

Pemodelan dan Peramalan Deret Waktu Studi kasus: Suhu Permukaan Laut di Selatan Jawa Timur

Hutomo Atman Maulana^{1,*}, Muliah², Maria Zefaya Sampe³, Farrah Hanifah⁴

¹Tadris Matematika, Fakultas Tarbiyah dan Keguruan, UIN Imam Bonjol Padang, Indonesia

^{2,3,4}Matematika, FMIPA, Institut Teknologi Bandung, Indonesia

Email: hutomomaulana@uinib.ac.id

Received: May 2017; Accepted: August 2017; Published: October 2017

Abstrak

Suhu permukaan laut merupakan salah satu komponen penting yang dapat menentukan potensi kekayaan laut. Penelitian ini bertujuan untuk memodelkan dan meramalkan data deret waktu suhu permukaan laut dengan menggunakan suatu metode Box-Jenkins. Data yang digunakan adalah suhu permukaan laut di Selatan Jawa Timur (Januari 1983-Desember 2013) dengan ukuran sampelnya sebanyak 372. Namun yang akan digunakan untuk pemodelan adalah sebanyak 360, yaitu dari Januari 1983 hingga Desember 2012, sedangkan data pada Tahun 2013 akan digunakan untuk forecasting. Berdasarkan hasil penelitian diperoleh model yang sesuai adalah SARIMA(1,0,0)(1,0,1)¹² dengan $Y_t = 0,010039 + 0,734220Y_{t-1} + 0,014893Y_{t-12} - (0,734220)(0,014893)Y_{t-13} + 0,940726e_{t-12}$ dan MSE sebesar 0,07888096.

Kata kunci: Suhu permukaan laut, deret waktu, Metode Box-Jenkins.

Abstract

The sea surface temperature is one of the important components that can determine the potential of the sea. This research aims to model and forecast time series data of sea surface temperature by using a Box-Jenkins method. Data in this research are the sea surface temperatures in the South of East Java (January 1983-December 2013) with sample size of 372. 360 data will be used for modeling which is from January 1983 to December 2012, and data in 2013 will be used for forecasting. Based on the results of analysis time series, the appropriate models is SARIMA(1,0,0) (1,0,1)¹² where can be written as $Y_t = 0,010039 + 0,734220Y_{t-1} + 0,014893Y_{t-12} - (0,734220)(0,014893)Y_{t-13} + 0,940726e_{t-12}$ with MSE of 0.07888096.

Keywords: Sea surface temperature, time series, Box-Jenkins method

*Corresponding author.

Peer review under responsibility UIN Imam Bonjol Padang.

© 2017 UIN Imam Bonjol Padang. All rights reserved.

ISSN: 2580-6726 (print), 2598-2133 (online)

PENDAHULUAN

Fakta menunjukkan bahwa Indonesia merupakan negara maritim dan kepulauan terbesar (17.504 pulau) di dunia serta memiliki garis pantai sepanjang 104.000 km. Perikanan merupakan salah satu sektor yang penting bagi perekonomian Indonesia yang memiliki kepulauan terluas di dunia, dimana sekitar 70% wilayahnya terdiri dari perairan. Hal tersebut tentu memberikan dampak bagi Indonesia yang memiliki potensi pada sektor perikanan yang sangat besar, baik ditinjau dari kuantitas maupun diversivitas. Dengan potensi yang ada tersebut seharusnya meletakkan sektor perikanan menjadi salah satu sektor riil yang potensial di Indonesia. Sumber daya perikanan yang dimiliki oleh Indonesia beragam dan berpotensi diantaranya budidaya laut dan tambak atau payau yang mengarah untuk kemajuan perekonomian Indonesia.

Berlimpahnya kekayaan laut yang terdapat di Indonesia diikuti dengan besarnya potensi sumberdaya dan jenis ikan seperti ikan pelagis besar dan kecil, ikan demersal, udang, lobster, cumi-cumi dan lainnya. Berdasarkan data Kementerian Kelautan dan Perikanan 2011, potensi lestari sumberdaya ikan laut Indonesia sebesar 6,52 juta ton. Volume produksi perikanan Indonesia baik dari perikanan tangkap maupun perikanan budidaya terus meningkat setiap tahunnya. Produksi perikanan tangkap tahun 2006 sebesar 4,8 juta ton dan meningkat menjadi 5,7 juta ton pada tahun 2011 (BPS, 2012). Rata-rata kenaikan produksi perikanan dirasakan cukup lambat hanya sekitar 3,2

persen. Produksi dari perikanan budidaya berbeda dengan perikanan tangkap, rata-rata kenaikan produksi perikanan budidaya 2006-2011 adalah 25,62 persen (Peneliti Badan Kebijakan Fiskal Kementerian Keuangan RI, 2014). Volume produksi perikanan budidaya tahun 2006 sebesar 2,68 juta ton, meningkat hampir 3 kali pada tahun 2011 yaitu 7,92 juta ton.

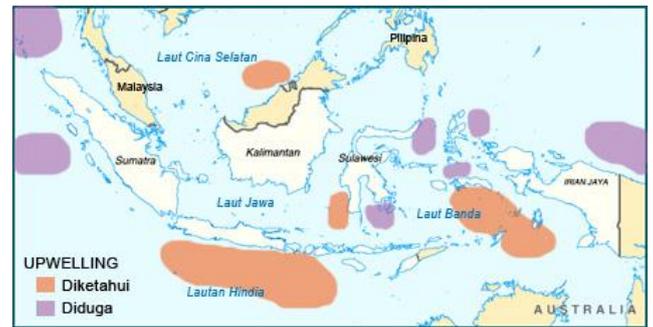
Salah satu provinsi yang berpotensi pada sektor perikanan adalah Jawa Timur yang merupakan pusat industrialisasi perikanan di Indonesia. Kontribusi sub sektor perikanan dari Provinsi Jawa Timur adalah sebanyak 25% dari kebutuhan perikanan nasional, sebanyak 285 ribu ton dengan ini provinsi Jawa Timur memiliki beberapa komoditi unggulan. Perairan selatan Jawa, khususnya selatan Jawa Timur terkenal dengan potensi perikanannya yang besar (lihat Gambar 1). Berdasarkan data yang diperoleh oleh Biro Pusat Statistik tahun 2013, Produksi Ikan dan Sub Sektor Perikanan untuk provinsi Jawa Timur untuk penangkapan perikanan laut sebanyak 381.574 ton dan untuk budidaya ikan laut sebesar 580,683 ton.

