

**EVALUASI MODEL *LONG-SHORT TERM MEMORY* (LSTM) DAN
SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE
(SARIMA) UNTUK PREDIKSI CURAH HUJAN DI WILAYAH SUMATRA**

Bimbing Putra¹, Rista Hernandi Virgianto^{2*}

^{1,2} Program Studi Klimatologi, Sekolah Tinggi Meteorologi
Klimatologi Dan Geofisika, Tangerang Selatan
bimbing.tobing@gmail.com¹, rista.hernandi@gmail.com^{*2}

Received 15 Oktober 2023; revised 25 Oktober 2023; accepted 30 Oktober 2023.

ABSTRAK

Kebutuhan akan informasi cuaca dan iklim masih sangat besar. Perubahan iklim menyebabkan permasalahan pertanian yang kompleks, prakiraan cuaca dan iklim yang akurat sangat diperlukan. Prediksi cuaca dan iklim dengan tingkat akurasi yang tinggi dan konsisten merupakan tantangan di wilayah tropis. Berbagai model prediksi cuaca telah dikembangkan dalam dunia meteorologi di seluruh dunia, namun model tersebut tidak tepat merepresentasikan parameter-parameter yang sangat dinamis seperti Indonesia. Banyak cara yang dapat dilakukan untuk memprediksi besarnya curah hujan yang akan datang pengembangan model SARIMA dan LSTM. Model SARIMA merupakan pengembangan dari model ARIMA dengan penambahan komponen musiman. LSTM ditemukan sebagai model kompetitif untuk data musiman. Data *blend* curah hujan ZOM terpilih di wilayah Sumatra dibentuk menjadi data *training* dan data uji. Model LSTM dan SARIMA dibangun dengan bantuan Bahasa pemrograman Python. Hasil model diverifikasi dengan RMSE, Bias, MAPE, dan Pearson. Terdapat pula modifikasi data menjadi tabel kontingensi multi kategori untuk melakukan verifikasi perhitungan *proportion of correct* dan *Peirce Skill Score*. Hasil model SARIMA memiliki nilai statistik kesalahan absolut yang lebih kecil, namun bias lebih besar dibandingkan dengan model LSTM. Hasil model LSTM cukup mampu menangkap pola data latih. Hasil model SARIMA memiliki nilai yang cenderung *flat*. Kedua model berada pada tingkat akurasi yang sama. Perhitungan *Peirce Skill Score* menunjukkan model LSTM memiliki jumlah ZOM dengan kategori baik lebih banyak dibandingkan dengan model SARIMA.

Kata kunci: LSTM, model curah hujan, SARIMA, verifikasi, ZOM.

ABSTRACT

The demand for weather and climate information remains substantial. Climate change poses complex agricultural issues, underscoring the need for accurate weather and climate forecasts. Achieving highly accurate and consistent weather and climate predictions is a challenge in tropical regions. Numerous weather prediction models have been developed worldwide, yet these models often fail to adequately represent the dynamic parameters specific to Indonesia. Various approaches can be taken to predict upcoming rainfall, such as developing SARIMA and LSTM models. SARIMA enhances the ARIMA model by adding seasonal components, while LSTM has emerged as a competitive model for seasonal data. Rainfall data blending from selected Seasonal Zones (ZOM) in Sumatra is divided into training and test sets. LSTM and SARIMA models are constructed using Python programming language. Model outcomes are verified using metrics like RMSE, Bias, MAPE, and Pearson correlation. Additionally, data is transformed into multi-category contingency tables to verify the calculations for the proportion of correct forecasts and Peirce Skill Score. The SARIMA model produces smaller absolute error statistical values, but exhibits larger bias compared to LSTM. LSTM effectively captures training data patterns. The SARIMA model's outcomes are somewhat flat. Both models exhibit similar levels of accuracy. Peirce Skill Score calculations reveal that the LSTM model outperforms the SARIMA model in terms of the number of ZOMs falling into the "good" category.

Keywords: LSTM, rainfall model, SARIMA, verification, seasonal rainfall zones (ZOM).

PENDAHULUAN

Ramalan cuaca atau prakiraan cuaca (weather forecast) adalah perkiraan keadaan cuaca yang memungkinkan akan terjadi di suatu wilayah (Wirjohamidjojo, 1985). Badan Meteorologi Klimatologi dan Geofisika (BMKG) adalah lembaga resmi yang mengeluarkan prakiraan cuaca dan iklim di Indonesia. 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, dan terus berubah karena interaksi antara komponen dan faktor eksternalnya (KLHK, 2017b). BMKG menerbitkan prakiraan iklim berupa informasi iklim dasarian, bulanan, tahunan, dan musiman.

Kebutuhan akan informasi cuaca dan iklim masih sangat besar karena berpengaruh pada kehidupan sehari-hari, sebagai contoh prediksi curah hujan sangat penting bagi bidang yang membutuhkan seperti bidang pertanian,

transportasi dan industri (Dwiratna dkk., 2013). Pada bidang pertanian, curah hujan sebagai salah satu acuan oleh petani dalam menentukan kecocokan jenis tanaman yang akan dibudidayakan atau waktu masa tanamnya (Rofiq dkk., 2020). Perubahan iklim menyebabkan permasalahan pertanian yang kompleks seperti kegagalan panen, kerusakan lahan pertanian, kekeringan baik areal dan intensitas, serta hama dan penyakit baru dan resistan (Hilman dkk., 2019). Pada bidang transportasi, hambatan akibat intensitas curah hujan sangat erat dengan kelancaran jalur transportasi. Pada bidang ekonomi dan industri, intensitas penyinaran matahari sangat penting untuk usaha kecil dan menengah (tradisional), sebagai contoh usaha pembuatan makanan kering dan material jamur atau batubata (KLHK, 2017a). Dengan demikian prakiraan cuaca dan iklim yang akurat sangat diperlukan.

Prediksi cuaca dan iklim dengan tingkat akurasi yang tinggi dan konsisten merupakan tantangan di wilayah tropis (Gadgil dan Srinivasan, 2011). Iklim Indonesia tidak hanya dipengaruhi oleh iklim regional dan global tetapi juga iklim lokal (Aldrian dkk., 2003). Setiap wilayah di Indonesia telah memiliki keluaran prakiraan iklim namun keluaran yang dihasilkan tidak dapat menghasilkan akurasi yang tinggi (Septiadi, 2008). Berbagai model prediksi cuaca telah dikembangkan dalam dunia meteorologi di seluruh dunia, baik model dinamis maupun model statistik (Kadarsah, 2010), namun model tersebut tidak tepat merepresentasikan parameter-parameter di khatulistiwa yang sangat dinamis seperti Indonesia (Megalina, 2014). Keefektifan metode prakiraan curah hujan bergantung pada berbagai faktor, termasuk wilayah spesifik, kondisi atmosfer setempat, data yang tersedia, dan skala waktu prakiraan (Nisa Ayunda dkk., 2021).

Curah hujan memiliki nilai rata-rata dan variasi pola yang tidak konstan sepanjang waktu. Besarnya curah hujan tidak dapat ditentukan secara pasti, namun nilainya bisa diperkirakan hingga nilai selisih kesalahan kecil. Data historis curah hujan yang panjang dimanfaatkan untuk membangun model prediksi, kemudian dapat di perkirakan besarnya curah hujan di masa yang akan datang melalui model prediksi tadi (Harlina dan Usman, 2020). Banyak cara yang dapat dilakukan untuk memprediksi besarnya curah hujan yang akan datang

menggunakan beberapa pendekatan statistik seperti moving average, SARIMA, dan algoritma LSTM. Model Autoregresif Integrated Moving Average (ARIMA) adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. Peramalan jangka panjang ARIMA memiliki ketepatan kurang baik. Peramalan biasanya akan cenderung flat (mendatar/konstan) untuk periode yang cukup panjang (Hendrawan, 2012). Model SARIMA merupakan pengembangan dari model ARIMA penambahan komponen musiman (Sinay dkk., 2020).

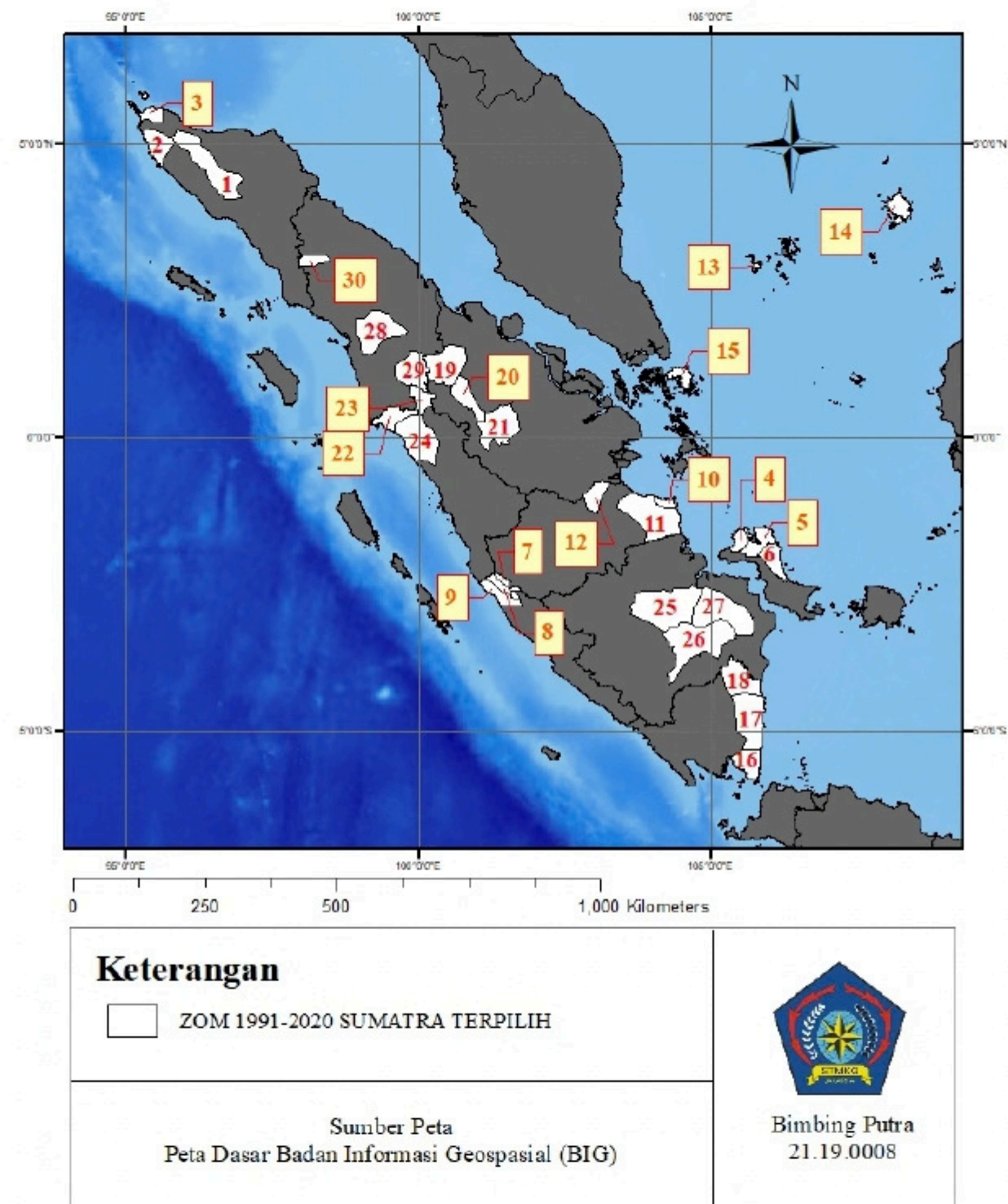
Long-short term memory (LSTM) adalah salah satu algoritma perhitungan data time series dengan data yang digunakan seperti curah hujan, algoritma ini cukup terbukti handal dalam memprediksi beberapa tahun terakhir ini (Sharmila dan Subramani, 2016). LSTM dirancang oleh Hochreiter dan Schmidhuber (1997) yang merupakan varian dari jaringan saraf berulang untuk menangkap pola data jangka panjang. Dengan ini, LSTM ditemukan sebagai model kompetitif untuk data musiman, seperti curah hujan. Salah satu keunggulan penting algoritma ini adalah kemampuan mengingat dari sekuens long term (ukuran data) yang sulit dicapai dengan teknik fitur tradisional dan algoritma ini juga dapat menggunakan ukuran data yang lebih besar dengan menggunakan semua informasi data sebagai masukannya (Desmonda dkk., 2018).

