

## PERBANDINGAN METODE PEMULUSAN EKSPONENSIAL GANDA HOLT DAN METODE ARIMA BOX-JENKINS

Agustini Tripena<sup>1</sup>, Yosita Lianawati<sup>2</sup>, Antonius Ary Setyawan<sup>3</sup>, Carolina Ety Widjayanti<sup>4</sup>

<sup>1</sup>Universitas Jenderal Soedirman

<sup>2</sup>Sekolah Tinggi Ilmu Komputer (STIKOM) Yos Sudarso Purwokerto

<sup>3</sup>Sekolah Tinggi Ilmu Komputer (STIKOM) Yos Sudarso Purwokerto

<sup>4</sup>Sekolah Tinggi Ilmu Komputer (STIKOM) Yos Sudarso Purwokerto

[agustinitripena17@gmail.com](mailto:agustinitripena17@gmail.com), [yosita.lianawati@stikomvos.ac.id](mailto:yosita.lianawati@stikomvos.ac.id), [arysetpr@stikomvos.ac.id](mailto:arysetpr@stikomvos.ac.id),  
[carolinaety@stikomvos.ac.id](mailto:carolinaety@stikomvos.ac.id)

### Abstrak

Penduduk merupakan faktor dominan dalam pelaksanaan pembangunan serta penduduk menjadi pelaksana tujuan pembangunan. Pemerintah perlu meramalkan jumlah penduduk untuk menentukan tujuan kebijakan ekonomi. Peramalan jumlah penduduk merupakan hal yang penting bagi pemerintah untuk menentukan arah kebijakan perekonomian bangsa. Data populasi yang pada dasarnya merupakan data deret waktu dapat dianalisis dengan metode rata-rata bergerak, metode pemulusan eksponensial, dan metode ARIMA. Metode pemulusan eksponensial yang digunakan adalah pemulusan eksponensial ganda Holt karena data jumlah penduduk berpola trend linier. Penelitian difokuskan pada peramalan populasi dengan menggunakan pemulusan eksponensial ganda Holt dan ARIMA. Pemulusan eksponensial ganda Holt menggunakan nilai parameter 0,3 dan 0,99. Nilai parameter yang digunakan adalah 0,05 dan 0,9. Pemulusan eksponensial ganda Holt, diperoleh nilai jumlah penduduk untuk tahun 2020 sebesar 1.614.339 jiwa, tahun 2021 sebesar 1.626.379 jiwa, tahun 2022 sebesar 1.641.519 jiwa. ARIMA hasil nilai ramalan jumlah penduduk tahun 2020 sebanyak 1.618.258 jiwa, tahun 2021 sebanyak 1.628.259 jiwa, tahun 2022 sebanyak 1.643.417 jiwa. Nilai  $\alpha=0,99$  dan  $\gamma=0,05$  memberikan MSE terkecil pada nilai 15175922,58. Model ARIMA (1, 1, 1) nilai MSE terkecil 17373940. Setelah dibandingkan, metode Holt pemulusan eksponensial ganda sama efektifnya dengan metode ARIMA untuk meramalkan jumlah penduduk di Kabupaten Banyumas.

**Kata kunci:** Pemulusan eksponensial, ganda Holt, penduduk, deret waktu, ARIMA, MSE.

### 1. PENDAHULUAN

Pengendalian kuantitas penduduk diarahkan pada keserasian, keselarasan, dan keseimbangan antara jumlah, struktur dan komposisi, pertumbuhan dan persebaran penduduk yang ideal sesuai daya dukung dan daya tampung serta kondisi perkembangan sosial ekonomi dan budaya. Peningkatan kualitas penduduk dilakukan melalui perbaikan kondisi penduduk dengan pengadaan sarana, fasilitas serta kesempatan untuk memperoleh pendidikan, pelayanan kesehatan, dan perluasan kesempatan kerja. Sedangkan pengarahannya mobilitas penduduk lebih terfokus pada persebaran penduduk yang optimal atau merata, sehingga memberikan peluang terciptanya sentra-sentra kegiatan ekonomi baru yang pada gilirannya akan meningkatkan kesempatan kerja.

Peramalan (*forecasting*) jumlah penduduk di masa yang akan datang menjadi hal yang penting, khususnya bagi pemerintah [1,4]. Pemerintah perlu mengetahui ramalan jumlah penduduk di masa yang akan datang guna membantu dan menunjang kegiatan sosial ekonomi di Indonesia. Data jumlah penduduk pada dasarnya merupakan data deret waktu (*time series*) [2,3]. Untuk menganalisis data deret waktu tersebut, ada beberapa metode yang dapat digunakan, antara lain metode rata-rata bergerak (*moving average*), metode pemulusan eksponensial (*exponential smoothing*) salah satunya

pemulusan eksponensial ganda dari Holt, dan metode ARIMA (*autoregressive integrated moving average*) [7,9,12].

## 2. DASAR TEORI

### 2.1. Trend

Trend adalah suatu gerakan yang menunjukkan arah perkembangan secara umum (kecenderungan secara menaik atau menurun dalam jangka yang panjang). Trend jangka panjang dari data deret waktu biasanya mengikuti pola-pola tertentu. Beberapa dari data tersebut bergerak ke atas, data lainnya menurun, data lainnya tetap pada tempat yang sama selama satu periode tertentu. Trend seringkali diperkirakan mempunyai garis lurus. Persamaan untuk menggambarkan garis trend adalah sebagai berikut [5,9,11].

$$\hat{Y} = a + bt \quad (1)$$

$$b = \frac{\sum tY - \frac{(\sum Y)(\sum t)}{n}}{\sum t^2 - \frac{(\sum t)^2}{n}} \quad (2)$$

$$a = \frac{\sum Y}{n} - b\left(\frac{\sum t}{n}\right) \quad (3)$$

$\hat{Y}$  = nilai proyeksi dari variabel Y pada nilai tertentu; a = nilai perpotongan (intersep) dari Y. Intersep ini merupakan nilai Y ketika t=0; b = kemiringan rata-rata dalam  $\hat{Y}$  untuk setiap perubahan dari satu unit t.; t = nilai waktu yang dipilih

