

Prediksi Kecepatan Angin Menggunakan *Gated Recurrent Unit (GRU)* dengan Estimasi Ketidakpastian Monte Carlo Dropout pada Data BMKG Tanjung Perak

Alvino Hadiyan Pradipta¹, Muhammad Rafli Feandika Nugroho², Maretta Fairuz Luthfia Winoto Putri³, Alfian Rizaldy Pratama⁴, Shindi Shella May Wara⁵, Muhammad Nasrudin⁶

Fakultas Ilmu Komputer, Program Studi Sains Data
Universitas Pembangunan Nasional Veteran Jawa Timur
23083010090@student.upnjatim.ac.id

Abstrak

Keterbatasan metode prediksi konvensional dalam memodelkan dependensi temporal dan ketidakpastian prediksi mendorong pengembangan pendekatan berbasis deep learning. Penelitian ini bertujuan mengembangkan model prediksi kecepatan angin menggunakan metode *Gated Recurrent Unit (GRU)* pada data meteorologi yang berasal dari stasiun pengamatan BMKG Tanjung Perak. Penelitian ini dilakukan karena metode prediksi sebelumnya masih memiliki keterbatasan dalam menangkap pola temporal dan dependensi jangka panjang pada data time series, serta umumnya belum mengakomodasi ketidakpastian hasil prediksi. *GRU* dipilih karena mampu memodelkan dependensi temporal secara efisien, sedangkan simulasi Monte Carlo digunakan untuk menghasilkan beberapa skenario prediksi dan mengestimasi interval kepercayaan. Data yang digunakan mencakup parameter kecepatan angin dengan interval waktu tertentu. Hasil evaluasi menunjukkan bahwa model menunjukkan performa yang baik untuk memprediksi kecepatan angin secara akurat, dengan nilai MAE sebesar 0,37, RMSE sebesar 0,50, MAPE sebesar 5,78%, dan R^2 sebesar 0,986. Dengan demikian, model yang dikembangkan dapat menjadi solusi dalam analisis dan peramalan data *time series* meteorologi secara komprehensif.

Kata kunci: Prediksi Kecepatan Angin, Analisis Time Series, *Gated Recurrent Unit (GRU)*, Simulasi Monte Carlo, Data Meteorologi

Abstract

*The limitations of conventional forecasting methods in modeling temporal dependencies and forecast uncertainty have driven the development of deep learning-based approaches. This study aims to develop a wind speed forecasting model using the *Gated Recurrent Unit (GRU)* method on meteorological data from the BMKG Tanjung Perak observation station. This study was conducted because previous prediction methods still have limitations in capturing temporal patterns and long-term dependencies in time series data, and generally do not accommodate the uncertainty of prediction results. *GRU* was chosen because it is capable of modeling temporal dependencies efficiently, while Monte Carlo simulation was used to generate several prediction scenarios and estimate confidence intervals. The data used includes wind speed parameters at specific time intervals. The evaluation results show that the model shows good performance in predicting wind speed accurately, with an MAE of 0.37, an RMSE of 0.50, a MAPE of 5.78%, and an R^2 of 0.986. Thus, the developed model can serve as a solution for comprehensive analysis and forecasting of meteorological time series data.*

Keywords: Wind Speed Prediction, Time Series Analysis, *Gated Recurrent Unit (GRU)*, Monte Carlo Simulation, Meteorological Data

PENDAHULUAN

Peningkatan keselamatan navigasi di lautan merupakan tujuan utama bagi komunitas maritim di seluruh dunia (Minarto & Santoso, 2023). Keselamatan pelayaran menjadi aspek penting dalam aktivitas maritim, terutama di Pelabuhan Tanjung Perak Surabaya yang memiliki lalu lintas kapal padat. Ketidakpastian cuaca maritim yang sulit diprediksi secara akurat kerap menyebabkan insiden pelayaran yang mengancam keselamatan jiwa dan menimbulkan kerugian ekonomi yang signifikan. Dalam konteks ini, keselamatan pelayaran ditentukan oleh kesiapan armada, kompetensi operator, serta kemampuan sistem menyediakan informasi risiko yang cepat, mudah dipahami, dan dapat ditindaklanjuti. Salah satu faktor yang paling konsisten memengaruhi risiko pelayaran adalah kondisi lingkungan laut terutama angin, gelombang, arus, serta fenomena cuaca lain yang dapat memperburuk stabilitas kapal dan meningkatkan probabilitas insiden (Jordi & Narulita, 2026). Oleh karena itu, sistem prediksi dan evaluasi risiko yang tepat dan responsif diperlukan agar pengambilan keputusan dalam kegiatan pelayaran dapat dilakukan dengan benar.