Hasil penelitian ini diharapkan dapat menggambarkan hasil prakiraan iklim model SARIMA dan LSTM di wilayah Sumatra beserta evaluasinya. Hasil prakiraan model yang lebih baik berdasarkan tingkat akurasi dan galat minimum diharapkan dapat menjadi prakiraan iklim terstandar dan fungsional. Informasi iklim dasarian dan bulanan dengan akurasi yang baik dapat mendukung berbagai sektor yang tergantung akan prakiraan informasi iklim.

METODE PENELITIAN

Bagian metode penelitian akan membahas secara lengkap lokasi dan data serta bagaimana metode yang akan digunakan dalam penelitian ini dan menyimpulkan dalam bentuk diagram alir. Dalam penelitian ini wilayah penelitian yaitu 30 zona musim (ZOM) terpilih periode 1991-2020 di wilayah Sumatra.

EVALUASI MODEL LONG-SHORT TERM MEMORY (LSTM) DAN SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) UNTUK PREDIKSI CURAH HUJAN DI WILAYAH SUMATRA



Gambar 1. Lokasi penelitian ZOM terpilih wilayah Sumatra.

Dari 30 ZOM terpilih, terdapat masing-masing 3 ZOM terpilih pada setiap provinsi. Dengan keterangan nomor pada gambar sebagai berikut:

Tabel 1. Keterangan nomor pada peta penelitian

Nomor	ZOM	Nomor	ZOM
1	ACEH 01	16	LAMPUNG 01
2	ACEH 02	17	LAMPUNG 02
3	ACEH 03	18	LAMPUNG 03
4	BABEL 01	19	RIAU 01
5	BABEL 02	20	RIAU 02
6	BABEL 03	21	RIAU 03
7	BENGKULU 01	22	SUMBAR 01
8	BENGKULU 02	23	SUMBAR 02
9	BENGKULU 03	24	SUMBAR 03
10	JAMBI 01	25	SUMSEL 01
11	JAMBI 02	26	SUMSEL 02
12	JAMBI 03	27	SUMSEL 03
13	KEPRI 01	28	SUMUT 01

14	KEPRI 02	29	SUMUT 02
15	KEPRI 03	30	SUMUT 03

Berdasarkan distribusi spasial dan jeda waktu, data yang digunakan untuk data periode percobaan dan data periode uji adalah data curah hujan dasarian hasil *blend* BMKG wilayah Sumatra, dengan data *blend* periode 1991-2020 sebagai data periode percobaan, dan 2021-2022 sebagai periode data uji.

Model SARIMA meliputi beberapa komponen yaitu; *Autoregressive* (AR) term: Komponen AR menunjukkan ketergantungan nilai saat ini pada nilai-nilai sebelumnya dalam *series* waktu. Jumlah term AR ditandai dengan p, misalnya AR(p). *Integrated* (I) term: Komponen I menunjukkan jumlah diferensiasi yang diperlukan untuk membuat data menjadi stasioner. Diferensiasi dilakukan untuk menghilangkan *tren* atau pola musiman dalam data. Jumlah term I ditandai dengan d, misalnya I(d). *Moving Average* (MA) term: Komponen MA menunjukkan ketergantungan antara nilai saat ini dan nilai-nilai residual (selisih antara nilai saat ini dengan nilai prediksi) pada periode sebelumnya. Jumlah term MA ditandai dengan q, misalnya MA(q).

Seasonal AR (SAR) term: Komponen SAR menunjukkan ketergantungan nilai saat ini pada nilai-nilai sebelumnya dalam satu musim atau periode waktu yang sama di tahun sebelumnya. Jumlah term SAR ditandai dengan P, misalnya SAR(P). Seasonal MA (SMA) term: Komponen SMA menunjukkan ketergantungan antara nilai saat ini dan nilai-nilai residual dalam satu musim atau periode waktu yang sama di tahun sebelumnya. Jumlah term SMA ditandai dengan Q, misalnya SMA(Q) (Sinay dkk., 220).

Model SARIMA menambahkan komponen musiman pada setiap komponen ARIMA. Model ARIMA dibuat dengan ordo (p,d,q) yang merupakan model campuran antara model *autoregressive* (AR) yang berordo p kemudian *moving average* (MA) berordo q yang mengalami pembedaan (*differencing*) sebanyak d kali. Model regresi dari ordo p (Box dan Jenkins, 1970), AR (p), menyatakan bahwa nilai pengamatan pada periode ke - t (Y_t) adalah hasil regresi dari nilai pengamatan sebelumnya selama p periode (Nuryadi, 2006). Plot data series harus menggunakan asumsi stasioner, apabila ada data yang akan diolah

belum stasioner maka perlu dilakukan pembedaan, yang selanjutnya dapat diduga nilai awal dari ordo d. Namun pembedaan tidak perlu dilakukan apabila data yang akan diolah sudah stasioner (Soetanto dan Ulfah, 2006).

Musiman digambarkan sebagai suatu ritme yang berulang dalam jangka waktu yang tetap atau tertentu. Notasi umum SARIMA adalah (Kafara dkk., 2017):

$$\text{SARIMA: } (pdq) (PDQ) s$$

(pdq) = Bagian yang tidak musiman dari model

(PDQ) = Bagian musiman dari model

s = Jumlah periode per musim

Perhitungan matematis dalam model LSTM yang dilakukan pada setiap tahap diberikan sebagai berikut Hochreiter dan Schmidhuber (1997):

$$\begin{aligned} \text{forget gate:} \quad \mathbf{f}_t &= \sigma(\mathbf{W}_f \mathbf{x}_t + \mathbf{R}_f \mathbf{h}_{t-1} + \mathbf{b}_f) &= \sigma(\bar{\mathbf{f}}_t) \\ \text{input gate:} \quad \mathbf{i}_t &= \sigma(\mathbf{W}_i \mathbf{x}_t + \mathbf{R}_i \mathbf{h}_{t-1} + \mathbf{b}_i) &= \sigma(\bar{\mathbf{i}}_t) \\ \text{candidate state:} \quad \mathbf{z}_t &= \tanh(\mathbf{W}_c \mathbf{x}_t + \mathbf{R}_c \mathbf{h}_{t-1} + \mathbf{b}_c) &= \tanh(\bar{\mathbf{z}}_t) \\ \text{cell state:} \quad \mathbf{c}_t &= \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \mathbf{z}_t \\ \text{cell gate:} \quad \mathbf{o}_t &= \sigma(\mathbf{W}_o \mathbf{x}_t + \mathbf{R}_o \mathbf{h}_{t-1} + \mathbf{b}_o) &= \sigma(\bar{\mathbf{o}}_t) \\ \text{output:} \quad \mathbf{h}_t &= \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \end{aligned}$$

Kinerja model diukur dengan *root mean squared error* (RMSE), bias, *Mean Absolute Percentage Error* (MAPE), Korelasi Pearson, *Proportion of Correct* (PC), dan *Peirce Skill Score* (PSS). Konstruksi dari arsitektur dilakukan secara iteratif sampai dengan RMSE terendah ditemukan. RMSE merupakan nilai rata-rata dari jumlah kuadrat kesalahan, juga dapat menyatakan ukuran besarnya kesalahan yang dihasilkan oleh suatu model prakiraan (Armstrong, 2001). RMSE dapat dinyatakan dengan rumus :

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - x_i)^2}$$

y = curah hujan observasi

x = curah hujan output model

n = Jumlah data

Bias merupakan selisih rata-rata antara prediksi dengan observasinya, jika bias bernilai positif maka prediksi *overestimated* (di atas nilai prediksinya) dan apabila negatif berarti nilai prediksi *underestimated*. Nilai bias semakin mendekati angka nol akan semakin baik (Walther dan Moore, 2005).

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n (y_i - x_i)$$

x = nilai CH observasi

y = nilai CH model

n = Jumlah data

Mean Absolute Percentage Error (MAPE) mengukur tingkat kesalahan relatif dalam bentuk persentase antara nilai prediksi dan nilai observasi (data aktual). MAPE juga dengan *Mean Absolute Percentage Deviation* (MAPD) (Rumus Statistik, 2021). Rumus untuk menghitung MAPE adalah sebagai berikut:

$$\text{MAPE (\%)} = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - y_i}{x_i} \right| \times 100$$

x = nilai CH observasi

y = nilai CH model

n = Jumlah data

Korelasi pearson digunakan untuk mengetahui seberapa besar pengaruh dan hubungan antara satu variabel dengan variabel lainnya. Dalam penelitian ini korelasi pearson digunakan untuk mencari kekuatan hubungan antara nilai curah hujan hasil dari model dan observasi. Korelasi pearson dapat dirumuskan dengan rumus sebagai berikut (Kustituantanto dan Rudy, 1994):

$$r_{xy} = \frac{n \sum xy - \sum x \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \cdot \sqrt{n \sum y^2 - (\sum y)^2}}$$

x = nilai curah hujan observasi

y = nilai curah hujan model

n = Jumlah data

Proportion of Correct (PC) adalah metode yang digunakan untuk mengukur tingkat akurasi atau ketepatan suatu model prediksi atau klasifikasinya.

Akurasi adalah ukuran deskriptif seberapa dekat nilai prakiraan dengan aktual. *Proportion of Correct* secara umum sama dengan *Percent of Correct* Keduanya digunakan secara untuk menggambarkan ketepatan metrik evaluasi yang sama, yaitu proporsi atau persentase memberikan gambaran tentang tingkat keberhasilan prediksi atau proporsi prediksi yang benar. Finley dalam Stephenson (2000) menilai akurasi prediksi dengan mempertimbangkan koefisien pencocokan sederhana berdasarkan 'proporsi' dari total prediksi yang 'benar' untuk hasil yang sesuai dan penolakan. Ketepatan pengukuran disebut sebagai PC. Nilai PC bervariasi dari 0 hingga 1, jika Semakin dekat dengan satu, semakin akurat. prakiraan meningkat untuk memprediksi nilai aktualnya (Muharsyah, 2018).

