

Peramalan Curah Hujan Sebagai Pendukung Kalender Tanam Padi di Pos Kedungadem Bojonegoro Menggunakan ARIMA, *Support Vector Regression* dan *Genetic Algorithm-SVR*

Kiki Wulan Suci dan Irhamah

Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember (ITS)
Jl. Arief Rahman Hakim, Surabaya 60111 Indonesia
e-mail : kwulansuci@gmail.com dan irhamahn@gmail.com

Abstrak—Kabupaten Bojonegoro dikenal sebagai lumbung padi di Jawa Timur. Sebanyak 33,31% lahan Bojonegoro digunakan sebagai lahan sawah. Pada beberapa tahun terakhir produksi padi Kabupaten Bojonegoro mengalami fluktuasi yang salah satunya disebabkan oleh iklim ekstrem. Iklim ekstrem dapat mengakibatkan musim kemarau panjang serta adanya hujan ekstrem yang mengakibatkan petani mengalami kerugian. Oleh karena itu, keberhasilan produksi padi sangat bergantung pada informasi mengenai data curah hujan yang tersusun dalam kalender tanam. Dalam penelitian ini dilakukan peramalan curah hujan dasarian di Pos Kedungadem menggunakan metode ARIMA, *Support Vector Regression* (SVR) dan *Genetic Algorithm-Support Vector Regression* (GA-SVR). Berdasarkan hasil analisis metode GA-SVR menghasilkan peramalan yang lebih akurat yang ditunjukkan oleh nilai RMSE dan SMAPE yang lebih kecil dibandingkan metode lain. Hasil kalender tanam padi berdasarkan hasil *fore-cast* pada bulan Juli 2016-Desember 2016 menunjukkan bahwa kebutuhan air untuk penanaman padi sawah tidak dapat tercukupi, dimana hasil *forecast* menunjukkan rata-rata curah hujan kurang dari 145 mm. Petani dapat mengganti dengan menanam palawija yang membutuhkan sedikit air. Jika petani tetap melakukan tanam padi, maka petani dan pemerintah harus memastikan tersedianya cadangan air dengan menggunakan sistem pengairan dan tidak mengandalkan pada turunnya hujan.

Kata Kunci—ARIMA, Curah Hujan, *Genetic Algorithm*, *Support Vector Regression* (SVR).

I. PENDAHULUAN

KABUPATEN Bojonegoro dikenal sebagai lumbung padi di Jawa Timur. Sebanyak 33,31% lahan Bojonegoro digunakan sebagai lahan sawah [1]. Pada beberapa tahun terakhir produksi padi Kabupaten Bojonegoro mengalami fluktuasi. Salah satu penyebab fluktuasi produksi padi adalah iklim ekstrem. Iklim ekstrem dapat mengakibatkan musim kemarau yang panjang serta adanya hujan ekstrem yang mengakibatkan petani mengalami kerugian [2].

Penelitian ini mengambil studi kasus di wilayah selatan Bojonegoro yaitu di pos pengamatan curah hujan Kedungadem. Kedungadem merupakan kecamatan di Bojonegoro Selatan dengan tipe sawah dominan adalah sawah tadah hujan yang mengandalkan turunnya air hujan dalam produksi padi. Pada kecamatan Kedungadem prosentase luas sawah tadah hujan mencapai 98% dari total lahan sawah [3]. Oleh

karena itu, keberhasilan produksi padi sangat bergantung pada informasi mengenai data curah hujan. Peramalan curah hujan yang akurat akan dapat mengurangi kerugian petani akibat kegagalan panen.

Peramalan curah hujan menggunakan ARIMA telah beberapa dilakukan. Namun, penggunaan ARIMA sering kali belum mampu memberikan residual yang berdistribusi normal, belum mampu mengakomodasi adanya data *outlier* serta tidak mampu memodelkan *time series* yang nonlinier [4]. Padahal data curah hujan seringkali menghasilkan data yang non linier. Oleh karena itu, akan digunakan metode *Support Vector Regression* (SVR). SVR dapat menangkap kasus non linieritas dengan menambahkan fungsi kernel. Pada penelitian ini digunakan fungsi kernel *Radial Basis Function* (RBF) yang sering digunakan dan menghasilkan kinerja yang lebih bagus [5]. Metode SVR dapat mengatasi kasus *overfitting* serta SVR adalah metode yang tidak memerlukan asumsi. Metode SVR juga dapat memberikan hasil yang baik walaupun hanya dengan data yang sedikit [6] [7]. Salah satu permasalahan yang ketika menggunakan SVR dengan menggunakan algoritma *grid search* adalah penentuan parameter model yang optimal, dika-rekan *grid search* hanya memilih parameter yang optimal pada *range* parameter yang ditetapkan. Oleh karena itu, penelitian ini akan digunakan *Genetic Algorithm* (GA) untuk optimasi nilai parameter SVR dengan harapan dapat menghasilkan nilai akurasi yang lebih bagus.

Penelitian dengan menggunakan metode SVR dan GA pernah dilakukan oleh Msiza, Nelwamondo dan Marwala (2008) untuk meramalkan permintaan air dengan hasil penelitian yaitu hasil prediksi yang dihasilkan oleh ANN dan SVR memberikan hasil error yang lebih kecil [8]. Tahun 2011 Gu, Zhu dan Jiang menggunakan GA-SVR untuk meramalkan harga perabotan rumah dimana metode GA-SVR menghasilkan peramalan yang lebih akurat dibandingkan menggunakan *grid algorithm* [9]. Penelitian kali ini akan menerapkan metode ARIMA, SVR dan GA-SVR untuk meramalkan curah hujan di Kabupaten Bojonegoro. Kriteria perbandingan akurasi ramalan yang digunakan adalah nilai RMSE dan SMAPE. Berdasarkan pada hasil peramalan periode selanjutnya dengan metode terbaik, akan dibuat susunan kalender tanam padi untuk tipe sawah tadah hujan.

