

Peramalan Curah Hujan Harian Kabupaten Jember Dengan Jaringan Saraf Tiruan Dan *General Circulation Model*

Abduh Riski⁽¹⁾, Ahmad Kamsyakawuni⁽²⁾, Cahya Ramadhani Azhar⁽³⁾

^{1,2,3} Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Jember

Jalan Kalimantan No. 37, Jember, Jawa Timur, Indonesia

e-mail: riski.fmipa@unej.ac.id⁽¹⁾

ABSTRAK

Curah hujan memiliki peran penting di beberapa bidang seperti pertanian dan pengairan. Oleh sebab itu diperlukan model peramalan untuk mengetahui curah hujan di masa yang akan datang. Model peramalan dapat dibentuk menggunakan jaringan saraf tiruan (JST) *backpropagation*. Hasil akurasi peramalan JST diukur dengan MAE, korelasi dan RMSE. Data lokal sebagai data target model merupakan data rata-rata curah hujan harian dari 73 stasiun di wilayah kabupaten Jember mulai dari Oktober 2019 hingga Desember 2020. Data global sebagai data input model menggunakan data *Global Circulation Model* (GCM) model CSIRO-MK3-6-0 dengan eksperimen RCP 2.6. Data GCM direduksi menggunakan *principal component analysis* (PCA) guna menghindari multikolinieritas pada data. Penelitian ini mengkombinasikan jumlah neuron sebesar 10 hingga 100 neuron dan dua fungsi aktivasi pada model JST. Berdasarkan hasil penelitian, model terbaik yang digunakan untuk peramalan adalah model JST dengan 100 neuron dan fungsi aktivasi biner dengan MAE sebesar 6,1205, korelasi sebesar -0,0125, dan RMSE sebesar 9,0251. Hasil peramalan curah hujan harian kabupaten Jember untuk bulan Januari 2021 adalah terjadi curah hujan tertinggi pada hari ke-19 sebesar 10,0471 mm/hari dan curah hujan terendah terdapat pada hari ke-2 sebesar 1,3106 mm/hari.

Kata kunci: curah hujan harian; *general circulation model*; jaringan saraf tiruan; peramalan

ABSTRACT

Rainfall is essential in several fields, such as agriculture and irrigation. Therefore, a forecasting model is needed to determine future rainfall. Forecasting models can be formed using artificial neural networks (ANN) *backpropagation*. MAE, correlation, and RMSE measure the results of ANN forecasting accuracy. Local data as the model target data is the average daily rainfall data from 73 stations in the Jember regency from October 2019 to December 2020. Global data as the input data model uses the *Global Circulation Model* (GCM) data model CSIRO-MK3-6-0 with the RCP 2.6 experiment. GCM data is reduced using *principal component analysis* (PCA) to avoid multicollinearity. This study combines the number of neurons of 10 to 100 neurons and two activation functions in the ANN model. Based on the research results, the best model used for forecasting is the ANN model with 100 neurons and a binary activation function with an MAE of 6.1205, a correlation of -0.0125, and an RMSE of 9.0251. The results of forecasting the daily rainfall for the Jember regency for January 2021 are that the highest rainfall occurred on the 19th day at 10.0471 mm/day, and the lowest rainfall occurred on the second day at 1.3106 mm/day.

Keywords: daily rainfall; *general circulation models*; artificial neural networks; forecasting

PENDAHULUAN

Indonesia merupakan negara beriklim tropis dan hanya memiliki dua musim yang didasarkan atas banyaknya curah hujan, kedua musim tersebut adalah musim hujan dan musim kemarau. Salah satu bidang yang bergantung pada kondisi hujan adalah bidang pertanian, karena tinggi rendahnya curah hujan dapat mempengaruhi hasil panen. Kabupaten Jember merupakan daerah yang mayoritas penduduknya bergerak di bidang pertanian. Untuk meminimalisir masalah tersebut, perlu adanya model peramalan yang dapat digunakan untuk meramalkan intensitas curah hujan secara lokal di Kabupaten Jember. Salah satu cara yang dapat dilakukan untuk membuat model peramalan adalah melalui pendekatan data iklim global menggunakan data *General Circulation Model* (GCM).