Seiring perkembangan teknologi, pemanfaatan kecerdasan buatan (artificial intelligence), khususnya deep learning, semakin banyak digunakan untuk meningkatkan akurasi prediksi data time series (Islam & Wahabi, 2025). Salah satu metode yang umum digunakan adalah Recurrent Neural Network (RNN), namun model ini memiliki keterbatasan dalam mempelajari ketergantungan jangka panjang akibat masalah vanishing gradient. Untuk mengatasi hal tersebut, dikembangkan arsitektur Long Short-Term

Memory (LSTM) dan Gated Recurrent Unit (GRU), yang mampu menangkap pola temporal secara lebih efektif. GRU memiliki keunggulan berupa struktur yang lebih sederhana dan efisiensi komputasi yang lebih baik dibandingkan LSTM (Arwansyah et al, 2024). Selain itu, ketidakpastian hasil prediksi dapat diestimasi menggunakan Simulasi Monte Carlo sehingga menghasilkan informasi prediksi yang lebih komprehensif dibandingkan prediksi tunggal.

Penelitian terdahulu menunjukkan bahwa model deep learning, khususnya GRU dan LSTM, memiliki kemampuan yang baik dalam memprediksi data deret waktu. Arwansyah et al. (2024) dan Ramadila et al. (2026) menunjukkan bahwa model GRU mampu menghasilkan akurasi prediksi yang tinggi pada berbagai data time series dan maritim. Selain itu, Sadeghi dan Samadi (2022) serta Murase et al. (2025) menunjukkan bahwa pendekatan Monte Carlo Dropout dapat meningkatkan reliabilitas prediksi sekaligus mengestimasi ketidakpastian hasil prediksi. Temuan-temuan tersebut menunjukkan bahwa kombinasi deep learning dan pendekatan probabilistik berpotensi menghasilkan prediksi yang lebih akurat dan informatif.

Penelitian ini mengintegrasikan model Gated Recurrent Unit (GRU) dengan simulasi Monte Carlo untuk prediksi kecepatan angin pada data meteorologi maritim BMKG Tanjung Perak. Berbeda dari pendekatan yang umumnya menghasilkan prediksi titik (*point forecast*), metode yang diusulkan tidak hanya berfokus pada akurasi prediksi, tetapi juga mengestimasi ketidakpastian melalui pembentukan berbagai skenario prediksi dan interval kepercayaan. Berdasarkan research gap yang telah diidentifikasi,

penelitian ini bertujuan mengembangkan model prediksi kecepatan angin menggunakan Gated Recurrent Unit (GRU) yang dikombinasikan dengan simulasi Monte Carlo untuk mengestimasi ketidakpastian prediksi. Model dievaluasi menggunakan metrik MAE, RMSE, MAPE, dan R^2 pada data meteorologi BMKG Tanjung Perak .

Tinjauan Pustaka

Prediksi kecepatan angin merupakan kajian penting dalam meteorologi untuk mendukung keselamatan aktivitas maritim. Perubahan kondisi cuaca, seperti kecepatan angin, gelombang, dan arus laut, dapat memengaruhi stabilitas pelayaran sehingga diperlukan metode prediksi yang akurat. Salah satu metode yang banyak digunakan adalah Gated Recurrent Unit (GRU) karena mampu mempelajari pola temporal pada data time series. Selain itu, Simulasi Monte Carlo dapat digunakan untuk mengestimasi ketidakpastian hasil prediksi melalui berbagai skenario yang mungkin terjadi.

Tabel 1. Penelitian Terdahulu

Peneliti	Metode	Hasil
Arwansyah et al. (2024)	GRU, LSTM, RNN, ARIMA	GRU memberikan performa terbaik dibanding metode lain
Ramadila et al. (2026)	GRU dan LSTM	Kedua model memiliki akurasi yang baik
Sadeghi Tabas & Samadi (2022)	RNN, LSTM, GRU dengan variational Bayesian/Monte Carlo Dropout	Pendekatan Monte Carlo Dropout mampu mengukur ketidakpastian sekaligus meningkatkan performa prediksi pada model RNN, LSTM, dan GRU
Murase et al. (2025)	LSTM dan GRU dengan Monte Carlo Dropout	Monte Carlo Dropout efektif dalam mengkuantifikasi ketidakpastian prediksi dan meningkatkan kualitas hasil prediksi

Berdasarkan Tabel 1, model deep learning, khususnya GRU dan LSTM, telah menunjukkan performa yang baik dalam berbagai penelitian prediksi data time series. Arwansyah et al. (2024) menunjukkan bahwa Deep GRU mampu menghasilkan akurasi tinggi dengan nilai R^2 mendekati 1 dengan error yang sangat rendah. Selain itu, Sadeghi dan Samadi (2022) serta Murase et al. (2025) menunjukkan bahwa pendekatan Monte Carlo Dropout dapat mengestimasi ketidakpastian hasil prediksi sekaligus meningkatkan reliabilitas model. Temuan-temuan tersebut menunjukkan bahwa kombinasi model deep learning dan pendekatan probabilistik berpotensi menghasilkan prediksi yang akurat sekaligus informatif.