II. TINJAUAN PUSTAKA

A. Autoregressive Moving Average (ARIMA)

Metode peramalan yang biasanya sangat baik digunakan untuk melakukan peramalan jangka pendek adalah metode ARIMA. Secara umum model ARIMA (p, d, q) di-tulis sebagai berikut [10].

$$\phi_p(B)(1-B)^d Y_t = \theta_0 + \theta_q(B) a_t \quad (1)$$

Sebelum menggunakan ARIMA terdapat asumsi yang harus dipenuhi yaitu data harus stasioner terhadap mean dan varians [10] [11]. Setelah data dapat memenuhi asumsi tersebut maka dilakukan identifikasi model ARIMA dilakukan dengan melihat plot ACF dan PACF [10].

Tahapan selanjutnya adalah melakukan estimasi parameter. Dalam penelitian ini, digunakan metode estimasi parameter Conditional Least Square (CLS). Kemudian di-lakukan uji signifikansi parameter dengan menggunakan uji t [12], jika parameter tidak signifikan maka dilakukan iden-tifikasi model ARIMA kembali. Setelah didapatkan para-meter yang signifikan, kemudian dilakukan pengujian asu-msi residual yaitu white noise dan distribusi normal. Pe-menuhan asumsi residual yang white noise menggunakan uji L-jung Box dan pemenuhan asumsi normalitas residual menggunakan uji Kolmogorov-Smirnov [10] [13].

B. Support Vector Regression (SVR)

SVR merupakan bagian dari Support Vector Machine (SVM) yang diperkenalkan oleh Vapnik (1995) yang digunakan untuk kasus regresi dan prediksi. Konsep SVR didasarkan pada pada risk minimization, yaitu mengestimasi suatu fungsi dengan cara meminimalkan batas atas dari generalization error, sehingga SVR mampu mengatasi over-fitting [7]. Pada kasus nonlinier, fungsi regresi setelah transformasi menjadi seperti berikut:

$$f(x) = w \cdot \Phi(x) + b \quad (2)$$

Dengan w merupakan vector pembobot, $\Phi(x)$ merupakan fungsi yang memetakan x dalam suatu dimensi dan b merupakan bias. Untuk mengevaluasi seberapa baik fungsi regresi, maka digunakan fungsi ϵ -insetive loss sebagai berikut.

$$L_\epsilon(y, f(x)) = \begin{cases} 0 & \text{untuk } |y - f(x)| \leq \epsilon \\ |y - f(x)| - \epsilon & \text{dan sebaliknya} \end{cases} \quad (3)$$

Fungsi ϵ -insetive loss digunakan untuk mengukur risiko empiris, risiko empiris diukur berdasarkan persamaan (3), selisih output/ target dengan hasil estimasi. Oleh karena itu parameter ϵ harus diatur untuk meminimalisir risiko empiris. Disamping itu, juga harus meminimalisir norma Euc-lidean dari beban yang linier $\|w\|$ yang mana berhubungan dengan kemampuan generalisasi dari model SVR yang di training [6], sehingga permasalahan regresi dapat dinyatakan seperti masalah optimasi kuadratik berikut ini:

$$L(w, \xi) = \frac{1}{2} \|w\|^2 + C[\sum_{i=1}^n (\xi_{2i} + \xi_{2i}')] \quad (4)$$

$$\text{subject to } \begin{cases} y_i - w * \Phi(x_i) - b \leq \epsilon + \xi_i \\ w * \Phi(x_i) + b - y_i \leq \epsilon + \xi_i^* \\ \phi_i, \xi_i^* \geq 0 \end{cases} \quad (5)$$

Koefisien C menyatakan penalti yang mendeterminasikan trade-off antara keempirisan dengan kesalahan generalisasi, sehingga besarnya nilai C tersebut perlu diatur [6]. Untuk menyelesaikan permasalahan pada optimasi kuadratik pada persamaan (4), dapat menggunakan dual Lagrangian:

$$f(x_i) = (w \Phi(x_i) + b) = \sum_{j=1}^n \alpha_j K(x_i, x_j) + b \quad (6)$$

dimana $K(x_i, x_j)$ merupakan fungsi kernel. Jika fungsi kernel yang digunakan Radial Basis Function (RBF) maka per-lu mengatur parameter γ sebagai berikut.

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (7)$$

Berdasarkan uraian mengenai SVR dapat dilihat bahwa terdapat tiga parameter yaitu C, ϵ , γ dan fungsi kernel sangat penting dalam menentukan tingkat keakuratan dari prediksi.

C. Genetic Algorithm (GA)

Genetic Algorithm (GA) pertama kali dikembangkan oleh John Holland dari Universitas Michigan tahun 1975. GA adalah sebuah algoritma optimasi metaheuristik yang berdasarkan populasi solusi dan menggunakan mekanisme spesifik yang terinspirasi dari evolusi secara biologi seperti individu (kromosom), reproduksi, mutasi, rekombinasi, se-leksi, ketahanan fitness [14].