GCM merupakan alat untuk memperkirakan perubahan iklim global di masa depan yang diukur berdasarkan peningkatan konsentrasi gas rumah kaca di atmosfer. Informasi dari GCM disajikan dalam bentuk *grid*, dimana tiap *grid* memiliki nilai dan persamaan dari parameter yang berbeda tergantung pada resolusi dari penyedia luaran GCM [1]. Model GCM menggunakan luaran *grid-grid* untuk menduga perubahan unsur-unsur cuaca. Data GCM memiliki banyak dimensi yang kemungkinan besar terjadi korelasi antar *grid* dalam domain dan multikolinearitas antar variabel. Semakin besar domain dan semakin banyak variabel yang digunakan, hal ini juga berpengaruh pada kompleksitas model. Sehingga perlu adanya pereduksian variabel tanpa mengurangi variansi data tersebut. *Principal component analysis* (PCA) merupakan metode yang dapat digunakan untuk mereduksi variabel-variabel yang saling berkorelasi. Variabel yang dihasilkan melalui PCA menghasilkan variabel-variabel yang tidak saling berkorelasi. Proporsi keragaman yang dianggap cukup mewakili total keragaman data jika keragaman kumulatif mencapai 70% s/d 80%. Penelitian oleh [2] menghasilkan tingkat akurasi lebih baik menggunakan metode PCA dibanding dengan metode HMM dalam mengenali pola wajah.

Salah satu metode *machine learning* yang dapat digunakan untuk membentuk model peramalan curah hujan harian adalah jaringan saraf tiruan (JST) [3]. Metode pembelajaran JST berbentuk sistem yang memproses informasi dan mempunyai cara kerja seperti jaringan saraf biologi. *Backpropagation* merupakan algoritma yang biasa digunakan dalam pembentukan model JST. Algoritma *backpropagation* ini mengurangi kesalahan atau error dengan menyesuaikan hasil pelatihan berdasarkan target yang ingin dicapai. Algoritma ini memiliki 3 tahap pelatihan, propagasi maju, propagasi mundur, dan modifikasi bobot.

Beberapa penelitian tentang peramalan curah hujan menggunakan JST seperti peramalan curah hujan bulanan kota Palembang menggunakan JST *Backpropagation* menghasilkan akurasi peramalan yang baik dengan tingkat error sebesar 0,2528 [4]. JST juga telah digunakan untuk meramalkan curah hujan bulanan daerah Jember dan menghasilkan RMSE 83,487 [5]. Berdasarkan hal tersebut, pada artikel ini dilakukan peramalan curah hujan harian Kabupaten Jember dengan reduksi variabel GCM menggunakan PCA sebagai metode pra-pemrosesan data dan metode pelatihan menggunakan JST *Backpropagation*. Batasan penelitian ini menggunakan *grid* GCM yang berada pada *latitude*: -19,5852° LS hingga 0,9337° LU dan *longitude*: 103,125° BT hingga 123,750° BB.

METODE PENELITIAN

Terdapat dua jenis data yang digunakan pada artikel ini yaitu data input dan data target. Data input adalah data GCM dari CMIP5 dengan model CSIRO-MK3-6-0, eksperimen RCP 2.6, dan resolusi GCM sebesar $1,875^{\circ} \times 1,875^{\circ}$ [6]. Data target merupakan data aktual dari 78 stasiun yang tersebar di wilayah kabupaten Jember mulai dari Oktober 2019 hingga Desember 2020. Seluruh data dibagi menjadi dua bagian yaitu data pelatihan sebanyak 366 data (80%) dan data pengujian sebanyak 92 data (20%).

Banyaknya domain GCM dapat memungkinkan terjadinya multikolinieritas antar variabel input. PCA dapat mengatasi masalah multikolinieritas. Reduksi variabel grid GCM oleh PCA akan menghasilkan variabel baru yang disebut *principal component* (PC). Variabel baru yang dibuat tetap mempertimbangkan varian dari data GCM. PC pertama merupakan kombinasi linier dengan maksimum variansi [7]. Variabel PC ini digunakan sebagai input ke dalam algoritma *backpropagation*.