### 2.2. Metode Pemulusan Eksponensial Ganda Holt

Peramalan dengan pemulusan eksponen sederhana dilakukan jika data tidak mengandung komponen trend dan musiman, sedangkan jika mengandung komponen trend tetapi tidak mengandung komponen musiman, maka harus digunakan metode Holt [3,11]. Pemulusan eksponensial ganda dari Holt adalah metode pemulusan eksponensial dengan dua kali pemulusan yang dikemukakan oleh Holt [5,9]. Pada prinsipnya, metode pemulusan eksponensial dari Holt serupa dengan metode pemulusan eksponensial dari Brown kecuali bahwa Holt tidak menggunakan rumus pemulusan berganda secara langsung. Holt memuluskan nilai trend dengan parameter yang digunakan pada deret yang asli. Ramalan dari pemulusan eksponensial ganda dari Holt untuk periode m ke depan adalah [5,12,15]:

$$S_t = \alpha Z_t + (1 - \alpha)(S_{t-1} + b_{t-1}) \quad (4)$$

$$b_t = \gamma(S_t - S_{t-1}) + (1 - \gamma) b_{t-1} \quad (5)$$

$$F_{t+m} = S_t + b_t m \quad (6)$$

dimana  $S_t$  = nilai pemulusan eksponensial pada periode ke t,  $b_t$  = nilai trend pada periode ke t. Persamaan (4) menyesuaikan  $S_t$  secara langsung untuk trend periode sebelumnya, yaitu  $b_{t-1}$ , dengan menambahkan nilai pemulusan yang terakhir yaitu  $S_{t-1}$ . Hal ini membantu untuk menghilangkan kelambatan dan menempatkan ke dasar perkiraan nilai data saat ini. [11,12]. Inisialisasi adalah penentuan nilai awal yang digunakan dalam peramalan pemulusan eksponensial. Proses inisialisasi untuk pemulusan eksponensial linier Holt memerlukan dua taksiran yaitu  $S_1$  dan  $b_1$ . Taksiran  $S_1$  dan  $b_1$  dapat mengambil dari persamaan garis trend. Selain itu, taksiran-taksiran tersebut dapat diperoleh dari :

- i. Untuk inisiasi  $S_1$ , ambil  $S_1 = Z_1$
- ii.  $b_1 = Z_2 - Z_1$  (7)

$$\text{iii. } b_1 = \frac{(Z_2 - Z_1) + (Z_3 - Z_2) + (Z_4 - Z_3)}{3} \quad (8)$$

$b_1$  = taksiran kemiringan (slope) setelah data diplot.

### 2.3. Perhitungan Kesalahan Peramalan

Untuk mengukur error (kesalahan) peramalan biasanya digunakan *Mean Absolute Error (MAE)* dan *Mean Square Error (MSE)*. *Mean Absolute Error (MAE)* adalah rata-rata nilai absolute error dari kesalahan meramal. Rumusnya [6,9,12]:

$$MAE = \frac{\sum |Z_t - F_t|}{n} \quad (9)$$

*Mean Square Error (MSE)* adalah rata-rata dari kesalahan yang dikuadratkan. Rumusnya [6,9,12]:

$$MSE = \frac{\sum (Z_t - F_t)^2}{n} \quad (10)$$

dimana,  $Z_t$  = data sebenarnya ;  $F_t$  = data ramalan dihitung dari model yang digunakan pada waktu tahun  $t$ ;  $n$  = banyaknya data. Kriteria : metode yang dipilih adalah metode yang mempunyai *MAE* dan *MSE* paling kecil

### 2.4. Metode ARIMA Box-Jenkins

Konsep dasar analisis deret waktu yang digunakan dalam peramalan dengan metode ARIMA [4,8] akan dibahas sebagai berikut.

#### 2.4.1. Kestasioneran

Deret waktu dikatakan stasioner jika tidak ada perubahan kecenderungan dalam rata-rata dan perubahan variansi. Dengan kata lain, deret waktu yang stasioner adalah relatif tidak terjadi kenaikan ataupun penurunan nilai secara tajam pada data (fluktuasi data berada pada sekitar nilai rata-rata yang konstan) [6,7,8,12].

#### 2.4.2 Differencing

Metode differencing adalah membentuk suatu data baru yang diperoleh dengan cara mengurangi nilai pengamatan pada waktu  $t$  dengan nilai pengamatan pada waktu sebelumnya. *Differencing* digunakan dalam peramalan jika data belum stasioner. Sebagai contoh untuk nilai data sebenarnya  $Z_1, Z_2, \dots, Z_n$  adalah data yang tidak stasioner maka untuk pembedaan pertama (*first differences*) adalah [6,7,8,9]:

$$W_t = Z_t - Z_{t-1} \text{ atau } W_t = (1 - B) Z_t \quad (11)$$

Jika pembedaan pertama belum menghasilkan runtun waktu yang stasioner, maka dilakukan pembedaan kedua. Pembedaan kedua dari runtun waktu  $Z_1, Z_2, \dots, Z_n$  didefinisikan sebagai :

$$\begin{aligned} Z_{t-1} &= (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) \\ &= Z_t - 2Z_{t-1} + Z_{t-2}, \text{ dengan } t = 3, 4, \dots, n \end{aligned} \quad (12)$$

Pembedaan kedua adalah pembedaan pertama dari hasil pengurangan pembedaan pertama. Pembedaan kedua biasanya menghasilkan data yang stasioner. Jika pembedaan kedua belum menghasilkan data yang stasioner, maka diperlukan transformasi lain, misalnya transformasi data musiman

### 2.5. Fungsi Autokorelasi (FAK)

Menurut Aswi dan Sukarna, koefisien autokorelasi adalah suatu fungsi yang menunjukkan besarnya korelasi (hubungan linier) antara pengamatan pada waktu ke- $t$  (dinotasikan dengan  $Z_t$ ) dengan pengamatan pada waktu-waktu yang sebelumnya (dinotasikan dengan  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k}$ ). Untuk suatu data deret waktu  $Z_1, Z_2, \dots, Z_n$  maka nilai fungsi autokorelasinya adalah sebagai berikut [3,10,14]:

$$r_k = \frac{\sum_{t=b}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=b}^n (Z_t - \bar{Z})^2} \tag{13}$$

dimana  $\bar{Z} = \sum_{t=b}^n Z_t$

**2.5.1. Fungsi Autokorelasi Parsial (FAKP)**

Menurut [3,12], fungsi autokorelasi parsial adalah suatu fungsi yang menunjukkan besarnya korelasi parsial antara pengamatan pada waktu ke-t (dinotasikan dengan  $Z_t$ ) dengan pengamatan pada waktu-waktu sebelumnya (dinotasikan dengan  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k}$ ). Autokorelasi parsial digunakan untuk mengukur tingkat keeratan (*Association*) antara  $Z_t$  dan  $Z_{t-k}$ , apabila pengaruh dari lag waktu (time lag) 1, 2, 3, ..., k-1 dianggap terpisah.