Metode untuk proses memverifikasi prakiraan model multi kategori dimulai dengan tabel kontingensi yang menunjukkan frekuensi prakiraan dan pengamatan di berbagai kotak, dengan plot yang terbagi untuk setiap kategori. Tabel kontingensi multi kategorial digunakan untuk menentukan PC. Tabel ini juga digunakan untuk perhitungan nilai PSS yang menjadi metode verifikasi yang menunjukkan nilai keunggulan (*skill*) model. Data model dan data aktual diubah ke dalam bentuk tabel kontingensi. Kategori nilai curah hujan yang digunakan mengikuti prakiraan curah hujan dasarian BMKG. Kategori untuk curah hujan kategorial meliputi:

Tabel 2. Kategori curah hujan dasarian

Kelas kuantitatif CH (mm)	Kategori CH
0-10	RENDAH
11--20	
21-50	
51-75	SEDANG
76-100	
101-150	
151-200	TINGGI
201-300	
>300	SANGAT TINGGI

Tabel kontingensi untuk curah hujan kategorial adalah sebagai berikut:

Tabel 3. Kontingensi untuk curah hujan kategorial

FREKUENSI (<i>i,j</i>)		OBSERVASI (O)									JUMLAH
		0-10	11--20	21-50	51-75	76-100	101-150	151-200	201-300	>300 (K)	
PRAKIRAAN (F)	0-10	nP ₁	0	0	0	0	0	0	0	0	NF ₁
	11--20	0	nP ₂	0	0	0	0	0	0	0	NF ₂
	21-50	0	0	nP ₃	0	0	0	0	0	0	NF ₃
	51-75	0	0	0	nP ₄	0	0	0	0	0	NF ₄
	76-100	0	0	0	0	nP ₅	0	0	0	0	NF ₅
	101-150	0	0	0	0	0	nP ₆	0	0	0	NF ₆
	151-200	0	0	0	0	0	0	nP ₇	0	0	NF ₇
	201-300	0	0	0	0	0	0	0	nP ₈	0	NF ₈
	>300 (K)	0	0	0	0	0	0	0	0	nP ₉	NF ₉
JUMLAH		NO ₁	NO ₂	NO ₃	NO ₄	NO ₅	NO ₆	NO ₇	NO ₈	NO ₉	N

Pada tabel ini $N(F_i, O_j)$ menunjukkan jumlah prakiraan dalam kategori i yang memiliki aktual dalam kategori j (*hit*), $N(F_i)$ menunjukkan jumlah prakiraan dalam kategori i , $N(O_j)$ menunjukkan jumlah aktual dalam kategori j , dan N adalah jumlah perkiraan. Pendekatan distribusi ini bertujuan memverifikasi peramalan dengan memeriksa hubungan antara elemen-elemen dalam tabel kontingensi multi kategori. Hal ini jelaskan oleh Murphy dan Winkler (1987), Murphy (1989), dan Brooks dan Doswell, (1996) dalam laman *Centre for Australian Weather and Climate Research* (CAWCR). Kelebihan dari pendekatan distribusi adalah sifat kesalahan model prakiraan dapat lebih mudah didiagnosis. Kekurangan adalah lebih sulit untuk memadatkan hasil menjadi satu angka.

PC dirumuskan sebagai jumlah proporsi prakiraan yang benar dibagi dengan jumlah data observasi dalam setiap kategori. PC dapat dirumuskan sebagai berikut:

$$PC = \frac{1}{N} \sum_{i=1}^k n(F_i, O_j)$$

Rumus ini dapat disederhanakan dengan membuat tabel kontingensi baru dimana semua elemen F_i dan O_j telah dibagi jumlah *hit* (N), sehingga rumus PC menjadi:

$$PC = \sum_k P_{ii} \quad i = 1, \dots, k$$

P_{ii} = *Proportion Correctness* (elemen diagonal) kategori i

k = banyaknya kategori

F_j = *hit* prakiraan kategori i

O_j = *hit* observasi kategori j

Peirce Skill Score (PSS) adalah metrik untuk mengukur kualitas dan akurasi dari peramalan atau prediksi cuaca terhadap data observasi (data sebenarnya). PSS mengukur sejauh mana peramalan cuaca atau model iklim lebih baik daripada peramalan acak (*random forecast*) dalam memprediksi kejadian cuaca. Skor PSS memiliki rentang dari -1 hingga 1. Sebagian kecil dari perkiraan yang benar adalah karena peluang acak untuk perkiraan yang tidak bias. PSS menghindari bias yang mungkin muncul dalam metrik seperti akurasi yang hanya mempertimbangkan prediksi yang benar tanpa mempertimbangkan tingkat kesulitan dalam memprediksi kejadian yang jarang terjadi. Perhitungan metrik PSS menyerupai perhitungan *Heidke Skill Score* (HSS), namun sedikit berbeda pada perhitungan pembagiannya. PSS mampu mengukur kemampuan tanpa terganggu oleh base rate. PSS cocok untuk mengukur distribusi bersama, yaitu untuk menampilkan akurasi prediksi potensial itu sendiri (Pelliccioni dan Cotroneo, 2009).

Berikut rumus PSS (CAWCR, 2009):

$$PSS = \frac{\frac{1}{N} \sum^k (A_{ii}) \frac{1}{N^2} \sum^k N(F_i) N(O_i)}{1 - \frac{1}{N^2} \sum^k (N(O_i))^2}, i = 1, \dots, k$$

Sama seperti PC rumus ini dapat disederhanakan dengan membuat tabel kontingensi baru dimana semua elemen F_i dan O_j telah dibagi jumlah *hit* (N), sehingga rumus PSS menjadi

$$PSS = \sum (NO_i)^2 \quad i = 1, \dots, k$$

A_{ii} = Jumlah kejadian *hit* kategori i

NF_i = Jumlah prakiraan kategori i

NO_i = Jumlah observasi kategori i

N = Jumlah seluruh kejadian

k = banyaknya kategori

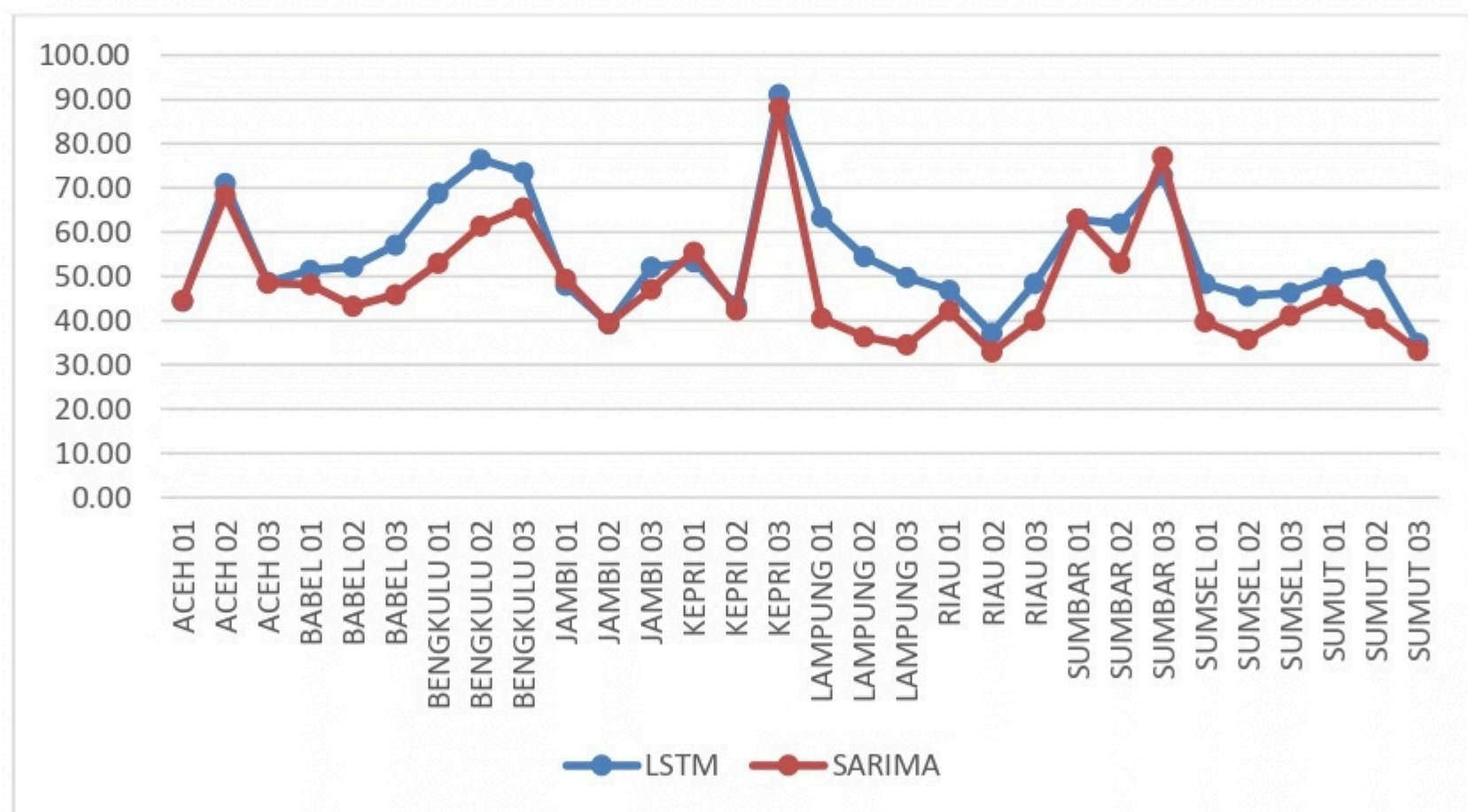
Nilai PSS dan HSS cenderung mirip atau mendekati nilai yang sama. Keduanya merupakan metrik evaluasi yang digunakan untuk mengukur kinerja

model peramalan atau prediksi dengan berfokus pada kategori klasifikasi atau distribusi spasial. Kategori nilai HSS (Muharsyah, 2018) digunakan untuk kategori PSS pada penelitian ini dimana $-1 \leq PSS < -0,05$ sebagai kategori buruk, $-0,05 \leq PSS \leq 0,05$ Tidak Ada, dan $0.05 < PSS \leq 1$ ada atau memiliki keunggulan.

HASIL PENELITIAN DAN PEMBAHASAN

Hasil verifikasi kedua model secara berurutan dimulai dari perhitungan nilai galat, arah kesalahan, korelasi, akurasi, dan keunggulan. Setiap metode dapat mendeskripsikan parameter perhitungan kesalahan yang berbeda, namun hasil verifikasi merupakan nilai yang berhubungan.

Grafik perbandingan RMSE dibuat dalam skala aktual (mm). Rata-rata nilai pada Model LSTM berada pada 54,83 mm sedangkan model SARIMA berada pada 48,55 mm. Nilai RMSE terbesar untuk kedua model berada pada ZOM Kepri 03 dengan nilai 91,20 mm untuk LSTM dan 88,15 mm untuk SARIMA. Nilai RMSE terkecil untuk model LSTM berada pada ZOM Sumut 03 dengan nilai 34,87 dan untuk model SARIMA berada pada ZOM Riau 02 dengan nilai 32,90.



