holt winter multiplicative*, *holt winter additive*, *time series regression*, *hybrid time series*, ARIMA, dan ARIMAX. Berdasarkan MAPE in sampel model time series terbaik untuk meodelkan eksistensi cokelat di Indonesia adalah ARIMAX (1,0,0) sedangkan untuk United States adalah Hybrid Time Series Regression-ARIMA(2,1,[10]). Untuk peramalan eksistensi cokelat di Indonesia model terbaik dalam peramalan yaitu ARIMA (([11],[12]),1,1) dan Naïve Seasonal. Berbeda pada model peramalan terbaik untuk eksistensi cokelat di United States yaitu Hybrid Naïve Seasonal-SARIMA

*Badan Pusat Statistik Provinsi Sulawesi Selatan

Email address: igbn.diksa@bps.go.id



(2,1,0)(0,0,1)¹² Hybrid Time SeriesRegression- ARIMA(2,1,[10]), Time Series Regression , Winter Multiplicative, ARIMAX ([3],0,0).

Kata Kunci: Cokelat, Calender Varians, Seasonal, MAPE

1. PENDAHULUAN

Cokelat adalah olahan yang dihasilkan dari suatu bahan baku yaitu biji dan lemak kakao. Cokelat memiliki kandungan senyawa yang berfungsi menenangkan ketegangan saraf, mampu mengontrol tekanan darah dan juga kolesterol. Orang yang mengonsumsi cokelat ini cenderung akan lebih rileks, sehingga akan berpengaruh baik kepada kesehatan tubuh manusia. Selain itu, manfaat dari mengonsumsi cokelat ini juga dapat mengontrol kadar gula darah pada penderita [7].

Dalam perkembangannya, cokelat menjadi minuman dan makanan yang disukai anak-anak, remaja maupun orang dewasa. Cokelat menjadi sesuatu yang dapat membuat orang ketagihan. Cokelat menjadi sesuatu yang dapat membuat orang ketagihan ketika sudah mengenal dan merasakan nikmatnya. Di setiap negara, cokelat bahkan menjadi sesuatu yang bukan hanya sekedar dipasarkan, bahkan dapat menjadi daya tarik wisata dan telah dijadikan museum [3].

Untuk sebagian kalangan, cokelat dapat dijadikan sebagai simbol kasih sayang atau sebagai bentuk perhatian lebih bagi orang-orang tertentu, terutama pada saat Valentine atau White Day. Petani kakao Indonesia kebanjiran permintaan jelang perayaan hari kasih sayang atau valentine setiap tahunnya. Karena cokelat menjadi bahan baku pembuatan kue, maka konsumsi cokelat meningkat juga di hari raya Idul Fitri yang mana masyarakat banyak membuat kue sebagai persiapan perayaan hari raya Idul Fitri. Namun hal ini berbeda yang terjadi di United States yang mana tradisi berbagi kue cokelat dilaksanakan di hari raya Natal.

Untuk memantau eksistensi cokelat ini dapat melalui pergerakan data di google trend. Pada zaman digital ini Google Trends digunakan untuk memahami perubahan secara sosial dan prediksi [6]. Google Trends merupakan salah satu produk Google Inc yang berfokus pada trends pencarian atau penelusuran pada laman google dalam kurun waktu tertentu. Penggunaannya menjadi signifikan, meskipun datanya sederhana. Studi menjelaskan dalam penelitiannya mengevaluasi kemampuan prediksi data pencarian internet dengan menggunakan google dengan proxy menggunakan google trend menghasilkan bahwa hasilnya 40 persen lebih baik untuk menentukan pergerakan positif atau negatif sebuah objek dibandingkan dengan menggunakan data secara actual [5].

Bila dikaitkan dengan eksistensi cokelat yang berlangsung disuatu periode maka google trend menjadi alternative dalam melihat fenomena ini. Di era digital ini banyak orang yang mencari informasi di Internet dan juga mengunggah apa yang menjadi trending saat itu. Misalkan pada waktu valentine, hari raya Idul Fitri, Natal dan lainnya, beberapa masyarakat akan banyak mencari informasi tentang cokelat seperti kue minuman dan lainnya sebagai literature mereka dalam mempersiapkan hari raya tersebut. Beberapa orang juga pasti mengunggah beberapa opini dan hasil karya mereka terkait cokelat pada waktu persiapan dan tepat pada waktu perayaan hari itu. Kebiasaan masyarakat ini sering dilaksanakan secara sama dalam tahun ke tahun. Dengan kata lain data pada google trend dapat menjadi alternative sebagai pendekatan eksistensi cokelat dari waktu ke waktu.

Peramalan pada eksistensi cokelat ini bisa menjadi peluang bisnis dalam kegiatan marketing mereka ketika bebisnis cokelat atau kebijakan oleh pemerintah tentang penyediaan cokelat di dalam negeri dan ekspor cokelat ke luar negeri. Oleh karena itu penelitian ini bertujuan untuk meramalkan eksistensi cokelat dari google trend yang mana penggunaan cokelat oleh masyarakat ini mengalami fluktuasi mengikuti ritme *calendar varians* dan musiman dengan

menggunakan metode peramalan klasik. Metode ini memiliki beberapa keunggulan yaitu mudah dalam pembentukan modelnya sehingga lebih cepat dalam pembentukannya dan fleksibel dan dapat mewakili rentang yang lebar dari karakter deret waktu yang terjadi dalam jangka pendek [8].

2. TINJAUAN PUSTAKA

2.1 Metode Naïve

Metode Naïve merupakan metode peramalan yang paling sederhana. Peramalan pada metode Naïve didasarkan pada nilai pengamatan tepat sebelumnya. Karena itu, Taylor *et al.* (2006) menyatakan bahwa metode naïve sering digunakan oleh para praktisi, sehingga menjadi benchmark untuk menunjukkan kebaikan metode yang diusulkan. Secara matematis, peramalan menggunakan metode Naïve pada data yang stasioner, pada data dengan trend linear, serta pada data dengan trend eksponensial masing-masing dapat dituliskan sesuai persamaan 1 dan 2 [1].