Tingkat produktivitas perairan dipengaruhi oleh berbagai faktor, dimana tiga faktor utamanya adalah ketersediaan nutrien, cahaya, dan temperatur. Nutrien dibutuhkan oleh fitoplankton untuk dapat melakukan proses fotosintesis, sedangkan temperatur mempengaruhi laju fotosintesis dan distribusi plankton di suatu kolom perairan. Fitoplankton sendiri merupakan produsen utama di perairan laut

yang berperan dalam menghasilkan produktivitas primer. Kandungan nutrisi yang tinggi yang umumnya ditemukan di lapisan dalam dengan tingkat pencahayaan yang rendah, sehingga diperlukan suatu proses tertentu untuk menjamin ketersediaan nutrisi di lapisan permukaan air laut (Kemili dan Putri, 2012).

Salah satu penyebab meningkatnya produktivitas perairan yakni dikarenakan adanya upwelling yang terjadi di perairan tersebut. Upwelling merupakan fenomena gerakan naiknya massa air dingin yang kaya unsur hara dari lapisan dalam ke permukaan suatu perairan. Tingginya kadar unsur hara tersebut merangsang perkembangan fitoplankton dan meningkatkan kesuburan perairan disekitarnya (Lalli dan Parsons, 1993 dalam Lidiawati, 2014).

Daerah upwelling selalu dihubungkan dengan meningkatnya populasi ikan di perairan tersebut, sehingga lokasi upwelling merupakan area yang ideal untuk menangkap ikan (Lidiawati, 2014). Salah satu tanda suatu daerah upwelling adalah temperatur permukaan laut yang lebih rendah dibandingkan daerah sekitarnya. Dengan melihat perubahan temperatur permukaan laut, akan bisa dilihat durasi terjadinya upwelling, puncak upwelling, kapan upwelling berakhir. Dengan mengetahui waktu kejadian upwelling, eksploitasi sumber daya hayati dapat dimaksimalkan. Sehingga peramalan (*forecasting*) suhu permukaan laut merupakan hal yang sangat penting dalam sektor perikanan yang mempengaruhi produktivitas primer di selatan laut Jawa Timur.



Gambar 1. Lokasi upwelling di Indonesia yang sudah diketahui dan baru diduga.

(Sumber: <http://belajar.kemdiknas.go.id>)

Adapun data suhu permukaan selatan laut Jawa Timur merupakan tipe data seasonal yang diperoleh dari Januari 1981 sampai dengan Desember 2013, sehingga pada penelitian ini akan digunakan model time series dengan pendekatan Box-Jenkins yang mengaplikasikan model Seasonal Autoregressive Integrated Moving Average (SARIMA). Metode ini dipopulerkan oleh George Box dan Gwilym Jenkins sekitar tahun 1970an dimana pemilihan metode ini disesuaikan pada aspek waktu, pola data, tipe model sistem yang diamati, tingkat keakuratan yang diinginkan dan sebagainya. Pendekatan dengan SARIMA memberikan pendekatan yang paling baik karena data suhu permukaan air laut tersebut memiliki fluktuasi tinggi dan sifat musiman. Model SARIMA sering digunakan untuk meramalkan karena disesuaikan dengan pola data yang paling cocok dari sekelompok data, dengan memanfaatkan sepenuhnya data masa lalu dan sekarang dari variabel *dependent* untuk melakukan peramalan jangka pendek yang akurat.

LANDASAN TEORI

Pemodelan Deret Waktu

Analisis deret waktu pada dasarnya digunakan untuk melakukan analisis data yang mempertimbangkan pengaruh waktu. Data dikumpulkan secara periodik berdasarkan urutan waktu, bisa dalam jam, hari, minggu, bulan, kuartal dan tahun. Analisis deret waktu dapat dilakukan untuk membantu dalam menyusun perencanaan ke depan. Untuk menentukan metode peramalan pada data deret waktu perlu diketahui pola dari data tersebut sehingga peramalan data dapat dilakukan dengan metode yang sesuai. Pola data dapat dibedakan menjadi empat jenis, yaitu pola musiman, siklis, trend, dan irregular. Pola musiman merupakan aktuasi dari data yang terjadi secara periodik dalam kurun waktu satu tahun, seperti triwulan, kuartalan, bulanan, mingguan, atau harian. Untuk data model stokastik terdapat beberapa model yang dapat digunakan seperti AR, MA, ARMA, ARIMA, SARIMA dan lainnya. Jika data mempunyai pola musiman, maka metode yang lebih tepat adalah *Seasonal Autoregressive Integrated Moving Average* (SARIMA).