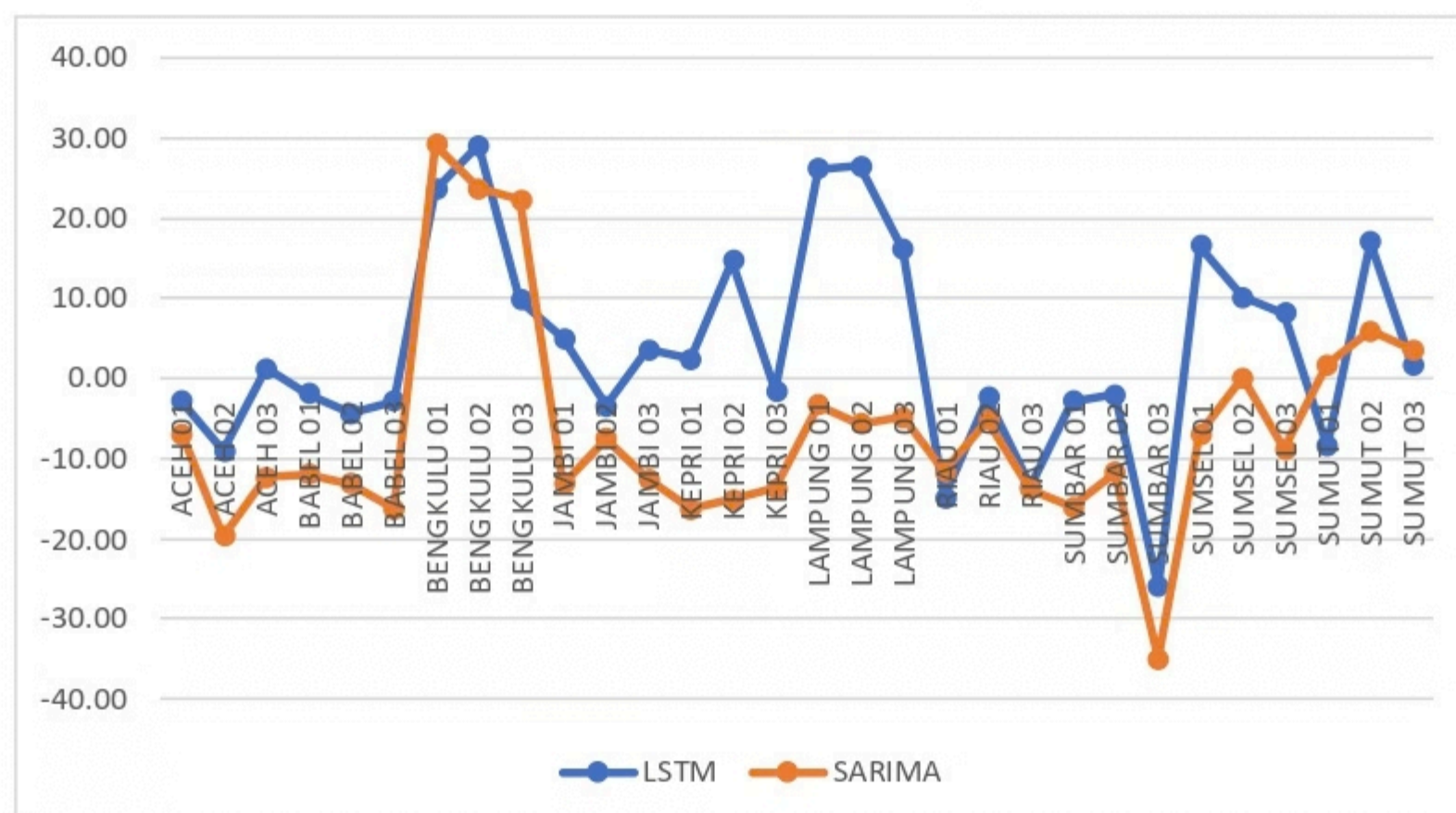
Gambar 2. Grafik perbandingan RMSE (mm)

Secara keseluruhan nilai RMSE kedua model terbilang mirip. Grafik menunjukkan pada beberapa ZOM nilai RMSE model lebih tinggi dan pada sebagian yang lain nilai RMSE lebih rendah hingga membentuk pola yang sama. Kemiripan pola ini menunjukkan beberapa ZOM memiliki nilai aktual yang cukup

sulit diprediksi oleh kedua model, dan sebagian yang lain lebih mudah diprediksi.

Perbedaan nilai yang mencolok antara beberapa ZOM dapat diartikan bahwa model dapat menangkap dengan baik pola curah hujan saat RMSE rendah dan sebaliknya saat RMSE tinggi. Peluang nilai curah hujan yang acak pada kejadian nyata dapat sewaktu-waktu keluar dari pola iklimnya yang mengakibatkan model tidak mampu memprediksi nilai aktual meski telah menunjukkan pola yang sesuai.

Grafik perbandingan bias dibuat dalam skala aktual (mm). Bias terbesar (*overestimated*) untuk model LSTM berada pada ZOM Bengkulu 02 dengan nilai 29,03 mm, dan untuk model SARIMA berada pada ZOM Bengkulu 01 dengan nilai 29,19 mm. Bias mines terbesar (*underestimate*) untuk kedua model berada pada ZOM yang sama yaitu SUMBAR 03 dengan nilai -25,94 untuk model LSTM dan -34,98mm untuk model SARIMA. Rata-rata nilai bias kedua model adalah 3,86 mm (*overestimated*) untuk model LSTM dan -6,49 mm (*underestimate*) untuk model SARIMA.



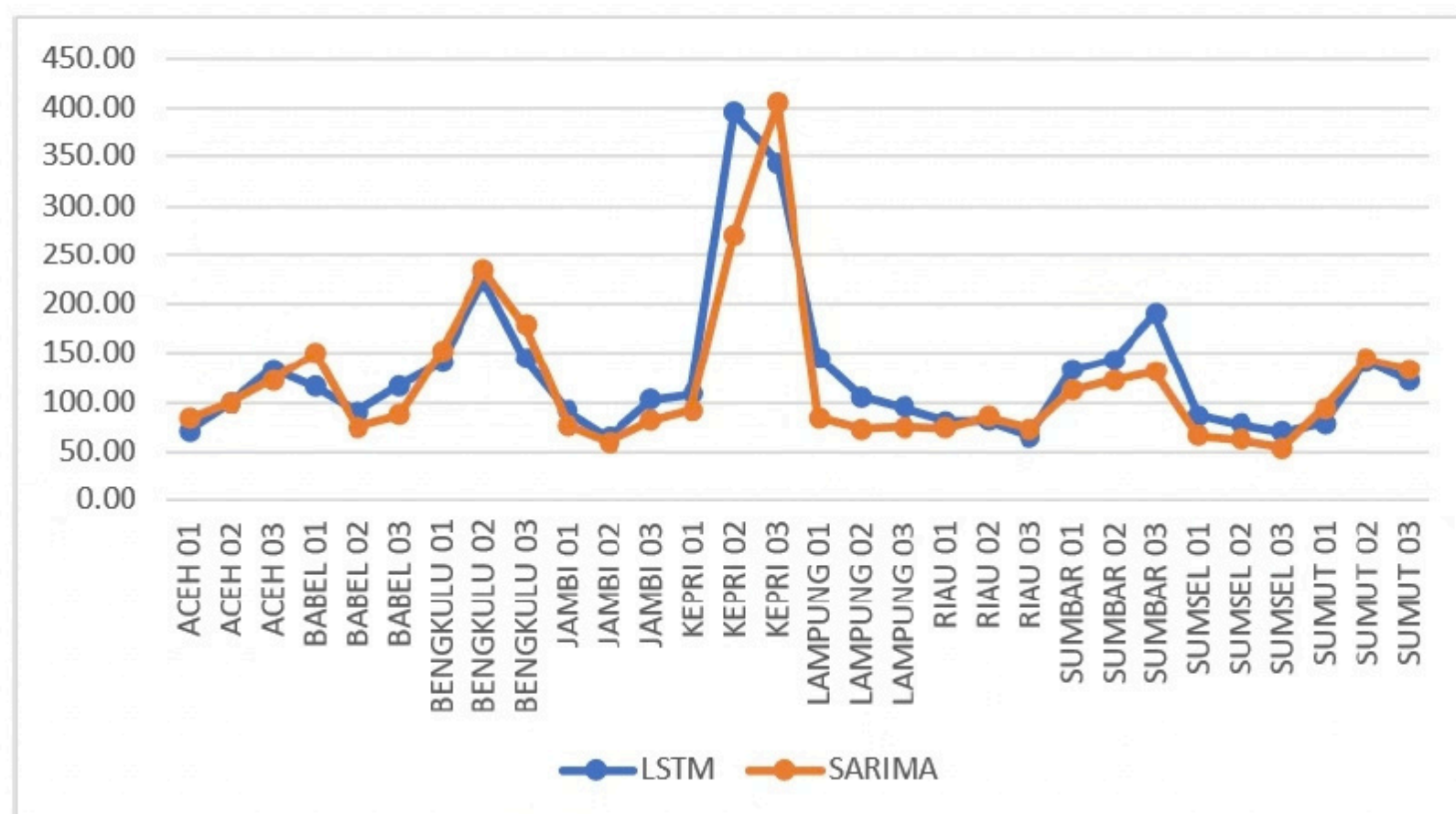
Gambar 3. Grafik perbandingan Bias (mm)

Secara keseluruhan model SARIMA memiliki bias yang lebih besar baik secara *overestimated* maupun *underestimate*. Model SARIMA memiliki hasil yang cenderung *flat* sehingga apabila nilai prakiraan berdistribusi lebih besar ataupun lebih kecil dari data aktual maka bias akan bernilai besar. Model SARIMA memiliki arah kesalahan negatif yang berkebalikan dengan model

LSTM.

Hampir sama dengan RMSE grafik bias menunjukkan pada beberapa ZOM, nilai bias kedua model *overestimated* dan pada sebagian yang *underestimated*. Nilai ini membentuk pola yang sama antar kedua model. Pola ini menunjukkan beberapa ZOM memiliki nilai aktual yang cukup sulit diprediksi oleh kedua model, dan sebagian yang lain lebih mudah diprediksi. Pola ini juga menunjukkan sifat hujan beberapa ZOM memiliki hubungan terhadap arah kesalahan nilai prakiraan model. Bias tidak mengukur besarnya kesalahan. Tidak mengukur korespondensi antara prakiraan dan pengamatan, yaitu, dimungkinkan untuk mendapatkan skor sempurna untuk prakiraan buruk jika ada kesalahan kompensasi. Metode bias hanya digunakan untuk melihat bagaimana arah kesalahan dari model prakiraan.