Model untuk data *trend* sederhana :

$$\hat{Y}_{t+1} = Y_t + (Y_t - Y_{t-1}) \text{ bersifat aditif (Trend Linier)} \quad (1)$$

$$\hat{Y}_{t+1} = Y_t \frac{Y_t}{Y_{t-1}} \text{ bersifat mutliplikatif (Trend Eksponensial)} \quad (2)$$

Jika data time series memuat unsur *seasonal*, s , maka peramalan pada metode Naïve didasarkan pada pengamatan pada periode yang sama pada *seasonal* sebelumnya. Jika memuat pola *double seasonal*, dimana $s_1 < s_2$, maka peramalan pada metode Naïve didasarkan pada periode *seasonal* terpanjang, yaitu s_2 . Secara matematis, peramalan menggunakan metode *seasonal* Naïve dan *double seasonal* Naïve dapat dituliskan sebagai berikut.

$$\hat{Y}_t(h) = Y_{t+h-s} \quad (3)$$

$$\hat{Y}_t(h) = Y_{t+h-s_2} \quad (4)$$

2.2 Double Exponential Smoothing

Metode *Double Exponential Smoothing* digunakan ketika data menunjukkan adanya trend. Metode *Double Exponential Smoothing* dari Holt menjadikan nilai trend menjadi *smooth* dengan parameter yang berbeda dari parameter yang digunakan pada deret yang asli [7]. Ramalan dari metode *Double Exponential Smoothing* dari Holt didapatkan dengan menggunakan dua konstanta pemulusan α dan γ (dengan nilai masing-masing antara 0 hingga 1). Peramalan dengan metode *Double Exponential Smoothing* ini dilakukan dengan mengkombinasikan berbagai kemungkinan nilai dari dua konstanta pemulusan α dan γ pada tiga persamaan berikut hingga diperoleh kombinasi yang menghasilkan ramalan yang paling optimal.

$$S'_t = \alpha X_t + (1 - \alpha)(S'_{t-1} + b_{t-1}) \quad (8)$$

$$b_t = \gamma(S'_t - S'_{t-1}) + (1 - \gamma)b_{t-1} \quad (9)$$

$$F_{t+m} = S'_t + b_t m \quad (10)$$

Dimana :

S'_t : Nilai *single exponential smoothing* periode ke t

S'_{t-1} : Nilai *single exponential smoothing* periode ke $t-1$

α : Nilai parameter exponential smoothing ($0 < \alpha < 1$)

- X_t : Data aktual pada periode ke-t
 b_t : Pemulusan trend pada periode ke t
 b_{t-1} : Pemulusan trend pada periode ke t-1
 γ : Nilai parameter pemulusan trend $0 < \gamma < 1$
M : Periode ke depan yang akan diramalkan
 F_{t+m} : Nilai peramalan untuk (t+m) periode ke depan

Proses inialisasi untuk pemulusan eksponensial ganda dari Holt memerlukan dua taksiran, yaitu S_t dan b_t . proses inialisasi, diawali dengan memilih $S_1 = X_1$ sedangkan untuk taksiran trend didapat dari rumus $b_1 = X_2 - X_1$. Adapun konstanta pemulusan α dan γ dalam metode peramalan *double exponential smoothing* berperan sebagai faktor pembobotan. Nilai α dan γ bernilai dekat dengan satu, ramalan terbaru akan menyertakan penyesuaian yang besar untuk setiap kesalahan yang terjadi pada ramalan sebelumnya. Sebaliknya jika α dan γ dekat dengan nol, ramalan terbaru akan sangat mirip dengan nilai yang lama [10].

2.3 Dekomposisi

Metode dekomposisi merupakan suatu metode peramalan yang menggunakan empat komponen utama dalam meramalkan nilai masa depan. Keempat komponen tersebut antara lain trend, musiman, siklus dan error. Metode dekomposisi dilandasi oleh asumsi bahwa data yang ada merupakan gabungan dari beberapa komponen, secara sederhana digambarkan sebagai berikut [11].

$$Data = Pola + error = f(trend, siklus, musiman) + error \quad (11)$$

Asumsi diatas artinya terdapat empat komponen yang mempengaruhi suatu deret waktu, yaitu tiga komponen yang dapat diidentifikasi karena memiliki pola tertentu, yaitu: trend, siklus dan musiman. Persamaan umum matematis dari pendekatan dekomposisi adalah

$$X_t = f(T_t, S_t, C_t, I_t) \quad (12)$$

Dimana :

- X_t = nilai deret berkala (data *actual*) pada periode t;
 T_t = komponen trend (trend) pada periode t;
 S_t = komponen musiman (*seasonal*) pada periode t;
 C_t = komponen siklus (*cyclic*) pada periode t;
 I_t = komponen kesalahan tidak beraturan (*irregular*) pada periode t;
t = periode (*time*)

Dalam metode dekomposisi terdapat model dekomposisi aditif dan multiplikatif. Model dekomposisi aditif dan multiplikatif dapat digunakan untuk meramalkan faktor trend, musiman dan siklus. Metode dekomposisi rata-rata sederhana berasumsi pada model aditif yang secara matematis dapat ditulis:

$$Y_t = T_t + S_t + C_t + I_t \quad (13)$$

Sedangkan metode dekomposisi pada data bergerak (dekomposisi klasik) berasumsi bahwa pada model multiplikatif yang secara matematis dapat ditulis [13].

$$Y_t = T_t \cdot S_t \cdot C_t \cdot I_t \quad (14)$$

2.4 Holt - Winters

Metode Holt – Winters merupakan perkembangan dari metode pemulusan eksponensial sederhana yang menggunakan tiga konstanta pemulusan yaitu konstanta untuk pemulusan

keseluruhan level, pemulusan kecenderungan (trend) dan pemulusan musiman (seasonal) . Model Holt Winter ini menggunakan dua pendekatan, yaitu :

1. Metode *Holt Winter Multiplicative* yang digunakan untuk variasi data musiman dari data runtun waktu yang mengalami peningkatan atau penurunan (fluktuasi). Nilai ramalan (\hat{Y}_{t+k}) untuk periode (t+k) yang ditinjau pada akhir periode ke- t dari model ini adalah

$$\hat{Y}_{t+k} = (L_t + kT_t) S_{t+k-e}, \quad (15)$$

Dengan nilai pemulusan yang digunakan sebagai berikut :

- a. Pemulusan Keseluruhan (*level*)

$$L_t = \alpha \frac{Y_t}{S_{t-e}} + (1 - \alpha)(L_{t-1} + T_{t-1}), \quad (16)$$

- b. Pemulusan Kecenderungan (*trend*)

$$T_t = \beta (L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (17)$$

- c. Pemulusan Musiman (*seasonal*)

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-e} \quad (18)$$

Dengan $0 \leq \alpha, \beta, \gamma \leq 1, S_{t-c}$ nilai estimasi faktor musiman, c adalah panjang musiman dan $k = 1, 2, \dots, c$.