Model Deret Waktu

Pemilihan model deret waktu yang sesuai dengan data akan menghasilkan keakuratan dalam peramalan (*forecasting*). Setiap model memiliki ciri masing-masing sehingga berdasarkan ciri tersebut kita jadikan acuan sebagai penentuan model yang cocok dari data. Beberapa model deret waktu tersebut adalah:

1. Model AR(p)

Deret waktu Y_t merupakan suatu proses AR (*Auto Regressive*) berorde p atau AR (p) dapat dinyatakan sebagai:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t$$

Dengan kata lain, Y_t merupakan kombinasi linier p buah nilai-nilai sebelumnya ditambah dengan galat pada saat t . Variabel galat e_t diasumsikan saling bebas terhadap $Y_{t-1}, Y_{t-2}, \dots, Y_{t-p}$ dan berdistribusi normal dengan rata-rata nol dan variansi σ^2 .

2. Model MA(q)

Model MA (*Moving Average*) berorde q merupakan suatu deret Y_t yang terbentuk dari galat pada waktu t dan galat-galat pada waktu sebelumnya yang diberi bobot, sehingga dapat dinyatakan sebagai:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

3. Model ARMA(p, q)

Model ARMA (*Auto Regressive Moving Average*) berorde p dan q merupakan proses deret waktu yang dibangun dari penggabungan antara AR (p) dan MA (q) dengan bentuk persamaan:

$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \dots - \theta_q e_{t-q}$$

4. Model ARIMA(p, d, q)

Model ARIMA dilakukan pada data yang *didifferencing* sehingga data telah stasioner model ARIMA (p, d, q) merupakan gabungan dari model ARMA (p, q) dan proses *differencing*, yaitu:

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)e_t$$

5. Model *Seasonal* ARIMA(p, d, q)(P, D, Q)^s

Musiman (*Seasonal*) adalah kecenderungan suatu deret waktu mengulangi pola gerak dalam

periode musim, biasanya satu tahun untuk data bulanan. Model *Seasonal ARIMA* merupakan model ARIMA yang digunakan untuk menyelesaikan proses deret waktu tersebut yang terdiri dari dua bagian, yaitu bagian tidak musiman (*non-seasonal*) dan bagian musiman (*seasonal*). Secara umum bentuk model *Seasonal ARIMA*(p,d,q)(P,D,Q)^s adalah:

$$\phi_p(B)\Phi_p(B^s)(1-B)^d(1-B^s)^D Y_t = \theta_q(B)\Theta_q(B^s)e_t$$

dengan:

p,d,q = orde AR, *differencing*, dan MA non-musiman

P,D,Q = orde AR, *differencing*, dan MA musiman

$\phi_p(B)$ = parameter AR non-musiman,
 $1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$

$\Phi_p(B^s)$ = parameter AR musiman,
 $1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_p B^{ps}$

$(1-B)^d$ = orde *differencing* non-musiman

$(1-B^s)^D$ = orde *differencing* musiman

$\theta_q(B)$ = parameter MA non-musiman,
 $1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$

$\Theta_q(B^s)$ = parameter MA musiman,
 $1 - \Theta_1 B^s - \Theta_2 B^{2s} - \dots - \Theta_q B^{qs}$

e_t = residual (galat) pada waktu ke- t

METODE PENELITIAN

Jenis Penelitian

Adapun jenis penelitian yang digunakan adalah jenis deskriptif dengan pendekatan studi kasus. Metode deskriptif digunakan untuk mendeskripsikan atau menguraikan permasalahan yang berkaitan dengan penentuan

model SARIMA yang sesuai dengan data suhu permukaan selatan laut Jawa Timur.

Waktu dan Tempat Penelitian

Penelitian dilaksanakan pada 18 Maret - 27 Mei 2015 bertempat di Bandung.

Target/Subjek Penelitian/Populasi dan Sampel

Pada penelitian ini, obyek yang digunakan adalah suhu permukaan perairan selatan laut Jawa Timur yang memiliki pengaruh terhadap produktivitas perairan Jawa Timur khususnya pada sektor perikanan.

Sumber Data

Data yang akan dipakai dalam penelitian ini adalah data suhu permukaan air laut di wilayah selatan Jawa Timur mulai Januari tahun 1983 sampai dengan Desember 2013 yang diperoleh dari Penelitian "Kajian Dampak Perubahan Iklim Terhadap Dinamika Upwelling sebagai Dasar untuk Memperkirakan Pola Migrasi Ikan Tuna di Perairan Selatan Jawa-Nusa Tenggara Barat dengan Menggunakan Model Transpor Temperatur Laut" Laporan Riset dan Inovasi KK. Institut Teknologi Bandung, Bandung oleh Ningsih dkk (2013).

Teknik Analisis Data

Penelitian ini menggunakan metode yang mampu mengakomodasi kejadian ekstrim yakni metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Model SARIMA merupakan pengembangan model ARIMA. Berikut ini adalah langkah-langkah seasonal pada SARIMA :

- 1) Mengidentifikasi seasonal data dengan menggunakan ACF/PACF pada *seasonal lags*
- 2) Melakukan *differencing* pada data sesuai dengan *season* yang diambil *differencing* pada *season* digunakan untuk

menghilangkan *seasonality* dikarenakan ada kemungkinan data yang dipakai membutuhkan *differencing*.

Tahapan Pemodelan Deret Waktu

Asumsi terpenting dalam analisis deret waktu adalah kestasioneran. Ide dari kestasioneran adalah perilaku data yang tidak berubah terhadap waktu dalam suatu proses (Cryer, 2008). Jenis kestasioneran ada dua macam, yaitu:

1. Stasioner kuat

Jika distribusi gabungan Y_1, Y_2, \dots, Y_n sama dengan distribusi gabungan dari $Y_{1-k}, Y_{2-k}, \dots, Y_{n-k}$, ditulis sebagai:

$$F(y_{t_1}, y_{t_2}, \dots, y_{t_n}) = F(y_{t_1-k}, y_{t_2-k}, \dots, y_{t_n-k})$$

dengan t_1, t_2, \dots, t_n dan \forall lag k .