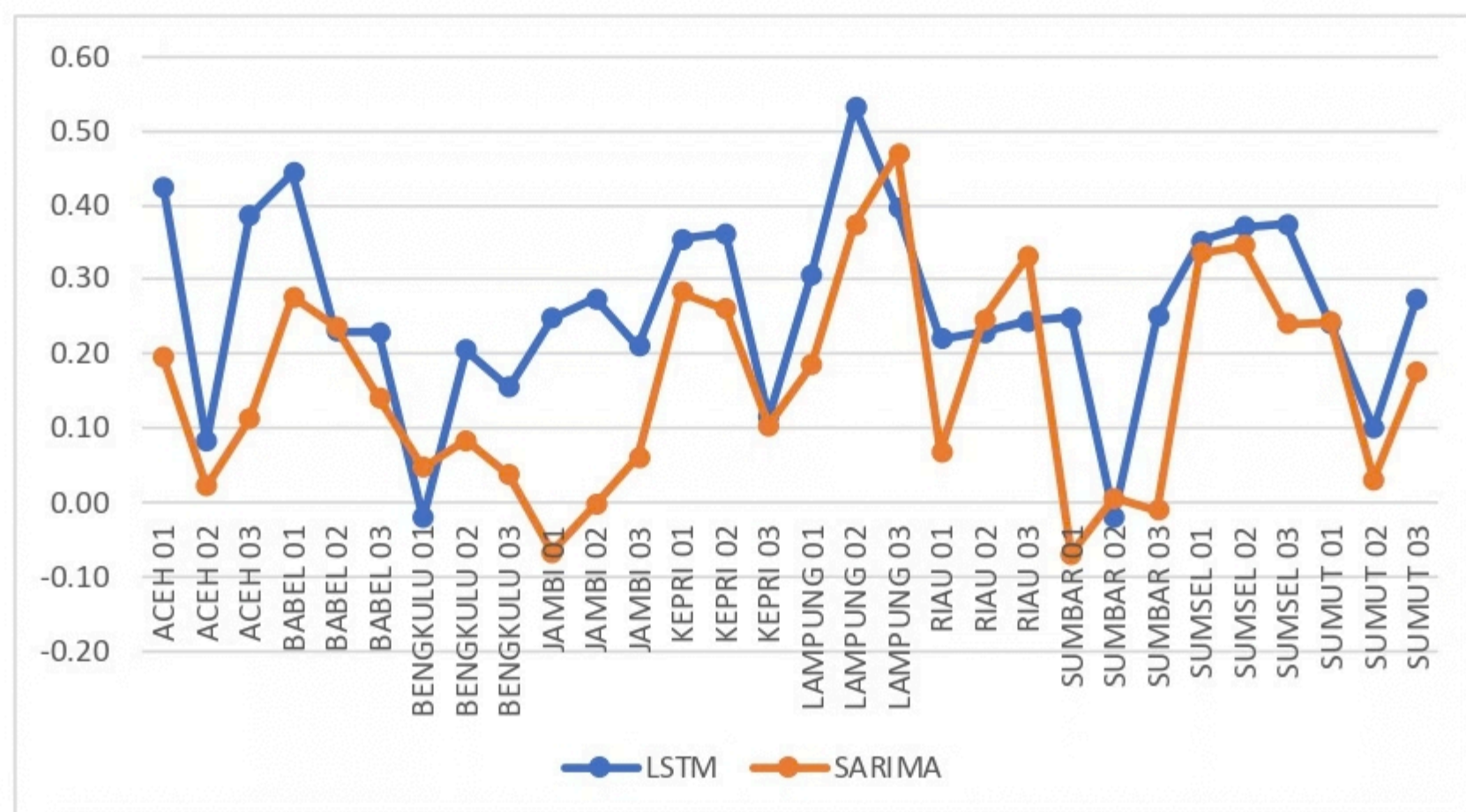
Grafik perbandingan MAPE dibuat dalam skala persen (%). MAPE terbesar untuk model LSTM berada pada ZOM Kepri 02 dengan nilai 396,07%, dan untuk model SARIMA berada pada ZOM Kepri 03 dengan nilai 405,38 %. MAPE terkecil untuk model LSTM berada pada ZOM Jambi 02 dengan nilai 64,44 %, dan untuk model SARIMA berada pada ZOM Sumsel 03 dengan nilai 54 %. Rata-rata nilai MAPE kedua model adalah 128,54 % untuk LSTM, dan 118,44 % untuk SARIMA.



Gambar 4. Grafik perbandingan MAPE (%)

Kedua MAPE model berada lebih dari 100 % untuk nilai rataannya. MAPE membentuk kesesuaian pola nilai antar kedua model, dimana ZOM Kepri 02 dan 03 merupakan yang paling mencolok karena memiliki nilai melebihi 270% untuk kedua model. RMSE dan MAPE menggunakan selisih absolut sebagai bobot penilaian. Pada RMSE selisih yang lebih besar akan lebih berdampak besar dalam peningkatan nilai. MAPE mempertimbangkan data aktual sebagai pembagi selisih. Ketika nilai data aktual curah hujan dasarian rendah namun hasil model menunjukkan nilai yang tinggi akan sangat berpengaruh terhadap peningkatan nilai MAPE. Nilai MAPE membentuk pola dimana dapat dikatakan beberapa ZOM memiliki nilai aktual yang cukup sulit diprediksi oleh kedua model, dan sebagian yang lain lebih mudah diprediksi.

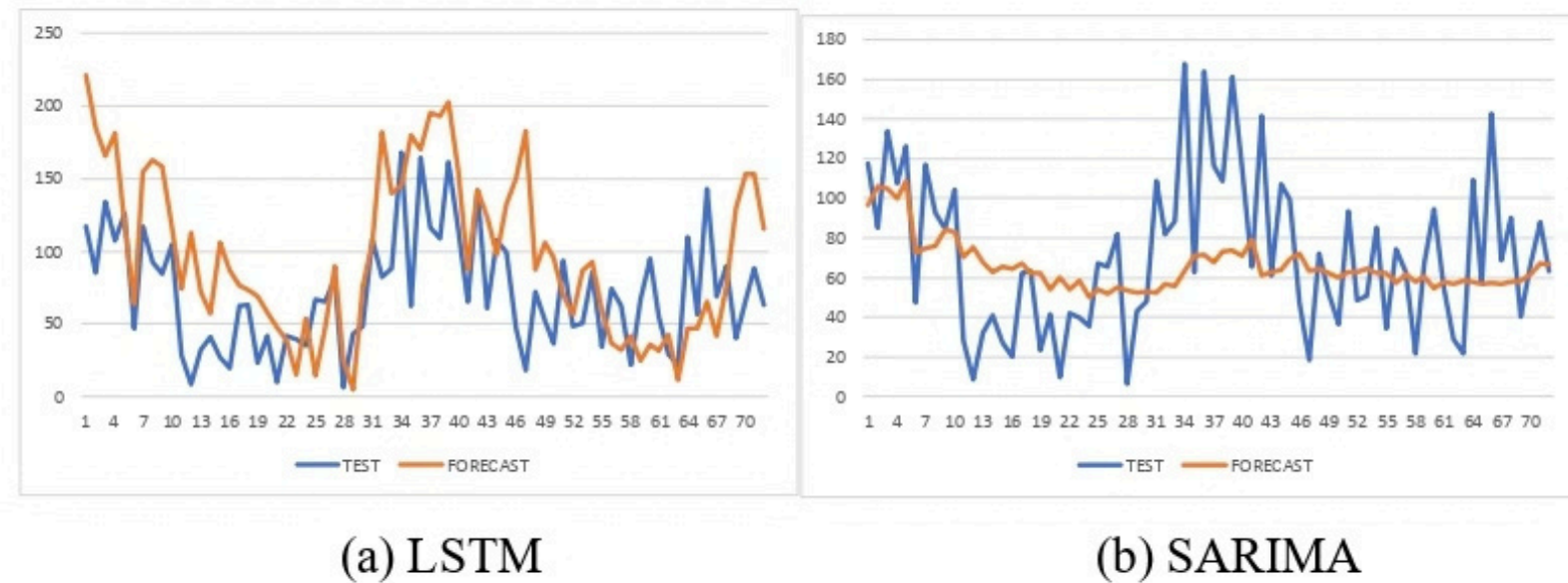
Grafik nilai korelasi per-ZOM bervariasi. Nilai korelasi ini menunjukkan seberapa mirip kesesuaian nilai antara model dan observasi. Nilai rata-rata untuk model LSTM 0,261 yang masuk dalam kategori korelasi lemah, dan untuk model SARIMA 0,159 yang masuk dalam kategori sangat lemah (Sugiyono, 2015).



Gambar 5. Grafik perbandingan nilai korelasi Pearson setiap ZOM

Dari kedua model cukup sulit untuk mendapatkan nilai korelasi melebihi 0,5. Nilai korelasi memiliki kemiripan di beberapa beberapa ZOM. Beberapa ZOM pada kedua hasil model terdapat pada nilai korelasi negatif, yang berarti berkorelasi buruk (berlawanan). Pola Nilai ini menunjukkan model dapat menangkap pola curah hujan dasarian selama 2 tahun dengan baik di beberapa

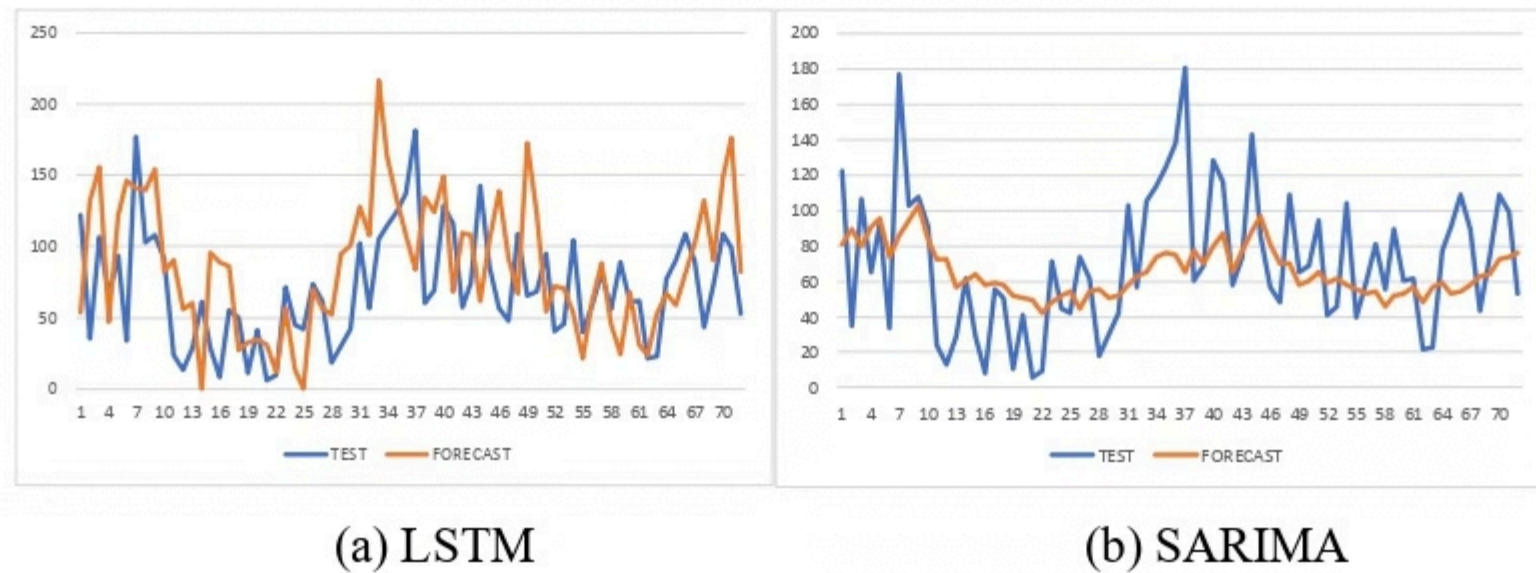
ZOM dan beberapa yang lain lebih sulit diprediksi polanya. Sebagai contoh, korelasi dengan nilai tertinggi dan terendah di setiap model akan ditelaah secara terpisah.