Sebuah solusi yang dibangkitkan dalam GA disebut kromosom, sedangkan kumpulan kromosom-kromosom disebut populasi. Berikut ilustrasi kromosom dengan tiga gen pada metode SVR dengan fungsi kernel RBF.

Parameter	Cost	γ	ϵ
Kromosom	9,01	0,05	1,01

Gambar 1. Ilustrasi Kromosom dengan Tiga Gen

Fitness individu dalam Genetic Algorithm adalah nilai fungsi objektif untuk fenotipe. Nilai fitness tidak hanya menunjukkan bagaimana solusi yang baik, tetapi juga berhubungan dengan seberapa dekat kromosom pada solusi optimum. Nilai fitness yang sering digunakan adalah RMSE [15].

Berikut adalah beberapa proses dalam Genetic Algorithm [15].

1. Seleksi adalah suatu proses untuk mendapatkan kromosom-kromosom berkualitas pada generasi berikutnya. Pemilihan kromosom berkualitas dalam seleksi dilakukan dengan memperhatikan nilai fitness. Menurut Gen dan Cheng (2000), metode seleksi yang paling dikenal ada-lah seleksi Roulette Wheel [15].
2. Crossover adalah suatu proses mengawinkan dua kromosom orang tua (pindah silang). Sebelum dilakukan crossover ditentukan probabilitas crossover (P_c), dimana nilai P_c mendekati 1. Crossover dilakukan jika nilai bilangan random kurang dari P_c .
3. Mutasi adalah suatu proses mengubah susunan gen dalam suatu kromosom untuk menghindari premature kon-vergen ketika nilai fitnessnya belum optimal. Sebelum dilakukan mutasi ditentukan probabilitas mutasi (P_m), dimana nilai P_m mendekati nol. Mutasi dilakukan jika nilai bilangan random kurang dari P_m .
4. Elitism adalah proses pengopian individu baru agar individu yang memiliki estimasi parameter yang menghasilkan nilai fitness terbaik nilai fitness terbaik tidak hilang selama proses evolusi.

D. Pemilihan Model Terbaik

Pemilihan model terbaik melalui pendekatan in sample dan out sample dengan menggunakan RMSE (Root Mean Square) dan SMAPE yang dirumuskan sebagai berikut [10].

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} = \sqrt{MSE} \quad (13)$$

$$SMAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{(Y_t + \hat{Y}_t)/2} \right| \times 100 \quad (14)$$

n = banyaknya data.

Y_t = Data *actual* pada waktu ke- t .

\hat{Y}_t = Data *actual* pada waktu ke- t .

E. Curah Hujan

Curah hujan merupakan ketinggian air hujan yang terkumpul dalam tempat yang datar, tidak menguap, tidak meresap, dan tidak mengalir. Satuan yang digunakan dalam mengukur curah hujan adalah millimeter atau mm. Berdasarkan intensitas, curah hujan dibedakan menjadi 3 yaitu hujan sedang yang berada diantara 20 dan 50 mm perhari, hujan lebat berada diantara 50 dan 100 mm perhari, dan hujan sangat lebat berada diatas 100 mm perhari [16].

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan dalam penelitian ini adalah data curah hujan dasarian yang berasal dari Pos pengamatan curah hujan di Kedungadem Kabupaten Bojonegoro yang di-ambil di BMKG Karangploso Malang. Periode data adalah mulai bulan Januari 2000 hingga Juni 2016. Adapun data periode bulan Januari 2000 hingga Desember 2015 digunakan sebagai data *in sample* dan data bulan Januari 2016- Juni 2016 digunakan sebagai data *out sample*.

B. Variabel Penelitian

Berikut ini adalah variabel bebas dan variabel tak bebas yang digunakan dalam penelitian ini.

1. Variabel tak bebas (Y_t) adalah data curah hujan dasarian yang dijadikan data *actual/target*.
2. Variabel bebas (X) adalah data curah hujan dasarian pada waktu sebelumnya ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-k}$).

C. Langkah Analisis

Langkah analisis yang digunakan dalam penelitian ini sebagai berikut.

1. Melakukan analisis statistika deskriptif dan plot data curah hujan dasarian di Pos Kedungadem.
2. Analisis ARIMA, peramalan curah hujan dengan ARIMA dilakukan sesuai prosedur sebagai berikut
 - a. Identifikasi kestasioneran data berdasarkan plot *Box-Cox* dan Uji *Augmented Dickey Fuller* (ADF). Jika data belum stasioner dalam varians maka dilakukan transformasi dan jika belum stasioner dalam *mean*, maka dilakukan *differencing*.
 - b. Identifikasi model ARIMA sementara berdasarkan plot ACF dan PACF.
 - c. Melakukan estimasi parameter dan pengujian signifikansi parameter model. Jika parameter belum signifikan maka kembali ke langkah 2b.
 - d. Melakukan *diagnostic checking* pada residual yakni pengujian asumsi *white noise* dan normalitas residual. Jika terdapat asumsi yang belum terpenuhi maka kembali ke langkah 2b.
3. Uji non linieritas Terasvirta