$$r_{kk} = \begin{cases} r_1 \\ r_k - \frac{\sum_{j=1}^{k-1} r_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} r_j} \end{cases}$$

(14)

Dengan  $r_{kj} = r_{k-1,j} - r_{kk} r_{k-1,k-j}$  untuk  $j=1,2,\dots,k-1$ .

**2.5.2. White Noise Process**

Suatu proses  $\{a_t\}$  dinamakan *proses white noise* (proses yang bebas dan identik) jika bentuk peubah acak yang berurutan tidak saling berkorelasi dan mengikuti distribusi tertentu. Rata-rata  $E(a_t) = \mu_a$  dari proses ini diasumsikan bernilai nol dan mempunyai variansi yang konstan yaitu  $var(a_t) = \sigma_a^2$ . Sehingga dapat disimpulkan suatu deret waktu disebut *white noise process* jika rata-rata dan variansinya konstan dan saling bebas[4,7,10].

**2.6. Peramalan dengan Autoregressive Integrated Moving Average (ARIMA)**

Metode ARIMA merupakan metode yang dikembangkan oleh George Box dan Gwilym Jenkins sehingga nama mereka sering disinonimkan dengan proses ARIMA yang diterapkan untuk analisis data dan peramalan data deret waktu [6,7,10,12].

**2.6.1. Model AR (Autoregressive)**

[6,7,12] menyatakan bahwa model AR adalah model yang menggambarkan nilai deret waktu saat ini (dilambangkan dengan  $Z_t$ ) sebagai fungsi dari nilai deret waktu yang lalu (dilambangkan dengan  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$ ). Secara umum model AR mempunyai bentuk sebagai berikut:

$$Z_t = \delta + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t \tag{15}$$

dimana :  $Z_t$  = nilai variabel *dependent* pada waktu  $t$ ;  $\delta$  = intersep atau nilai konstan

$\phi_1, \phi_2, \dots, \phi_p$  = parameter model (*Autoregressive*) AR;  $Z_{t-p}$  = variabel *dependent* yang dalam hal ini merupakan lag (beda waktu) dari variabel *dependent* pada satu periode sebelumnya

$a_t$  = residual pada waktu  $t$

**2.6.2. Model MA (Moving Average)**

[6,7,8] menyatakan bahwa model MA merupakan model yang menggambarkan ketergantungan peubah  $Z_t$  terhadap nilai-nilai galat pada waktu sebelumnya yang berurutan. Secara umum model MA mempunyai bentuk sebagai berikut:

$$Z_t = \delta + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (16)$$

dimana :  $Z_t$  = nilai variabel *dependent* pada waktu  $t$ ;  $\delta$  = intersep atau nilai konstan

$a_{t-1}, a_{t-2}, \dots, a_{t-q}$  = nilai residual sebelumnya;  $a_t$  = residual pada waktu  $t$ ;  $\theta_1, \theta_2, \dots, \theta_q$  = parameter model MA . Perbedaan model MA dengan model AR terletak pada jenis variabel *independent*. Jika variabel pada model AR adalah nilai sebelumnya dari variabel *independent* maka pada model MA yang menjadi variabel *independent* adalah nilai residual pada periode sebelumnya.

### 2.7. Model ARIMA

[6, model AR dan MA yang dikombinasikan untuk menghasilkan model ARIMA mempunyai bentuk umum sebagai berikut:

$$Z_t = \delta + \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + a_t + \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (17)$$

Dengan penggabungan ini diharapkan model ARIMA dapat mengakomodasi pola data yang tidak diidentifikasi secara sendiri-sendiri oleh model MA atau AR. Orde dari model ARIMA ditentukan oleh jumlah periode variabel *independent* baik dari nilai sebelumnya dari variabel *independent* maupun nilai residual periode sebelumnya [6.7,12,14]

### 2.8. Penaksiran Parameter dan Pemeriksaan Diagnostik

Parameter dari dugaan model awal ARIMA hasil indentifikasi ditaksir sehingga didapatkan besaran koefisien model. Metode *Least Squares* adalah suatu metode yang dilakukan dengan cara mencari nilai parameter yang meminimumkan jumlah kuadrat kesalahan (selisih antara nilai aktual dan ramalan). Atau dapat dituliskan, jika diberikan data deret waktu yang telah stasioner yakni  $Z_b, Z_{b+1}, \dots, Z_n$  maka metode ini mencari nilai parameter dengan meminimumkan  $\sum_{t=b}^n a_t^2$ .

Pemeriksaan diagnostik (*diagnostic checking*) dapat dibagi ke dalam dua bagian, yaitu uji kesignifikanan parameter dan uji kesesuaian model (meliputi uji asumsi *white noise* dan distribusi normal) [9,12,13].

#### 2.8.1. Uji kesignifikanan parameter

Model ARIMA yang baik dapat menggambarkan suatu kejadian adalah model yang salah satunya menunjukkan bahwa penaksiran parameternya signifikan berbeda dengan nol. Secara umum, misalkan  $\theta$  adalah suatu parameter pada model ARIMA Box-Jenkins dan  $\hat{\theta}$  adalah nilai taksiran dari parameter tersebut, serta  $SE(\hat{\theta})$  adalah standar error dari nilai taksiran  $\hat{\theta}$  dimana [7,12,17] :

$$SE(\hat{\theta}) = \sqrt{\frac{SSE}{n-n_p}} = \sqrt{\frac{\sum_{t=1}^n (Z_t - \bar{Z}_t)^2}{n-n_p}} \quad (18)$$

Uji kesignifikanan parameter dapat dilakukan dengan tahapan sebagai berikut :

Hipotesis :  $H_0 : \hat{\theta} = 0$ ;  $H_1 : \hat{\theta} \neq 0$

Statistik uji 
$$t = \frac{\hat{\theta}}{SE(\hat{\theta})} \quad (19)$$

Daerah penolakan : Tolak  $H_0$  jika  $|t| > t_{\alpha/2}$ ;  $df = n - n_p$ ,  $n_p$  = banyaknya parameter dan  $n$  = banyaknya data. Atau dengan menggunakan *p-value*, yakni tolak  $H_0$  jika nilai- $p < \alpha$  .