2. Stasioner lemah

Jika μ (mean) dan σ^2 (variansi) selalu konstan setiap waktu (t).

Dikarenakan sulitnya untuk menggambarkan distribusi gabungan yang merupakan syarat dari stasioner kuat, maka dalam penelitian ini cukup hanya dengan asumsi stasioner lemah saja.

a. Stasioner terhadap mean

Uji yang digunakan untuk melihat kestasioneran terhadap mean adalah uji akar unit. salah satu metode pengujian adalah dengan *Augmented Dickey-Fuller Test*. dimana menggunakan persamaan berikut:

$$\nabla Y_t = (\rho - 1)Y_{t-1} + \mu_t = \delta Y_{t-1} + \mu_t$$

$H_0: \delta = 0$ (data memiliki akar unit)

$H_1: \delta \neq 0$

Jika kita tidak menolak $H_0: \delta = 0$, maka $\rho = 1$. Artinya data memiliki akar unit, dimana data time series Y_t tidak stasioner.

b. Stasioner terhadap variansi

Uji Levene digunakan untuk melihat apakah data memiliki variansi yang sama:

$$H_0: \sigma_1 = \sigma_2 = \dots = \sigma_i, \quad i \in N$$

$$H_1: \text{terdapat paling tidak satu pasang } \sigma_i \neq \sigma_j, \quad i, j \in N.$$

Jika kita tidak menolak $H_0: \sigma_1 = \sigma_2 = \dots = \sigma_i, \quad i \in N$, artinya data memiliki variansi yang sama atau konstan, dimana data time series Y_t stasioner.

Asumsi terpenting lainnya dalam analisis deret waktu adalah hubungan autokorelasi. Autokorelasi adalah hubungan kelinieran antara satu variabel dengan dirinya sendiri. Secara matematis, autokorelasi melibatkan nilai mean dan fungsi autokovariansi yang didenisikan sebagai:

$\gamma_{t,s} = Cov(Y_t, Y_s)$ untuk $t, s = 0, \pm 1, \pm 2, \dots$, yaitu:

$$\begin{aligned} Cov(Y_t, Y_s) &= E[(Y_t - \mu_t)(Y_s - \mu_s)] \\ &= E(Y_t - Y_s) - \mu_t \mu_s \end{aligned}$$

sehingga diperoleh fungsi autokorelasi yang didefinisikan sebagai:

$\rho_{t,s} = Corr(Y_t, Y_s)$ untuk $t, s = 0, \pm 1, \pm 2, \dots$ adalah:

$$\begin{aligned} Corr(Y_t, Y_s) &= \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t) Var(Y_s)}} \\ &= \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t} \gamma_{s,s}}} \end{aligned}$$

$$\rho_k = \frac{\sum_{t=1}^{n-k} (Y_t - \mu)(Y_{t-k} - \mu)}{\sum_{t=1}^n (Y_t - \mu)^2} \quad \text{dengan } k = 0, 1, 2, \dots$$

Nilai kovariansi dan korelasi menunjukkan hubungan linier dari dua variabel acak, sehingga apabila $\rho_{t,s}$ mendekati ± 1 maka hubungan liniernya kuat antar dua variabel. Namun, jika $\rho_{t,s}$ mendekati nol maka hubungan liniernya lemah, dan apabila $\rho_{t,s} = 0$, maka Y_t, Y_s tidak berkorelasi (tidak memiliki hubungan linier).

Analisis deret waktu memerlukan suatu tahapan pemodelan yang sistematis. Salah satu metode analisis deret waktu yang biasa digunakan adalah metode George E. P. Box dan Gwilym M. Jenkins yang dikenal sebagai metode Box-Jenkins. Dalam metode ini memiliki tiga iterasi utama. Adapun iterasinya adalah:

1. Identifikasi model

Pada tahap ini, proses identifikasi model dengan menggunakan prinsip parsimony (model sederhana) dengan jumlah parameter yang sedikit. Proses ini diawali dengan membuat plot ACF (*auto correlation function*) dan PACF (*partial auto correlation function*). Identifikasi dilakukan dengan mengamati perilaku dari ACF dan PACF.

Tabel 1. Perilaku ACF dan PACF Model AR(p), MA(q), dan ARMA(p,q)

Model	AR(p)	MA(q)	ARMA(p,q)
ACF	Menurun secara eksponensial atau membentuk gelombang sinus teredam	Cut off setelah lag-q	Menurun secara eksponensial atau membentuk gelombang sinus teredam
PACF	Cut off setelah lag-p	Menurun secara eksponensial atau membentuk gelombang sinus teredam	Menurun secara eksponensial atau membentuk gelombang sinus teredam

Data yang tidak menunjukkan perilaku ACF dan PACF sesuai dengan tabel diatas, maka cenderung menunjukkan perilaku tidak stasioner. Data yang tidak stasioner harus dilakukan proses transformasi terlebih dahulu untuk menjadikan stasioner. Salah satu proses stasioner yang paling sederhana adalah *differencing*. *Differencing* adalah proses pengurangan data pada waktu t dengan data waktu t-1, dengan tujuan data yang telah *differencing* menjadi stasioner. *Differencing* pertama, kedua, dan selanjutnya dari data observasi Y_t yaitu :

$$d = 1: \quad W_t = \nabla Y_t = Y_t - Y_{t-1}$$

$$d = 2: \quad W_t = \nabla^2 Y_t = \nabla(\nabla Y_t) = Y_t - 2Y_{t-1} + Y_{t-2}$$

2. Penaksiran Parameter

Pada tahap ini, dilakukan estimasi atau taksiran terbaik dari parameter yang tidak diketahui. Terdapat tiga cara untuk melakukan penaksiran parameter yaitu, dengan menggunakan metode momen, least square, atau maximum likelihood. Pada tahapan ini juga dilihat AIC (*Akaike's Information Criterion*), yaitu suatu kriteria yang diperkenalkan oleh Akaike pada tahun 1973 dengan mempertimbangkan banyaknya parameter dalam model. AIC digunakan sebagai salah satu penentu dalam pemilihan model terbaik. Kriteria AIC dapat dirumuskan sebagai berikut:

$$AIC = -2\ln(L) + 2M$$

dengan M merupakan jumlah dari parameter (p dan q), dan L adalah nilai maksimum dari fungsi likelihood untuk estimasi model. Model yang lebih baik akan memberikan nilai AIC yang minimum.

