

DETEKSI ANOMALI DATA PREDIKSI UNTUK MENINGKATKAN AKURASI HASIL PERAMALAN DATA CURAH HUJAN

ANOMALY DETECTION OF PREDICTED DATA TO IMPROVE THE ACCURACY OF RAINFALL DATA FORECASTING RESULTS

Agus Qomaruddin Munir^{1*}, Farida Nuraini², Evrita Lusiana Utari³, Yusriadi⁴

^{1,2,4}Teknologi Informasi, Fakultas Sains dan Teknologi, Universitas Respati Yogyakarta

³Teknik Elektro, Fakultas Sains dan Teknologi, Universitas Respati Yogyakarta

¹agusqmnr@respati.ac.id, ²faridaunriyo@gmail.com, ³evrita_lusiana@yahoo.com, ⁴yusri.stbk@gmail.com

*penulis korespondensi

Abstrak

Air merupakan kebutuhan pokok bagi makhluk hidup. Beberapa sumber mata air salah satunya adalah berasal dari air hujan. Untuk dapat memenuhi kecukupan kebutuhan pokok tanaman, lahan pertanian harus memiliki kecukupan air agar memperoleh hasil tanam yang berkualitas. Curah hujan merupakan salah satu indikator untuk mengetahui perkiraan kecukupan air tanah pada lahan pertanian. Ketersediaan air tanah yang dihasilkan oleh hujan dapat digunakan untuk mengetahui pola tanam pada suatu wilayah. Namun demikian, ketersediaan data curah hujan biasanya bergantung pada data real kondisi iklim saat ini, hal inilah yang menyulitkan pihak terkait menentukan klasifikasi pola tanam dimasa mendatang. Deteksi anomali data prediksi merupakan kegiatan penelitian yang bertujuan untuk mendapatkan identifikasi awal data peramalan dengan tujuan penggunaan data peramalan yang digunakan untuk melakukan prediksi agar mendapatkan hasil yang baik dengan kriteria tingkat error yang rendah dan memiliki akurasi yang memadai. Model prediksi curah hujan yang akurat dibutuhkan untuk mengatasi permasalahan pergeseran pola hujan. Model prediksi curah hujan untuk penentuan pola tanam direkomendasikan oleh *Food and Agriculture Organization* (FAO) seperti regresi linier yang hingga sekarang masih banyak digunakan. Kelemahan model ini adalah inkonsistensi peramalan jangka pendek. Deteksi anomali merupakan bagian dari penelitian pengembangan model baru prediksi curah hujan menggunakan metode *seasonal autoregressive integrated moving average* (SARIMA). Data prediksi yang digunakan adalah data curah hujan dengan kurun waktu tahun 2010 sampai dengan 2020 dari tujuh stasiun pengumpul curah hujan di Kabupaten Sleman digunakan sebagai data *training* untuk memprediksi curah hujan dimasa mendatang. Hasil luaran dari deteksi anomali adalah berupa hasil normalisasi data prediksi yang selanjutnya akan diproses pada tahap prediksi curah hujan pada bulan tertentu dan mendapatkan akurasi yang memadai.

Kata kunci : Deteksi Anomali; Prediksi; *Seasonal Autoregressive Integrated Moving Average*; Pola Tanam

Abstract

Water is a basic need for living things. Some sources of water, one of which comes from rainwater. To meet the adequacy of the basic needs of crops, agricultural land must have sufficient water to obtain quality crops. Rainfall is one indicator to determine the estimated adequacy of groundwater on agricultural land. The availability of groundwater produced by rain can determine cropping patterns in an area. However, the availability of rainfall data usually depends on accurate data on current climatic conditions. This makes it difficult for related parties to determine the classification of cropping patterns in the future. Prediction data anomaly detection is a research activity that aims to obtain early identification of forecasting data to use forecasting data to make predictions to get good

results with low error rate criteria and adequate accuracy. Accurate rainfall prediction models are needed to overcome the problem of shifting rain patterns. Rainfall prediction models for determining cropping patterns are recommended by the Food and Agriculture Organization (FAO), such as linear regression, which is still widely used today. The weakness of this model is the inconsistency of short-term forecasts. Anomaly detection is part of research to develop a new model of rainfall prediction using the seasonal autoregressive integrated moving average (SARIMA) method. Prediction data used is rainfall data from 2010 to 2020 from seven rainfall collection stations in Sleman Regency, which is used as training data to predict rainfall in the future. The output of anomaly detection is in the form of normalization of predictive data, which will then be processed at the rainfall prediction stage in a specific month and obtain adequate accuracy.