#### 2.8.2. Uji kesesuaian model

Uji kesesuaian model meliputi kecukupan model (uji apakah sisanya white noise) dan uji asumsi distribusi normal.

i. Uji sisa *white noise*

Secara ringkas, uji sisa white noise dapat dituliskan [7,9,12,17]:

Hipotesis :  $H_0$  : model sudah memenuhi syarat cukup (sisa memenuhi syarat *white noise*)

$H_1$  : model belum memenuhi syarat cukup (sisa memenuhi syarat *white noise*)

Statistik uji, yaitu statistik uji Ljung –Box atau Box-Pierce Modified :

$$Q^* = n(n + 2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k} \quad (20)$$

dimana :  $n$  = jumlah sampel;  $k$  = lag (selisih waktu);  $K$  = lag maksimum;  $\hat{\rho}_k$  = nilai autokorelasi time lag  $k$  dan  $\hat{\rho}_k^2$  diperoleh dari

$$\hat{\rho}_k^2 = \frac{\sum_{t=1}^{n-k} (\hat{a}_t - \bar{\hat{a}})(\hat{a}_{t+k} - \bar{\hat{a}})}{\sum_{t=1}^n (\hat{a}_t - \bar{\hat{a}})^2} \quad (21)$$

Daerah penolakan : Tolak  $H_0$  jika  $Q^* > \chi_{\alpha;df=K-m}^2$ .  $K$  berarti pada lag  $K$  dan  $m$  adalah jumlah parameter yang ditaksir dalam model.

ii. Uji Asumsi Distribusi Normal

Uji asumsi ini bertujuan untuk mengetahui apakah data telah memenuhi asumsi kenormalan atau belum. Salah satu cara yang dapat ditempuh untuk melakukan uji asumsi kenormalan ini adalah uji Kolmogorof Smirnov dengan menggunakan pedoman pengambilan keputusan sebagai berikut. Jika nilai-p < 0,05, data tidak berasal dari populasi yang berdistribusi normal. Jika nilai-p  $\geq$  0,05, data berasal dari populasi yang berdistribusi normal.

### 2.9. Ketepatan Metode Peramalan

Pada metode peramalan, ketepatan dipandang sebagai kriteria penolakan untuk memilih suatu metode peramalan. Ukuran ketepatan peramalan yang akan digunakan dalam pembahasan ini adalah *MSE (Mean Square Error)*. Menurut [4,12], *MSE (Mean Square Error)* didefinisikan sebagai:

$$MSE = \frac{\sum_{t=1}^n (Z_t - \hat{Z}_t)^2}{n} \quad (22)$$

Model terbaik memiliki nilai *MSE* yang lebih kecil jika dibandingkan dengan model lainnya

### 2.10. Uji Beda Dua Rata-Rata Model Pemulusan Eksponensial Ganda Holt dan Model ARIMA Box-Jenkins

Pengujian ini bertujuan untuk mengevaluasi apakah terdapat perbedaan hasil peramalan yang signifikan antara kedua model yang digunakan. Ada dua kemungkinan yang dapat terjadi pada uji beda dua rata-rata ini. Pertama, terdapat perbedaan yang nyata (signifikan). Model yang hasil ramalannya mendekati data aktual akan direkomendasikan untuk digunakan dalam peramalan jumlah penduduk di masa yang akan datang. Kedua, tidak terdapat perbedaan yang nyata [11,13,15,16]. Uji beda dua rata-rata dapat dilakukan dengan uji  $t$ . uji  $t$  digunakan jika kedua sampel *independent*, ukuran sampel kecil ( $n \leq 30$ ) dan variansi dari kedua populasi tidak diketahui. Asumsi dalam uji ini adalah kedua populasi mendekati distribusi normal. Prosedur pengujiannya adalah sebagai berikut.

i. Hipotesis :  $H_0 : \mu_1 - \mu_2 = D_0 ; H_1 : \mu_1 - \mu_2 \neq D_0$

ii. Statistik Uji: 
$$t = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{\sqrt{s_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} \quad (23)$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (24)$$

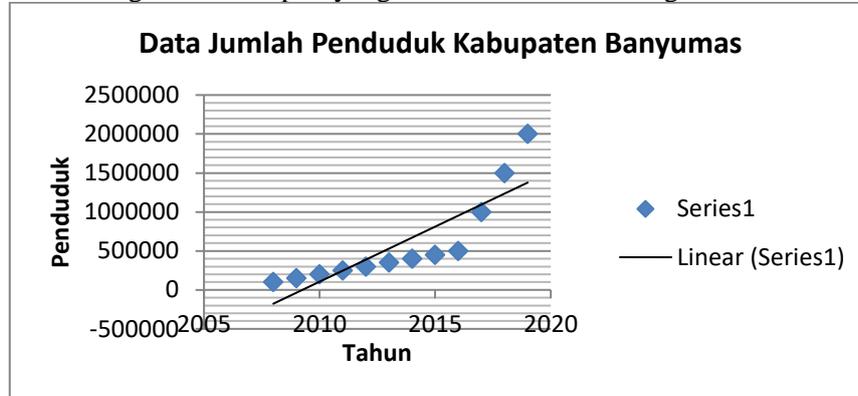
iii. Daerah penolakan,  $H_0$  ditolak jika  $|t| > t_{\frac{\alpha}{2}, n_1 + n_2 - 2}$

Keterangan :  $\bar{y}_1$  = rata-rata dari model pemulusan eksponensial ganda Holt;  $\bar{y}_2$  = rata-rata dari model ARIMA Box-Jenkins;  $s_1^2$  = variansi dari model pemulusan eksponensial Ganda Holt  $s_2^2$  = variansi dari model ARIMA Box-Jenkins

**3. Pembahasan**

**3.1. Diagram Scatter Plot**

Langkah pertama dalam menentukan nilai trend adalah dengan membuat diagram scatter plot. Diagram scatter plot yang dihasilkan adalah sebagai berikut.