Gambar 6. Grafik Korelasi Pearson ZOM Lampung 02 (a) LSTM, (b) SARIMA

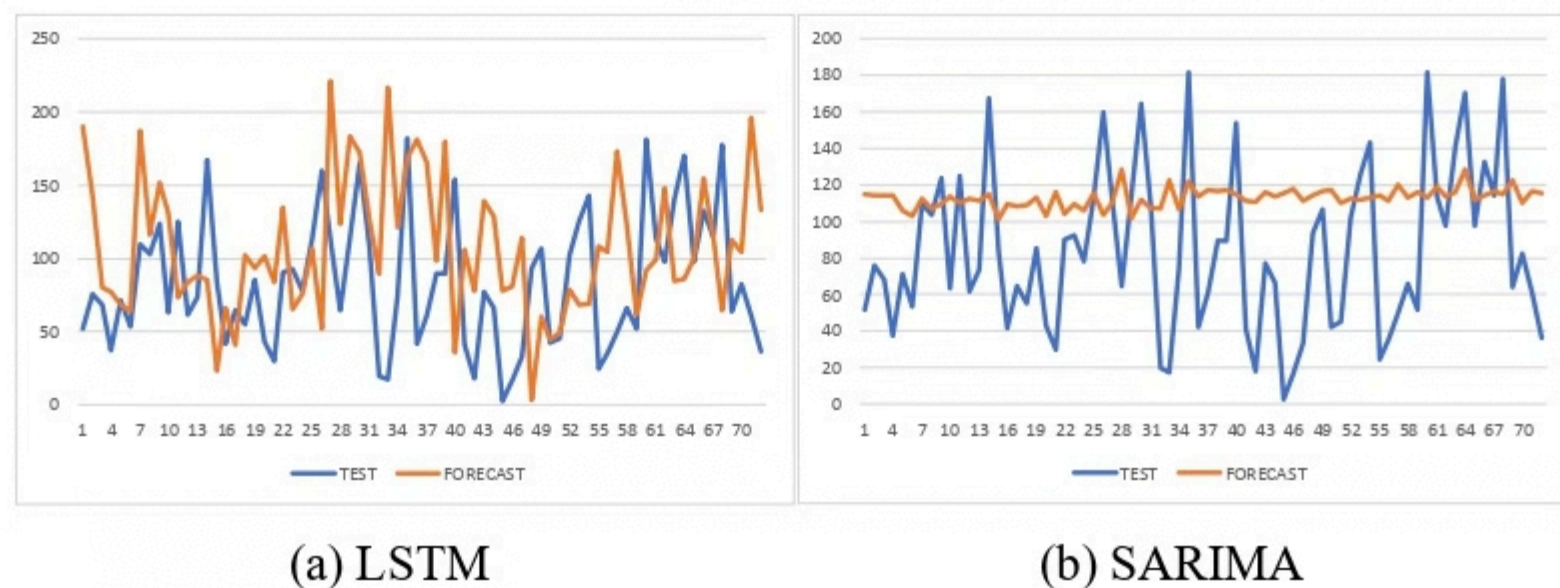
Model LSTM pada ZOM Lampung 02 memiliki nilai korelasi 0,53 yang merupakan nilai korelasi terbesar, sedangkan model SARIMA pada ZOM yang sama memiliki nilai 0,38 kedua nilai ini masuk dalam kategori korelasi sedang. Model LSTM dapat menangkap pola dengan sangat baik pada ZOM ini, namun beberapa nilai ekstrem dan pola detail perdasarian tidak dapat tunjukkan oleh model ini. Dari korelasi ini dapat dikatakan LSTM mampu menangkap pola data dalam masa *training*, namun masih belum cukup baik dalam menginterpretasikan nilai curah hujan dasarian secara detail dalam hasil model. Model SARIMA tidak menunjukkan nilai yang jauh berbeda, hal ini menunjukkan pada ZOM Lampung 02 kedua model dapat menangkap pola dalam masa *training*.

Perbedaan paling mencolok pada kedua model ialah hasil keluaran model LSTM dapat menginterpretasikan nilai yang beragam dan tetap memiliki variasi nilai beragam, sedangkan pada model SARIMA variasi nilai yang ditunjukkan sangat kecil meski masih mampu mengikuti pola data aktual. Melalui grafik ini dapat dikatakan bahwa SARIMA mengambil rata-rata nilai tengah data pada pola.



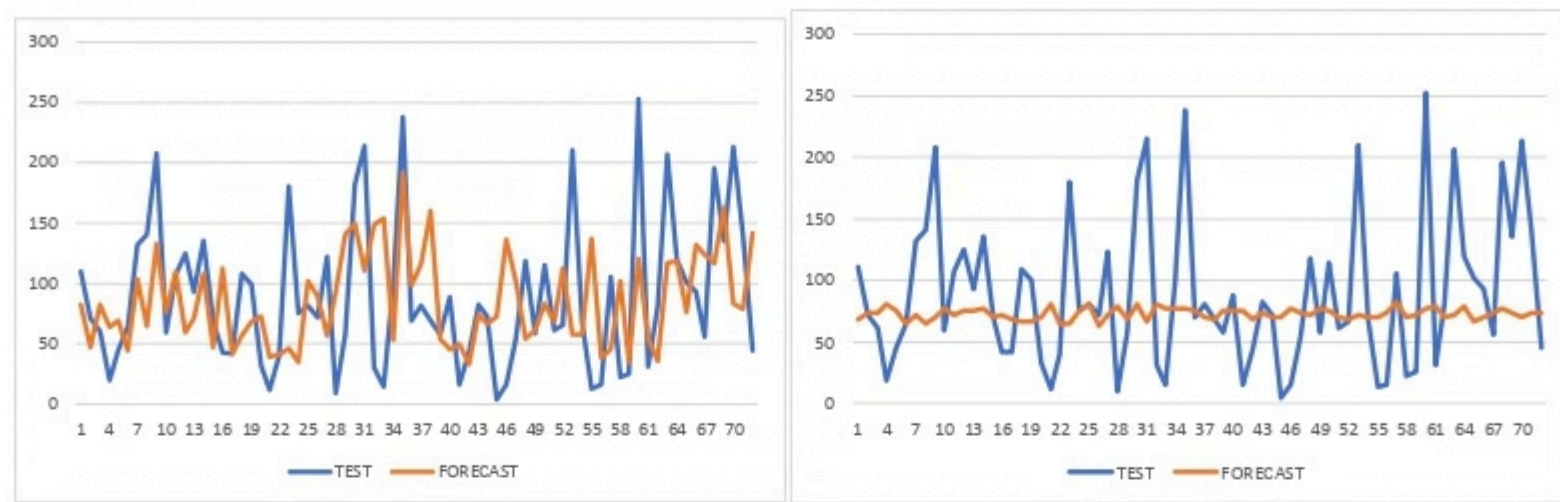
Gambar 7. Grafik Korelasi Pearson ZOM Lampung 03 (a) LSTM, (b) SARIMA

Berikutnya merupakan perbandingan grafik korelasi dengan nilai korelasi tertinggi SARIMA, yaitu pada ZOM Lampung 03. Model SARIMA memiliki nilai korelasi 0,47 sedangkan pada model LSTM 0,40. Kedua model masih memiliki korelasi yang terbilang baik dimana model mampu menangkap pola data *training* namun cukup sulit dalam interpretasi hasilnya secara detail.



Gambar 8. Grafik Korelasi Pearson ZOM Bengkulu 01 (a) LSTM, (b) SARIMA

Grafik berikut merupakan korelasi terendah untuk model LSTM pada ZOM Bengkulu 01 dengan korelasi -0,02, dan pada ZOM yang sama SARIMA memiliki korelasi 0,05. LSTM sebenarnya telah cukup mampu dalam menangkap pola model namun pada ZOM berikut model tidak dapat meinterpretasikan hasil dengan baik. Model SARIMA memiliki hasil keluaran yang *flat* (mendatar). Meski memiliki nilai korelasi yang lebih baik, nilai model hanya berada pada nilai tengah data aktual, dan cenderung overestimated yang mana hal ini menjawab nilai bias pada model SARIMA untuk ZOM Bengkulu 1 yaitu sebesar 29,19 mm.



(a) LSTM

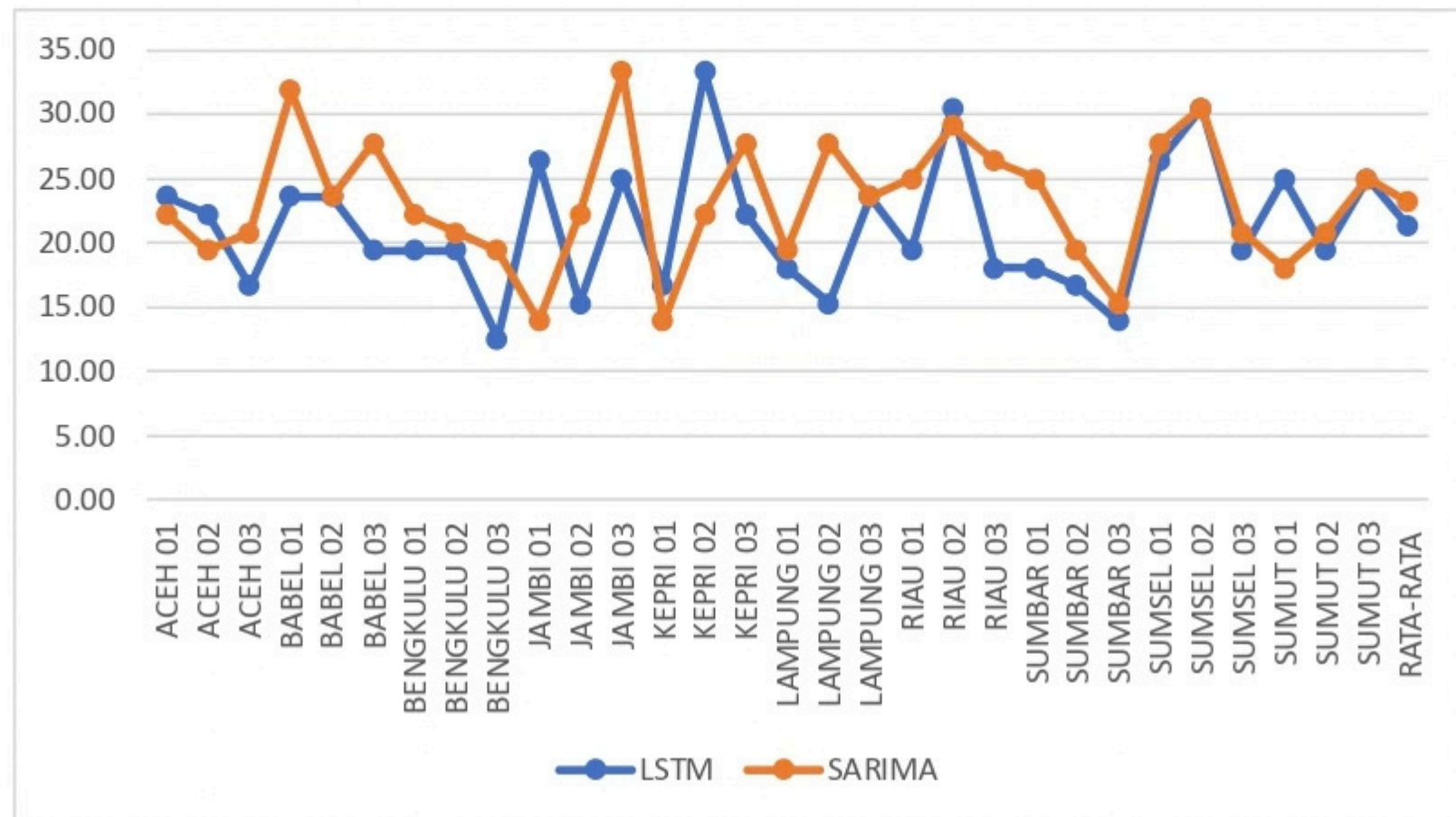
(b) SARIMA

Gambar 9. Grafik Korelasi Pearson ZOM Sumbar 01 (a) LSTM, (b) SARIMA

Korelasi terendah untuk model SARIMA terdapat pada ZOM Sumbar 01 dengan korelasi $-0,07$. Terlihat hasil keluaran model SARIMA sangat *flat* dan dalam pandangan visual, tampak kurang memiliki kesamaan. Pada ZOM yang sama model LSTM memiliki korelasi sebesar $0,25$, dimana model masih dapat menangkap pola namun masih cukup sulit untuk interpretasikan nilai keluarannya. Secara umum model LSTM mampu menangkap pola data dalam masa *training*, namun masih belum cukup baik dalam menginterpretasikan nilai curah hujan dasarian secara detail dalam hasil model.