Meskipun demikian, berdasarkan penelitian - penelitian terdahulu masih terdapat beberapa keterbatasan. Penelitian Arwansyah et al. (2024) dan Ramadila et al. (2026) berfokus pada peningkatan akurasi prediksi tanpa mempertimbangkan ketidakpastian hasil prediksi yang dapat memengaruhi proses pengambilan keputusan. Sebaliknya, penelitian Sadeghi dan Samadi (2022) serta Murase et al. (2025) telah mengakomodasi estimasi ketidakpastian, namun penerapannya masih terbatas pada bidang hidrologi dan medis. Berdasarkan hasil kajian literatur, studi yang mengombinasikan model Gated Recurrent Unit (GRU) dan Simulasi Monte Carlo dalam prediksi kecepatan angin berbasis data meteorologi maritim masih sangat terbatas untuk ditemukan, sehingga penelitian pada bidang ini masih memiliki peluang untuk dikembangkan lebih lanjut. Oleh karena itu, penelitian ini mengusulkan integrasi GRU dan Simulasi Monte Carlo untuk menghasilkan prediksi kecepatan

angin yang tidak hanya akurat, tetapi juga mampu merepresentasikan tingkat ketidakpastian hasil prediksi melalui berbagai skenario prediksi. Kebaruan penelitian ini terletak pada penerapan pendekatan tersebut pada data meteorologi BMKG Tanjung Perak sehingga dapat memberikan informasi prediksi yang lebih komprehensif dan mendukung keputusan pada sektor maritim.

LANDASAN TEORI

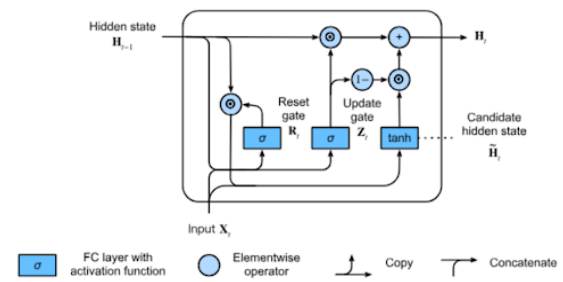
Prediksi Kecepatan Angin

Prediksi kecepatan angin merupakan proses memperkirakan kondisi angin berdasarkan data historis. Dalam meteorologi maritim, prediksi ini penting untuk mendukung keselamatan pelayaran dan aktivitas pesisir karena perubahan kecepatan angin dapat mempengaruhi kondisi gelombang, stabilitas kapal, serta kelancaran aktivitas maritim. Oleh karena itu, diperlukan metode yang mampu memprediksi kecepatan angin secara akurat sebagai dasar pengambilan keputusan.

Analisis *Time Series*

Analisis *time series* merupakan metode analisis data yang memperhatikan urutan waktu dalam setiap pengamatannya. Data *time series* memiliki karakteristik khusus karena nilai pada waktu tertentu dapat dipengaruhi oleh nilai pada waktu sebelumnya. Dalam penelitian ini, data kecepatan angin termasuk ke dalam data *time series* karena tersusun berdasarkan waktu pengamatan tertentu. Melalui analisis *time series*, pola tren, fluktuasi, dan perubahan kecepatan angin dapat dipahami secara lebih sistematis sehingga dapat digunakan sebagai dasar dalam proses pemodelan prediksi.

Gated Recurrent Unit (GRU)



Gambar 1. Arsitektur Model GRU

Gated Recurrent Unit (GRU) merupakan salah satu pengembangan Recurrent Neural Network (RNN) yang dirancang untuk memproses data berurutan, seperti time series. GRU mampu menangkap ketergantungan jangka panjang serta mengatasi masalah vanishing gradient pada RNN konvensional. Secara matematis, mekanisme aliran informasi pada sel GRU di setiap langkah waktu t diatur melalui empat tahapan persamaan utama yang diformulasikan sebagai berikut [10]:

- Update Gate : Gerbang ini mengontrol keseimbangan antara memori lama dan proporsi informasi dari hidden state sebelumnya yang akan dipertahankan menuju hidden state yang baru, dirumuskan melalui persamaan:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z)$$

- Reset Gate: Gerbang ini mengatur seberapa banyak informasi masa lalu yang perlu diabaikan oleh model sebelum membentuk representasi kandidat status yang baru, dengan persamaan:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r)$$

- Kandidat Hidden State: Fase ini merepresentasikan pemrosesan informasi adaptif baru yang dihasilkan berdasarkan input saat ini dan hidden state sebelumnya yang telah difilter

secara ketat oleh reset gate dirumuskan sebagai:

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_h)$$

- d. Hidden State Pembaruan Akhir: Tahap ini merupakan hasil akhir keluaran pada waktu ke-t, di mana nilainya adalah kombinasi antara hidden state sebelumnya dan kandidat hidden state yang baru, dengan proporsi yang dikendalikan penuh oleh update gate:

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$$

Simulasi Monte Carlo

Metode Simulasi *Monte Carlo* merupakan metode yang menggunakan sampel acak berulang untuk merepresentasikan dalam suatu sistem (Rais et al, 2025). Dalam penelitian ini, metode tersebut digunakan untuk mendukung hasil prediksi GRU dengan menghasilkan berbagai skenario prediksi sehingga dapat memberikan informasi mengenai rentang kemungkinan dan tingkat ketidakpastian kecepatan angin.

Mean Absolute Error (MAE)

MAE digunakan untuk mengukur rata-rata penyimpangan absolut antara hasil prediksi dan nilai aktual. Semakin kecil nilai MAE, semakin baik kemampuan model dalam menghasilkan prediksi yang mendekati kondisi sebenarnya. Pada penelitian ini, MAE digunakan untuk mengevaluasi rata-rata kesalahan model dalam memprediksi kecepatan angin.

Root Mean Squared Error (RMSE)

RMSE digunakan untuk mengukur tingkat kesalahan antara hasil prediksi dan data aktual dengan memberikan bobot lebih besar pada kesalahan yang tinggi. Menurut Matondang et al. (2026), RMSE dihitung dari akar rata-rata kuadrat selisih antara nilai aktual dan nilai prediksi. Metrik ini

penting karena kecepatan angin memiliki fluktuasi yang cukup tinggi. Semakin kecil nilai RMSE, semakin baik akurasi prediksi model.

Mean Absolute Percentage Error (MAPE)

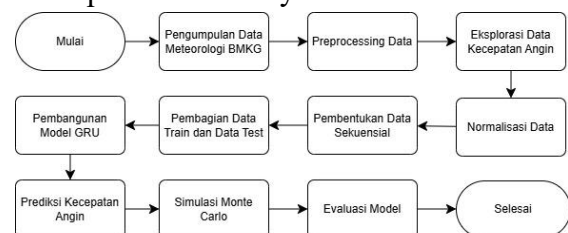
Menurut Arifuddin et al. (2025) Mean Absolute Percentage Error (MAPE) digunakan untuk mengukur akurasi prediksi dalam bentuk persentase sehingga tidak bergantung pada skala data yang digunakan. Semakin rendah nilai MAPE, semakin baik akurasi prediksi model. Nilai MAPE di bawah 10% umumnya menunjukkan tingkat akurasi yang sangat baik.

R² Score

R² Score digunakan untuk mengukur kemampuan model dalam menjelaskan variasi data aktual. Nilai R² yang semakin mendekati 1 menunjukkan bahwa model semakin baik dalam merepresentasikan pola data. Pada penelitian ini, R² Score digunakan untuk mengevaluasi kemampuan model GRU dalam menjelaskan variasi kecepatan angin berdasarkan data historis.

METODE PENELITIAN

Alur penelitian harus dibuat sebagai pedoman agar setiap tahapan penelitian dapat dilakukan secara sistematis, terarah, dan sesuai dengan tujuan yang telah ditetapkan sebelumnya.



Gambar 2.. Alur Penelitian

Berikut adalah penjelasan dari alur penelitian tersebut :