2. Metode *Holt Winter Aditive* digunakan untuk variasi data musiman dari data runtun waktu yang konstan. Pada akhir periode ke-t, nilai ramalan (\hat{Y}_{t+k}) untuk periode (t + k) diperoleh dari persamaan

$$\hat{Y}_{t+k} = L_t + kT_t + S_{t+k-c} \quad (19)$$

- a. Pemulusan Keseluruhan (level)

$$L_t = \alpha(Y_t - S_{t-c}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (20)$$

- b. Pemulusan Kecenderungan (trend)

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (21)$$

- c. Pemulusan Musiman (seasonal)

$$S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-c} \quad (22)$$

Dalam penggunaan metode peramalan ini, diperlukan nilai awal. Model pemulusan Winter dapat digunakan dengan mengambil secara sebarang beberapa nilai awal yaitu :

$$L_c = \frac{1}{c}(Y_1 + Y_2 + \dots + Y_c) \quad (23)$$

$$T_c = \frac{1}{K} \left(\frac{Y_{c+1} - Y_1}{c} + \frac{Y_{c+2} - Y_2}{c} + \dots + \frac{Y_{c+k} - Y_k}{c} \right) \quad (24)$$

Dengan c adalah panjang musiman dan K merupakan konstanta pembagi terhadap panjang musiman [11].

2.5 Time Series Regression

Regresi time series terdiri dari variabel dependen yang dipengaruhi oleh variabel independen yang telah diketahui. Berikut persamaan umum regresi (Draper & Smith, 1998) [14].

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + a_t \quad (25)$$

Koefisien parameter regresi pada persamaan di atas diperoleh dari perhitungan sebagai berikut

$$\beta = (X'X)^{-1}(X'Y) \quad (26)$$

dengan

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}, \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} \quad (27)$$

Model regresi time series dari data yang memiliki tren mengikuti persamaan berikut :

$$Y_t = \beta_1 t + \beta_2 D_{1,t} + \dots + \beta_j D_{it} + \beta_{j+1} t D_{1,t} + \dots + \beta_{j+k} t D_{it} + a_t \quad (28)$$

Dengan a_t adalah error yang telah memenuhi asumsi *white noise*. Pada model dengan pola musiman terdapat beberapa variabel dummy dengan asumsi bahwa ada M musiman (bulanan) per tahun dengan dummy musimannya ($M_{1,t}, M_{2,t}, \dots, M_{12,t}$) yang dapat dituliskan sebagai berikut.

$$Y_t = \alpha_1 M_{1,t} + \alpha_2 M_{2,t} + \alpha_3 M_{3,t} + \dots + \alpha_{12} M_{12,t} + a_t \quad (29)$$

Dengan dummy $M_{1,t}, M_{2,t}, \dots, M_{12,t}$ yang menyatakan pola musiman dengan periode dua belas bulan maka terdapat dua belas variabel dummy, satu variabel untuk satu bulan. Periode waktu berupa variasi kalender juga dapat dilakukan permodelan dengan menggunakan regresi. Model variasi kalender merupakan model time series yang digunakan untuk meramalkan data berdasarkan pola musiman dengan periode bervariasi. Model regresi untuk data dengan variasi kalender mengikuti persamaan sebagai berikut.

$$Y_t = \hat{\delta}_1 L_{1,t} + \hat{\delta}_2 L_{2,t} + \hat{\delta}_3 L_{3,t} + \hat{\delta}_4 L_{4,t} + a_t \quad (30)$$

Dan

$$Y_t = \hat{\gamma}_1 L_{1,t-1} + \hat{\gamma}_2 L_{2,t-1} + \hat{\gamma}_3 L_{3,t-1} + \hat{\gamma}_4 L_{4,t-1} + a_t \quad (31)$$

Dengan variabel dummy adalah $L_{1,t}, L_{2,t}, L_{3,t}$, dan $L_{4,t}$ adalah efek hari raya yang terjadi pada minggu ke-1 sampai minggu ke-4. Untuk variabel dummy ($L_{1,t-1}, L_{2,t-1}, L_{3,t-1}$ dan $L_{4,t-1}$) adalah efek persiapan sebulan sebelum hari raya. Sementara a_t adalah error yang telah memenuhi asumsi *white noise* [12].

2.6 Autoregressive Integrated Moving Average (ARIMA)

Model Box-Jenkin merupakan salah satu teknik peramalan model data time series yang hanya berdasarkan perilaku data variabel yang diamati. Model ini secara teknis dikenal sebagai model autoregressive integrated moving average (ARIMA). Model Box-Jenkins ini terdiri dari beberapa model yaitu : autoregressive (AR), moving average (MA), autoregressive-moving average (ARMA), dan autoregressive integrated moving average (ARIMA). Dengan menggunakan metode Box-Jenkins, maka akan mempermudah menentukan orde (p,d,q) pada model ARIMA sehingga dapat mengetahui model apa yang cocok untuk data time series yang akan kita gunakan [13].

