

PEMODELAN *DEEP LEARNING* UNTUK PREDIKSI CURAH HUJAN SKALA *SUBSEASONAL TO SEASONAL*

FERRY LODEWIK LUMALESSIL



**PROGRAM MAGISTER ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**

@Hak cipta milik IPB University

IPB University



IPB University
— Bogor Indonesia —

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University



PERNYATAAN MENGENAI TESIS DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa tesis dengan judul “Pemodelan *Deep Learning* untuk Prediksi Curah Hujan Skala *Subseasonal to Seasonal*” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir tesis ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2026

Ferry Lodewik Lumalessil
G6501231029

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

RINGKASAN

FERRY LODEWIK LUMALESSIL. Pemodelan *Deep Learning* untuk Prediksi Curah Hujan Skala *Subseasonal to Seasonal*. Dibimbing oleh AZIZ KUSTIYO, AGUS BUONO dan AKHMAD FAQIH.

Indonesia memiliki potensi pertanian yang besar, tetapi keberhasilan produksi sangat dipengaruhi oleh variabilitas curah hujan karena kondisi terlalu basah maupun terlalu kering dapat memicu banjir, kekeringan, gangguan tanaman, perubahan kalender tanam, dan penurunan hasil produksi. Oleh karena itu, informasi prediksi cuaca yang akurat diperlukan untuk mendukung pengambilan keputusan dan langkah antisipasi di sektor pertanian. Salah satu skala prediksi yang berkembang adalah *Subseasonal to Seasonal* (S2S), yaitu rentang prediksi sekitar dua minggu hingga dua bulan yang menjembatani prediksi cuaca dan iklim serta mendukung kebutuhan operasional dan strategis, seperti peringatan dini bencana, pengelolaan irigasi, dan perencanaan pertanian.

Prediksi S2S masih menjadi tantangan karena jarak waktu prediksi yang panjang, tetapi keterampilannya dapat dipengaruhi oleh variabilitas atmosfer, kelembaban tanah, tutupan salju, dan kondisi laut. Pengembangan prediksi S2S dapat memanfaatkan *output Global Climate Model* (GCM), yang merepresentasikan sistem iklim dalam bentuk *grid* tiga dimensi dan menyediakan prediksi *ensemble* sekitar 30 hingga 60 hari dari berbagai pusat penyedia data, sehingga dapat digunakan sebagai prediktor dalam pembangunan model prediksi curah hujan skala S2S.

Penelitian terdahulu menunjukkan bahwa metode statistik dan *deep learning* mampu meningkatkan keterampilan prediksi curah hujan pada skala S2S. Namun, sebagian penelitian masih berfokus pada evaluasi *output* GCM, koreksi bias, *downscaling*, atau penggunaan informasi spasial dan temporal secara terpisah. Kondisi tersebut menjadi dasar untuk mempertimbangkan pendekatan yang menggabungkan pola spasial variabel atmosfer ECMWF dengan informasi temporal curah hujan. Penelitian ini bertujuan membangun model prediksi curah hujan skala S2S berbasis *deep learning* menggunakan arsitektur *hybrid* CNN–LSTM dan CNN–GRU dengan memanfaatkan *output* GCM ECMWF sebagai prediktor dan data curah hujan beresolusi tinggi sebagai prediktan pada wilayah Pulau Jawa.

Tahapan penelitian meliputi praproses data, pengelompokan spasial dengan *K-Means*, penyesuaian *lead time*, pembagian data pelatihan, validasi, dan pengujian, serta optimasi *hyperparameter* menggunakan *grid search*. CNN digunakan untuk mengekstraksi fitur spasial dari tujuh variabel atmosfer ECMWF, sedangkan LSTM dan GRU digunakan untuk memodelkan dinamika temporal curah hujan lokal. Penelitian ini membangun model prediksi curah hujan skala S2S berbasis *deep learning* melalui pendekatan hibrida CNN–LSTM dan CNN–GRU, dengan data prediktor berupa tujuh variabel atmosfer ECMWF S2S, yaitu skt, t200, t850, u200, u850, v200, dan v850, serta data prediktan berupa curah hujan CHIRPS.