1. Pengumpulan Data Meteorologi BMKG : Data penelitian diperoleh dari Stasiun Meteorologi Maritim BMKG Tanjung Perak Surabaya melalui prosedur permintaan data resmi. Dataset mencakup periode 2019-2025 dengan interval pencatatan setiap 3 jam untuk merekam dinamika cuaca di wilayah pelabuhan. Pada penelitian ini, parameter kecepatan angin (wind speed) dalam satuan knots digunakan sebagai variabel utama.
2. Preprocessing Data : Fase preprocessing bertujuan mengubah dataset meteorologi mentah menjadi data yang siap digunakan dalam pemodelan deep learning. Proses ini meliputi konversi waktu observasi ke format datetime, pemeriksaan nilai hilang, deteksi outlier, serta analisis statistik deskriptif untuk memastikan kualitas dan konsistensi data kecepatan angin.
3. Eksplorasi Data Kecepatan Angin : Data kecepatan angin dieksplorasi secara visual untuk mengidentifikasi tren temporal dan pola fluktuasi. Analisis mencakup visualisasi time series dengan garis acuan berupa nilai rata-rata (mean) keseluruhan dataset. Tahap ini bertujuan memahami karakteristik data, termasuk tren, musiman, dan outlier, sebelum dilanjutkan ke tahapan berikutnya.
4. Normalisasi Data : Normalisasi data kecepatan angin dilakukan menggunakan MinMaxScaler untuk mengubah nilai data ke rentang 0-1. Proses ini bertujuan menjaga kestabilan pelatihan model GRU serta mengurangi pengaruh perbedaan skala data. Setelah dinormalisasi, dataset siap ditransformasikan menjadi data sekuensial sebagai masukan model.
5. Pembentukan Data Sekuensial : Data kecepatan angin ditransformasikan ke dalam format *supervised learning* menggunakan metode sliding window menggunakan metode sliding window dengan ukuran 24 time steps yang merepresentasikan 72 jam data historis. Transformasi ini diperlukan agar model GRU dapat mempelajari pola temporal dari urutan data masa lalu untuk memprediksi nilai pada periode berikutnya. Selanjutnya, data sekuensial yang terbentuk digunakan pada tahap pembagian data latih dan data uji.
6. Pembagian Data Train dan Data Test : Dataset sekuensial yang telah terbentuk selanjutnya dipisahkan menjadi dua subset utama, yaitu data latih (training data) dan data uji (testing data), dengan proporsi pembagian sebesar 80% berbanding 20%. Proses pemisahan data ini bertujuan untuk mempertahankan integritas dan kesinambungan urutan waktu dari struktur *time-series*. Sebanyak 80% observasi awal dialokasikan secara khusus sebagai data latih untuk memberikan rentang informasi historis yang memadai bagi model GRU dalam mempelajari dependensi temporal kecepatan angin. Sementara itu, sisa 20% observasi terakhir diperuntukkan sebagai data uji untuk mengukur kapabilitas dan tingkat akurasi prediksi model saat dihadapkan pada skenario data riil yang belum pernah dipelajari sebelumnya.
7. Pembangunan Model GRU : Arsitektur GRU dibangun dengan dua *hidden state* untuk menangkap dependensi temporal

kecepatan angin. Lapisan pertama terdiri dari 64 unit untuk mengekstraksi pola cuaca berdimensi tinggi, diikuti lapisan kedua dengan 32 unit untuk kompresi informasi dan penyaringan fitur relevan. Antar lapisan diterapkan Dropout sebesar 20% dengan setting $\text{training_Dropout}=\text{True}$ untuk simulasi Monte Carlo. Model dioptimalkan menggunakan Adam dengan loss function Mean Squared Error (MSE), dan dilatih selama 20 epoch. Penghentian pelatihan pada epoch ini didasarkan pada konvergensi validation loss (0,0004-0,0005), yang menunjukkan stabilitas dan mencegah overfitting.

8. Prediksi Kecepatan Angin : Tahap prediksi dieksekusi dengan memanfaatkan model Gated Recurrent Unit (GRU) yang telah selesai dilatih untuk memperkirakan nilai kecepatan angin pada sekumpulan data uji
9. Mengingat observasi masukan sebelumnya telah melalui tahap pra pemrosesan, hasil keluaran awal dari model masih berada dalam bentuk skala ternormalisasi. Oleh karena itu, penerapan transformasi balik (inverse transform) diwajibkan untuk mengonversi kembali nilai prediksi tersebut ke dalam parameter dan satuan aslinya, yakni knots. Pengembalian skala dasar ini sangat krusial agar keluaran model dapat dibandingkan secara matematis maupun visual terhadap kondisi data angin aktual di lapangan. Nilai tebakan deterministik tunggal dari tahap inilah yang kemudian dijadikan sebagai garis acuan (baseline) sebelum diekspansi menjadi berbagai kemungkinan skenario melalui simulasi Monte Carlo
10. Simulasi Monte Carlo : Tahap selanjutnya menerapkan Simulasi Monte Carlo menggunakan teknik Monte Carlo Dropout. Implementasi dilakukan dengan mengaktifkan lapisan Dropout pada fase inference untuk menghasilkan variasi prediksi stokastik. Simulasi dilakukan sebanyak 100 iterasi, menghasilkan 100 sampel prediksi independen pada setiap langkah waktu. Prakiraan titik tunggal (point estimate) ditentukan dari rata-rata keseluruhan 100 sampel tersebut sebagai ensemble mean. Estimasi ketidakpastian model diperoleh melalui perhitungan standar deviasi dari distribusi sampel prediksi, mencerminkan variabilitas prediksi akibat dropout stokastik. Confidence Interval 95% dibentuk dengan mengalikan standar deviasi dengan nilai kritis distribusi normal ($z = 1,96$), kemudian menambahkan dan mengurangkan hasil perkalian tersebut terhadap nilai rata-rata prediksi. Operasi ini menghasilkan batas bawah dan batas atas interval kepercayaan untuk fluktuasi kecepatan angin.
11. Evaluasi Model : Tahap akhir penelitian dilakukan untuk mengukur akurasi prediksi dan keandalan sistem secara kuantitatif. Kinerja prediksi deterministik dievaluasi menggunakan MAE, RMSE, MAPE, dan R^2 . Karena model menerapkan pendekatan probabilistik, kualitas estimasi ketidakpastian dari simulasi Monte Carlo juga dinilai melalui rata-rata standar deviasi dan Coefficient of Variation (CV). Selain itu, efektivitas model GRU divalidasi dengan membandingkan performanya terhadap

model baseline deret waktu, yaitu LSTM dan RNN.