Keywords: Anomaly Detection; Prediction; Seasonal Autoregressive Integrated Moving Average; Cropping Pattern

1. PENDAHULUAN

Iklim didefinisikan sebagai ukuran rata-rata dan variabilitas kuantitas yang relevan dari variabel tertentu, seperti temperatur, curah hujan atau angin), pada periode waktu tertentu, yang merentang dari bulanan hingga tahunan atau jutaan tahun. Perubahan iklim menurut ruang ini dibedakan berdasarkan perubahan iklim secara lokal dan global. Dalam beberapa dekade terakhir telah terjadi pergeseran iklim atau musim dari periode perubahan musim yang biasanya (sirkulasi muson). Kondisi ini selalu berubah-ubah secara tidak beraturan dalam kurun waktu yang acak, sehingga sulit untuk memprediksi cuaca secara tepat.

Iklim berubah secara terus menerus karena interaksi antara komponen-komponennya dan faktor eksternal seperti erupsi vulkanik, variasi sinar matahari, dan faktor-faktor disebabkan oleh kegiatan manusia. Hujan merupakan salah satu proses alam yang terjadi karena interaksi komponen faktor eksternal. Hujan mengalami proses statistik dan diperlukan analisis untuk dapat menerapkan teori peramalan dan analisis frekuensi [1]. Analisis frekuensi curah hujan diperlukan untuk menentukan jenis sebaran atau distribusi. Menemukan distribusi probabilitas yang paling cocok untuk curah hujan harian maksimum tahunan memiliki implikasi praktis dan teoritis yang kuat untuk dapat dilakukan proses peramalan.

Peramalan atau *forecasting* merupakan bagian penting dari pengambilan keputusan manajemen untuk mengurangi ketergantungannya pada hal-hal yang belum pasti [2]. Peramalan data kejadian merupakan salah satu sub ilmu *Artificial Neural Network* (ANN) [3] yang mampu memberikan solusi terhadap pengetahuan data yang sulit. Pada teknik peramalan kecukupan data untuk observasi menjadi hal penting sebagai pendukung keberhasilan kegiatan meramalkan suatu kejadian. Dalam melakukan peramalan diperlukan relevansi antar data dan faktor yang tepat untuk menghasilkan nilai peramalan dengan presisi lebih baik [4].

Terdapat beberapa metode statistik, baik yang bersifat univariat misalnya ARIMA (*Autoregressive Integrated Moving Average*) dan SARIMA (*Seasonal Autoregressive Integrated Moving Average*) maupun multivariat khususnya VAR (*Vector Autoregression*), yang bisa digunakan untuk melakukan peramalan curah hujan yang akan turun di suatu wilayah tertentu pada periode waktu tertentu pula. Pendekatan metode statistik runtun waktu, baik yang bersifat univariat maupun yang bersifat multivariat, merupakan pendekatan yang relatif baru pada sistem-sistem yang bersifat natural/alami. Pendekatan statistik univariat dan multivariat, juga banyak dimanfaatkan untuk melakukan peramalan sistem-sistem alami/natural, terutama untuk prediksi/peramalan curah

hujan. bersama tim BMKG (Badan Meteorologi, Klimatologi, dan Geofisika) melakukan penelitian di Purbalingga, Jawa Timur [5].

BMKG (Badan Meteorologi, Klimatologi, dan Geofisika) Kabupaten Mojokerto menggunakan metode ARIMA meramalkan curah hujan di tahun 2012 dengan memanfaatkan data curah hujan di waktu-waktu yang lampau. Meski demikian, pendekatan ini masih memiliki nilai galat/MAPE (*Mean Absolute Percentage Error*) yang relatif tinggi (22,13%) serta curah hujan hasil prediksi belum terkait dengan posisi geografis tertentu [6]. Peneliti lain Chen dan Zhu (2013) juga telah melakukan penelitian yang mempelajari penggunaan beberapa teknik runtun waktu secara simultan untuk peramalan curah hujan bulanan [7]. Teknik-teknik peramalan curah hujan yang digunakan dalam penelitian ini adalah MA (*Moving Average*), VR (*Vector Regression*), dan kombinasinya, kemudian dibandingkan dengan peramalan menggunakan SARIMA (*Seasonal Autoregression Integrated Moving Average*) yang bersifat musiman. Selain itu peneliti lain Mulyani dkk (2019) meneliti prediksi curah hujan bulanan dengan menggunakan beberapa prediktor di Stasiun Meteorologi Jatiwangi Majalengka [8]. Prediksi curah hujan bulanan untuk tahun 2019 dengan menggunakan data parameter cuaca harian periode tahun 2018 di Kabupaten Majalengka telah dilakukan dengan menggunakan metode persamaan regresi linier berganda namun dari hasil akurasi masih tergolong pada tingkat belum memadai.

Beberapa penelitian fokus pada proses peramalan data, baik data yang diperoleh secara primer maupun sekunder. Metode yang digunakan juga beragam, dari pendekatan metode statistik murni hingga pendekatan *neural network* ilmu komputer. Pembahasan hasil pada beberapa penelitian fokus terhadap pembahasan hasil [5],[6],[7],[8] terkait dengan data yang digunakan, periode waktu yang dipakai, dan capaian hasil. Namun demikian, pembahasan pada tahap awal (*preprocessing*) data untuk meningkatkan capaian akurasi peramalan belum dibahas secara mendalam. Tahapan penelitian yang dilakukan pada jurnal ini yaitu mengupayakan pada tahapan proses awal sebelum dilakukan proses peramalan untuk mendeteksi adanya anomali data prediksi untuk meningkatkan akurasi hasil peramalan.