Gambar 1. Scatter diagram untuk data penduduk Kab. Banyumas

Gambar 1 menggunakan *scatter* diagram dari akan dicari garis yang mendekati titik di dalam diagram. Dari *scatter* diagram dapat disimpulkan bahwa data banyaknya penduduk naik dari tahun ke tahun.

**3.2. Persamaan Garis Trend**

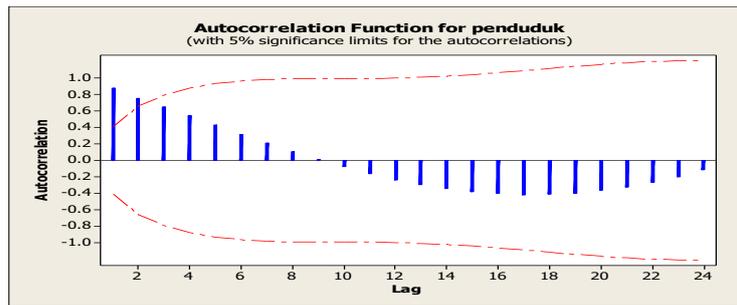
Perhitungan diperoleh persamaan garisnya :  $\hat{Y} = 1288988,70 + 12793,46t$ ; Berdasarkan Gambar 1 dan dari persamaan garis  $\hat{Y}$ , data banyaknya penduduk di Kabupaten Banyumas cenderung berpola trend linier, sehingga metode Pemulusan Eksponensial dapat digunakan untuk menghitung pemulusan data jumlah penduduk di Kabupaten Banyumas. Hasil pemulusan dan taksiran nilai trendnya dapat ditentukan dengan menentukan  $\alpha$  dan  $\gamma$  yang mempunyai kisaran nilai antara 0 sampai 1. Berikut ini diberikan contoh perhitungan pemulusan dan nilai trend dengan menggunakan  $\alpha = 0,3$  dan  $\gamma = 0,25$ . Taksiran untuk nilai  $S_1 = Z_1$  dan  $b_1$  dapat mengambil dari persamaan garis trend,  $b_1 = 12793,46$ . ;  $S_t = 0,3Z_t + 0,7(S_{t-1} + b_{t-1})$  ;  $b_t = 0,25(S_t - S_{t-1}) + 0,75b_{t-1}$  ;  $F_{t+m} = S_t + b_t m$ . Ketika  $\alpha = 0,3$  dan  $\gamma = 0,05$  diperoleh nilai MSE = 16656306,29. Selanjutnya dicoba nilai-nilai  $\alpha$  dan  $\gamma$  yang lain sedemikian sehingga memberikan nilai MSE yang minimum. Kisaran nilai MSE yang besar menunjukkan pentingnya peranan  $\alpha$  dan  $\gamma$  dalam menentukan kesalahan yang dihasilkan. Setelah dilakukan percobaan dengan mengganti nilai-nilai  $\alpha$  dan  $\gamma$ , dipilih nilai  $\alpha = 0,99$  dan  $\gamma = 0,05$  karena memberikan nilai MSE terkecil yaitu sebesar 15175922,48.

**3.3. Penentuan Nilai Ramalan**

Langkah selanjutnya adalah menentukan nilai ramalan untuk  $F_{26}, F_{27}, F_{28}$  (tahun 2020, 2021, dan 2022).  $F_{t+m} = S_t + b_t m$ , dengan  $t = 25$  dan  $m = 1$ ;  $F_{25+1} = S_{25} + b_{25}$   
 $F_{26} = 1614339,321$ ; Untuk  $t = 25$  dan  $m = 2$ ;  $F_{25+2} = S_{25} + b_{25} \cdot (2)$ .  $F_{27} = 1626378,985$ ; Untuk  $t = 25$  dan  $m = 3$ ;  $F_{25+3} = S_{25} + b_{25} \cdot (3)$  ;  $F_{28} = 1641519,218$ . Jadi diperoleh nilai ramalan untuk tahun 2020 adalah 1614339 jiwa, tahun 2021 sebanyak 1626379 jiwa dan tahun 2022 sebanyak 1641519 jiwa.

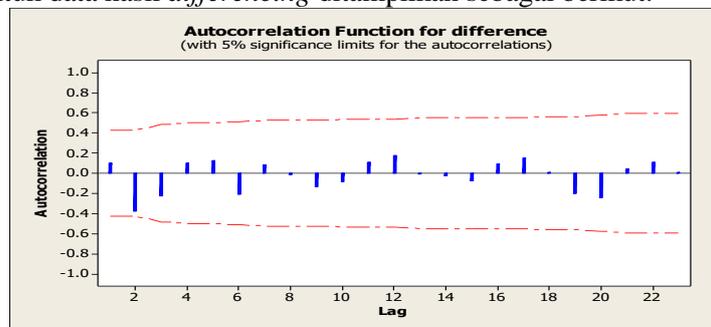
**3.4. Identifikasi Model pada Metode ARIMA**

Langkah awal dalam identifikasi model adalah menguji kestasioneran data, karena model ARIMA Box-Jenkins hanya berlaku untuk data yang stasioner. Data penduduk Kabupaten Banyumas Kestasioneran data juga dapat diidentifikasi dari diagram FAK-nya. Diagram FAK untuk data penduduk Kabupaten Banyumas ditampilkan sebagai berikut.



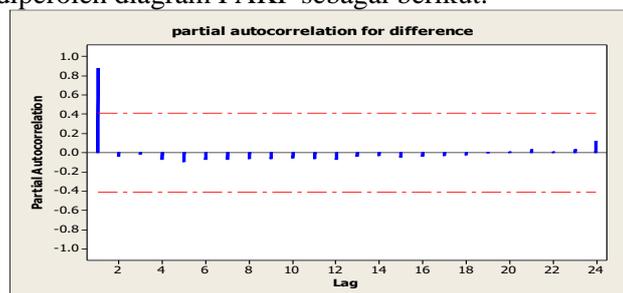
Gambar 2 diagram FAK untuk data penduduk Kab. Banyumas

Berdasarkan Gambar 2. diagram deret waktu diketahui bahwa terjadi kenaikan pada data, sedangkan pada diagram FAK diketahui bahwa nilai autokorelasi untuk data penduduk Kabupaten Banyumas turun secara lambat. Dapat disimpulkan bahwa data yang digunakan belum stasioner dalam rata-rata. Langkah selanjutnya adalah melakukan *differencing* terhadap data. Setelah *differencing* dilakukan, data baru hasil *differencing* diuji kembali kestasionerannya. Diagram deret waktu dan FAK untuk data hasil *differencing* ditampilkan sebagai berikut.