4. Analisis dengan metode *Support Vector Regression* (SVR).
 - a. Menyusun model regresi dengan variable bebas (X) adalah data curah hujan dasarian pada waktu sebelumnya yang diperoleh dari penguraian model ARI-MA yang memiliki parameter signifikan pada langkah 2.
 - b. Menentukan fungsi kernel yang digunakan.
 - c. Menentukan *range* nilai parameter C , ϵ dan γ .
 - d. Melakukan pemodelan dengan SVR berdasarkan pada *range* nilai parameter.
 - e. Melakukan pengecekan terhadap solusi yang didapatkan. Model terbaik adalah model yang memiliki nilai parameter paling optimum yang memiliki nilai *error* terkecil, jika kondisi ini belum terpenuhi maka proses diulang dari langkah 4d.
 - f. Melakukan peramalan berdasarkan model SVR terbaik serta menghitung RMSE dan SMAPE.
5. Analisis dengan metode *Genetic Algorithm-Support Vector Regression* (GA-SVR).
 - a. Menyusun kromosom dengan membangkitkan 100 kromosom. Kromosom yang dibangkitkan terdiri dari tiga gen yang menunjukkan parameter SVR dengan menggunakan fungsi kernel RBF. Nilai inisial kromosom berasal dari nilai parameter model SVR yang telah didapatkan pada langkah 4.
 - b. Menentukan nilai *fitness*.
 - c. Melakukan proses seleksi *roulette wheel*.
 - d. Melakukan proses pindah silang (*crossover*).
 - e. Melakukan proses mutasi.
 - f. Melakukan proses *elitism*.
 - g. Melakukan pergantian populasi lama dengan generasi baru dengan cara memilih kromosom terbaik berdasarkan nilai *fitness*.
 - h. Melakukan pengecekan terhadap solusi yang telah didapatkan. Solusi dikatakan telah mencapai kriteria apabila nilai *fitness* terbaik telah konvergen jika kondisi ini belum terpenuhi maka proses *Genetic Algorithm* diulang dari langkah 5c.
 - i. Melakukan estimasi parameter dengan metode GA kemudian memasukkan parameter ke algoritma SVR dan menghitung RMSE dan SMAPE
6. Melakukan perbandingan metode ARIMA, SVR dan GA-SVR berdasarkan nilai RMSE dan SMAPE untuk memilih metode peramalan terbaik.
7. Melakukan *forecast* 6 bulan selanjutnya dengan metode terbaik dan dilakukan penyusunan kalender tanam padi berdasarkan hasil *forecast* yang digunakan pada tipe sawah tadah hujan

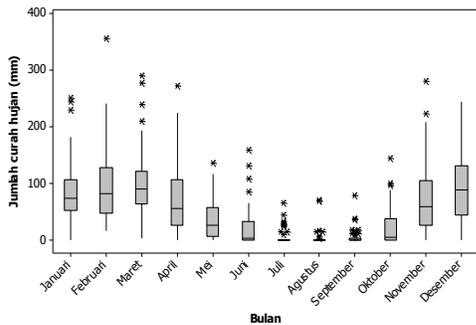
IV. HASIL DAN PEMBAHASAN

A. Deskripsi Curah Hujan Pos Kedungadem

Untuk melihat kondisi curah hujan tiap bulan pada Pos Kedungadem dapat dilihat secara visual dengan *boxplot* pada Gambar 2. berikut.

Gambar 2. menunjukkan bahwa rata-rata curah hujan yang cukup tinggi terjadi pada bulan Januari, Februari, Maret, April, November dan Desember. Pada bulan Juli, Agustus dan September memiliki rata-rata curah hujan yang rendah, yang mengindikasikan bahwa pada bulan tersebut terjadi musim

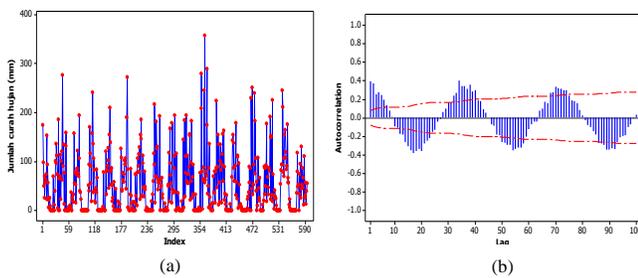
kemarau dan musim pancaroba. Bulan Januari, Februari, Maret, April dan November terdapat outlier yang tinggi yang menunjukkan terjadinya hujan dengan intensitas yang tinggi.



Gambar 2. Boxplot Curah Hujan Pos Kedungadem Tiap Bulan

B. Peramalan Curah Hujan Menggunakan ARIMA

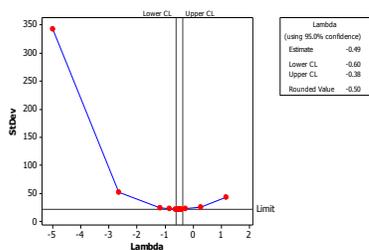
Pada peramalan dengan ARIMA langkah pertama yang harus dilakukan adalah identifikasi terhadap pola data untuk mengetahui apakah data sudah stasioner terhadap varians dan mean. Dalam menggunakan ARIMA, perlu melihat *ti-me series plot* dari data tersebut untuk mengetahui pola dan karakteristik dari data yang ditunjukkan pada Gambar 3.(a) berikut.



Gambar 3. Time Series Plot (a) dan Plot ACF (b) Data Curah Hujan Pos Kedungadem

Berdasarkan Gambar 3.(a) menunjukkan bahwa curah hujan dasarian di Pos Kedungadem belum stasioner terhadap mean dan varians karena fluktuasi antara titik satu dengan lainnya bervariasi. Plot ACF pada Gambar 3.(b) menunjukkan data memiliki pola musiman (*seasonal*).

Langkah selanjutnya adalah dilakukan analisis dengan melihat plot *Box-Cox* untuk melihat apakah data sudah stasioner dalam varians seperti pada Gambar 4.