Selanjutnya dibangun model arsitektur jaringan dengan mengkombinasikan neuron dan fungsi aktivasi ke dalam dua *hidden layer*. Fungsi aktivasi yang digunakan adalah fungsi aktivasi biner dengan rentang nilai $[0,1]$ dan fungsi aktivasi bipolar dengan rentang nilai $[-1,1]$. Pembelajaran berhenti apabila epoch telah mencapai 1000 atau target minimal error mencapai 10^{-6} . Model terbaik adalah model yang memiliki MAE dan RMSE yang terkecil dan korelasi terbesar. MAE dan RMSE yang kecil menunjukkan kemiripan antara data target dan hasil dari pembelajaran. Model JST terbaik kemudian dipilih untuk digunakan sebagai model peramalan.

HASIL DAN PEMBAHASAN

A. Konstruksi Data

Titik wilayah Kabupaten Jember diambil di pusat kota Jember pada *longitude* $113,7047581^{\circ}$ dan *latitude* $-8,17704765^{\circ}$. Kabupaten Jember terletak di antara grid baris ke-6 kolom ke-6 (*grid_{6,6}*), dengan *longitude* $112,500^{\circ}$ BT s/d $114,375^{\circ}$ BB dan *latitude* $-8,3937^{\circ}$ LS s/d $-10,2589^{\circ}$ LU. Penentuan ukuran domain *grid* yang digunakan berukuran $n \times n$ yaitu ukuran domain *grid* 1×1 , 3×3 , 5×5 hingga 11×11 . Ukuran-ukuran domain *grid* tersebut menentukan banyaknya variabel input. Gambaran wilayah dalam bentuk grid dapat dilihat pada Gambar 1.



Gambar 1. Domain Grid GCM

B. Pra-Pemrosesan Data

Domain *grid* yang digunakan sebagai variabel input berukuran 1×1 , 3×3 , 5×5 hingga 11×11 . Banyak variabel input beragam sesuai dengan *grid* GCM yang digunakan. Berdasarkan Tabel 1 persentase varian kumulatif ukuran domain *grid* 3×3 adalah sebesar $\geq 98\%$, dengan PC yang tereduksi menjadi 6 variabel. Peneliti menetapkan PC minimum yang memenuhi persentase varian kumulatif $\geq 98\%$ sebagai PC terpilih. Persentase varian kumulatif setiap ukuran domain *grid* dapat dilihat pada Tabel 1. Jumlah PC masing-masing domain *grid* pada Tabel 1 merupakan jumlah variabel input yang akan digunakan pada proses pelatihan. Pada ukuran domain *grid* 3×3 dengan variabel input sebanyak 6 PC mampu menjelaskan sebesar 98,55% varian dari data GCM.

Tabel 1. Persentase Varian Kumulatif PC Setiap Ukuran Domain *Grid*

Ukuran Domain <i>Grid</i>	Jumlah PC	Persentase Varian Kumulatif (%)
1×1	1	100
3×3	6	98,55
5×5	15	98,31
7×7	27	98,00
9×9	44	98,05
11×11	65	98,03

C. Pelatihan Model JST

Hasil model pelatihan dengan fungsi aktivasi biner dengan jumlah iterasi 1000 dan jumlah neuron pada lapisan tersembunyi: 10, 20, 30, sampai dengan 100 memberikan nilai yang beragam. Nilai RMSE, MAE dan koefisien korelasi (COR) pelatihan dan pengujian menggunakan fungsi aktivasi biner dan bipolar dapat dilihat pada Tabel 2.