Nilai PC dinyatakan dalam persen (%). Nilai rata-rata PC kedua model adalah $21,30\%$ untuk LSTM dan $23,19\%$ untuk SARIMA. Nilai PC tertinggi untuk model LSTM adalah PC pada ZOM Kepri 02 sebesar $33,33\%$, dan untuk model SARIMA berada pada ZOM Jambi 03 dengan nilai yang sama. Nilai rata-rata berada pada kategori rendah dan sangat rendah merujuk pada penelitian Muharsyah (2018), meskipun pada penelitian tersebut digunakan untuk prakiraan curah hujan bulanan.

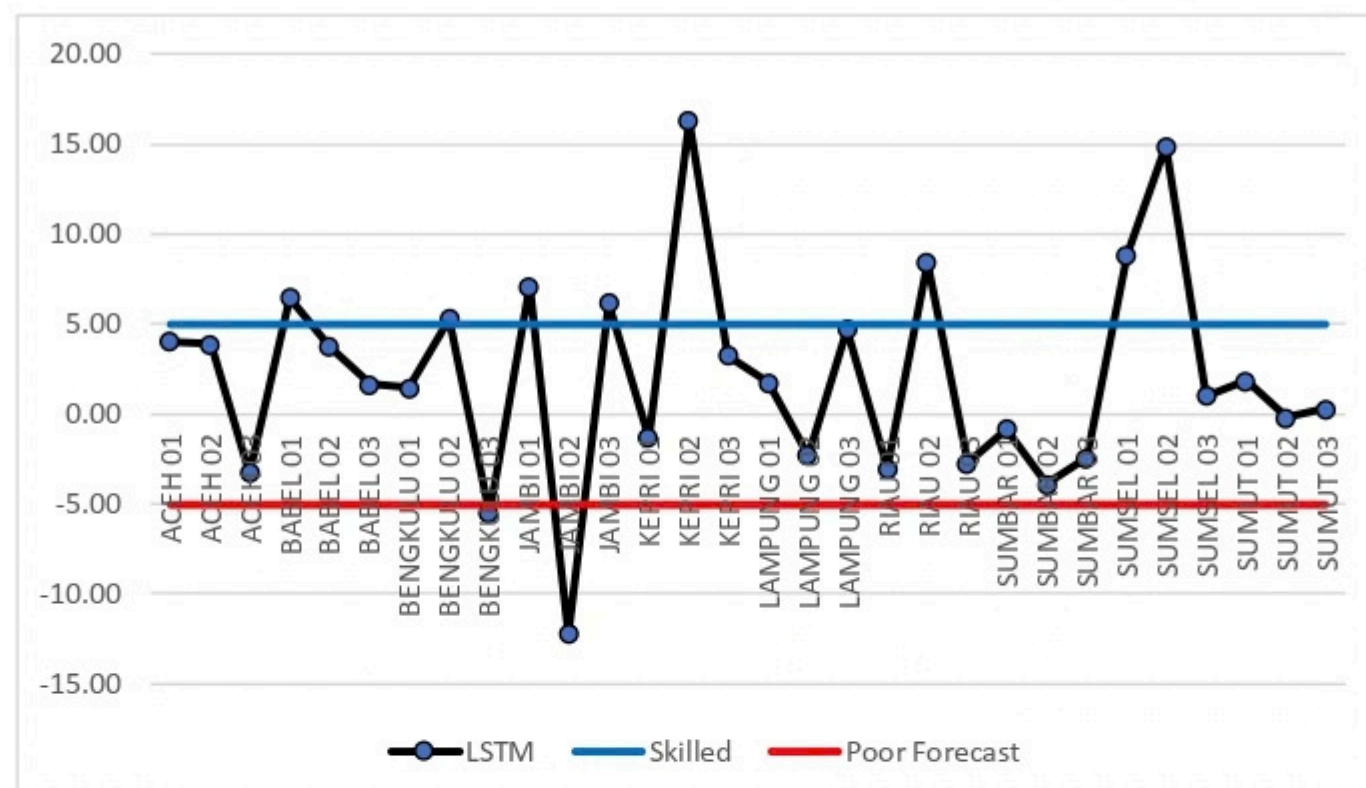
EVALUASI MODEL LONG-SHORT TERM MEMORY (LSTM) DAN SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA) UNTUK PREDIKSI CURAH HUJAN DI WILAYAH SUMATRA



Gambar 10. Grafik perbandingan nilai PC (%)

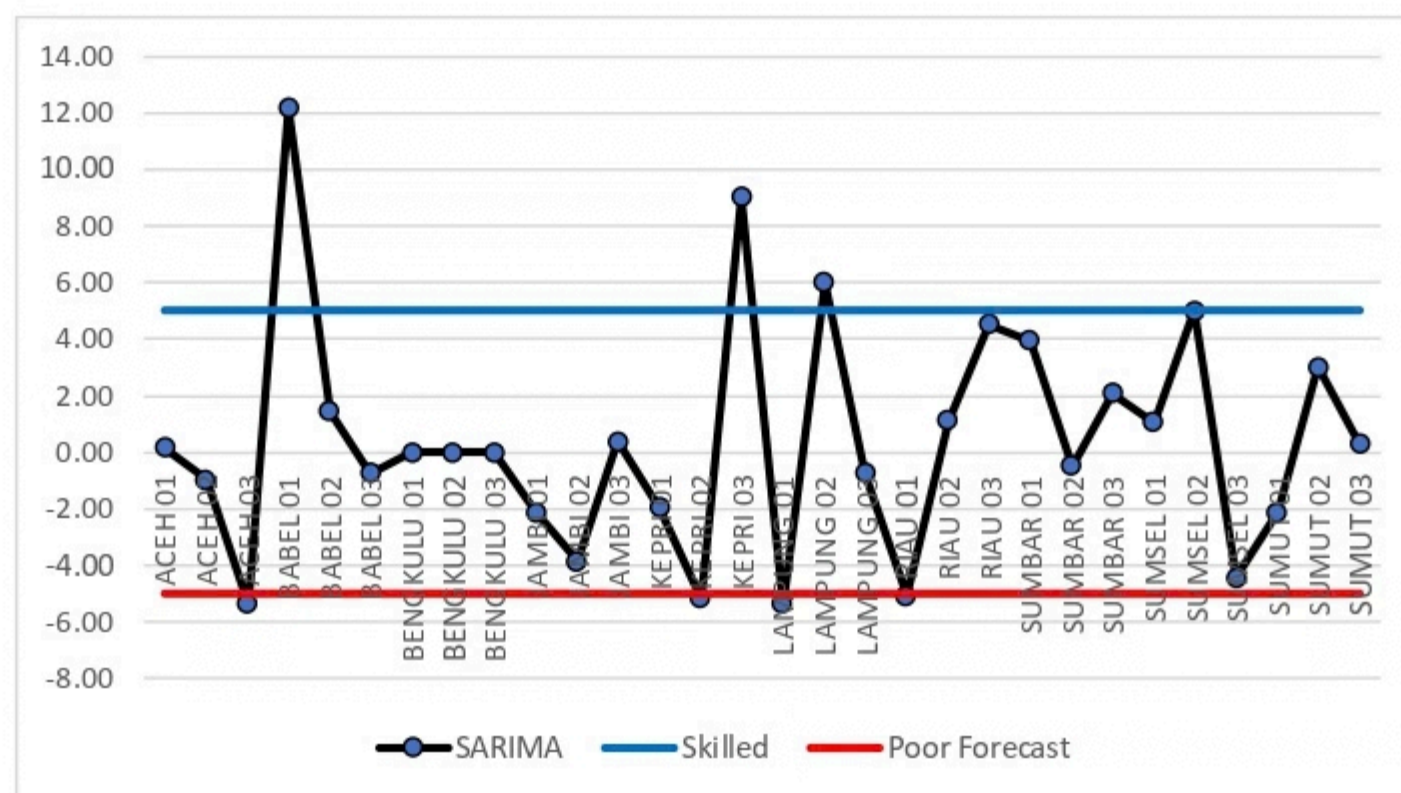
SARIMA memiliki nilai PC yang lebih tinggi dibandingkan LSTM, perbedaan nilai keduanya sangat kecil dan masih berada pada kategori yang sama. Sehingga secara keseluruhan hasil PC kedua model dianggap sama.

PSS rata-rata untuk model LSTM adalah 2,11 %. Nilai PSS dengan kategori baik dan memiliki keunggulan (*skilled*) berada diatas garis biru, yaitu ZOM dengan nilai PSS diatas 5%. Model LSTM memiliki 8 ZOM dengan kategori baik. Jumlah ZOM dengan kategori tanpa keunggulan untuk model LSTM sebanyak 20 ZOM. Jumlah ZOM dengan kategori buruk untuk model LSTM sebanyak 2 ZOM. Model LSTM memiliki lebih banyak ZOM dengan kategori baik dibandingkan dengan kategori buruk (*poor forecast*).



Gambar 11. Nilai PSS (%) pada model LSTM

PSS rata-rata untuk model SARIMA adalah 0,39 %. Model SARIMA memiliki 4 ZOM dengan kategori baik. Jumlah ZOM dengan kategori tanpa keunggulan untuk sebanyak 22 ZOM, dan jumlah ZOM dengan kategori buruk sebanyak 4 ZOM. Model SARIMA memiliki jumlah ZOM dengan kategori baik dan buruk yang sama, hal ini menunjukkan secara keseluruhan model LSTM memiliki hasil PSS yang lebih baik dibandingkan dengan model SARIMA baik dari nilai rata-rata maupun jumlah kategori ZOM.



Gambar 12. Nilai PSS (%) pada model SARIMA

Verifikasi model prediksi curah hujan dasarian di wilayah Sumatera menggunakan sumber data (ZOM) dan metode yang berbeda menghasilkan karakteristik yang berbeda pada tiap hasilnya. Berikut merupakan pembahasan bagaimana hasil verifikasi pada masing-masing model.

Nilai rata-rata RMSE model SARIMA masuk dalam kategori kecil. Fluktuasi nilai serupa ditunjukkan pada model LSTM namun nilai RMSE pada model SARIMA cenderung lebih kecil. Nilai RMSE yang lebih rendah menunjukkan model memiliki kesalahan yang lebih kecil dalam memprediksi curah hujan di lokasi tersebut. Nilai RMSE lebih tinggi di beberapa ZOM, hal ini dapat menandakan beberapa ZOM memiliki karakteristik pola curah hujan yang lebih sulit untuk diprediksi sehingga galatnya lebih besar. Hasil serupa

ditunjukkan pada pengukuran MAPE yang sama-sama termasuk pengukuran metrik galat absolut.