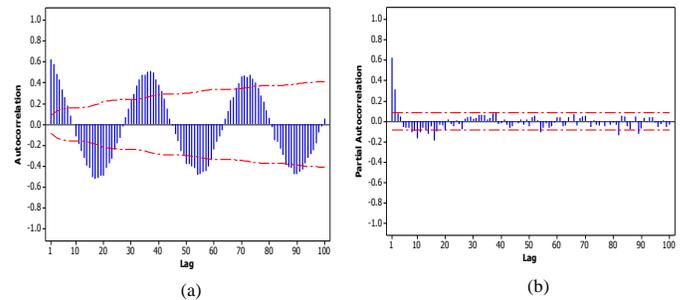


Gambar 4. Box Plot Curah Hujan Pos Kedungadem

Berdasarkan Gambar 4. diketahui bahwa data curah hujan dasarian di Pos Kedungadem belum stasioner terhadap varians. Hal ini ditunjukkan dengan nilai *Lower CL* dan *Upper CL* yang

belum melewati angka 1 atau nilai *rounded va-lue* (λ) yang belum bernilai satu. Dikarenakan nilai λ sebesar -0,5 maka data curah hujan akan ditransformasikan $1/\sqrt{Y_t}$ untuk menstabilkan varians.

Selanjutnya dilakukan pengujian stasioneritas data terhadap mean dengan menggunakan uji *Augmented Dickey Fuller* (ADF). Pada uji ADF data hasil transformasi didapatkan *P-value* sebesar 0,01. Nilai ini menunjukkan *Pva-lue* kurang dari α , sehingga tolak H_0 , maka disimpulkan bahwa data transformasi sudah stasioner terhadap mean. Langkah selanjutnya dilakukan identifikasi model ARIMA sementara berdasarkan pada plot ACF dan plot PACF data hasil transformasi sebagai berikut.



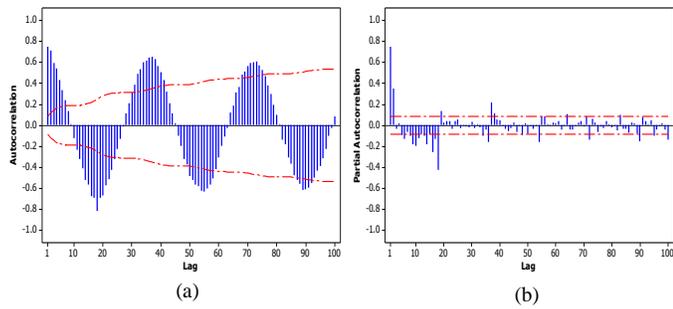
Gambar 5. Plot ACF (a) dan Plot PACF (b) Data Hasil Transformasi

Plot ACF pada Gambar 5. menunjukkan bahwa plot memiliki pola *dies down*. Pada plot PACF diperoleh informasi bahwa terdapat beberapa lag yang keluar dari batas *significance limit* yang menunjukkan bahwa lag tersebut signifikan. Berdasarkan plot ACF dan PACF diperoleh beberapa dugaan model ARIMA seperti pada Tabel 1.

Tabel 1. menunjukkan bahwa tidak terdapat model dugaan ARIMA yang dapat memenuhi asumsi residual *white noise* dan berdistribusi normal. Dikarenakan tidak ada model yang memenuhi asumsi residual, maka analisis selanjutnya akan dilakukan *differencing* pada data hasil transformasi. Pada *time series plot* data transformasi pada Gambar 3. (a) dihasilkan plot yang belum stasioner terhadap mean serta pada Gambar 3. (b) menunjukkan bahwa pola data merupakan pola musiman. Oleh karena itu akan dilakukan *differencing* musiman lag 18 pada data hasil transformasi. Plot ACF dan plot PACF setelah dilakukan *differencing* musiman 18 terdapat pada Gambar 6.

Tabel 1.
Beberapa Model Dugaan ARIMA Tanpa *Differencing*

Model ARIMA	Signifikansi Parameter	Residual White Noise	Normalitas Residual
ARIMA (2,0,0)	Signifikan	Tidak Terpenuhi	Tidak normal
ARIMA ([1,2,8],0,0)	Tidak Signifikan	-	-
ARIMA ([1,2,17],0,0)	Signifikan	Tidak Terpenuhi	Tidak normal
ARIMA ([1,2,10,16],0,0)	Tidak Signifikan	-	-
ARIMA ([8,10],0,0)	Signifikan	Tidak Terpenuhi	Tidak normal
ARIMA ([1,8],0,0)	Signifikan	Tidak Terpenuhi	Tidak normal



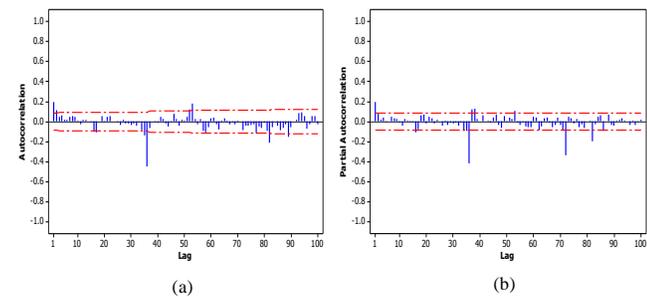
Gambar 6. Plot ACF (a) dan Plot PACF (b) Data Transformasi Setelah Differencing Musiman 18

Plot ACF pada Gambar 6.(a) memiliki pola *dies down*. Gambar 6.(b) menunjukkan plot PACF memiliki beberapa lag yang keluar dari batas *significance limit*. Berikut beberapa model dugaan ARIMA berdasarkan plot ACF dan plot PACF.