Tabel 2. Hasil Pelatihan dan Pengujian dengan Fungsi Aktivasi Biner dan Bipolar

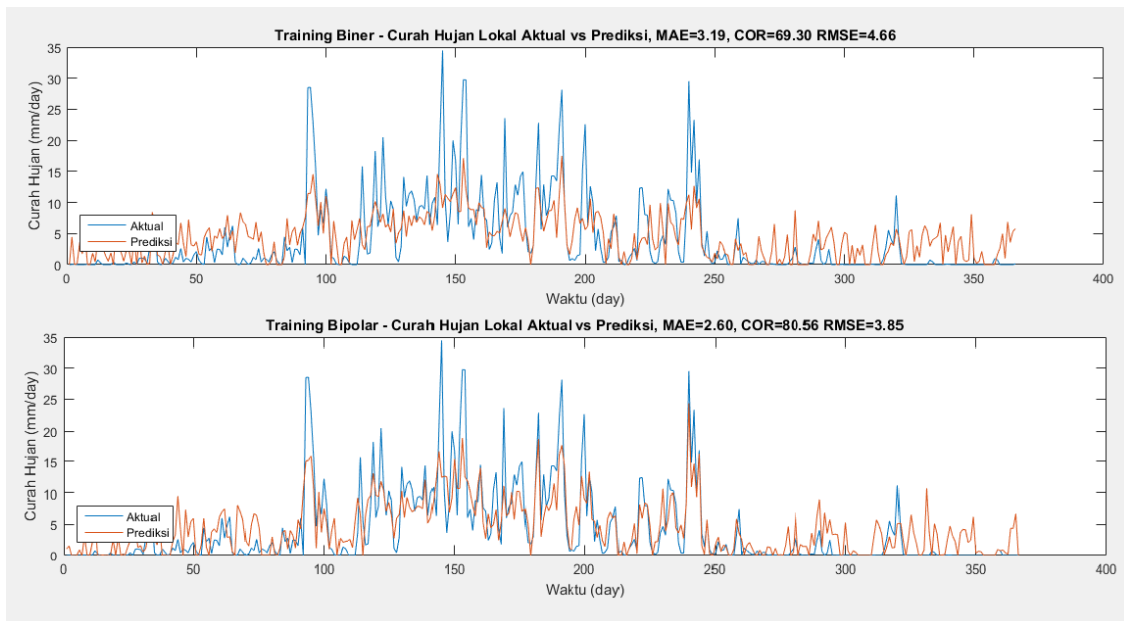
Ukuran Domain	Neuron	Fungsi Aktivasi	Pelatihan			Pengujian		
			MAE	COR	RMSE	MAE	COR	RMSE
1×1	90	Biner	4,0639	0,3775	5,8693	6,1376	0,0509	8,8596
3×3	90	Biner	3,9840	0,3844	5,8524	6,1711	-0,0778	8,9706
5×5	40	Biner	3,7824	0,4929	5,5446	5,9585	0,0335	9,0930
7×7	70	Biner	3,6154	0,5423	5,3605	5,9891	0,0623	8,9220
9×9	90	Biner	3,5122	0,5930	5,1593	6,0246	0,1123	8,8052
11×11	100	Biner	3,1900	0,6930	4,6633	6,1205	-0,0125	9,0251
1×1	90	Bipolar	3,9233	0,4006	5,8070	6,1837	0,0287	8,9209
3×3	80	Bipolar	3,7240	0,4647	5,6225	6,3585	-0,0902	9,4232
5×5	50	Bipolar	3,9445	0,4174	5,7765	5,7826	0,1343	8,7424
7×7	50	Bipolar	3,2685	0,6227	5,0157	5,9219	0,1730	8,6167
9×9	90	Bipolar	2,9285	0,7016	4,5898	6,3321	0,1116	9,0666
11×11	100	Bipolar	2,5984	0,8056	3,8519	6,6568	-0,0110	9,3085

Dari Tabel 2 nilai MAE, koefisien korelasi (COR), dan RMSE tiap domain *grid* berbeda-beda. Dari hasil pada Tabel 2 neuron yang lebih banyak cenderung menghasilkan MAE yang