2. METODE PENELITIAN

Metode yang digunakan untuk melakukan peramalan data curah hujan menggunakan metode *seasonal autoregressive integrated moving average (SARIMA)*.

2.1 Peramalan Runtun Waktu

Peramalan dilakukan untuk memprediksi kejadian di masa yang akan datang. Salah satu jenis peramalan adalah peramalan runtun waktu (*time series*). Analisis *time series* memiliki dua tujuan, yaitu untuk mengidentifikasi fenomena yang dapat digambarkan melalui urutan kejadian berdasarkan hasil pengamatan dan meramalkan kejadian di masa yang akan datang sesuai dengan variabel *time series* yang digunakan [9].

Definisi lain, menurut Makridakis dan Wheelwright (1989), peramalan adalah suatu kegiatan untuk memperkirakan secara sistematis nilai-nilai sebuah variabel di masa yang akan datang berdasarkan nilai yang diketahui dari variabel tersebut atau variabel yang berhubungan di masa lalu dan sekarang. Menurut Mulyana (2004) pengertian peramalan (*forecasting*) dapat dibedakan dengan prakiraan (*prediction*) [10]. Peramalan adalah proses penaksiran nilai data berdasarkan sebuah model hubungan fungsional antar nilai data sedangkan prakiraan adalah estimasi nilai data tanpa memperhatikan hubungan antar nilai data.

Peramalan memprediksi kejadian di masa yang akan datang berdasarkan pola kejadian yang pernah ada pada runtun waktu sebelumnya yang dicatat secara periodik (misalnya dalam bulan atau

tahun). Menurut Bob Jenkins dalam Ramasubramanian (1998), untuk tujuan prediksi dibutuhkan minimal 50 nilai hasil pengamatan masa lalu (historis) [11]. Saat ini telah banyak metode pendugaan berbasis statistik berkembang seperti regresi linier, stokastik dan runtun waktu. Setiap metode memiliki kelebihan dan keterbatasan sesuai dengan tujuan pemanfaatannya. Peramalan jenis *time series* memiliki asumsi bahwa pola data di masa lampau bisa digunakan untuk memprediksi data di masa yang akan datang. Berdasarkan jangka waktunya, peramalan time series dapat dibagi dalam tiga jenis, yaitu:

- Peramalan jangka panjang lebih dari 2 tahun.
- Peramalan jangka menengah 3 bulan sampai 2 tahun.
- Peramalan jangka pendek antara 0 bulan sampai 3 bulan.

Runtun waktu, memiliki 4 komponen penting yang menggambarkan karakteristik dari data yang akan diramalkan memiliki kecenderungan, berupa 1) Tren Sekuler yaitu keadaan data yang mengalami fluktuasi dari waktu ke waktu. 2) Variasi musiman, yakni fluktuasi yang muncul setiap tahun atau memiliki pola naik dan turun yang relatif tetap dari waktu ke waktu. 3) Variasi siklis, yakni variasi yang berulang setelah jangka waktu tertentu. 4) Variasi ireguler atau variasi random, yakni variasi data yang tidak teratur (irregular) yang sulit diprediksi.

Menurut penelitian Gharbi dkk (2011), peramalan gejala demam menggunakan variabel iklim pada time series. Pendekatannya menggunakan interval waktu 1 tahun, 3 bulan, dan 1 bulan agar didapat hasil perbandingan yang paling akurat [12]. Hasil yang paling akurat adalah pendekatan dengan interval waktu 3 bulan dan variabel iklim yang terbaik untuk memprediksi.

2.2 Seasonal Autoregressive Integrated Moving Average (SARIMA)

Model SARIMA adalah bentuk ARIMA yang memiliki pola *seasonal*/musiman. Untuk mendapatkan pola musiman diperlukan syarat bahwa telah terjadi perulangan data dalam periode waktu yang berbeda. Suatu deret $\{Z_t\}$ dikatakan memiliki periode variabel musiman dan tidak musiman dinotasikan dalam bentuk model ARIMA berikut:

$$\Phi_p(B)(1-B)^d Z_t = \theta_0 + \theta_q(B)b_t \quad (1)$$

Untuk lebih mudah melihat korelasi antar periode, dapat direpresentasikan sebagai model ARIMA berikut:

$$\Phi_p(B^s)(1-B^s)^d b_t = \theta_0(B^s) \alpha_t \quad (2)$$

Dengan

$$\Phi_p(B^s) = 1 - \Phi_1(B^s) - \Phi_1 B^{2s} - \dots - \Phi_p B^{ps} \quad (3)$$