Seasonal Autoregressive Integrated Moving Average (SARIMA) adalah pengembangan dari model ARIMA pada data runtun waktu yang memiliki pola musiman. Notasi SARIMA adalah: SARIMA $(p,d,q) (P,D,Q)^S$. Dengan p,d,q : bagian yang tidak musiman dari model; $(P,D,Q)^S$: bagian musiman dari model; S : jumlah periode per musim. Rumus umum SARIMA dapat dituliskan: $\Phi_p B^S \Theta_q (1-B)^d (1-B^S)^D Z_t = \theta_q(B) \Theta_q(B^S) a_t$. Dengan: $\Phi_p(B) = AR$ non-seasonal; $(1-B)^d = differencing$ non-seasonal; $(1-B^S)^D = differencing$ seasonal; $\theta_q(B) = MA$ non-seasonal; $\Theta_q(B^S) = MA$ seasonal. Untuk proses SARIMA yang umum, fungsi peramalan jangka panjang bergantung pada pola musiman individual paling akhir dalam data percobaan, dan peramalan terdiri atas penambahan periodik pola dan komponen modifikasi aditif (pertambahan) yang cenderung sama dengan nol secara eksponensial secara cepat karena horizon bertambah besar [13].

2.7. ARIMAX

Model ARIMA dengan tambahan variabel *dummy* disebut model ARIMAX. Variabel tersebut merupakan variabel *dummy* untuk efek variasi kalender saja atau variabel *dummy* untuk efek variasi kalender aja atau variabel *dummy* untuk efek variasi kalender dan efek deterministic trends [4]. Model untuk *dummy* efek variasi kalender adalah

$$Y_t = \beta_1 t + \beta_2 D_{1,t} + \dots + \beta_j D_{j,t} + \beta_{j+1} t D_{1,t} + \dots + \beta_{j+k} t D_{j,t} + \alpha_1 M_{1,t} + \alpha_2 M_{2,t} + \alpha_3 M_{3,t} + \dots + \alpha_{12} M_{12,t} + \hat{\delta}_1 L_{1,t} + \hat{\delta}_1 L_{1,t} + \hat{\delta}_1 L_{1,t} + \hat{\delta}_1 L_{1,t} + \hat{\gamma}_1 L_{1,t-1} + \hat{\gamma}_2 L_{2,t-1} + \hat{\gamma}_3 L_{3,t-1} + \hat{\gamma}_4 L_{4,t-1} + \frac{\theta_q(B)}{\Phi_p(B)} a_t \quad (32)$$

Keterangan :

- $t, D_{1,t}, \dots, tD_{1,t}, \dots$: variabel *dummy* trend
 $L_{1,t-1}, \dots, L_{4,t}$: variabel *dummy* untuk Idul Fitri
 $M_{1,t}, \dots, M_{12,t}$: variabel *dummy* bulan

2.8 Hybrid Model Time Series

Dalam model *hybrid* ini, data deret waktu diasumsikan terdiri dari komponen *nonlinear* dan *linear* atau linear dengan non linear [2]. Secara umum model data deret waktu yang merupakan kombinasi komponen *non linear* dan *linear* secara sistematis ditulis sebagai berikut

$$Y_t = N_t + L_t \quad (33)$$

Dengan Y_t merupakan data deret waktu pada waktu ke- t , N_t merupakan komponen *nonlinear* waktu ke- t dan L_t merupakan komponen linear waktu ke- t . tahap pertama dalam membentuk model *hybrid nonlinear – linier* atau sebaliknya adalah membentuk komponen nonlinear. Komponen nonlinear dimodelkan dengan menggunakan model non linier yang mana input data yang digunakan merupakan input data deret waktu (*sampel in*). Hasil prediction dari model non linier selanjutnya disebut sebagai komponen nonlinier. Kemudian residual dari model non linier diasumsikan tidak memenuhi asumsi *white noise*. Residual dari model komponen nonlinear menggunakan non linier dituliskan sebagai

$$\varepsilon_t = Z_t - \hat{N} \quad (34)$$

Dengan ε_t merupakan residual model *non linier* waktu ke- t dan \hat{L}_t merupakan komponen linear waktu ke- t hasil ramalan dari nilai prediksi model *non linier*. Setelah komponen nonlinear terbentuk, langkah selanjutnya adalah membentuk komponen linear. Komponen linear dimodelkan menggunakan model *linier* dengan data yang digunakan merupakan residual dari

model *non linier*. Selanjutnya untuk mendapatkan nilai hybrid menurut zhang(2003) [16] menggabungkan komponen non linier dengan linier yaitu model *non linier – linier* dapat dinyatakan sebagai

$$\hat{Z}_t = \hat{N}_t + \hat{L}_t \quad (35)$$

Dengan \hat{Z}_t merupakan model peramalan *hybrid nonlinier linier*, \hat{N}_t merupakan komponen nonlinear waktu ke- t hasil ramalan dari model *non linier* dan \hat{L}_t komponen linear waktu ke- t hasil peramalan $\hat{\epsilon}_t$ [15]

2.9 Mean Absolute Percentage Error

MAPE adalah ukuran ringkasan yang paling sering digunakan untuk mengevaluasi keakuratan prakiraan populasi. Ukuran ini memberikan petunjuk seberapa besar kesalahan peramalan dibandingkan dengan nilai sebenarnya. MAPE digunakan untuk mengukur ketepatan nilai dugaan model yang dinyatakan dalam bentuk rata-rata persentase absolut kesalahan dan lebih banyak digunakan untuk perbandingan pada data-data yang mempunyai skala interval waktu berbeda. Hasil peramalan dikatakan semakin akurat jika nilai MAPE semakin kecil [9]. Secara matematis dapat ditulis sebagai berikut

$$MAPE = \left(\frac{1}{n} \sum_{t=1}^n \left| \frac{y(t) - y'(t)}{y(t)} \right| \right) \times 100 \quad (36)$$

dengan y_t adalah nilai observasi pada waktu ke- t , y'_t adalah nilai peramalan pada waktu ke- t dan n adalah banyaknya observasi. Model yang baik memiliki nilai MAPE sesuai kriteria pada Tabel 1.