Model yang dibuat mampu menghasilkan prediksi mengikuti struktur operasional rilis ECMWF, yaitu dari *lead time* 0 hingga *lead time* 45, sehingga setiap rilis dapat menyediakan prakiraan curah hujan hingga 46 hari ke depan. Hasil evaluasi menunjukkan bahwa performa model bervariasi secara spasial di Pulau

Jawa, dengan kecenderungan hasil yang lebih baik pada beberapa titik di wilayah Jawa bagian tengah hingga timur. Pada model CNN–LSTM, performa terbaik cenderung ditemukan di wilayah Jawa bagian tengah, yang ditunjukkan oleh RMSE relatif rendah sebesar 3,286, korelasi tinggi sebesar 0,944, BSS sebesar 0,414, dan CRPSS tertinggi sebesar 0,775. Secara spasial, wilayah Jawa bagian tengah hingga timur umumnya menunjukkan RMSE lebih rendah, korelasi tinggi, serta BSS dan CRPSS yang relatif baik, sedangkan wilayah Jawa bagian barat, terutama bagian barat laut hingga ujung barat, masih menunjukkan RMSE lebih tinggi, korelasi lebih rendah, dan BSS yang sebagian bernilai negatif. Pada model CNN–GRU, performa yang relatif baik juga terkonsentrasi di wilayah Jawa bagian tengah, khususnya bagian tengah hingga selatan, yang ditunjukkan oleh RMSE sebesar 2,939, korelasi sebesar 0,954, BSS sebesar 0,406, dan CRPSS sebesar 0,631. Sebaliknya, kinerja yang lebih rendah terlihat pada sebagian wilayah Jawa bagian barat serta bagian timur hingga tenggara, yang ditandai oleh RMSE lebih tinggi dan BSS rendah atau negatif. Evaluasi per klaster juga menunjukkan bahwa CNN–LSTM memiliki CRPSS lebih tinggi pada sebagian besar klaster, sedangkan CNN–GRU menunjukkan variasi RMSE dan BSS yang lebih besar antarwilayah. Hal ini diperkuat oleh distribusi metrik pada seluruh titik, dengan CNN–LSTM memiliki median RMSE yang lebih rendah, korelasi dan BSS yang cenderung lebih tinggi, serta CRPSS yang secara jelas lebih tinggi dibandingkan CNN–GRU. Secara umum, CNN–LSTM menunjukkan performa yang lebih stabil dibandingkan CNN–GRU, dengan wilayah Jawa bagian tengah menjadi area yang cenderung memberikan hasil prediksi terbaik pada kedua model.

Kata kunci: *convolution neural network, curah hujan, gated recurrent unit, long short-term memory, subseasonal to seasonal*

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

SUMMARY

FERRY LODEWIK LUMALESSIL. Deep Learning Modeling for Subseasonal to Seasonal Rainfall Forecasting. Supervised by AZIZ KUSTIYO, AGUS BUONO and AKHMAD FAQIH.

Indonesia has great agricultural potential, but the success of production is greatly influenced by the variability of rainfall because conditions that are too wet or too dry can trigger floods, droughts, crop disturbances, changes in the planting calendar, and decreased production yields. Therefore, accurate weather prediction information is needed to support decision-making and anticipatory measures in the agricultural sector. One of the forecast scales that is developing is Subseasonal to Seasonal (S2S), which is a prediction range of about two weeks to two months that bridges weather and climate predictions and supports operational and strategic needs, such as disaster early warning, irrigation management, and agricultural planning.

S2S prediction is still a challenge due to the long prediction time range, but its skills can be affected by atmospheric variability, soil moisture, snow cover, and ocean conditions. The development of S2S predictions can utilize the output of the Global Climate Model (GCM), which represents the climate system in the form of a three-dimensional grid and provides ensemble predictions of about 30 to 60 days from various data provider centers, so that it can be used as a predictor in the construction of S2S-scale rainfall prediction models.

Previous research has shown that statistical methods and deep learning are able to improve rainfall prediction skills on the S2S scale. However, some studies still focus on evaluating GCM outputs, bias correction, downscaling, or using spatial and temporal information separately. These conditions are the basis for considering an approach that combines the spatial patterns of the ECMWF atmospheric variables with temporal precipitation information. This study aims to build a deep learning-based S2S scale rainfall prediction model using the hybrid architecture of CNN-LSTM and CNN-GRU by utilizing the output of GCM ECMWF as a predictor and high-resolution rainfall data as predictors in the Java Island region.