Dan

$$\theta_q(B^s) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_p B^{qs} \quad (4)$$

adalah persamaan polynomial dalam Bs. jika akar-akar dari polynomial-polynomial tersebut berada di luar lingkaran unit dan $\{\alpha_t\} = 0$, maka proses tersebut adalah proses *white noise*. Dengan mengkombinasikan persamaan (2) dan persamaan (4), diperoleh model SARIMA, yaitu:

$$\Phi_p(B^s) \Phi_p(B)(1-B)^d (1-B^s)^D Z_t = \theta_q(B) \theta_q(B^s) \alpha_t \quad (5)$$

Dengan:

$$\begin{aligned} \Phi_p(B) & : \text{Parameter AR Non Musiman} \\ \theta_q(B) & : \text{Parameter MA Non Musiman} \end{aligned}$$

$\Phi_p(B^s)$:	Parameter AR Musiman
$\Theta_q(B^s)$:	Parameter MA Musiman
$(1-B)^d$:	<i>Differencing</i> Non Musiman
$(1-B^s)^D$:	<i>Differencing</i> Musiman

Langkah-langkah yang dilakukan dalam mendapatkan hasil SARIMA sama dengan langkah-langkah ARIMA. Setelah melakukan perhitungan ARIMA/*Box Jenkins*, rumus ditambahkan dengan bagian musiman dari modelnya. Bagian musiman tersebut didapatkan dari plot ACF dan PACF dari lag 12 hingga lag 48. Penentuan *lag* tersebut karena periode yang digunakan adalah 12 bulan atau 1 tahun. Secara umum, model SARIMA dinotasikan sebagai berikut [13]:

$$\text{SARIMA } (p,d,q)(P,D,Q)_s \quad (6)$$

Dengan:

(P,D,Q)	:	bagian musiman dari model
P	:	orde musiman untuk AR
Q	:	orde musiman untuk MA
s	:	jumlah periode per musim

2.3 Deteksi Anomali Data Peramalan

Deteksi anomali merupakan kegiatan awal penelitian guna melakukan identifikasi pada data awal yang memungkinkan terjadi *missing* data (hilangnya elemen data), data runtun waktu yang tidak lengkap, dan normalisasi data. Tahap ini dilakukan untuk mempersiapkan kebutuhan data dalam penelitian yang nantinya digunakan sebagai inputan pada proses-proses peramalan. Sebelum data dijadikan sebagai input pada model, data penelitian yang dikumpulkan dilakukan pra pemrosesan data, meliputi seleksi data, pencarian informasi yang hilang dan proses normalisasi data.

Proses peramalan data menggunakan metode *seasonal autoregressive integrated moving average (SARIMA)* dilakukan beberapa tahapan, antara lain: 1) Preprocessing Data, 2) Identifikasi Plot Data, 3) Penaksiran Parameter, 4) Pemeriksaan Diagnostik (*Diagnostic Checking*) dan 5) Proses Peramalan Data.

1. *Pre-Processing* Data, tahap pertama adalah *preprocessing* data penelitian yang terdiri dari data curah hujan bulanan, periode tahun 2015 sampai dengan 2020. Pra pemrosesan data curah hujan dilakukan dengan cara menjumlahkan data curah hujan harian ke dalam periode bulanan untuk lokasi di wilayah Kabupaten Sleman. Pra pemrosesan data curah hujan diawali dengan melakukan pengecekan data series bulanan mulai dari bulan Januari 2015 sampai dengan bulan Desember 2020. Data curah hujan terdiri dari 60 *record* kasus DBD selama tahun 2015 sampai dengan 2020 dengan jumlah 4 stasiun pengambilan data cuaca diseluruh wilayah kecamatan di Kabupaten Sleman, sehingga total *record* sejumlah 240 data curah hujan bulanan.

Tahap *pre-processing* data menjadi hal yang sangat penting dilakukan untuk dapat meningkatkan akurasi dalam melakukan peramalan. Pada sebuah tabel data terdiri dari baris dan kolom, baris pada sebuah tabel data merepresentasikan unit, atau disebut juga dengan observasi, dalam hal ini adalah jumlah curah hujan bulanan, sedangkan kolom pada sebuah tabel merepresentasikan variabel yang diukur dari suatu unit. Sebuah tabel data terdiri atas baris dan kolom dari n unit observasi dengan m variabel yang dapat diukur. Tapi dalam prakteknya, tidak selamanya akan diperoleh suatu matriks data yang lengkap.

Eksperimen data *pre-processing* dilakukan untuk semua *missing value* dari total data jumlah curah hujan. Metode yang diusulkan memerlukan input berupa data jumlah hujan bulanan dengan series data bulanan, namun penerapan metode diperlukan kecukupan data yang memiliki kecenderungan pola musiman atau *seasonal*. Kesulitan utama yang terjadi adalah ketika di suatu wilayah tidak terjadi hujan pada laporan tidak terekam dengan baik sehingga karakteristik data tidak dapat diidentifikasi polanya, maka dari itu perlu dilakukan pencatatan ulang bagi wilayah yang pada bulan tertentu tidak terjadi kasus dengan diinputkan angka nol (0), hal ini dilakukan agar dapat diketahui kinerja peramalan dan tingkat pencapaian akurasi yang memadai terhadap penerapan metode yang diusulkan. Contoh hasil dari proses pada 5 *record* awal data series curah hujan bulanan untuk beberapa wilayah di Kabupaten Sleman seperti pada Tabel 1.