Tabel 2. Model Dugaan ARIMA dengan Differencing Musiman 18

Model ARIMA	Signifikansi Parameter	Residual White Noise	Normalitas Residual
ARIMA (2,0,0) (2,1,0) ¹⁸	Signifikan	Tidak Terpenuhi	Normal
ARIMA ([1,2,16],0,0) (3,1,0) ¹⁸	Signifikan	Tidak Terpenuhi	Normal
ARIMA ([1,16,37],0,0) (3,1,0) ¹⁸	Signifikan	Tidak Terpenuhi	Normal
ARIMA ([1,2,6],0,0)(3,1,1) ¹⁸	Tidak Signifikan	-	-

Tabel 2. menunjukkan tidak terdapat model ARIMA di Pos Kedungadem dengan differencing musiman 18 yang dapat memenuhi asumsi residual yang *white noise*. Untuk analisis selanjutnya akan digunakan differencing musiman 36 terhadap data hasil transformasi, karena diduga data curah hujan dasarian musiman selama 12 bulan. Plot ACF dan plot PACF data transformasi setelah differencing musiman 36 terdapat pada Gambar 7.



Gambar 7. Plot ACF (a) dan Plot PACF (b) Data Transformasi Setelah Differencing Musiman 36

Plot ACF dan PACF pada Gambar 6. menunjukkan terdapat beberapa lag yang keluar dari batas *significance limit*. Berdasarkan plot ACF dan PACF tersebut dapat ditentukan beberapa model ARIMA sementara seperti pada Tabel 3.

Tabel 3. Model Dugaan ARIMA dengan Differencing Musiman 36

Model ARIMA	Signifikansi Parameter	Residual White Noise	Normalitas Residual
ARIMA (0,0,1) (0,1,1) ³⁶	Signifikan	Tidak Terpenuhi	Tidak Normal
ARIMA (0,0,2) (0,1,1) ³⁶	Signifikan	Terpenuhi	Tidak Normal
ARIMA (1,0,1)(0,1,1) ³⁶	Signifikan	Terpenuhi	Tidak Normal
ARIMA ([1,16,37],0,0) (0,1,1) ³⁶	Tidak Signifikan	-	-
ARIMA (1,0,[1,52])(0,1,1) ³⁶	Signifikan	Terpenuhi	Tidak Normal

Dari Tabel 3. didapatkan kesimpulan bahwa dari 5 model dugaan ARIMA terdapat 3 model yang memenuhi asumsi residual yang *white noise*. Namun, dari 3 model tersebut tidak ada yang memenuhi asumsi normalitas residual, sehingga akan dilakukan pendeteksian outlier. Namun, hasil pemodelan dengan penambahan outlier tidak memberikan hasil residual yang berdistribusi normal, dimana *P-value* se-besar <0,0100, sehingga tolak *H₀*. Histogram dari residual cenderung lebih runcing (leptokurtik) serta nilai kurtosisnya bernilai positif yang menunjukkan residual model ARIMA outlier tidak berdistribusi normal. Untuk mengetahui model ARIMA yang terbaik, maka menggunakan kriteria kebaikan model *in sample* dan *out sample* yang dt pada Tabel 4.

Tabel 4. Kriteria Kebaikan Model ARIMA

Model ARIMA	In sample		Out sample
	RMSE	RMSE	SMAPE
ARIMA (0,0,2)(0,1,1) ³⁶	58,0898	37,3650	45,5300
ARIMA (1,0,1)(0,1,1) ³⁶	57,9267	36,8041	45,6412
ARIMA (1,0,[1,52])(0,1,1) ³⁶	57,7702	33,9779	43,5077

A. Peramalan Curah Hujan Menggunakan SVR

Tahap awal sebelum masuk pada metode SVR yakni dilakukan pengujian nonlinieritas. Pengujian nonlinieritas digunakan untuk memilih fungsi kernel pada metode SVR. Dari uji Terasvirta didapatkan *P-value* sebesar $2,931.10^{-11}$, sehingga *P-value* kurang dari α ($\alpha=5\%$) yang berarti data curah hujan dasarian di Pos Kedungadem bersifat nonlinier, oleh karena akan digunakan fungsi kernel *Radial Basis Function* (RBF).

Pada peramalan dengan SVR akan menggunakan *lag-lag* input yang didapat dari penguraian model ARIMA pada analisis sebelumnya yang sudah memiliki parameter yang signifikan yang berjumlah 11 model. *Lag input* tersebut digunakan sebagai input (X) pada peramalan dengan menggunakan metode SVR.

Metode SVR dengan menggunakan fungsi kernel RBF terdapat 3 parameter yaitu C, γ dan ϵ . Namun, untuk mendapatkan hasil yang optimal parameter-parameter tersebut perlu diatur nilainya. Dari beberapa percobaan yang telah dilakukan didapatkan hasil yakni parameter C yang optimal pada range 0,01-10, parameter ϵ pada range 0,5-1 dan parameter γ pada range 0,05-0,1. Selanjutnya setelah didapatkan range parameter yang optimal, maka range tersebut akan digunakan untuk memodelkan ARIMA yang telah memiliki parameter yang signifikan dengan metode SVR. Berikut merupakan hasil perhitungan kebaikan model dengan menggunakan metode SVR.