Alat Dan Bahan Penelitian

Perangkat lunak dan perangkat keras digunakan untuk mendukung proses pengolahan data, pemodelan, dan evaluasi hasil prediksi penelitian. Perangkat lunak yang digunakan adalah Python dengan beberapa pustaka pendukung, termasuk Pandas dan NumPy untuk pengolahan data, visualisasi dengan Matplotlib dan Seaborn, normalisasi dan evaluasi model dengan Scikit-learn, dan pembuatan model Gated Recurrent Unit (GRU) dengan TensorFlow dan Keras. Selain itu, digunakan pula *Statsmodels* untuk mendukung analisis data deret waktu dan *Plotly* untuk visualisasi interaktif.

ANALISIS DATA DAN PENGEMBANGAN MODEL

1. Pengumpulan Data

Pengumpulan data meteorologi maritim dilakukan dengan mengekstraksi rekaman observasi historis secara langsung dari basis data stasiun pengamatan BMKG Tanjung Perak Surabaya yang mencakup periode pencatatan selama tujuh tahun, yang dihitung mulai dari 1 Januari 2019 hingga akhir tahun 2025. Rekaman data ditarik dengan interval per 3 jam untuk mencatat dinamika parameter cuaca secara kronologis. Berdasarkan hasil ekstraksi tersebut, diperoleh himpunan data deret waktu yang secara total terdiri dari 20.463 baris observasi dan 24 kolom fitur meteorologi. Dari keseluruhan struktur data yang terkumpul, parameter kecepatan angin (Wind Speed) dalam satuan knots kemudian diisolasi untuk dijadikan sebagai variabel target tunggal pada proses pemodelan.

2. Preprocessing Data

Tahap pra pemrosesan diawali dengan mengkonversi atribut waktu observasi ke tipe data datetime dan menetakannya sebagai indeks utama dataset. Data diurutkan secara kronologis untuk mempertahankan kontinuitas time-series. Inspeksi terhadap nilai kosong (missing value) menunjukkan bahwa seluruh observasi lengkap tanpa Null value yang perlu diimputasi. Tahap pembersihan data dilakukan dengan menghilangkan kolom meteorologi sekunder, sehingga dataset hanya menyisakan variabel target kecepatan angin (*WindSpeed*) untuk tahap pemodelan.

3. Analisis Eksploratif Data

Tahap analisis eksploratif data dimulai dengan menghitung parameter statistik deskriptif (minimum, maksimum, dan rata-rata) untuk memetakan sebaran distribusi kecepatan angin. Selanjutnya, data divisualisasikan dalam bentuk grafik deret waktu interaktif menggunakan Plotly, dengan garis acuan horizontal pada titik rata-rata untuk memandu pengamatan tren dan fluktuasi. Melalui analisis deskriptif dan visualisasi ini, karakteristik dinamika kecepatan angin diinspeksi sebelum dilanjutkan ke tahap normalisasi data.

4. Normalisasi Data

Tahap selanjutnya adalah normalisasi variabel target kecepatan angin menggunakan metode MinMaxScaler untuk mengonversi data asli ke rentang skala 0-1. Transformasi ini merupakan prasyarat agar GRU dapat memproses pembaruan bobot secara stabil tanpa terganggu oleh nilai ekstrem. Penyeragaman skala ini memfasilitasi konvergensi jaringan yang lebih

proporsional selama pelatihan. Setelah semua observasi berhasil diskalakan, data disiapkan untuk direstrukturisasi pada tahap pembentukan sekuensial.

5. Pembentukan Data Sekuensial

Variabel kecepatan angin yang telah dinormalisasi direstrukturisasi ke format supervised learning menggunakan sliding window dengan window size 24 time steps (72 jam historis). Setiap window mengelompokkan 24 data berurutan sebagai input yang dipetakan dengan satu observasi berikutnya sebagai target prediksi. Struktur array tiga dimensi ini mempertahankan dependensi temporal sebelum data dibagi menjadi data latih dan data uji.

6. Pembagian Dataset

Himpunan data sekuensial dibagi menjadi data latih (80%) dan data uji (20%) berdasarkan urutan waktu tanpa pengacakan untuk mempertahankan kesinambungan deret waktu. Data latih dengan proporsi lebih besar dialokasikan agar jaringan dapat mengekstraksi pola dependensi temporal. Data uji digunakan untuk memvalidasi generalisasi model terhadap data yang belum dipelajari sebelumnya.