Gambar 3 Diagram FAK untuk data hasil differencing

Berdasarkan Gambar 3. diagram deret waktu diketahui bahwa terjadi kenaikan nilai pada data, sedangkan berdasarkan diagram FAK diketahui bahwa nilai-nilai autokorelasi untuk data hasil *differencing* turun mendekati nol. Oleh karena itu, dapat disimpulkan bahwa data yang digunakan telah stasioner dalam rata-rata. Langkah selanjutnya yang harus ditempuh untuk menetapkan model ARIMA yang sesuai, dibuat diagram FAKP untuk data hasil *differencing*. Dengan menggunakan software Minitab 16. diperoleh diagram FAKP sebagai berikut.



Gambar 4. Diagram FAKP untuk data hasil differencing

Gambar 4. hasil dalam bentuk diagram autokorelasi parsial menunjukkan nilai FAKP pada lag 1 melewati garis yang terputus-putus. Dari diagram autokorelasi parsial tersebut dapat disimpulkan bahwa bentuk FAKP terpotong setelah lag 1, kemudian turun secara eksponensial. Berdasarkan bentuk FAK dan FAKP, diperoleh dugaan bahwa model yang sesuai untuk data ARIMA (1,1,1).

### 3.5. Penaksiran Parameter Dan Pemeriksaan Diagnostik

Setelah memperoleh dugaan model-model awal ARIMA (p,d,q), tahap selanjutnya adalah menaksir parameter dari model-model tersebut, sehingga diperoleh besaran koefisien modelnya. Setelah besaran koefisien model diperoleh, dilakukan pemeriksaan diagnostik yang meliputi uji signifikansi parameter dan uji kesesuaian model. Penaksiran parameter dan pemeriksaan diagnostik untuk dugaan model-model awal ARIMA (1,1,1) karena yang terbaik.

**3.5.1 ARIMA (1,1,1)**

Berdasarkan data penduduk Kabupaten Banyumas diperoleh nilai taksiran parameter untuk model ARIMA (1,1,1) sebagai berikut.

**Tabel 1. Output Minitab ARIMA (1,1,1)**

Final Estimates of Parameters				
Type	Coef	SE Coef	T	P
AR 1	0.8899	0.0069	128.971	0.000
MA 1	0.8344	0.1072	7.784	0.000
Differencing: 1 regular difference				
Number of observations: Original series 25, after differencing 24				
Residuals: SS = 382226673 (backforecasts excluded)				
MS = 17373940 DF = 22				
Modified Box-Pierce (Ljung-Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	11.0	*	*	*
DF	10	*	*	*
P-Value	0.361	*	*	*

Karena model awal yang diperoleh adalah ARIMA (1,1,1), maka parameter yang perlu diuji adalah parameter autoregressive  $\phi_1$  dan moving average  $\theta_1$ . Bentuk hipotesis yang akan diuji adalah :  $H_0 : \phi_1 = 0$  (parameter AR tidak cukup signifikan dalam model) dan  $H_1 : \phi_1 \neq 0$  (parameter AR cukup signifikan dalam model) dan  $H_0 : \theta_1 = 0$  (parameter MA tidak cukup signifikan dalam model);  $H_1 : \theta_1 \neq 0$  (parameter MA cukup signifikan dalam model)

**a. Hipotesis untuk AR(1)**

$H_0 : \phi_1 = 0; H_1 : \phi_1 \neq 0$ ; Taraf signifikansi  $\alpha = 5\%$

Statistik uji  $t = \frac{\hat{\phi}_1}{SE(\hat{\phi}_1)} = 128,97$

Daerah penolakan :  $H_0$  ditolak jika  $|t| > t_{0,025;22}$ ; Keputusan :  $|t| = 128,97 > 2,06390 = t_{\alpha/2;22}$  dan p-value  $0,000 < 0,025 = \frac{\alpha}{2}$  maka  $H_0$  ditolak pada taraf signifikansi  $\alpha = 5\%$ .

Kesimpulan : Taksiran parameter AR(1) tersebut signifikan berbeda dari nol.

**b. Hipotesis untuk MA(1)**

$H_0 : \theta_1 = 0; H_1 : \theta_1 \neq 0$  ; Taraf signifikansi  $\alpha = 5\%$

Statistik uji  $t = \frac{\hat{\theta}_1}{SE(\hat{\theta}_1)} = 7,7836$

Daerah penolakan :  $H_0$  ditolak jika  $|t| > t_{0,025;22}$ ; Keputusan :  $|t| = 7,7836 > 2,06390 = t_{\alpha/2;22}$  dan p-value  $0,000 < 0,025 = \frac{\alpha}{2}$  maka  $H_0$  ditolak pada taraf signifikansi  $\alpha = 5\%$ .

Kesimpulan : Taksiran parameter MA(1) tersebut signifikan berbeda dari nol

Setelah dilakukan pengujian signifikansi parameter, langkah selanjutnya adalah uji kesesuaian model. Uji kesesuaian model meliputi kecukupan model (uji asumsi white noise) dan uji asumsi distribusi normal.