Tabel 1. Data Jumlah Curah Hujan Tidak Lengkap

NO	Tahun	BULAN											
		1	2	3	4	5	6	7	8	9	10	11	12
1	2015	347.3	417.7	457.7	432.4	178.1	56.8	-	-	-	-	96.8	416.1
2	2016	7.9	16.0	13.7	2.6	1.9	6.4	3.2	6.4	4.8	6.4	17.7	10.9
3	2017	414.5	336.9	135.3	283.3	18.7	46.4	36.7	1.5	-	TTU	316.4	319.3
4	2018	412.4	517.6	195.5	244.4	245.8	226.4	78.1	0.6	0.9	43.3	407.8	329.5
5	2019	377.3	279.5	298	155.6	69.4	5.2	TTU	-	-	87.6	430.5	453.7

(Sumber: Badan Pusat Statistik, 2020)

Pada tabel **Error! Reference source not found.** ada beberapa kolom dengan isian berupa TTU, TTU ini berarti bahwa data jumlah curah hujan tidak terukur. Proses selanjutnya adalah melakukan penginputan data pada bagian bulan yang tidak terisi data dengan nilai nol (0) pada tabel Tabel 1 untuk memenuhi kebutuhan *pre-processing* data. Adapun hasilnya menjadi seperti pada Tabel 2.

Tabel 2. Data Jumlah Curah Hujan Setelah dilakukan Normalisasi.

NO	Tahun	BULAN											
		1	2	3	4	5	6	7	8	9	10	11	12
1	2015	347.3	417.7	457.7	432.4	178.1	56.8	0	0	0	0	96.8	416.1
2	2016	7.9	16.0	13.7	2.6	1.9	6.4	3.2	6.4	4.8	6.4	17.7	10.9
3	2017	414.5	336.9	135.3	283.3	18.7	46.4	36.7	1.5	0	0	316.4	319.3
4	2018	412.4	517.6	195.5	244.4	245.8	226.4	78.1	0.6	0.9	43.3	407.8	329.5
5	2019	377.3	279.5	298	155.6	69.4	5.2	0	0	0	87.6	430.5	453.7

(Sumber: Badan Pusat Statistik, 2020)

Setelah semua proses tersebut di atas dilakukan pada rentang waktu bulan Januari 2015 sampai dengan Desember 2020 maka data cukup memadai untuk dilanjutkan ke proses berikutnya.

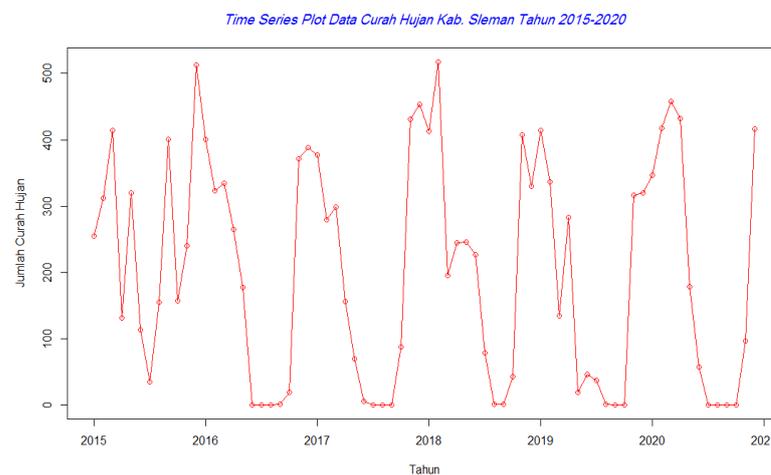
- Identifikasi Plot Data Series, setelah data bersih dan memiliki periode bulanan maka selanjutnya di lakukan identifikasi plot data series dan melakukan uji stasioneritas data sebagai syarat yang di gunakan dalam peramalan curah hujan menggunakan metode SARIMA. Untuk

mendapatkan gambaran perkiraan awal dari bentuk model yang sesuai terhadap data. Identifikasi plot data series dilakukan untuk mengetahui apakah data memiliki sifat *trend* (memiliki kecenderungan naik atau bisa juga turun pada periode tertentu), musiman maupun acak, maka data ditampilkan dalam bentuk plot runtun waktu. Melalui plot runtun waktu akan terlihat bahwa data tersebut memiliki kecenderungan stationer atau belum secara kasat mata. Misalkan diberikan contoh data jumlah curah hujan pada Tahun 2015 sampai dengan Tahun 2020 seperti pada Tabel 3.