3. Uji diagnosis residual

Tahap ini memfokuskan pada pengujian residual. Asumsi dari residual atau galatnya adalah bersifat *white noise*, yaitu $e_t \sim N(0, \sigma^2)$ dan tidak berkorelasi. Pengujian asumsi dilakukan melalui tahapan berikut:

a. Uji normalitas terhadap residual

Uji normalitas dilakukan untuk melihat apakah residual berdistribusi normal dan dilakukan plot residual untuk melihat apakah berfluktuasi di sekitar 0 sehingga dapat diasumsikan mean=0. Uji yang digunakan adalah uji Jarque-Bera.

H0: Residual berdistribusi normal

H1: residual tidak berdistribusi normal

Statistik uji:

$$JB = \frac{T}{6} \left(S^2 + \frac{(K-3)^2}{4} \right)$$

Dengan T = ukuran sampel, S = nilai *skewness*, dan K = nilai kurtosis

b. Uji Ljung-Box

Uji Ljung-Box dilakukan untuk melihat adalah autokorelasi dari residual.

H₀: $\rho_1 = \rho_2 = \dots = \rho_k = 0$

H₁: minimal terdapat satu $\rho_k \neq 0, k = 1, 2, \dots, K$

Statistik uji:

$$Q^* = n(n+2) \sum_{k=1}^h \frac{r_k^2}{n-k}$$

Jika $Q^* > \chi_{\alpha, \nu}^2$ dengan $\nu = h - m$ dan $m = \#$ parameter, maka H₀ ditolak.

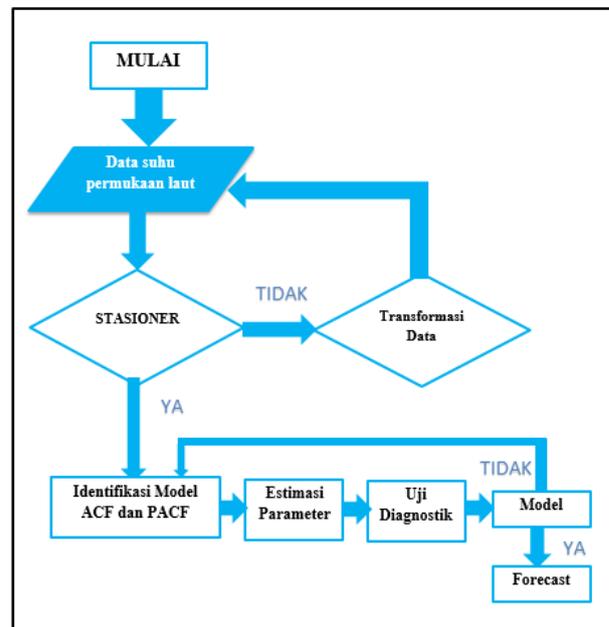
Alat dan Tahapan Pemodelan SARIMA

Alat yang digunakan dalam pengolahan data pada penelitian ini yakni software Eviews 8

untuk membantu menganalisis pemodelan SARIMA. Adapun tahapan-tahapannya adalah :

1. Mengidentifikasi modelnya terlebih dahulu.
2. Pendugaan parameter model.
3. Memeriksa nilai residual (sisaan).
4. Menggunakan model untuk peramalan jika model telah memenuhi syarat.

Adapun diagram alir rencana penelitian yang dilakukan yakni :

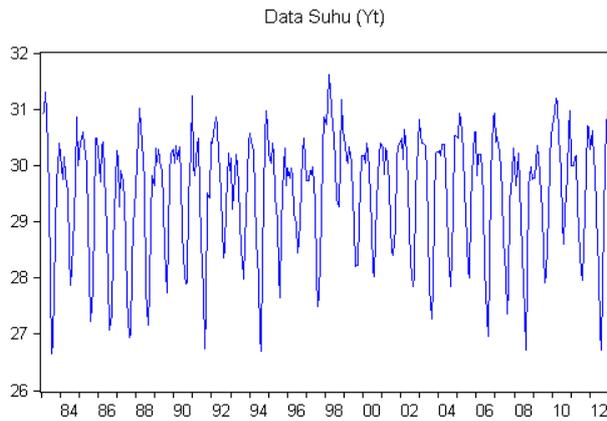


Gambar 2. Diagram alir rencana kerja yang akan dilakukan dalam penelitian

HASIL PENELITIAN DAN PEMBAHASAN

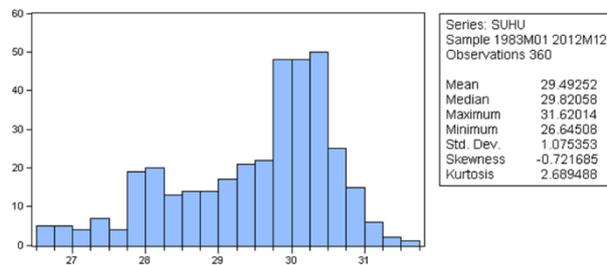
Statistika Deskriptif

Data yang digunakan adalah suhu permukaan laut di Selatan Jawa Timur (Januari 1983-Desember 2013) dengan ukuran sampelnya sebanyak 372. Namun yang akan digunakan untuk pemodelan adalah sebanyak 360, yaitu dari Januari 1983 hingga Desember 2012. Data pada Tahun 2013 akan digunakan untuk forecasting. Adapun plot data suhu permukaan selatan laut Jawa Timur disajikan pada gambar 3.