Tabel 5.
Kriteria Kebaikan Model SVR

Model Input	RMSE in sample	RMSE in sample	SMAPE out sample
$f(Y_{t-1}, Y_{t-2})$	53,4628	36,2386	41,2593
$f(Y_{t-1}, Y_{t-2}, Y_{t-17})$	51,4148	36,8883	44,2713
$f(Y_{t-8}, Y_{t-10})$	61,3685	38,5177	40,5517
$f(Y_{t-1}, Y_{t-8})$	55,3269	38,7053	45,0523
$f(Y_{t-1}, Y_{t-2}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-54}, Y_{t-55}, Y_{t-56})$	48,0288	36,3635	40,1549
$Y_{t-1}, Y_{t-2}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-56}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-74}, Y_{t-88}$	42,8774	44,0174	44,8926
$Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109}$	45,1607	36,5999	39,3409
Y_{t-36}	57,1360	32,1855	39,9169
Y_{t-36}	57,1360	32,1855	39,9169
$Y_{t-1}, Y_{t-36}, Y_{t-37}$	52,7867	41,2816	48,5682
$Y_{t-1}, Y_{t-36}, Y_{t-37}$	52,7867	41,2816	48,5682

B. Peramalan Curah Hujan Menggunakan GA-SVR

Setelah dilakukan peramalan menggunakan SVR, maka perlu dilakukan optimasi parameter SVR agar dihasilkan peramalan yang memiliki akurasi yang lebih baik. Pada GA-SVR akan digunakan 5 model SVR terbaik yang dihasilkan pada analisis sebelumnya. Langkah awal pada *Genetic Algorithm* adalah menginisialisasi kromosom sebanyak 100, dimana nilai inisial pada kromosom didapatkan dari parameter model SVR yang didapatkan dari analisis sebelumnya. Kromosom yang digunakan terdiri dari tiga gen yang menunjukkan *Cost* (C), γ dan ϵ . Berikut adalah ilustrasi kromosom dengan tiga gen.

	Cost	γ	ϵ
Kromosom	1,01	0,05	0,5

Gambar 8. Ilustrasi Kromosom Parameter SVR dengan Tiga Gen

Tabel 6.
Hasil Optimasi GA pada Model SVR

Model	RMSE in sample	RMSE out sample	SMAPE out sample
$f(Y_{t-1}, Y_{t-2})$	53,4059	36,2795	41,1911
$f(Y_{t-1}, Y_{t-2}, Y_{t-17})$	51,1191	36,3896	43,5163
$f(Y_{t-1}, Y_{t-2}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-54}, Y_{t-55}, Y_{t-56})$	47,7961	36,1657	40,4743
$f(Y_{t-1}, Y_{t-2}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-20}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-38}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-74}, Y_{t-88})$	42,8520	41,1462	44,3972
$f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$	42,9440	35,2800	39,1441

Nilai *fitness* yang digunakan pada penelitian ini adalah nilai kesalahan dalam peramalan yaitu RMSE. Nilai *fitness* merupakan acuan dalam tahapan *Genetic Algorithm* seperti melakukan proses seleksi, pindah silang (*crossover*), mutasi dan *elitism* sehingga didapatkan nilai *fitness* terbaik. Fungsi objektif yang ingin dicapai pada penelitian ini adalah untuk

meminimumkan nilai RMSE pada pada model SVR. Penelitian ini akan menggunakan probabilitas *crossover* (P_c) sebesar 0,8 dan probabilitas mutasi (P_m) sebesar 0,01. Tabel 6. menunjukkan hasil optimasi 5 model SVR terbaik yang didapatkan pada langkah sebelumnya.

C. Perbandingan Model Terbaik

Hasil perbandingan model terbaik dari masing-masing metode peramalan ditunjukkan pada Tabel 7.

Tabel 7.
Perbandingan Model Terbaik

Metode	Model	RMSE in sample	RMSE out sample	SMAPE out sample
ARIMA	ARIMA (1,0,[1,52]) (0,1,1) ³⁶	57,7702	33,9779	43,5077
SVR	$f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$	45,1607	36,5999	39,3409
GA-SVR	$f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$	42,9440	35,2800	39,1441

Tabel 7. menunjukkan bahwa model terbaik untuk memramalkan curah hujan Pos Kedungadem adalah model *input* $f(Y_{t-1}, Y_{t-16}, Y_{t-18}, Y_{t-19}, Y_{t-34}, Y_{t-36}, Y_{t-37}, Y_{t-52}, Y_{t-54}, Y_{t-55}, Y_{t-70}, Y_{t-72}, Y_{t-73}, Y_{t-88}, Y_{t-91}, Y_{t-109})$ yang di-dapatkan dari metode GA-SVR. Model ini memiliki kete-patan akurasi yang paling tinggi dibandingkan metode lain-nya. Tabel 8. menunjukkan *forecast* curah hujan 18 periode selanjutnya yakni pada bulan Juli 2016 sampai dengan De-seMBER 2016 menggunakan metode GA-SVR pada Pos Ke-dungadem.

Tabel 8.
Forecast 6 bulan Selanjutnya

No	Bulan	Forecast	No	Bulan	Forecast
1	Juli	36,2111	10	Oktober	13,2873
2	Juli	22,6570	11	Oktober	20,3639
3	Juli	26,7335	12	Oktober	29,5119
4	Agustus	13,8875	13	November	44,3715
5	Agustus	23,2287	14	November	67,2780
6	Agustus	19,4592	15	November	95,2588
7	September	12,0215	16	Desember	74,9463
8	September	21,9582	17	Desember	72,3197
9	September	17,3114	18	Desember	55,2144

D. Kalender Tanam Padi Bulan Juli 2016-Desember 2016

Hasil *forecast* curah hujan pada Tabel 8. menunjukkan bahwa curah hujan selama 6 bulan tersebut antara 50-100 mm dan kurang dari 50 mm. Susunan kalender tanam padi untuk wilayah Kedungadem Kabupaten Bojonegoro bulan Juli 2016 hingga Desember 2016 terdapat pada Gambar 9.