7. Pembangunan Model GRU

Arsitektur jaringan dirancang menggunakan Functional API untuk menerima data sekuensial dan meneruskannya ke lapisan GRU pertama berkapasitas 64 unit. Keluaran diproses melalui fungsi Dropout dengan rasio 20%, dikonfigurasi untuk tetap aktif (`training=True`) sebagai instrumen dasar simulasi Monte Carlo. Representasi fitur kemudian diproses pada lapisan GRU

kedua berkapasitas 32 unit, dilewatkan pada Dropout 20%, dan masuk ke lapisan Dense tunggal sebagai output layer. Model dikompilasi menggunakan optimizer Adam dengan loss function Mean Squared Error (MSE) untuk pembaruan bobot. Pelatihan dilakukan sebanyak 20 epoch dengan batch size 32 observasi. Selama pelatihan, 10% dari data latih dipisahkan sebagai validation set untuk memantau konvergensi loss.

8. Simulasi Monte Carlo

Tahap selanjutnya adalah penerapan simulasi Monte Carlo untuk membentuk estimasi ketidakpastian menggunakan Monte Carlo Dropout. Metode ini diimplementasikan dengan menjalankan proses inferensi model secara berulang sambil mempertahankan lapisan Dropout tetap aktif pada fase pengujian. Dengan aktivasi mekanisme ini, model melakukan pengambilan sampel iteratif untuk menghasilkan 100 prediksi independen pada setiap langkah waktu. Kumpulan sampel keluaran kemudian diagregasi untuk mengekstrak dua parameter utama: nilai rata-rata (mean) sebagai prediksi utama, dan standar deviasi sebagai representasi ketidakpastian prediksi. Terakhir, standar deviasi digunakan untuk membentuk pita ketidakpastian (uncertainty band) pada interval kepercayaan 95%, memetakan batas bawah dan batas atas fluktuasi kecepatan angin.

10. Evaluasi Model

Tahap akhir penelitian adalah mengevaluasi performa model untuk menilai tingkat akurasi dan reliabilitas prediksi. Evaluasi deterministik dilakukan menggunakan metrik MAE, RMSE, MAPE, dan R^2 dengan membandingkan hasil prediksi terhadap data aktual. Selain itu, evaluasi probabilistik diterapkan untuk

memverifikasi estimasi ketidakpastian melalui simulasi Monte Carlo, menggunakan rata-rata standar deviasi dan Coefficient of Variation (CV). Untuk menguji efektivitas arsitektur yang diusulkan, model GRU dibandingkan dengan model baseline, yaitu LSTM dan RNN. Hasil evaluasi ini digunakan untuk menilai kemampuan model dalam mengurangi kesalahan prediksi serta menghasilkan interval kepercayaan yang andal.

IMPLEMENTASI DAN PEMBAHASAN

1. Pengumpulan Data

Data meteorologi maritim dari BMKG Tanjung Perak Surabaya digunakan dalam penelitian ini. Data yang digunakan berbentuk *time series* dan memuat beberapa parameter meteorologi, seperti waktu pengamatan, tinggi gelombang, kecepatan angin, arah angin, serta parameter pendukung lainnya. Berdasarkan hasil pemuatan data, dataset terdiri dari 20.463 observasi dengan 24 kolom fitur. Proses prediksi kecepatan angin menggunakan model Gated Recurrent Unit (GRU) menggunakan data ini sebagai sumber utama.