Tabel 3. Sampel Data Jumlah Curah Hujan Tahun 2015 - 2020

No	Tahun 2015		Tahun 2016		Tahun 2017		Tahun 2018		Tahun 2019		Tahun 2020	
1	Jan	254.3	Jan	400.7	Jan	377.3	Jan	412.4	Jan	414.5	Jan	347.3
2	Feb	311.6	Feb	322.7	Feb	279.5	Feb	517.6	Feb	336.9	Feb	417.7
3	Maret	413.6	Maret	334.7	Maret	298	Maret	195.5	Maret	135.3	Maret	457.7
4	April	131.2	April	264.6	April	155.6	April	244.4	April	283.3	April	432.4
5	Mei	319.3	Mei	334.7	Mei	69.4	Mei	245.8	Mei	18.7	Mei	178.1
6	Juni	113.1	Juni	0	Juni	5.2	Juni	226.4	Juni	46.4	Juni	56.8
7	Juli	34.5	Juli	0	Juli	0	Juli	78.1	Juli	36.7	Juli	0
8	Agust	155.2	Agust	0	Agust	0	Agust	0.6	Agust	1.5	Agust	0
9	Sept	400.5	Sept	0.5	Sept	0	Sept	0.9	Sept	0	Sept	0
10	Okt	157.2	Okt	19	Okt	87.6	Okt	43.3	Okt	0	Okt	0
11	Nov	240	Nov	371.3	Nov	430.5	Nov	407.8	Nov	316.4	Nov	96.8
12	Des	512.3	Des	388.5	Des	453.7	Des	329.5	Des	319.3	Des	416.1

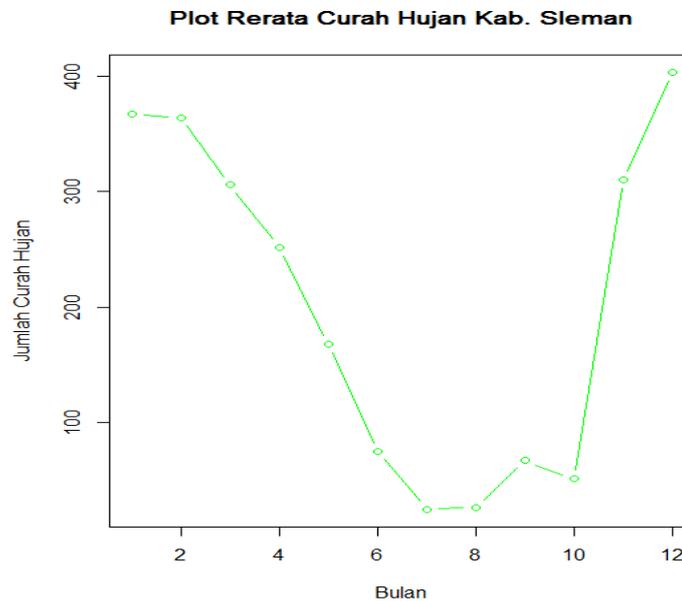
Pengecekan lebih lanjut adalah uji stasioneritas. Apabila hasil menunjukkan data belum stationer, maka harus dilakukan *differencing*. Asumsi awal bahwa data memiliki kecenderungan stationer, maka perlu dilakukan representasi data secara keseluruhan yaitu dengan membuat plot data penelitian. Untuk mendapatkan gambaran implementasi terdapat pada Gambar 1.



Gambar 1. Plot Data Jumlah Curah Hujan Tahun 2015-2020

3. Penaksiran Parameter, penaksiran parameter digunakan untuk mendapatkan model sementara untuk selanjutnya dilakukan uji parameter pada model sementara adalah sebagai berikut:
 1. Model 1 : SARIMA (1,1,1)(1,1,0)₆
 2. Model 2 : SARIMA (1,1,1)(2,1,0)₆
 3. Model 3 : SARIMA (2,1,2)(1,1,0)₆
 4. Model 4 : SARIMA (2,1,2)(2,1,0)₆

4. Hasil Peramalan dari penaksiran parameter pada langkah 3 adalah sebagai berikut:



Gambar 2. Nilai Hasil Prediksi Curah Hujan Periode 12 Bulan

3. PEMBAHASAN

Peramalan yang digunakan untuk meramalkan data curah hujan ini adalah peramalan kuantitatif yang didasarkan atas penggunaan analisis pola hubungan antara variabel yang akan diperkirakan dengan variabel waktu, yang merupakan deret waktu/time series. Peramalan kuantitatif hanya dapat digunakan apabila terdapat tiga kondisi sebagai berikut:

- a. Adanya informasi tentang keadaan yang lain.
- b. Informasi tersebut dapat dikuantifikasikan dalam bentuk data.
- c. Dapat diasumsikan bahwa pola yang lalu akan berkelanjutan pada masa yang akan datang.