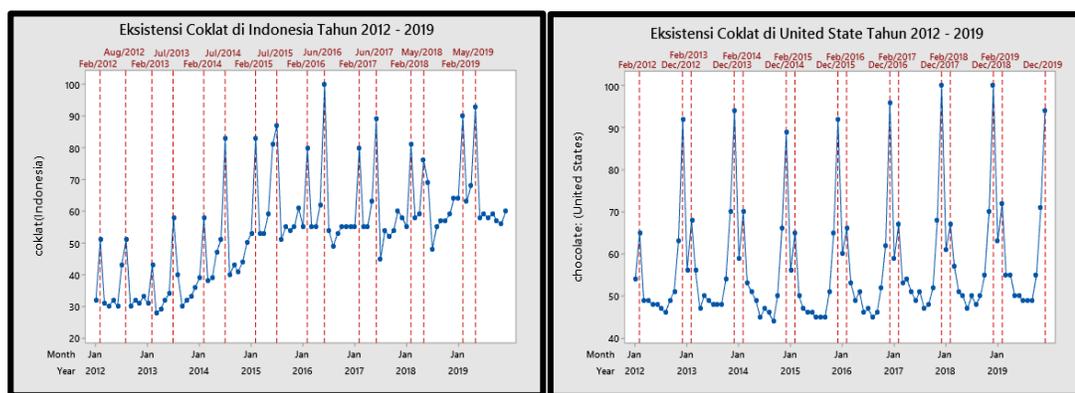
Tabel 1. Kriteria MAPE

Nilai MAPE (%)	Kriteria
< 10	Kemampuan peramalan sangat baik
10-20	Kemampuan peramalan baik
20-50	Kemampuan peramalan cukup
>50	Kemampuan peramalan buruk

3. HASIL DAN PEMBAHASAN

3.1 Perkembangan Data Cokelat Indonesia dan United States Tahun 2012 – 2019

Dalam penelitian ini, data yang digunakan adalah data bulanan pencarian akan eksistensi cokelat dengan kategori makanan dan minuman di Indonesia dan United States. Data yang digunakan adalah data sekunder yang mana berasal dari website *www.googletrend.com*. Dalam pembentukan model data dibagi menjadi dua bagian yaitu data training dan data testing. Data training kedua Negara tersebut dimulai dari bulan Januari 2012 sampai Desember 2017 dengan total data yang digunakan adalah 72 data. Untuk data testing didapat dari bulan Januari 2018 sampai Desember 2019 sebanyak 24 data.

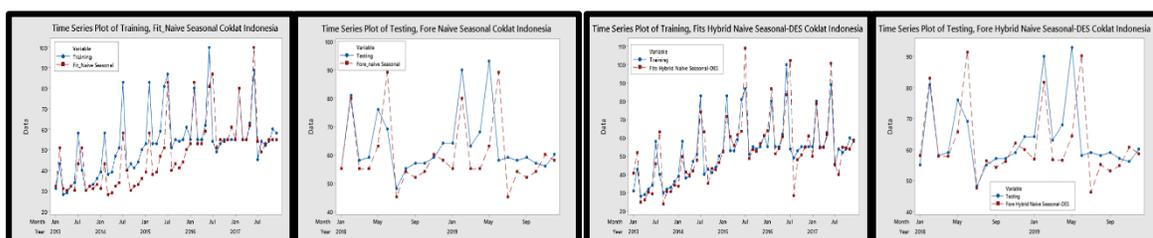


Gambar 1. Time series plot eksistensi cokelat pada negara Indonesia dan United States

Berdasarkan gambar 1, maka dapat dikatakan bahwa data cokelat antara Negara Indonesia memiliki kesamaan yaitu bersifat fluktuatif. Namun karakteristik fluktuatif dua Negara tersebut memiliki perbedaan. Terjadinya perbedaan tiap tahun atau lonjakan tiap tahun yang berbeda bulan terjadinya sepanjang tahun 2012-2019 pada Negara Indonesia karena konsumsi cokelat paling banyak mengikuti hari valentine yang musiman dan hari raya Idul Fitri. Namun berbeda dengan United States yang lonjakannya terjadi pada bulan yang sama di setiap tahunnya. Hal tersebut dikarenakan konsumsi cokelat tertinggi oleh masyarakat berada pada hari Valentine dan hari raya Natal yang mana terjadi pada tanggal dan bulan yang sama di sepanjang tahunnya.

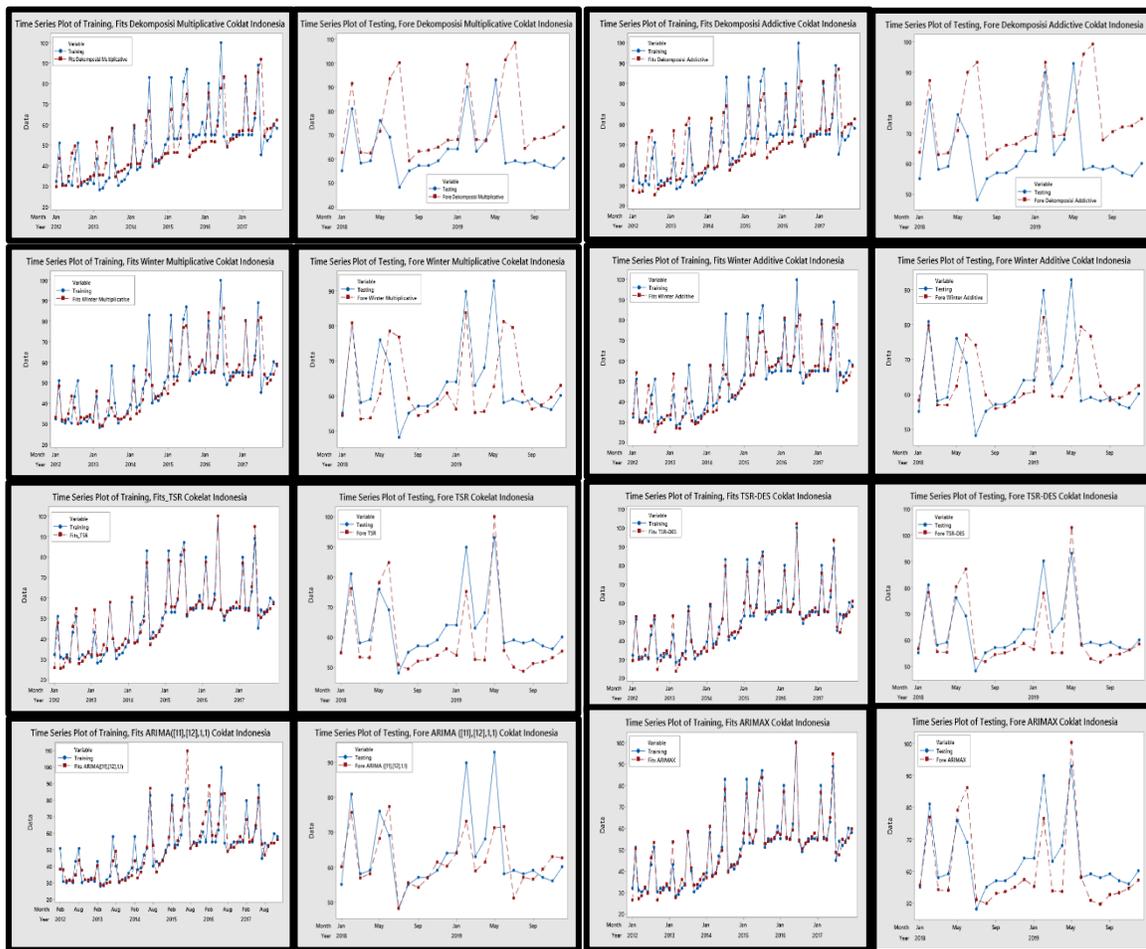
3.2 Peramalan Eksistensi Cokelat Di Indonesia