**3.5.2. Uji asumsi white noise dapat dituliskan sebagai berikut.**

- i. Hipotesis :  $H_0 : \rho_1 = \rho_2 = \dots = \rho_k = 0; H_1 : \text{minimal ada satu } \rho_j \neq 0, j = 1, 2, \dots, k$   
Hipotesis nol menyatakan bahwa sisa (residual) memenuhi asumsi white noise, sedangkan hipotesis alternatif mengatakan bahwa sisa tidak memenuhi asumsi white noise.

ii. Statisti Uji : Ljung-Box statistics (Modified Box-Pierce) :  $Q^* = n(n + 2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k}$

Tabel 2 Modified Box-Pierce (Ljung-Box) Chi Square Statistics

Modified Box-Pierce (Ljung-Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	11.0	*	*	*
DF	10	*	*	*
P-Value	0.351	*	*	*

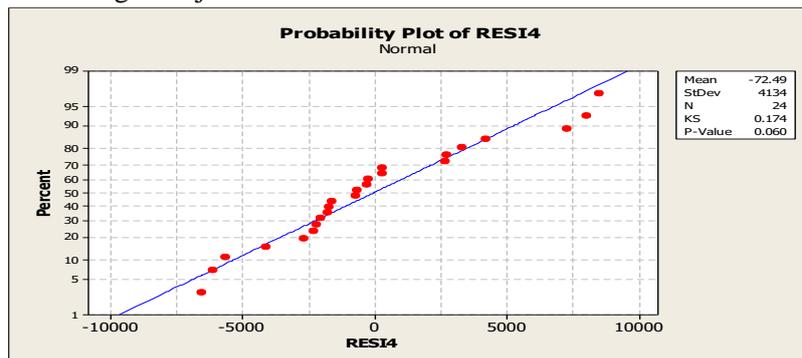
Berdasarkan Tabel 2 nilai statistik Ljung-Box yang diperoleh adalah pada lag 12. Pada Tabel 2 diperoleh nilai  $Q^*$  untuk lag 12 adalah 11,0; sedangkan  $\chi^2_{0,05;12-1} = 19,6751$  (diperoleh dari tabel  $\chi^2$ ). Ringkasan hasil Ljung-Box dapat dilihat pada Tabel 3 berikut.

Tabel 3 Ringkasan hasil Ljung-Box

Lag (K)	Df(K-m)	Statistik Ljung-Box ( $Q^*$ )	$\chi^2_{\alpha;df}$	p-value
12	11	11,0	19,6751	0,351

Tabel 3 menunjukkan bahwa pada lag 12 nilai  $Q^* = 11,0 < \chi^2_{0,05;11} = 19,6751$  artinya bahwa sampai lag 12, kesimpulan yang dapat diambil adalah tidak ada korelasi antara sisa pada lag t dengan sisa lag 12. Hal ini berarti sisa telah memenuhi syarat white noise. Apabila menggunakan p-value, kesimpulan dapat dibuat tanpa menggunakan tabel  $\chi^2$ , hanya cukup membandingkan p-value tersebut dengan  $\alpha$  yang digunakan dalam hipotesis tersebut. Kriteria kesimpulannya yaitu menolak  $H_0$  apabila  $p < \frac{\alpha}{2}$  dan menerima  $H_0$  apabila  $p > \frac{\alpha}{2}$ . Berdasarkan Tabel 3, p-value untuk statistik Ljung-Box pada lag 12 adalah 0,351. Nilai tersebut lebih besar dari  $\frac{\alpha}{2} = 0,025$ . Dengan demikian, hipotesis  $H_0$  diterima, artinya sisa telah memenuhi asumsi *white noise*. Selanjutnya dilakukan uji distribusi normal dengan bentuk hipotesis sebagai berikut.

$H_0$  : sisa berdistribusi normal;  $H_1$  : sisa tidak berdistribusi normal. Dengan menggunakan software Minitab 16 diperoleh diagram uji distribusi normal untuk sisa berikut ini.



Gambar 5. Uji distribusi normal untuk sisa

Daerah penolakan :  $H_0$  ditolak jika p-value  $< \frac{\alpha}{2}$ , Keputusan : karena p-value = 0,060  $> 0,025 = \frac{\alpha}{2}$  maka  $H_0$  diterima. Kesimpulan : sisa berdistribusi normal. Berdasarkan uraian diatas, dapat disimpulkan bahwa model telah memenuhi semua uji yang dilakukan, sehingga model ARIMA (1,1,1) dapat digunakan untuk memodelkan data penduduk Kabupaten Banyumas.

### 3.6. Pemilihan Model Terbaik

Setelah diuji kesignifikanan parameter dan uji diagnostik, langkah selanjutnya adalah memilih model terbaik dengan kriteria MSE yakni memilih model dengan nilai MSE yang paling kecil. Berdasarkan hasil tersebut, diketahui bahwa model ARIMA (1, 1, 1) menghasilkan nilai MSE yang terkecil. Oleh karena itu, model terbaik berdasarkan MSE adalah model ARIMA (1, 1, 1) dengan nilai MSE sebesar 17373940 .

**3.7. Peramalan Jumlah Penduduk**

Secara matematis, bentuk umum model ARIMA (1, 1, 1) dapat dituliskan sebagai berikut.

$$(1 - \phi_1 B)(1 - B)\dot{Z}_t = a_t(1 - \theta B)$$

$$\dot{Z}_t - B\dot{Z}_t - \phi_1 B\dot{Z}_t + \phi_1 B^2\dot{Z}_t = a_t(1 - \theta B)$$

$$(Z_t - \mu) - (Z_{t-1} - \mu) - \phi_1(Z_{t-1} - \mu) + \phi_1(Z_{t-2} - \mu) = a_t - \theta a_{t-1}$$

Berdasarkan data jumlah penduduk dan data residual , dengan menggunakan persamaan tersebut, ramalan jumlah penduduk untuk tahun 2020:  $\hat{Z}_{26} = 1618258,187$  ; 2021;  $\hat{Z}_{27} = 1628259,401$  ; dan 2022:  $\hat{Z}_{28} = 1643416,893$ . Jadi diperoleh nilai ramalan untuk tahun 2020 adalah 1618258 jiwa, tahun 2021 sebanyak 1628259 jiwa dan tahun 2022 sebanyak 1643417 jiwa.