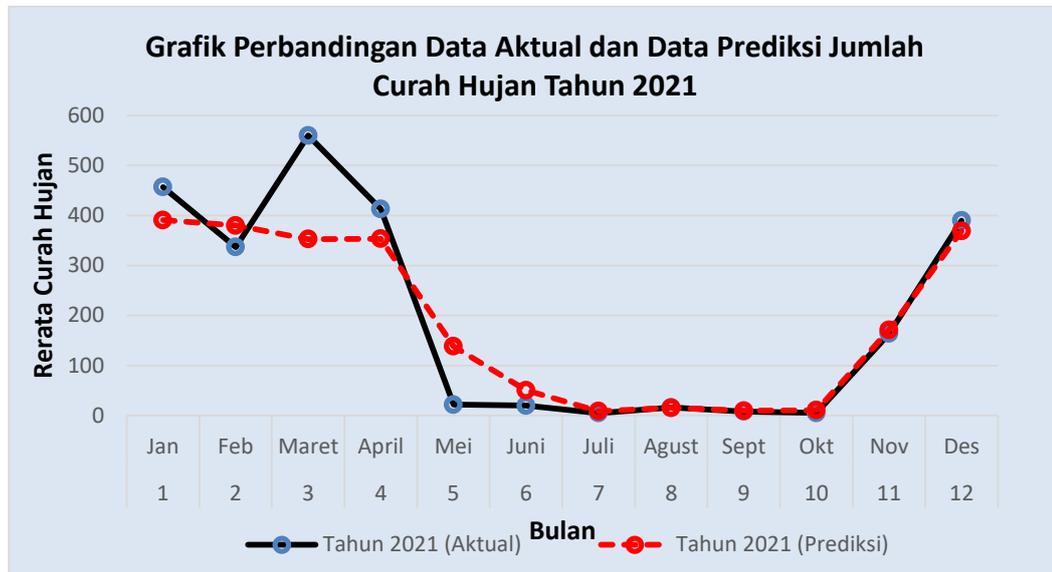
Dalam penelitian mengenai peramalan data curah hujan ini berpedoman pada beberapa langkah, seperti kajian pustaka, identifikasi permasalahan dan metode, pengambilan data di BMKG Kab. Sleman, pengerjaan analisis peramalan, melakukan pengecekan terhadap akurasi dan tingkat *error*, pembuatan laporan hasil penelitian, dan dokumentasi. Pada sub pembahasan ini dilakukan uji akurasi hasil peramalan dengan cara membandingkan data aktual curah hujan tahun 2021 dan data hasil peramalan tahun 2021. Data hasil prediksi dan data aktual terdapat pada Tabel 4.

Tabel 4. Perbandingan Data Aktual dan Data Prediksi Curah Hujan Tahun 2021

No	Bulan	Tahun 2021 (Aktual)	Tahun 2021 (Prediksi)	Error
1	Jan	457	390,8	66,2
2	Feb	337	380,1	-43,1
3	Maret	560	352,8	207,2
4	April	413	353,4	59,6
5	Mei	22	138,4	116,4
6	Juni	20	50,7	-30,7
7	Juli	5	8,84	-3,84
8	Agust	16	15,6	0,4
9	Sept	8,11	9,71	-1,6

10	Okt	5,3	10,64	-5,34
11	Nov	164	171,2	-7,2
12	Des	390	369,3	20,7

Jika disajikan dalam bentuk grafik ada pada Gambar 3.



Gambar 3. Grafik Aktual dan Prediksi Data Curah Hujan periode 12 Bulan.

Pada Gambar 3 terlihat plot data aktual dan data prediksi rerata curah hujan pada periode bulan Januari sampai dengan Desember 2021 memiliki kemiripan plot dengan data aktual pada periode yang sama. Pada bulan Januari sampai dengan Maret 2021 data prediksi memiliki kesamaan pola pada bulan Januari, Februari, April dan Mei, sedangkan untuk Maret dan Juni pola prediksi agak sedikit berbeda. Sedangkan untuk bulan Juli sampai dengan Desember data prediksi memiliki pola yang mirip, dengan demikian akurasi data prediksi cukup memadai.

Berdasarkan hasil prakiraan curah hujan yang diperoleh pada bulan Januari, Februari, Maret, April dan Mei terjadi curah hujan yang relatif tinggi, dengan kondisi curah hujan yang relatif tinggi sangat menentukan kecocokan dan optimalisasi pembudidayaan tanaman pertanian, misalnya padi. Selanjutnya untuk bulan Juni, Juli, Agustus, September dan Oktober memiliki kecenderungan bulan kering atau curah hujan sangat sedikit sehingga dapat dilakukan optimalisasi budidaya tanaman tembakau, kedelai, jagung, kacang tanah, kacang hijau, ubi kayu dan ubi jalar.