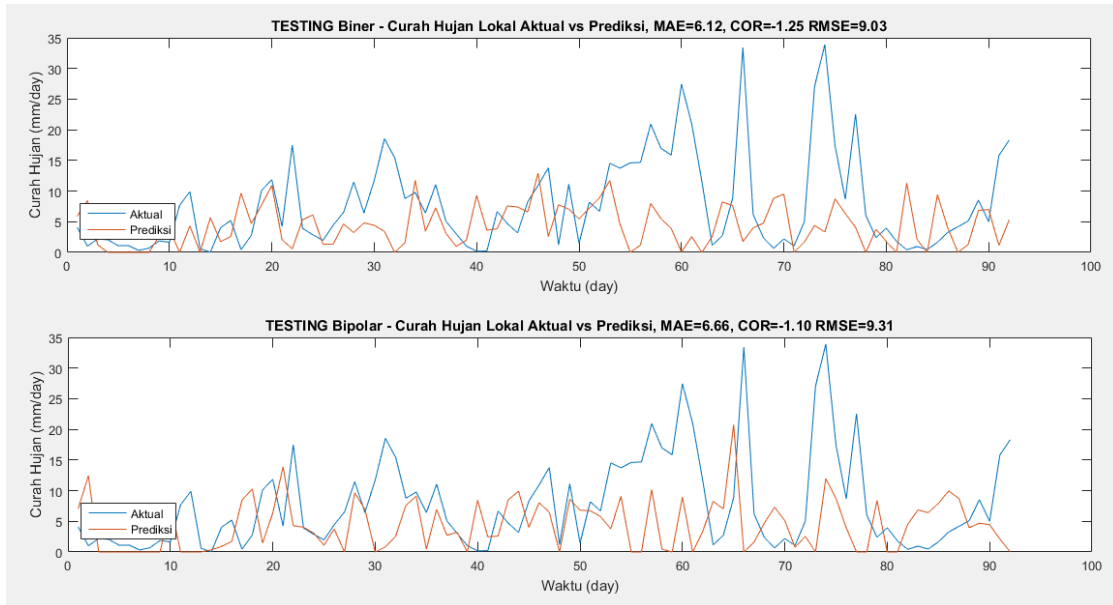
rendah pada pelatihan dan pengujian. Nilai terendah MAE dan RMSE pada domain grid 1×1 dengan 100 neuron untuk pelatihan sebesar 4,06 dan 5,85, akan tetapi dibandingkan 90 neuron, MAE dan RMSE yang dihasilkan menjadi lebih tinggi yaitu sebesar 6,1685 dan 8,9949. Sehingga model terbaik untuk domain *grid* 1×1 menggunakan model dengan neuron 90.

D. Analisis dan Pemilihan Model Terbaik

Pemilihan model JST pada fungsi aktivasi biner mempertimbangan nilai MAE dan RMSE pada pelatihan dan pengujian. Berdasarkan Tabel 2, nilai MAE dan RMSE pelatihan pada domain *grid* 11×11 menghasilkan nilai MAE dan RMSE yang terkecil. Pada pengujiannya diperoleh MAE dan RMSE dari fungsi aktivasi biner lebih kecil dari fungsi aktivasi bipolar. Nilai MAE dan RMSE fungsi aktivasi biner 6,1205 dan 9,0251 sementara pada fungsi aktivasi bipolar sebesar 6,6568 dan 9,3085. Grafik pelatihan dan pengujian model JST *grid* 11×11 dengan fungsi aktivasi biner dan bipolar dapat dilihat pada Gambar 2 dan Gambar 3. Pada kedua gambar tersebut terlihat bahwa hasil pelatihan curah hujan harian Kabupaten Jember menggunakan fungsi aktivasi biner maupun fungsi aktivasi bipolar memiliki pola curah hujan yang hampir sama. Sehingga model JST dengan *grid* 11×11 menjadi ukuran domain yang terpilih (model terbaik) untuk digunakan sebagai model peramalan, dengan fungsi aktivasi biner dan 100 neuron.