Nilai MAPE SARIMA lebih dari 100% untuk beberapa ZOM, ini menunjukkan bahwa rata-rata kesalahan absolut dalam prediksi melebihi nilai curah hujan yang sebenarnya. Situasi seperti ini dapat mengisyaratkan model peramalan mungkin tidak cocok atau tidak efektif dalam memprediksi curah hujan di ZOM tersebut. Sebagai Faktor *non-linier* lingkungan tropis seringkali memiliki variasi yang *non-linier* dan kompleks dalam curah hujan. Jika model peramalan tidak mampu menangkap pola *non-linier* ini, maka kesalahan prediksi dapat menjadi sangat besar. Beberapa perubahan dalam curah hujan tropis (Sumatra) dapat dipengaruhi oleh faktor-faktor eksternal seperti perubahan iklim, fenomena alam, dan lainnya. Model akan memiliki galat yang besar apabila model tidak mampu menangkap faktor eksternal ini. Keseluruhan untuk penilaian metrik kesalahan absolut menunjukkan model SARIMA memiliki kategori yang lebih baik dibanding model LSTM.

Pada pengukuran bias nilai menunjukkan sifat hujan beberapa ZOM memiliki hubungan terhadap arah kesalahan nilai prakiraan model. Bias hanya digunakan untuk melihat bagaimana arah kesalahan dari model prakiraan. Nilai bias model SARIMA lebih besar pada kedua arah dibandingkan model LSTM, ini menunjukkan bahwa model cenderung memiliki kecenderungan untuk salah memprediksi curah hujan dengan jumlah yang signifikan. Model SARIMA memiliki penyesuaian yang lebih agresif terhadap tren atau pola musiman yang tidak selalu tepat pada data curah hujan yang spesifik.

Nilai korelasi model SARIMA lebih kecil dibandingkan model LSTM. Hasil model SARIMA cenderung *flat*, dan hanya pada beberapa ZOM dengan memiliki pola yang baik. Pola ini lebih terlihat pada ZOM dengan karakteristik hujan musiman yang dominan, sehingga model lebih dapat mengikuti pola data latih. Data *flat* ini menjelaskan penyebab nilai galat absolut pada model SARIMA lebih kecil dibandingkan model LSTM namun memiliki bias yang lebih besar pada kedua arah. Kompleksitas data time series yang tinggi dan faktor lainnya dapat menjadi alasan utama model tidak dapat memprediksi dengan galat yang rendah.

Data curah hujan 10 harian jangka panjang (30 tahun) untuk wilayah Sumatra memiliki pola yang kompleks dan sulit untuk diprediksi. Namun tidak menutup kemungkinan bahwa terdapat faktor luar yang mempengaruhi nilai aktual sehingga model tidak dapat menginterpretasikan nilai dengan baik. Peluang nilai curah hujan yang acak pada kejadian nyata dapat sewaktu-waktu keluar dari pola iklimnya yang mengakibatkan model tidak mampu memprediksi nilai aktual meski telah menunjukkan pola yang sesuai.

Kategori nilai PC model SARIMA sama dengan model LSTM, namun nilai ini sedikit lebih tinggi untuk model SARIMA. Distribusi intensitas curah hujan yang sangat tidak seimbang dapat memberikan gambaran yang kurang akurat tentang performa model karena model mungkin akan cenderung memprediksi kategori mayoritas. Nilai PC dapat lebih tinggi apabila data prakiraan dan data aktual berada pada kategori yang sama meski sebaran data bersifat acak (*unskill forecast*), maka diperlukan perhitungan nilai *skill* (keunggulan) untuk melihat kemampuan model selain parameter keakuratan.

Model SARIMA memiliki jumlah kategori ZOM dengan keunggulan lebih sedikit dibandingkan model LSTM. Hasil model LSTM dan SARIMA memiliki sebaran rentang data yang berbeda namun nilai PC kedua model tidak jauh berbeda atau berada pada kategori yang sama. PSS membantu menjelaskan hasil akurasi yang sama dapat disebabkan kekurangan metode PC dalam membedakan peluang akurasi karena model memiliki keunggulan atau hanya sekedar *random forecast* ataupun berpaku pada nilai klimatologisnya. Secara keseluruhan untuk model SARIMA (dibandingkan model LSTM) dapat dikatakan hasil galat absolut lebih kecil, bias lebih besar, korelasi lebih rendah, akurasi setara, dan keunggulan lebih sedikit.

Nilai rata-rata RMSE model LSTM masuk dalam kategori kecil dan fluktuasinya serupa pada model SARIMA namun pada LSTM nilai ini sedikit lebih besar. Nilai RMSE lebih tinggi di beberapa ZOM, hal ini dapat menandakan beberapa ZOM memiliki karakteristik pola curah hujan yang lebih sulit untuk diprediksi sehingga galatnya lebih besar. Hasil serupa ditunjukkan pada pengukuran MAPE yang sama-sama termasuk pengukuran metrik galat absolut.

Nilai bias model LSTM lebih kecil pada kedua arah dibandingkan model SARIMA, ini menunjukkan bahwa model memiliki penyesuaian yang lebih baik terhadap tren atau pola musiman pada data curah hujan keseluruhan. Variabilitas nilai keluaran model LSTM secara umum telah mengikuti pola data uji, yang ditunjukkan nilai korelasi yang cukup pada beberapa ZOM, namun tingkat kesesuaian yang detail belum mampu ditunjukkan oleh hasil model. Berbeda dengan SARIMA dimana hasil model cenderung berada pada tengah nilai aktual, pada model LSTM beberapa ZOM meski dengan pola musiman yang tidak dominan tetap dapat diikuti oleh model LSTM. Meski demikian pola korelasi tidak konsisten terbaca pada semua ZOM. Pada beberapa ZOM, LSTM memiliki korelasi yang buruk yang umumnya juga diikuti oleh model SARIMA dengan kategori yang sama.

Kategori PC model LSTM sama dengan model SARIMA namun sedikit lebih rendah. Model LSTM memiliki jumlah ZOM dengan kategori keunggulan lebih banyak dan jumlah ZOM kategori keunggulan buruk lebih sedikit pada pengukuran PSS. Model LSTM memiliki keunggulan yang merupakan ketepatan relatif dari suatu prakiraan terhadap model prakiraan referensi (prakiraan tanpa referensi). Keunggulan juga diartikan sebagai peningkatan ketepatan prakiraan karena membaiknya sistem prakiraan tersebut.

Prakiraan yang tidak memiliki skill tidak akan bisa untuk meningkatkan nilai PSS dengan mengikuti nilai perkiraannya menuju klimatologi (nilai rata-rata). Skor keterampilan Peirce tidak bisa ditingkatkan dengan meramalkan kelas peristiwa tertentu ini adalah skor yang menyeluruh (Gandin dan Murphy, 1992). Secara keseluruhan untuk model LSTM (dibandingkan model SARIMA) dapat dikatakan hasil galat absolut lebih besar, bias lebih kecil, korelasi lebih tinggi, akurasi setara, dan keunggulan lebih banyak atau lebih baik.

SIMPULAN

1. Hasil verifikasi model SARIMA menunjukkan nilai perhitungan galat absolut (RMSE dan MAPE) yang lebih kecil dibandingkan pada model LSTM, namun nilai bias pada arah *overestimated* dan *underestimated* lebih besar. Nilai korelasi Pearson berada pada kategori sangat lemah. Pada uji ini juga menunjukkan sebagian besar hasil model cenderung *flat* dan tidak dapat menginterpretasikan nilai ekstrem, dan menjelaskan mengapa hasil galat absolut lebih kecil dibandingkan pada model LSTM namun memiliki bias yang lebih tinggi. Nilai PC berada pada kategori sangat kecil. Perhitungan PSS menunjukkan 4 ZOM memiliki keunggulan.
2. Hasil verifikasi model LSTM menunjukkan nilai perhitungan galat absolut (RMSE dan MAPE) yang lebih besar dibandingkan pada model SARIMA, namun nilai bias pada arah *overestimated* dan *underestimated* lebih kecil. Nilai korelasi Pearson berada pada kategori lemah. Hasil model dapat mengikuti pola data uji namun belum dapat menginterpretasikan nilai yang detail. Nilai PC berada pada kategori sangat kecil. Perhitungan PSS menunjukkan 8 ZOM memiliki keunggulan.

UCAPAN TERIMA KASIH

Kepada seluruh pihak yang telah berkontribusi dalam mendukung proses penelitian dari artikel ini baik secara langsung ataupun tidak, penulis mengucapkan banyak terima kasih.

DAFTAR PUSTAKA

- Armstrong, J. S. (2001). *Evaluating Forecasting Methods*. 443–472.
https://doi.org/10.1007/978-0-306-47630-3_20
- Brooks, H. E., & Doswell, C. A. (1996). A comparison of measures-oriented and distributions-oriented approaches to forecast verification. In *Weather and Forecasting* (Vol. 11, Issue 3). [https://doi.org/10.1175/1520-0434\(1996\)011<0288:ACOMOA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1996)011<0288:ACOMOA>2.0.CO;2)
- CAWCR. (2009). *Forecast Verification - Issues, Methods and FAQ*.
- Gandin, L. S., & Murphy, A. H. (1992). Equitable skill scores for categorical forecasts. *Monthly Weather Review*, 120(2). [https://doi.org/10.1175/1520-0493\(1992\)120<0361:ESSFCF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1992)120<0361:ESSFCF>2.0.CO;2)
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short Term Memory. *Neural Computation*, 9(8).
- Kafara, Z., Rumlawang, F. Y., & Sinay, L. J. (2017). PERAMALAN CURAH HUJAN DENGAN PENDEKATAN SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA). *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, 11(1).
<https://doi.org/10.30598/barekengvol11iss1pp63-74>
- Kustitunto, B., & Rudy, B. (1994). Statistika I. In *Gunadarma*.
[https://homel.vsb.cz/~bri10/Teaching/Statistika 1/uvod.PDF](https://homel.vsb.cz/~bri10/Teaching/Statistika%201/uvod.PDF)
- Muharsyah, R. (2018). Kajian Verifikasi Produk Prakiraan Curah Hujan Bulanan (2003-2012). *Jurnal Meteorologi Dan Geofisika*, 18(1), 33–44.
<https://doi.org/10.31172/jmg.v18i1.265>
- Murphy, A. H. (1989). *Skill scores and correlation coefficients in model verification*.
- Murphy, A. H., & Winkler, R. L. (1987). A general framework for forecast verification. *Monthly Weather Review*, 115(7).
[https://doi.org/10.1175/1520-0493\(1987\)115<1330:AGFFV>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<1330:AGFFV>2.0.CO;2)
- Nuryadi. (2006). *Model ARIMA Sebagai Alat Prediksi Curah Hujan*. Diklat Teknis Pengolahan dan Analisa Data Klimatologi dan Kualitas Udara. Badan Meteorologi dan Geofisika.
- Rumus Statistik. (2021). *Cara Menghitung MAPE (Mean Absolute Percentage Error) di Excel dan R*. <https://www.rumusstatistik.com/2021/05/cara-menghitung-mape-mean-absolute.html>
- Sinay, L. J., Lembang, F. K., Aulele, S. N., & Mustamu, D. (2020). Analisis Curah Hujan Bulanan Di Kota Ambon Menggunakan Model Heteroskedastisitas: Sarima-Garch. *Media Statistika*, 13(1), 68–79.
<https://doi.org/10.14710/medstat.13.1.68-79>
- Soetanto dan Ulfah, A. M. (2006). *Metode ARIMA, Modul Pelatihan*

Peningkatan Akurasi Prakiraan Musim. Badan Meteorologi dan Geofisika.

Walther, B. A., & Moore, J. L. (2005). The concepts of bias, precision and accuracy, and their use in testing the performance of species richness estimators, with a literature review of estimator performance. In *Ecography* (Vol. 28, Issue 6). <https://doi.org/10.1111/j.2005.0906-7590.04112.x>