**3.8. Perhitungan Uji Beda Dua Rata-Rata Model Pemulusan Eksponensial Ganda Holt dan Model ARIMA Box-Jenkins**

- i. Hipotesis:  $H_0 : \mu_1 - \mu_2 = D_0$  (rata-rata model pemulusan Eksponensial ganda Holt sama dengan model ARIMA Box-Jenkins);  $H_1 : \mu_1 - \mu_2 \neq D_0$  (rata-rata model pemulusan Eksponensial ganda Holt tidak sama dengan model ARIMA Box-Jenkins)

- ii. Statistik Uji

$$t = \frac{(\bar{y}_1 - \bar{y}_2) - D_0}{\sqrt{s_p^2 \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}} = -0,025$$

Daerah Penolakan:  $H_0$  ditolak jika  $|t| > t_{\frac{\alpha}{2}, 42}$

- iii. Keputusan: Berdasarkan perhitungan diperoleh nilai t hitung sebesar -0,025. Nilai ini kemudian akan dibandingkan dengan nilai t tabel.  $|t| = 0,024$ ;  $t_{\frac{\alpha}{2}, 42} = 2,015$

$|t| < t_{\frac{\alpha}{2}, 42}$  , maka diperoleh keputusan untuk menerima  $H_0$ . Kesimpulan: Rata-rata model penghalusan Eksponensial ganda Holt sama dengan model ARIMA Box-Jenkins

Hasil pengujian beda dua rata-rata menunjukkan bahwa peramalan dengan menggunakan model Pemulusan Eksponensial Ganda Holt dan ARIMA Box-Jenkins tidak memperlihatkan perbedaan yang nyata dengan derajat keyakinan 95%. Peramalan jumlah penduduk yang dikembangkan dalam penelitian ini, yaitu model Pemulusan Eksponensial Ganda Holt dan ARIMA Box-Jenkins, sama-sama memiliki akurasi yang tinggi. Dengan kata lain, baik model pertama maupun model kedua dapat digunakan secara mandiri untuk peramalan dengan tingkat akurasi yang sama.

**4. Kesimpulan**

Model deret waktu linier terbaik untuk data jumlah penduduk di Kabupaten Banyumas adalah ARIMA (1,1,1).

- a. Ramalan jumlah penduduk berdasarkan Metode Pemulusan Eksponensial Ganda Holt dan metode ARIMA Box-Jenkins untuk tahun 2020, 2021, dan 2022 adalah sebagai berikut.

Tabel 4. Hasil Ramalan Eksponensial Ganda Holt dan metode ARIMA Box-Jenkins untuk tahun 2020, 2021, dan 2022

Tahun	Hasil Ramalan	
	Metode Pemulusan Eksponensial	Metode ARIMA Box-Jenkins
2020	1.614.339 jiwa	1.618.258 jiwa
2021	1.626.379 jiwa	1.628.259 jiwa

2022	1.641.519 jiwa	1.643.417 jiwa
------	----------------	----------------

- b. Berdasarkan kriteria MSE, Metode Pemulusan Eksponensial Ganda Holt memberikan nilai MSE yang lebih kecil dibandingkan Metode ARIMA Box-Jenkins. Setelah dilakukan uji beda rata-rata model, ternyata rata-rata dari model pemulusan eksponensial Ganda Holt sama dengan rata-rata model ARIMA Box-jenkins. Kedua metode tersebut sama-sama efektif untuk meramalkan jumlah penduduk di Kabupaten Banyumas.

## 5. DAFTAR PUSTAKA

- [1] Aritonang R. 2002, Peramalan Bisnis, Jakarta : Ghalia Indonesia.
- [2] Assauri S. 1984, *Teknik dan Metode Peramalan (Penerapannya Dalam Ekonomi dan Dunia Usaha)*, Edisi Satu, Jakarta : LPFE - UI.
- [3] Aswi dan Sukarna, (2004), *Analisis Deret Waktu: Teori dan Aplikasi*, Makassar: Andira Publisher.
- [4] Abdullah, Lazim. 2012. ARIMA Model for Gold Bullion Coin Selling Prices Forecasting. *International Journal of Advances in Applied Sciences (IJAAS)*. 1(4): 153-158
- [5] Alli, P., Sundar, D., Devi B. Uma. 2013. An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50. *International Journal of Data Mining Knowledge Management Process (IJDMP)*. 3(1): 65-78.
- [6] Bowerman, Bruce L. dan O'Connell, Richard T. 1993, *Forecasting and Time Series: An Applied Approach 3<sup>rd</sup> edition*, California : Duxbury Press.
- [7] Box, G. E. P. & g. M.Jenkins. 1970. *Time Series Analysis*. California: Holden Day.
- [8] Ekata & Shivastav Anand Kumar. 2012. Applicability of Box Jenkins ARIMA Model in Crime Forecasting: A Case Study of Counterfeiting in Gujarat State. *Internasional Journal of Advanced Research in Computer Engineering Technology*. Vol I. ISSN 2278-1323
- [9] Gujarati, Damodar, 1979, *Basic Econometrics*, New York : McGraw-Hill International Book Company.
- [10] Hanke, John E. dan Reitsch, Arthur G. 1995, *Business Forecasting*, Washington : Prentice Hall.
- [11] Hermawan, Eddy. 2011. Perbandingan Metode Box-Jenkins dan Holt-Winters dalam Memprediksi Anomali Air OLR Pentad di Kawasan Barat Indonesia. *Jurnal Sains Dirgantara*. 9(1):25-35.
- [12] Makridakis, S., Wheelwright, S.C., dan McGee, V.E. 1992, *Metode dan Aplikasi Peramalan*, Jakarta : Erlangga.
- [13] Newbold, Paul, 1988, *Statistics for Business and Economics*, Edisi Kedua, New Jersey : Prentice Hall, Englewood Cliffs 07632, USA.
- [14] Phumchusri, Naragain & Udom, Patimaporn. 2014. A Comparison Study Between Time Series Model and ARIMA Model for Sales Forecasting of Distributor in Plastic Industry. *IOSR Journal of Engineering (IOSRJEN) Departement of Industrial Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok, Thailand*. 4(6): 32-38.
- [15] Pramita, Wahyu & Tanuwijaya, Haryanto. 2010. Penerapan Metode Exponential Smoothing Winter dalam Sistem Informasi Pengendalian Persediaan Produk dan Bahan Baku sebuah Cafe. UPN Veteran Yogyakarta. *Prosiding Seminar Nasional Informatika*. Surabaya: STIKOM Surabaya
- [16] Subagyo, P. 1986, *Forecasting Konsep dan Aplikasi*, Yogyakarta : BPFE-Yogyakarta.
- [17] Wei, William W. S. 2006, *Time Series Analysis: Univariate and Multivariate Methods 2<sup>nd</sup> edition*, New Jersey : Pearson Education.