Dalam penelitian ini peramalan eksistensi cokelat di Negara Indonesia dilakukan dengan menggunakan beberapa metode yaitu *naïve*, model *hybrid naïve seasonal* dengan *double exponential smoothing* (DES), *double exponential smoothing*, *dekomposisi multiplicative*, *dekomposisi additive*, *holt winter multiplicative*, *holt winter additive*, *time series regression*, *hybrid time series regression* dengan *double exponential smoothing*, ARIMA((11],[12],1,1), dan ARIMAX (1,0,0)



Jurnal Matematika, Statistika & Komputasi

I Gusti Bagus Ngurah Diksa



Gambar 2. Time series plot data aktual dengan hasil peramalan dari model peramalan eksistensi coklat di Indonesia

Berdasarkan gambar 2 terlihat bahwa di setiap metode peramalan yang digunakan memiliki gap antara nilai aktual dengan peramalan yang berbeda beda. Untuk menentukan metode terbaik dalam model dan peramalan digunakan nilai MAPE in Sampel dan MAPE out Sampel dalam penentuannya. Semakin kecil nilai MAPE semakin bagus model tersebut dalam meramalkan data eksistensi coklat Indonesia di *google trend*.

Untuk rincian lebih lengkap tentang penentuan model peramalan eksistensi coklat di Indonesia yang terbaik berdasarkan nilai MAPE Out Sampel maka dapat dilihat pada Tabel 2.

Tabel 2. Rangkuman nilai MAPE pada model peramalan Eksistensi Cokelat di Indonesia

Metode	MAPE		MAPE Out Sampel			
	In Sampel	1	3	6	12	24
Naïve	14.563267	0	2.13566056	9.8795688	7.70421649	12.2239887
Seasonal						

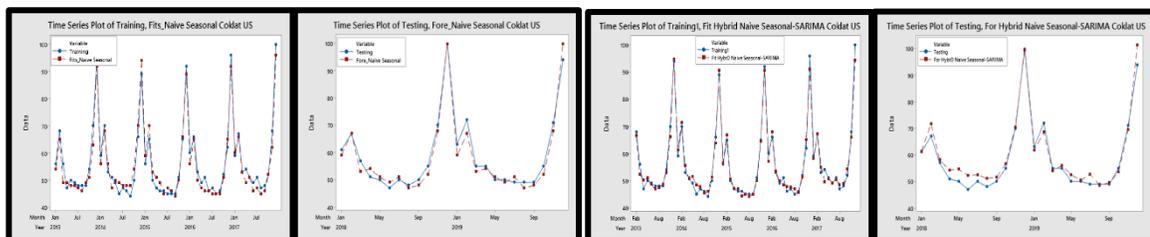
Jurnal Matematika, Statistika & Komputasi
I Gusti Bagus Ngurah Diksa

<i>Hybrid Naïve Seasonal-DES</i>	12.47860	5.4899455	2.72715246	9.4436857	6.52742872	11.0131415
<i>Dekomposisi Multiplikative</i>	12.73470	13.6922567	11.5072371	13.552660	19.5949191	22.0807183
<i>Dekomposisi Additive</i>	13.79441	15.7689571	10.6294516	12.748340	19.2726817	21.8041255
<i>Winter Multiplicative</i>	10.02263	1.28553890	3.13887072	8.8116228	11.3087966	13.1379684
<i>Winter Additive</i>	11.00158	5.92605270	3.14156876	7.1299436	9.79378557	11.2716967
<i>Time Series Regression</i>	5.404119	0.54569718	4.92256730	8.3934014	8.62801460	10.5927026
<i>TSR-DES</i>	5.315331	3.02961029	3.74195085	8.2087078	7.23309861	8.02301495
<i>ARIMA (([11],[12]),1,1)</i>	9.308358	9.15090909	5.94076705	6.9755138	4.87444427	7.52727867
<i>ARIMAX (1,0,0)</i>	4.736034	1.024247	4.310525	8.438568	8.080792	9.298933