The research stages include data preprocessing, spatial grouping with K-Means, lead time adjustment, training data sharing, validation, and testing, and hyperparameter optimization using grid search. CNN was used to extract spatial features from seven atmospheric variables of the ECMWF, while LSTM and GRU were used to model the temporal dynamics of local precipitation. This study builds a deep learning-based S2S scale rainfall prediction model through a hybrid approach of CNN-LSTM and CNN-GRU, with predictor data in the form of seven ECMWF S2S atmospheric variables, namely skt, t200, t850, u200, u850, v200, and v850, as well as predictive data in the form of CHIRPS rainfall.

The model is able to produce predictions following the operational structure of the ECMWF release, which is from lead time 0 to lead time 45, so that each release can provide rainfall forecasts for up to the next 46 days. The results of the evaluation showed that the performance of the model varied spatially on the island of Java, with a tendency for better results at several points in the central to eastern



Java region. In the CNN-LSTM model, the best performance tends to be found in the central Java region, which is shown by the relatively low RMSE of 3.286, high correlation of 0.944, BSS of 0.414, and highest CRPSS of 0.775. Spatially, the central to eastern Java region generally shows lower RMSE, high correlation, as well as relatively good BSS and CRPSS, while the western Java region, especially the northwest to the western part, still shows higher RMSE, lower correlation, and some negative BSS. In the CNN-GRU model, relatively good performance was also concentrated in the central Java region, especially the central to southern part, which was shown by RMSE of 2.939, correlation of 0.954, BSS of 0.406, and CRPSS of 0.631. On the other hand, lower performance was seen in parts of western Java as well as the eastern to southeastern part, which was characterized by higher RMSE and low or negative BSS. Evaluation per cluster also showed that CNN-LSTM had higher CRPSS in most clusters, while CNN-GRU showed greater variation in RMSE and BSS between regions. This is reinforced by the distribution of metrics across points, with CNN-LSTM having a lower median RMSE, correlation and BSS tending to be higher, and a clearly higher CRPSS than CNN-GRU. In general, CNN-LSTM showed a more stable performance than CNN-GRU, with the central Java region being the area that tends to provide the best prediction results in both models.

Kata kunci: *convolution neural network, gated recurrent unit, long short-term memory, rainfall, subseasonal to seasonal*



@Hak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2026
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

PEMODELAN *DEEP LEARNING* UNTUK PREDIKSI CURAH HUJAN SKALA *SUBSEASONAL TO SEASONAL*

FERRY LODEWIK LUMALESSIL

Tesis
sebagai salah satu syarat untuk memperoleh gelar
Magister pada
Program Studi Ilmu Komputer

**PROGRAM MAGISTER ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2026**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Tesis:
Hari Agung Adrianto, S.Kom., M.Si., Ph.D.



Judul Tesis : *Pemodelan Deep Learning untuk Prediksi Curah Hujan Skala Subseasonal to Seasonal.*

Nama : Ferry Lodewik Lumalessil

NIM : G6501231029

@Hak cipta milik IPB University

Disetujui oleh

Pembimbing 1:

Dr. Aziz Kustiyo, S.Si, M.Kom.



Pembimbing 2:

Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.



Pembimbing 3:

Dr. Akhmad Faqih, S.Si.



Diketahui oleh

Ketua Program Studi:

Dr. Toto Haryanto, S.Kom., M.Kom.

NIP 19821117201404 1 001



Dekan Sekolah Sains Data, Matematika, dan Informatika:

Prof. Dr. Ir. Agus Buono, M.Si., M.Kom.

NIP 19660702199302 1 001



Tanggal Ujian:
3 Agustus 2026

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas segala karunia-Nya sehingga penelitian yang berjudul “Pemodelan *Deep Learning* untuk Prediksi Curah Hujan Skala *Subseasonal to Seasonal*” dapat diselesaikan yang berlangsung sejak bulan Januari 2025 sampai bulan Desember 2025.