4. KESIMPULAN

Peramalan jumlah curah hujan di wilayah Kabupaten Sleman menggunakan data dasarian yaitu 2010-2020 untuk meramalkan jumlah curah hujan pada tahun 2021 menunjukkan bahwa curah hujan tertinggi terjadi pada bulan Januari yaitu 390,8 mm dan jumlah curah hujan terendah terjadi pada bulan Juli yaitu 8,84 mm. Selanjutnya dengan melihat hasil peramalan curah hujan dapat diklasifikasikan dalam 1 tahun menjadi 2 kategori yaitu bulan kering dan bulan basah. Pada saat bulan basah (curah hujan tinggi) yaitu kurun waktu bulan Januari sampai dengan Mei kecocokan budidaya tanaman pertanian adalah padi, sedangkan untuk bulan kering antara bulan Juni sampai dengan bulan Oktober adalah tanaman tembakau, kedelai, jagung, kacang tanah, kacang hijau, ubi kayu dan ubi jalar.

Deteksi anomali pada preprocessing data peramalan curah hujan efektif dapat meningkatkan hasil akurasi peramalan dibandingkan dengan yang tidak dilakukan preprocessing data. Penelitian lanjutan perlu ditambahkan deskripsi pemetaan spasial wilayah untuk mendapatkan gambaran secara riil agar dapat merencanakan secara terstruktur pola tanam yang efektif sesuai dengan kondisi spesifik wilayah dengan memanfaatkan data prediksi curah hujan.

DAFTAR PUSTAKA

- [1] T. C. Upomo and R. Kusumawardani. 2016., *Pemilihan Distribusi Probabilitas pada Analisa Hujan dengan Metode Goodness of Fit Test*. J. Tek. Sipil dan Perenc., vol. 18, no. 2, pp. 139–148, doi: 10.15294/jtsp.v18i2.7480.
- [2] Makridakis, S.Wheelwright., 1999, *Metode dan Aplikasi Peramalan Jilid 1* (Ir. Untung Sus Ardiyanto, M.Sc. & Ir. Abdul Basith, M.Sc Terjemah). *Edisi Kedua, Jakarta: Penerbit Erlangga*.
- [3] Sharda, R., 1994. Neural Networks For The MS/OR Analyst: An Application Bibliography. *Interfaces* 24 (2), 116–130.
- [4] Hii, Y. L., Zhu, H., Ng, N., Ng, L. C., & Rocklöv, J. 2012. *Forecast of Dengue Incidence Using Temperature and Rainfall*. *PLoS Neglected Tropical Diseases*, 6(11). doi:10.1371/journal.pntd.000190.
- [5] Tresnawati, R., 2010. *Prediksi Curah Hujan Bulanan Dengan Menggunakan Metode Kalman Filter Dengan Prediktor SST Nino 3.4 Diprediksi*, *Jurnal Meteorologi dan Geofisika* Vol. 11, No. 2, hal. 106-115.
- [6] Huda, A.M., Choiruddin, A., Budiarto, O., Sutikno, 2012, *Peramalan Data Curah Hujan Dengan Seasonal Autoregression Moving Average (SARIMA) Dengan Deteksi Outlier Sebagai Upaya Optimalisasi Produksi Pertanian Di Kabupaten Mojokerto*. *Seminar Nasional Kedaulatan Pangan dan Energi*, Fakultas Pertanian Universitas Trunojoyo, Madura.
- [7] Chen, X., Zhu, S., 2013. *Improved Hybrid Model Based on Support Vector Regression Machine for Monthly Precipitation Forecasting*. *Journal of Computers*, Vol. 8, No. 1, hal. 232-239.
- [8] Evi Dewi Sri Mulyani, Indah Septianingrum, Nisa Nurjanah, Reka Rahmawati, Syifa Nurhasani, Kiky Milky R. K., 2019. *Prediksi Curah Hujan Di Kabupaten Majalengka Dengan Menggunakan Algoritma Regresi*, Volume 8, Nomor 1, *Jurnal Sistem Informasi dan Teknologi Informasi (JUSITI)*, <https://doi.org/10.36774/jusiti.v8i1.602>.
- [9] Alsultanny, Y., 2012. *Successful Forecasting For Knowledge Discovery By Statistical Methods*. *Proceedings of the 9th International Conference on Information Technology, ITNG 2012*, hal.584–588.
- [10] Mulyana, 2004. *Analisis Deret Waktu*, Universitas Padjadjaran FMIPA Jurusan Statistika, Bandung.
- [11] Box, G. and Jenkins, G., 1976, *Time Series Analysis Forecasting and Control*, *Holden-Day, San Fransisco*.
- [12] M. Gharbi, P. Quenel, J. Gustave, S. Cassadou, G. L. Ruche, L. Girdary, and L. Marrama, Jun. 2011. “Time series analysis of dengue incidence in Guadeloupe, French West Indies: Forecasting models using climate variables as predictors,” *Bmc Infect. Dis.*, vol. 11, no. 1, p. 166.

- [13] Makridakis; Wheelwright & McGee. 1999. *Metode dan Aplikasi Peramalan*. Jakarta; Binarupa Aksara.