Gambar 2. Grafik Pelatihan Model dengan Fungsi Aktivasi Biner dan Bipolar



Gambar 3. Grafik Pengujian Model dengan Fungsi Aktivasi Biner dan Bipolar

E. Peramalan

Hasil pemodelan JST terbaik kemudian digunakan untuk meramalkan curah hujan harian di Kabupaten Jember selama 1 bulan penuh yaitu 1 Januari 2021 sampai dengan 31 Januari 2021. Parameter yang digunakan adalah 2 *hidden layer* dengan fungsi aktivasi biner, 1 *output layer* dengan fungsi aktivasi linier, 100 neuron pada hidden layer. Domain *grid* yang terbaik adalah *grid* dengan ukuran 11 × 11. Sebanyak 65 variabel yang mewakili 121 variabel dengan persentase varian kumulatif sebesar 98% sebagai variabel input. Hasil peramalan curah hujan harian Kabupaten Jember Januari 2021 dapat dilihat pada Tabel 3.

Tabel 3. Hasil Peramalan Curah Hujan Harian Kabupaten Jember Januari 2021

Hari	mm/hari	Hari	mm/hari	Hari	mm/hari
1	5.8603	11	8.5583	21	5.9703
2	1.3106	12	8.6513	22	8.9885
3	8.8921	13	6.9761	23	8.3052
4	2.6628	14	6.0941	24	6.6885
5	3.5648	15	5.9329	25	8.442
6	7.3438	16	3.7436	26	8.5982
7	2.5936	17	2.1324	27	7.9574
8	5.4087	18	6.2416	28	7.1832
9	6.8419	19	10.0471	29	6.0312
10	4.7952	20	9.0329	30	7.4895
				31	7.4289

SIMPULAN DAN SARAN

Model JST dengan fungsi aktivasi biner menggunakan dua *hidden layer* dan 100 neuron menjadi model terbaik untuk melakukan peramalan curah hujan harian Kabupaten Jember. Pemodelan tersebut diperoleh dengan menggunakan domain *grid* GCM 11×11 sebagai variabel input dan menghasilkan nilai MAE 6,1205, Koefisien Korelasi -0,0125, dan RMSE 9,0251.

UCAPAN TERIMAKASIH

Penelitian ini didanai oleh Lembaga Penelitian dan Pengabdian kepada Masyarakat (LP2M) Universitas Jember melalui Hibah Kelompok Riset dan Pengabdian Masyarakat Tahun 2023 No. 3287/UN25.3.1/LT/2023.

DAFTAR PUSTAKA

- [1] A. H. Wigena, "Pemodelan statistical downscaling dengan regresi projection pursuit untuk peramalan curah hujan bulanan," *Disertasi. IPB*, 2006.
- [2] A. R. Syakhala, D. Puspitaningrum, and E. P. Purwandari, "Perbandingan Metode Principal Component Analysis (PCA) Dengan Metode Hidden Markov Model (HMM) Dalam Pengenalan Identitas Seseorang Melalui Wajah," *Rekursif: Jurnal Informatika*, vol. 3, no. 2, Mar. 2016, doi: 10.33369/rekursif.v3i2.743.
- [3] L. V. Fausett, *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications*. London: Prentice-Hall, 1994.
- [4] I. Sofian and Y. Apriani, "Metode Peramalan Jaringan Saraf Tiruan Menggunakan Algoritma Backpropagation (Studi Kasus Peramalan Curah Hujan Kota Palembang)," May 2017.
- [5] A. Riski, A. F. Hadi, O. Tazkiyah, and D. Anggraeni, "Neural Network and Principal Component Analysis on Statistical Downscaling for Local Rainfall Forecasting," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 12, pp. 812–826, May 2020.
- [6] K. E. Taylor, R. J. Stouffer, and G. A. Meehl, "An Overview of CMIP5 and the Experiment Design," *Bull Am Meteorol Soc*, vol. 93, no. 4, pp. 485–498, Apr. 2012, doi: 10.1175/BAMS-D-11-00094.1.
- [7] R. A. Johnson and D. W. Wichern, *Applied Multivariate Statistical Analysis*, 6th ed. Upper Saddle River: Pearson Prentice Hall, 2007.