Berdasarkan Gambar 9. didapatkan informasi yaitu kebutuhan air untuk penanaman padi sawah sepanjang bulan Juli 2016 - Desember 2016 tidak dapat terpenuhi, karena penanaman padi sawah membutuhkan air rata-rata 145 mm pada musim hujan [17], sehingga penanaman padi harus di-bantu dengan pengairan dan tidak mengandalkan turunnya hujan. Pada bulan Juli hingga Oktober petani dapat meng-ganti padi dengan palawija, dikarenakan palawija mem-butuhkan air yang lebih sedikit sekitar 50 mm pada musim kemarau [17].

Juli			Agustus			September			Oktober			November			Desember		
1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Palawija												Tanam padi dengan bantuan pengairan					

Gambar 9. Kalender Tanam Padi Periode Juli 2016-Desember 2016

[9] Gu, J., Zhu, M, dan Jiang L. 2011. *Housing Price Forecasting Based on Genetic Algorithm and Support Vector Machine. Expert System with Application* 3383-3386.

[10] Wei, W.,W.,S. 2006. *Time Series Analysis Univariate and Multivariate Method 2nd Editions*. New York: Addison Wesley Publishing Company, Inc.

[11] Cryer, J. D. dan Chan, K. 2008. *Time Series Analysis with Application in R, 2nd Edition*. New York: Springer.

[12] Bowerman, B., O'Connell, R. T., dan Koehler, A. B. 2004. *Forecasting, Time Series, and Regression: An Applied Approach*. Boston: Cengage Learning, Inc.

[13] Daniel, W. W. 1989. *Statistika Nonparametrik Terapan*. Alih Bahasa: Alex Tri Kuncoro. Jakarta:PT Gramedia.

[14] Putri, T. 2016. *Peramalan Return Saham Bank Central Asia Menggunakan Self Exciting Threshold Autoregressive- Genetic Algorithm*. Penelitian. Surabaya: Institut Teknologi Sepuluh Nopember.

[15] Yuliani, W., Irhamah dan Prastyo D D. 2014. *Penaksiran Parameter Model ARIMA dengan Menggunakan Genetic Algorithm*. Penelitian. Surabaya: Institut Teknologi Sepuluh Nopember.

[16] Suriadikusumah, A. 2007. *Analisis Curah Hujan Perhitungan & Penggunaannya*. Bandung: Rakayasa Sains.

[17] Oldeman, L. R. 1980. *The Agroclimate Classification of Rice Growing Enviroment in Indonesia*.Philiphinas: IRRI.

V. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan hasil analisis yang sudah dilakukan, maka didapatkan beberapa kesimpulan untuk menjawab tujuan yang telah ditetapkan sebagai berikut.

1. Berdasarkan RMSE dan SMAPE maka peramalan de-ngan menggunakan GA-SVR menghasilkan akurasi pe-ramalan yang lebih baik.
2. Hasil kalender tanam padi untuk tipe sawah tadah hujan menunjukkan pada bulan Juli hingga Desember 2016 kebutuhan air untuk penanaman padi sawah tidak dapat dipenuhi. Petani dapat mengganti dengan menanam pa-lawija atau jika tetap dilakukan tanam padi, maka pe-tani dan pemerintah harus memastikan tersedianya ca-dangan air.

B. Saran

Saran untuk penelitian selanjutnya adalah dapat di-gunakan metode yang bisa memperhatikan adanya pengaruh dari variabel lain. Dalam metode SVR dapat dicoba meng-gunakan *input* dari data asli atau *input* lain tanpa melalui metode ARIMA, serta dapat digunakan pilihan *range* para-meter yang lain agar didapatkan peramalan yang lebih baik

DAFTAR PUSTAKA

[1] Badan Pusat Statistik. 2014. *Bojonegoro Dalam Angka Tahun 2014*. Bojonegoro: BPS.

[2] Badan Meteorologi, Klimatologi dan Geofisika (BMKG). 2016. *Tentang Meteorologi?*. http://meteojuanda.info/index.php?option=com_content&view=article&id=36&Itemid=34. Diakses pada tanggal 12 November 2016.

[3] Badan Pusat Statistik. 2016. *Bojonegoro Dalam Angka Tahun 2016*. Bojonegoro : BPS.

[4] Zhang, G. 2003. *Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model. Journal Neurocomputing* 50:159-175.

[5] Hong, W C., dkk. 2011. *Forecasting Urban Traffic Flow by SVR with continuous ACO. Appl. Math. Model.*vol. 35, no. 3, pp. 1282–1291.

[6] Zhao, W., Tao, T., dan Zio, E. 2015. *System Reliability Prediction by Support Vector Regression with Analytic Selection and Genetic Algorithm Parameters Selection. Applied Soft Computing*, 30, 792–802.

[7] Gunn, S.1998. *Support Vector Machines for Classification. Technical Report*.Southampton: University of Southampton.

[8] Msiza, I. S., Nelwamondo F. V. dan Marwala T., 2008. *Water Demand Prediction using Artificial Neural Networks and Support Vector Regression. Journal of Computers*, Vol. 3, No. 11, November 2008.