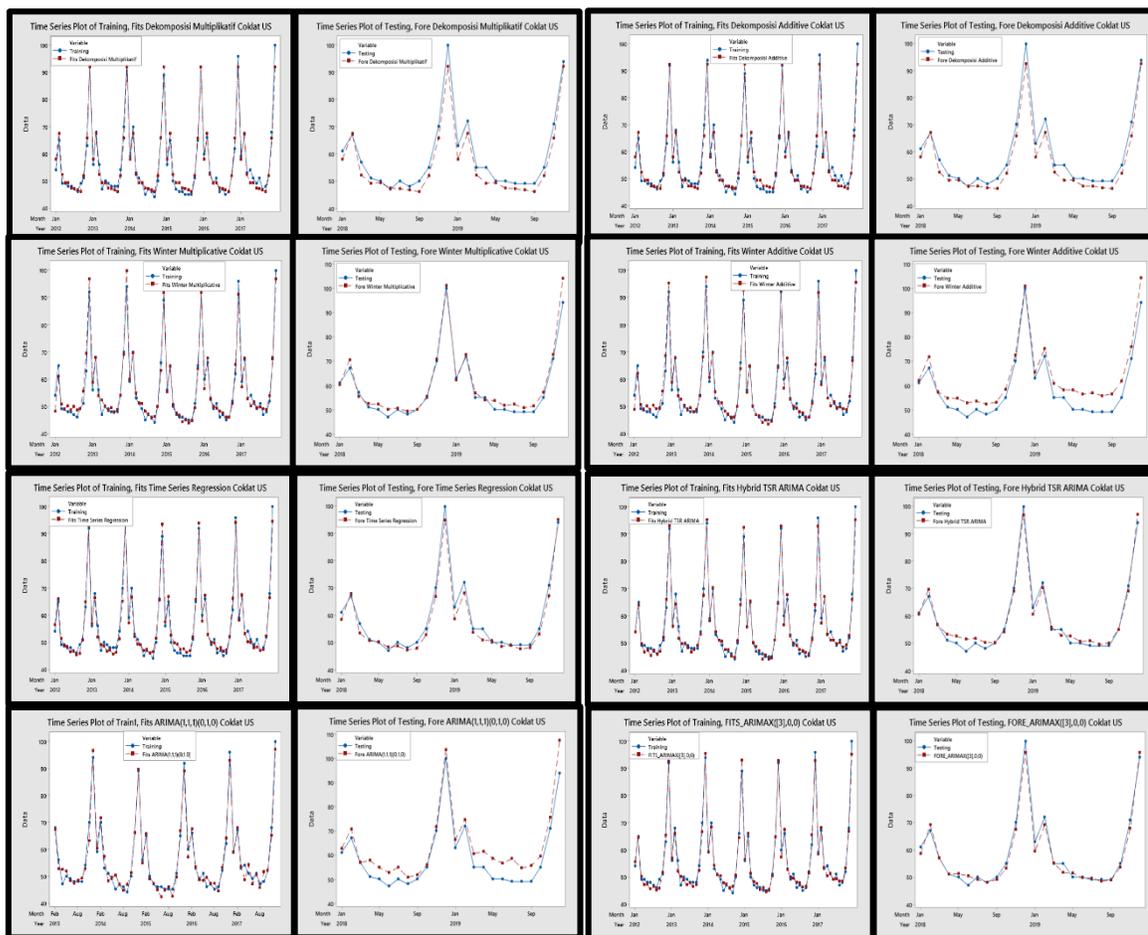
Berdasarkan tabel 1 dapat disimpulkan MAPE In Sampel yang merupakan cerminan fit nya suatu model paling bagus adalah model ARIMAX (1,0,0). Namun hal ini tidak berlaku pada hasil peramalan. Pada peramalan 1 bulan dan 3 bulan yang akan datang, model yang paling bagus digunakan adalah *Naïve Seasonal*. Bila digunakan peramalan 6, 12 dan 24 bulan yang akan datang, model yang paling bagus digunakan adalah ARIMA (([11],[12]),1,1). Oleh karena itu dapat dikatakan bahwa pergerakan eksistensi cokelat di Indonesia dipengaruhi oleh lag 11 dan 12 bulan. Eksisnya cokelat di Indonesia dapat dikatakan bersifat tahunan bila berdasarkan model yang telah didapatkan.

Peramalan Eksistensi Cokelat Di United States

Dalam penelitian ini peramalan eksistensi cokelat di Negara United States dilakukan dengan menggunakan beberapa metode yaitu *naïve*, model *hybrid naïve seasonal* dengan *double exponential smoothing* (DES), *double exponential smoothing*, *dekomposisi multiplikative*, *dekomposisi addictive*, *holt winter multiplicative*, *holt winter addictive*, *time series regression*, *hybrid time series regression* dengan ARIMA, SARIMA, ARIMAX. Untuk hasil peramalannya dapat dilihat pada gambar 3.



Jurnal Matematika, Statistika & Komputasi
I Gusti Bagus Ngurah Diksa



Gambar 3. Time series plot data aktual dengan hasil peramalan dari model peramalan eksistensi cokelat di United States

Untuk menentukan metode terbaik dalam model dan peramalan digunakan nilai MAPE in sampel dan MAPE out sampel dalam penentuannya. Semakin kecil nilai MAPE semakin bagus model tersebut dalam meramalkan data eksistensi cokelat United States di google trend. Untuk rincian lebih lengkap dapat dilihat pada tabel 2.

Tabel 3. Rangkuman nilai MAPE pada model peramalan Eksistensi Cokelat di United States

Metode	MAPE In Sampel	MAPE Out Sampel				
		1	3	6	12	24
<i>Naïve</i>	4.4843993	3.2786885	3.43207746	3.73898407	3.235743843	3.66017001
<i>Seasonal</i>						
<i>Hybrid Naïve</i>						
<i>Seasonal-SARIMA</i> $(2,1,0)(0,0,1)^{12}$	2.6586319	0.434067	3.138595	6.026587	4.588595	3.893592
<i>Dekomposisi</i>	3.0452791	4.9320600	4.66182494	3.26478825	4.621721398	5.05095384

<i>Multiplicative Dekomposisi Additive</i>	3.0413285	4.8895719	4.46830882	3.05456408	4.421525507	4.92436661
Winter Multiplicative	2.9426996	1.0706819	3.08570926	3.84350445	2.503704804	3.36161600
<i>Winter Additive</i>	2.9207854	1.5354242	3.02185611	6.34533207	5.858586922	8.27934376
Time Series Regression Hybrid Time Series	2.9444617	4.2486338	3.96850828	2.64400591	3.221921038	3.33715652
Regression- ARIMA (2,1,[10])	2.2714629	0.6892453	1.73209469	4.01024528	3.241218953	2.91406249
ARIMA (0,1,1)(0,1,0) ¹²	3.3646321	2.7583006	2.94982675	7.31836557	5.866590963	8.52066777
ARIMAX ([3],0,0)	2.4710098	3.9060655	2.48637291	2.87657196	2.696989788	2.63558575