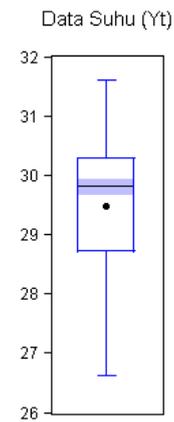
Gambar 3. Plot suhu permukaan laut di Selatan Jawa Timur (Januari 1983-Desember 2012)

Pada plot suhu pada gambar di atas dapat dilihat secara umum data stationer terhadap mean begitu juga dengan variansinya. Namun untuk lebih meyakinkan maka akan dilakukan beberapa uji statistik. Selanjutnya yang akan dilakukan adalah menganalisis statistika deskriptif yang bertujuan untuk memberikan deskripsi pada data. Statistika deskriptif menjelaskan berbagai karakteristik data seperti mean, variansi, rentang (range), nilai maksimum dan minimum.



Gambar 4. Statistika deskriptif suhu permukaan laut di Selatan Jawa Timur (Januari 1983-Desember 2012)

Berdasarkan statistika deskriptif pada gambar di atas dapat dilihat bahwa pada data nilai median lebih besar dari nilai mean, maka dapat disimpulkan data menceng menceng kiri (skew to the left). ini berarti terdapat beberapa data yang secara ekstrim terlalu kecil sehingga mengakibatkan kondisi $mean < median$. Selanjutnya perlu dilakukan pendeteksian pencilan (outliers) pada data. pencilan dapat mudah dideteksi dengan melihat dari boxplot. Box plot atau boxplot (juga dikenal sebagai diagram box-and-whisker) merupakan suatu box (kotak berbentuk bujur sangkar).



Gambar 5. Boxplot suhu permukaan laut di Selatan Jawa Timur (Januari 1983-Desember 2012)
Berdasarkan Gambar 5 terlihat bahwa tidak terdapat pencilan pada data yang digunakan, sehingga data dapat langsung digunakan.

1. Analisis Kestasioneran

Terdapat dua jenis kestasioneran yang akan dilihat pada data, yaitu kestasioneran terhadap mean dan variansi data.

a. Mean

Uji akar unit (unit root test) digunakan untuk melihat apakah data stationer terhadap mean.

Tabel 2. Uji Akar Unit Suhu Permukaan Laut

Null Hypothesis: SUHU has a unit root
Exogenous: Constant
Lag Length: 12 (Automatic - based on SIC, maxlag=16)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.959928	0.0018
Test critical values:		
1% level	-3.448998	
5% level	-2.869653	
10% level	-2.571161	

*MacKinnon (1996) one-sided p-values.

Berdasarkan Uji akar unit (unit root test) terlihat bahwa p-value sebesar 0.0018. Maka dengan α sebesar 5%, $p\text{-value} < \alpha$, H_0 ditolak. Sehingga dapat disimpulkan tidak terdapat akar unit pada data atau dengan kata lain data yang digunakan stasioner terhadap mean.

b. Variansi

Uji Levene digunakan untuk melihat kestasioneran terhadap variansi. Sebelum dilakukan pengujian maka data dibagi menjadi tiga kelompok, masing-masing kelompok terdiri dari 30 tahun dengan ukuran sampel 120. Jika data stasioner terhadap variansi maka nilai.

Berdasarkan Uji Levene didapat p-value sebesar 0.0873. Maka dengan α sebesar 1%, p-value > α , Ho tidak ditolak. Sehingga dapat disimpulkan atau dengan kata lain data yang digunakan stasioner terhadap variansi.

Tabel 3. Uji Kesamaan Variansi

Test for Equality of Variances Between Series
 Date: 05/17/15 Time: 08:58
 Sample: 1 120
 Included observations: 120

Method	df	Value	Probability
Bartlett	2	3.451965	0.1780
Levene	(2, 357)	2.454776	0.0873
Brown-Forsythe	(2, 357)	1.351643	0.2601

Category Statistics

Variable	Count	Std. Dev.	Mean Abs.	Mean Abs.
			Mean Diff.	Median Diff.
SERIES01	120	1.138578	0.946811	0.907298
SERIES02	120	0.968853	0.790770	0.762173
SERIES03	120	1.109682	0.927918	0.879720
All	360	1.075353	0.888500	0.849730

Bartlett weighted standard deviation: 1.074931

2. Analisis Korrelogram (ACF-PACF)

Kandidat model time series yang akan digunakan dilihat berdasarkan perilaku ACF dan PACF dari data.