Terima kasih penulis ucapkan kepada para pembimbing, Dr. Aziz Kustiyo, S.Si., M.Kom., Prof. Dr. Ir. Agus Buono, M.Si., M.Kom. dan Dr. Akhmad Faqih, S.Si. yang telah membimbing dan banyak memberi saran dan arahan dalam penulisan tesis ini, serta kepada moderator seminar Dr. Toto Haryanto, S.Kom., M.Kom. dan Hari Agung, S.Kom., M.Si., Ph.D. selaku penguji luar komisi pembimbing yang sudah memberikan masukan untuk penelitian ini, serta staf dosen dan bagian administrasi Program Magister Ilmu Komputer

Ungkapan terima kasih juga penulis sampaikan kepada Mama dan Papa, serta seluruh keluarga besar yang telah memberikan dukungan, doa, dan kasih sayangnya sehingga penulis dapat menyelesaikan studi ini. Selain itu, penghargaan dan terima kasih penulis sampaikan kepada Lembaga Pengelola Dana Pendidikan (LPDP) yang telah mensponsori studi penulis. Penulis juga menyampaikan terima kasih kepada Risma, teman-teman kontrakan ceria, teman-teman ILKOM 2023, keluarga besar Moluccas dan semua teman-teman seperjuangan di Bogor yang tidak sempat penulis sebutkan namanya atas dukungan partisipasinya selama penulis menjalani perkuliahan hingga penyelesaian penelitian ini.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan dan teknologi.

Bogor, Agustus 2026

Ferry Lodewik Lumalessil

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



DAFTAR ISI

DAFTAR TABEL	xi
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xi
PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	4
1.5 Ruang Lingkup	4
TINJAUAN PUSTAKA	
2.1 <i>Global Climate Model</i>	5
2.2 <i>Subseasonal to Seasonal</i>	6
2.3 <i>Convolutional Neural Network</i>	7
2.4 <i>Long Short-Term Memory</i>	9
2.5 <i>Gated Recurrent Unit</i>	11
METODE	
3.1 Area Studi	13
3.2 Tahapan Penelitian	13
3.3 Pengumpulan Data	14
3.4 Praproses Data	14
3.5 Pembagian Data	15
3.6 Pembuatan Model <i>Hybrid</i>	16
3.7 Evaluasi Model	17
3.8 Lingkungan Pengembangan	19
HASIL DAN PEMBAHASAN	
4.1 Pengumpulan Data	20
4.2 Praproses Data	21
4.3 Pembagian Data	23
4.4 Pemodelan	24
4.5 Hasil Pemodelan	25
4.6 Evaluasi Model	31
SIMPULAN DAN SARAN	
5.1 Simpulan	39
5.2 Saran	39
DAFTAR PUSTAKA	40
LAMPIRAN	45
RIWAYAT HIDUP	61

DAFTAR TABEL

1	Karakteristik utama basis data model S2S	7
2	Data <i>output</i> GCM	14
3	Arsitektur CNN, LSTM, dan GRU	17
4	Struktur data GCM S2S ECMWF	21
5	Proporsi pembagian data	24
6	Parameter dalam pemodelan LSTM dan GRU	25
7	Perbandingan metrik evaluasi aktual dan prediksi pada deret <i>lead time</i> pada satu titik	27

DAFTAR GAMBAR

1	Ilustrasi data GCM	5
2	Estimasi ketrampilan prakiraan berdasarkan rentang waktu	6
3	Arsitektur CNN	8
4	Ilustrasi proses konvolusi dan <i>pooling</i>	8
5	Fungsi <i>relu</i>	9
6	Arsitektur LSTM	9
7	Arsitektur GRU	11
8	Peta Pulau Jawa	13
9	Tahapan Penelitian	13
10	Pembagian dataset	16
11	Arsitektur CNN-LSTM dan CNN-GRU	16
12	Data GCM	20
13	Data curah hujan CHIRPS	21
14	Penentuan dan Klasterisasi Curah Hujan	22
15	Perbandingan data curah hujan dan hasil EMA	23
16	Contoh perbandingan aktual dan prediksi pada <i>lead time</i> 0	26
17	Contoh perbandingan aktual dan prediksi pada <i>lead time</i> 0 setelah pergeseran	28
18	Perbandingan prediksi sebelum <i>invers</i> EMA dan sesudah <i>invers</i> EMA	29
19	Contoh prediksi curah hujan 46 <i>lead time</i>	30
20	Skema penyusunan deret harian prediksi	32
21	Distribusi spasial evaluasi Model CNN-LSTM	33
22	Distribusi spasial evaluasi Model CNN-GRU	34
23	Perbandingan rata-rata performa CNN-LSTM dan CNN-GRU berdasarkan klaster	35
24	Perbandingan distribusi metrik Model CNN-LSTM dan CNN-GRU	37

DAFTAR LAMPIRAN

1	Evaluasi Model CNN-LSTM	46
2	Evaluasi Model CNN-GRU	48
3	Grafik Aktual dan Prediksi Setiap Klaster	51