Berdasarkan tabel 2 dapat disimpulkan MAPE in Sampel yang merupakan cerminan fit nya suatu model. Berdasarkan nilai MAPE in sampel maka model yang paling bagus adalah model *Hybrid Time Series Regression- ARIMA(2,1,[10])*. Namun hal ini tidak berlaku pada hasil peramalan. Pada peramalan 1 bulan model yang paling bagus digunakan adalah model *Hybrid Naïve Seasonal-SARIMA (2,1,0)(0,0,1)¹²* Peramalan untuk tiga bulan yang akan datang, model yang paling bagus digunakan adalah *Hybrid Time Series Regression- ARIMA(2,1,[10])*. Bila digunakan peramalan enam bulan yang akan datang maka model yang paling bagus adalah model peramalan *Time Series Regression*. Jika dilakukan peramalan untuk 12 bulan yang akan datang maka model yang terbaik adalah *Winter Multiplicative*. Untuk yang terakhir yaitu peramalan 24 bulan yang akan datang, model yang paling baik digunakan model peramalan *ARIMAX ([3],0,0)*.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan yang di ulas maka terdapat beberapa kesimpulan yang dapat diambil :

1. Data coklat dari google trend dalam kategori makanan dan minuman yang mencakup wilayah Indonesia dan United States memiliki persamaan yaitu sama sama berfluktuasi. Namun ke dua data tersebut memiliki perbedaan karakter, Indonesia dengan efek *calendar variation* dan United States dengan efek musiman atau *seasonal*.
2. Pada Indonesia, MAPE In Sampel yang merupakan cerminan fit nya suatu model paling bagus adalah model ARIMAX (1,0,0). Namun hal ini tidak berlaku pada hasil peramalan. Pada peramalan 1 bulan dan 3 bulan yang akan datang, model yang paling bagus digunakan adalah *Naïve Seasonal*. Bila digunakan peramalan 6, 12 dan 24 bulan yang akan datang, model yang paling bagus digunakan dalam peramalan eksistensi coklat di Indonesia adalah ARIMA (([11],[12]),1,1) yang berarti eksistensi coklat di Indonesia dipengaruhi oleh lag 11 dan 12 bulan.
3. Pada United States, bila dilihat dari sisi MAPE in sampel maka model yang paling bagus adalah model *Hybrid Time Series Regression- ARIMA(2,1,[10])*. Namun hal ini tidak

berlaku pada hasil peramalan. Pada peramalan 1 bulan model yang paling bagus digunakan adalah model *Hybrid Naïve Seasonal-SARIMA (2,1,0)(0,0,1)¹²* Peramalan untuk tiga bulan yang akan datang, model yang paling bagus digunakan adalah *Hybrid Time Series Regression- ARIMA(2,1,[10])*. Bila digunakan peramalan enam bulan yang akan datang maka model yang paling bagus adalah model peramalan *Time Series Regression*. Jika dilakukan peramalan untuk 12 bulan yang akan datang maka model yang terbaik adalah *Winter Multiplicative*. Untuk yang terakhir yaitu peramalan 24 bulan yang akan datang, model yang paling baik digunakan model peramalan *ARIMAX ([3],0,0)*.

DAFTAR PUSTAKA

- [1]. Chrisdayanti, B., 2015. *Peramalan kandungan Particulate Matter (PM10) dalam udara ambien kota Surabaya menggunakan Double Seasonal ARIMA (DSARIMA)* (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- [2]. Cryer, J. D., & Chan, K. S., 2008. *Time series analysis: with applications in R* (Vol. 2). New York: Springer.
- [3]. Halim, J. F., 2016. Implementasi Konsep Addiction in Catchiness pada Perancangan Interior “Rumah Cokelat” di Surabaya. *Intra*, 4(2), 339-351.
- [4]. Hanim, Y. M., 2015. *Penerapan Regresi Time Series dan ARIMAX untuk Peramalan Inflow dan Outflow Uang Kartal di Jawa Timur, DKI Jakarta dan Nasional* (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- [5]. Huang, M. Y., Rojas, R. R., & Convery, P. D., 2020. Forecasting stock market movements using google trend searches. *Empirical Economics*, 59(6), 2821-2839.
- [6]. Jun, S. P., Yoo, H. S., & Choi, S., 2018. Ten years of research change using Google Trends: From the perspective of big data utilizations and applications. *Technological forecasting and social change*, 130, 69-87.
- [7]. Lim, S., Tanudjaja, B. B., & Salamoon, D. K., 2015. Perancangan Kemasan Produk "Choco Molten" Online Bakery Instacake Sebagai Media Promosi. *Jurnal DKV Adiwarna*, 1(6), 12.
- [8]. Lorie, M., & Hagen, E. 2007. Predicting refill of the Patuxent and Occoquan reservoirs: A new iterative tool based on ARIMA. *ICPRB Rep*, 07-4.
- [9]. Makridakis, S., Wheelwright, S. C., & McGee, V. E., 2003. *Metode dan Aplikasi Peramalan*, Jilid 1 Edisi Revisi (terj.). *Binarupa Aksara. Jakarta*.
- [10]. Mun, M. K., & Chong, C. W., 2018. Forecasting movie demand using total and split exponential smoothing. *Jurnal Ekonomi Malaysia*, 52(2), 81-94.
- [11]. Rosalina, E., Sugiarto, S. and Gamal, M. D. H., 2018. Metode Peramalan Holt-Winter untuk Memprediksi Jumlah Pengunjung Perpustakaan Universitas Riau. *Riau: Universitas Riau*.
- [12]. Sadew, E. 2013. Perbandingan Beberapa Metode Time Series pada Peramalan Jumlah Kunjungan Wisatawan Mancanegara: Studi Kasus di Kabupaten Karimun, Provinsi Kepulauan Riau. *Researchgate DOI*, 10.
- [13]. Wirawan, I. G. P. N., & Utama, M. S., 2014. Model Box-Jenkins dalam Rangka Peramalan Produk Domestik Regional Bruto Provinsi Bali. *Buletin Studi Ekonomi*, 44257.
- [14]. Yuni, S., Talakua, M. W., & Lesnussa, Y. A., 2015. Peramalan Jumlah Pengunjung Perpustakaan Universitas Pattimura Ambon Menggunakan Metode Dekomposisi. *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, 9(1), 41-50.
- [15]. Zhang, G. P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.