Tabel 4. ACF dan PACF Sampel

Date: 05/17/15 Time: 08:40
 Sample: 1983M01 2012M12
 Included observations: 360

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.779	0.779	220.41	0.000
		2	0.356	-0.64...	266.45	0.000
		3	-0.07...	-0.12...	268.35	0.000
		4	-0.36...	-0.04...	316.97	0.000
		5	-0.48...	-0.13...	404.92	0.000
		6	-0.50...	-0.26...	499.79	0.000
		7	-0.46...	-0.16...	577.98	0.000
		8	-0.33...	0.019	619.16	0.000
		9	-0.09...	0.193	622.37	0.000
		1...	0.256	0.330	646.69	0.000
		1...	0.588	0.203	775.72	0.000
		1...	0.736	0.022	978.35	0.000
		1...	0.593	-0.20...	1110.3	0.000
		1...	0.250	-0.02...	1133.9	0.000
		1...	-0.11...	0.053	1138.6	0.000
		1...	-0.35...	0.011	1187.3	0.000
		1...	-0.46...	-0.05...	1268.5	0.000
		1...	-0.48...	-0.15...	1357.7	0.000
		1...	-0.44...	-0.04...	1434.6	0.000
		2...	-0.33...	0.023	1477.7	0.000
		2...	-0.11...	0.031	1482.6	0.000
		2...	0.222	0.152	1501.7	0.000
		2...	0.539	0.017	1614.1	0.000
		2...	0.674	-0.03...	1790.2	0.000
		2...	0.547	-0.04...	1906.5	0.000
		2...	0.215	-0.10...	1924.5	0.000
		2...	-0.13...	0.030	1931.7	0.000
		2...	-0.37...	-0.05...	1987.7	0.000
		2...	-0.47...	-0.03...	2076.9	0.000
		3...	-0.47...	-0.03...	2166.6	0.000
		3...	-0.42...	-0.04...	2239.1	0.000
		3...	-0.30...	0.073	2275.3	0.000
		3...	-0.06...	0.065	2277.1	0.000
		3...	0.244	-0.01...	2301.0	0.000
		3...	0.542	0.097	2418.9	0.000
		3...	0.670	0.023	2599.3	0.000

Berdasarkan Tabel 4 plot ACF tidak cut-off dan terlihat adanya pola musiman setiap lag 12. Oleh karena itu kemungkinan model yang akan digunakan adalah Seasonal ARIMA (p,d,q) (P,D,Q)S. Melihat perilaku ACF yang memiliki pola musiman (seasonal) maka perlu dilakukan differencing musiman 12.

Tabel 5. ACF dan PACF Sampel setelah Differencing Musiman 12

Date: 05/17/15 Time: 09:10
 Sample: 1983M01 2012M12
 Included observations: 348

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 0.712	0.712	177.73	0.000
		2 0.440	-0.13...	245.93	0.000
		3 0.237	-0.04...	265.69	0.000
		4 0.076	-0.07...	267.75	0.000
		5 -0.00...	0.018	267.76	0.000
		6 -0.00...	0.054	267.78	0.000
		7 0.009	0.014	267.81	0.000
		8 -0.00...	-0.06...	267.81	0.000
		9 -0.06...	-0.10...	269.28	0.000
		1... -0.15...	-0.12...	278.06	0.000
		1... -0.27...	-0.15...	304.76	0.000
		1... -0.36...	-0.12...	353.06	0.000
		1... -0.24...	0.263	375.06	0.000
		1... -0.12...	-0.01...	381.10	0.000
		1... -0.04...	-0.01...	381.74	0.000
		1... 0.018	-0.03...	381.85	0.000
		1... 0.050	0.029	382.76	0.000
		1... 0.002	-0.08...	382.76	0.000
		1... -0.05...	-0.01...	383.73	0.000
		2... -0.09...	-0.05...	386.72	0.000
		2... -0.14...	-0.15...	394.62	0.000
		2... -0.12...	0.031	400.00	0.000
		2... -0.11...	-0.14...	404.51	0.000
		2... -0.14...	-0.17...	412.81	0.000
		2... -0.11...	0.210	417.91	0.000
		2... -0.13...	-0.14...	424.73	0.000
		2... -0.14...	-0.02...	433.19	0.000
		2... -0.14...	-0.02...	441.16	0.000
		2... -0.07...	0.126	443.54	0.000
		3... 0.003	-0.03...	443.54	0.000
		3... 0.048	-0.03...	444.41	0.000
		3... 0.095	0.030	447.89	0.000
		3... 0.121	-0.09...	453.54	0.000
		3... 0.067	-0.06...	455.26	0.000
		3... 0.037	-0.03...	455.81	0.000
		3... 0.057	-0.02...	457.08	0.000

Berdasarkan differencing musiman di atas terlihat bahwa ACF cut-off setelah lag-3 dan PACF cut-off setelah lag-2 sehingga kandidat model yang digunakan adalah:

- ARIMA(2,0,1)(1,0,3)¹²
- ARIMA(1,0,1)(1,0,1)¹²,
- ARIMA(1,0,0)(1,0,3)¹²,
- ARIMA(1,0,0)(1,0,2)¹²,
- ARIMA(1,0,0)(1,0,1)¹²,
- ARIMA(1,0,0)(1,0,0)¹²,
- ARIMA(1,0,0)(2,0,2)¹²,
- ARIMA(1,0,0)(2,0,1)¹²,
- ARIMA(1,0,0)(2,0,0)¹²,
- ARIMA(0,0,1)(1,0,2)¹²,

- ARIMA(0,0,1)(1,0,1)¹²,
- ARIMA(0,0,1)(1,0,0)¹²,
- ARIMA(0,0,1)(2,0,2)¹²,
- ARIMA(0,0,1)(2,0,1)¹²,
- ARIMA(0,0,1)(2,0,0)¹²

3. Identifikasi Model

Pada tahap ini akan dilihat bagaimana signifikansi setiap model-model time series, korelasi dan distribusi dari residualnya (residual white noise).

Dari beberapa model di atas maka terdapat dua model yang memenuhi semua kriteria, yaitu:

1. ARIMA(1,0,1)(1,0,1)¹²

dengan $\theta_0 = 0.009855$, $\phi = 0.6715$,

$\Phi = 0.024196$, $\theta_1 = 0.137841$, $\Theta =$

-0.941793 , sehingga diperoleh

$$Y_t = 0.009855 + 0.6715 Y_{t-1} + 0.024196 Y_{t-12} - (0.6715)(0.024196) Y_{t-13} - 0.137841 e_{t-1} + 0.940726 e_{t-12} + (0.137841)(0.940726) e_{t-13}$$

2. ARIMA(1,0,0)(1,0,1)¹²

dimana $\theta_0 = 0.010039$, $\phi = 0.734220$,

$\Phi = 0.014893$, $\Theta = -0.940726$, se-

hingga diperoleh

$$Y_t = 0.010039 + 0.734220 Y_{t-1} + 0.014893 Y_{t-12} - (0.734220)(0.014893) Y_{t-13} + 0.940726 e_{t-12}$$

Forecasting

Forecasting dilakukan untuk tahun 2013 dengan menggunakan dua model terbaik, yaitu ARIMA(1,0,1)(1,0,1)¹² dan ARIMA(1,0,0)(1,0,1)¹² dan akan dibandingkan dengan data 197ctual paada tahun 2013 kemudian dihitung MSE (Mean Squared error) untuk setiap model.

Tabel 6. Identifikasi Model-model Seasonal ARIMA

No.	Model	AIC	Signifikansi parameter	Residual tidak berautokorelasi	Residual berdistribusi normal
1.	ARIMA(2,0,1)(1,0,3) ¹²	0.679070	signifikan	tidak	ya
2.	ARIMA(1,0,1)(1,0,1) ¹²	0.752506	signifikan	ya	ya
3.	ARIMA(1,0,0)(1,0,3) ¹²	0.697527	signifikan	tidak	ya
4.	ARIMA(1,0,0)(1,0,2) ¹²	0.763575	signifikan	tidak	ya
5.	ARIMA(1,0,0)(1,0,1) ¹²	0.758467	signifikan	ya	ya
6.	ARIMA(1,0,0)(1,0,0) ¹²	1.217458	signifikan	tidak	ya
7.	ARIMA(1,0,0)(2,0,2) ¹²	0.727629	signifikan	tidak	ya
8.	ARIMA(1,0,0)(2,0,1) ¹²	0.733783	signifikan	tidak	ya
9.	ARIMA(1,0,0)(2,0,0) ¹²	1.089230	signifikan	tidak	ya
10.	ARIMA(0,0,1)(1,0,2) ¹²	1.004536	signifikan	tidak	ya
11.	ARIMA(0,0,1)(1,0,1) ¹²	0.998584	signifikan	tidak	ya
12.	ARIMA(0,0,1)(1,0,0) ¹²	1.415179	signifikan	tidak	ya
13.	ARIMA(0,0,1)(2,0,2) ¹²	0.969290	signifikan	tidak	ya
14.	ARIMA(0,0,1)(2,0,1) ¹²	0.964614	signifikan	tidak	ya
15.	ARIMA(0,0,1)(2,0,0) ¹²	1.279127	signifikan	tidak	ya

Tabel 7. Perbandingan Data Aktual dan Forecasting

Periode	Data Aktual	Data Forecasting ARIMA(1,0,1)(1,0,1) ¹²	Data Forecasting ARIMA(1,0,0)(1,0,1) ¹²
Jan-13	30.555393	30.43172044	30.44142
Feb-13	30.698734	30.36424504	30.36846
Mar-13	30.619045	30.49970501	30.51268
Apr-13	30.431911	30.3186255	30.33205
May-13	29.995226	29.76671715	29.78104
Jun-13	29.495384	29.02190612	29.03603
Jul-13	28.600901	28.10643663	28.12147
Aug-13	28.080976	27.70145596	27.71609
Sep-13	28.429859	28.37346364	28.38403
Oct-13	30.13018	29.71363648	29.71581
Nov-13	30.437611	30.43955611	30.4426
Dec-13	30.478622	30.5753076	30.57614
MSE		0.08376852	0.07888096

Model yang memberikan MSE terkecil adalah $ARIMA(1,0,0)(1,0,1)^{12}$, sehingga model ini merupakan model yang terbaik.

SIMPULAN DAN SARAN

Simpulan

Dari hasil analisis deret waktu (*time series*) terhadap data suhu permukaan laut di selatan Jawa Timur dari Januari 1983 sampai Desember 2012 diperoleh :

1. Data mengikuti pola musiman dan model terbaik yang diperoleh adalah $ARIMA(1,0,0)(1,0,1)^{12}$ secara matematisnya yaitu:

$$Y_t = 0.010039 + 0.734220 Y_{t-1} + 0.014893 Y_{t-12} - (0.734220)(0.014893) Y_{t-13} + 0.940726 e_{t-12}$$

2. Hasil forecasting untuk tahun 2013 dengan menggunakan model $ARIMA(1,0,0)(1,0,1)^{12}$ didapatkan mendekati data aktualnya dengan MSE sebesar 0,07888096

Saran

Data yang digunakan pada penelitian ini tidak terdapat pencilan, sehingga penelitian selanjutnya diharapkan agar menggunakan data yang memiliki pencilan.

REFERENSI

- Box and Jenkins. 2008. *Time Series Analysis: Forecasting and Control Revised Edition*, Oakland, California :
- Holden-Day. Cryer, Jonathan D. and Kung-Sik Chan. 2008. *Time Series Analysis with Applications in R*, Second Edition, Iowa City : Springer.
- Lidiawati, L., 2014, *Dinamika dan Variabilitas Upwelling di Perairan Selatan Jawa Timur*, Disertasi, Program Studi Doktor Sains Kebumihan, Institut Teknologi Bandung.
- Ningsih, N. S., S. Hadi, I. Sofian, Kunarso, dan F. Hanifah, 2012, *Kajian Dampak Perubahan Iklim Terhadap Dinamika Upwelling sebagai Dasar untuk Memperkirakan Pola Migrasi Ikan Tuna di Perairan Selatan Jawa – Nusa Tenggara Barat dengan Menggunakan Model Transpor Temperatur Laut*. Laporan Riset dan Inovasi KK. Institut Teknologi Bandung, Bandung.