2. Preprocessing Data

Tahap *preprocessing* diawali dengan pemeriksaan struktur dataset untuk mengidentifikasi jumlah baris dan kolom, tipe data, serta keberadaan nilai kosong sebelum analisis dan pemodelan dilakukan. Hasil pemeriksaan struktur dataset ditunjukkan pada Gambar 3.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20463 entries, 0 to 20462
Data columns (total 24 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                ---
0   Time(UTC/GMT)                          20463 non-null object
1   Hsig(m)                                 20463 non-null float64
2   Hsig(Scale)                             20463 non-null object
3   Hmax(m)                                 20463 non-null float64
4   Hmax(Scale)                             20463 non-null object
5   WaveDir(deg)                           20463 non-null int64
6   WaveDir(compass)                       20463 non-null object
7   PrimSwell(m)                           20463 non-null float64
8   PrimSwell(Scale)                       20463 non-null object
9   PrimSwellDir(deg)                      20463 non-null int64
10  PrimSwellDir(compass)                   20463 non-null object
11  WindSea(m)                              20463 non-null float64
12  WindSea(Scale)                          20463 non-null object
13  WindSeaDir(deg)                         20463 non-null int64
14  WindSeaDir(compass)                     20463 non-null object
15  WavePeriod(s)                           20463 non-null int64
16  SurfCurrentSpd(cm/s)                    20463 non-null float64
17  SurfCurrentDir                           20463 non-null int64
18  SurfCurrentDir(compass)                 20463 non-null object
19  SeaSurfaceTemperature(°C)              20463 non-null float64
20  SeaSurfaceSalinity(PSU)                 20463 non-null float64
21  WindSpeed(knots)                        20463 non-null int64
22  WindDir(deg)                            20463 non-null int64
23  WindDir(compass)                        20463 non-null object
dtypes: float64(7), int64(7), object(10)
memory usage: 3.7+ MB
```

Gambar 3. Informasi Struktur Dataset

Berdasarkan Gambar 3, dataset terdiri dari 20.463 baris dan 24 kolom. Seluruh kolom memiliki 20.463 data non-null sehingga tidak terdapat nilai kosong. Tipe data yang digunakan meliputi float64, int64, dan object. Data numerik digunakan untuk parameter seperti tinggi gelombang, kecepatan angin, suhu permukaan laut, dan tekanan udara, sedangkan tipe object digunakan untuk data kategorikal, seperti skala dan arah mata angin. Hasil ini menunjukkan bahwa struktur dataset telah memadai dan siap digunakan pada tahap analisis selanjutnya..

3. Eksplorasi Data Kecepatan Angin

Pada tahap pra-pemodelan, dilakukan eksplorasi data kecepatan angin untuk mengidentifikasi pola temporal, termasuk tren, fluktuasi, dan nilai rata-rata selama periode pengamatan. Hasil visualisasi tren tersebut ditampilkan pada Gambar 4.



Gambar 4. Tren Kecepatan Angin

Berdasarkan Gambar 4, kecepatan angin menunjukkan perubahan yang dinamis selama periode pengamatan dengan rentang sekitar 2–20 knots dan rata-rata 9,50 knots yang ditandai garis putus-putus merah. Data memperlihatkan

fluktuasi naik-turun pada beberapa periode, termasuk peningkatan yang tampak pada awal 2025. Pola temporal yang berubah-ubah ini menunjukkan bahwa data sesuai untuk dianalisis menggunakan model deret waktu seperti GRU.

4. Normalisasi Data

Sebelum masuk ke tahap pemodelan, data kecepatan angin dinormalisasi menggunakan *MinMaxScaler* ke dalam rentang 0 sampai 1. Normalisasi ini dilakukan agar skala data menjadi lebih seimbang dan proses pelatihan model *GRU* dapat berjalan lebih stabil. Setelah dinormalisasi data kemudian siap dibentuk menjadi data sekuensial untuk kebutuhan model.

5. Pembentukan Data Sekuensial

Dalam proses ini, data kecepatan angin diubah menjadi format supervised learning dengan menggunakan teknik sliding window yang memiliki rentang waktu 24 jam. Dengan demikian, data historis dari 24 jam sebelumnya dapat digunakan sebagai variabel masukan untuk memprediksi nilai kecepatan angin pada periode berikutnya. Tahap ini penting karena model *GRU* membutuhkan data berurutan agar dapat mempelajari pola temporal dari data kecepatan angin.

6. Pembagian Data *Train* dan Data *Test*

Pada tahap ini, dataset dibagi menjadi dua bagian dengan perbandingan 80 banding 20. Bagian pertama digunakan sebagai data pelatihan untuk melatih model, sedangkan bagian kedua digunakan sebagai data pengujian untuk mengevaluasi kinerja model dalam memprediksi data yang belum pernah di uji sebelumnya.

7. Pembangunan Model *GRU*

Pada tahap pembangunan model, arsitektur GRU dilatih untuk mempelajari pola kecepatan angin dari data historis. Model terdiri atas dua lapisan GRU dengan 64 unit pada lapisan pertama dan 32 unit pada lapisan kedua. Pelatihan dilakukan selama 20 epoch menggunakan optimizer Adam dan fungsi loss Mean Squared Error (MSE). Hasil pelatihan ditunjukkan pada Gambar 5. .

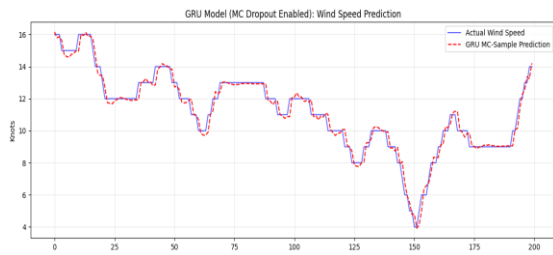


Gambar 5. Hasil Pelatihan Model GRU

Berdasarkan Gambar 5, nilai loss pelatihan menurun secara konsisten dari 0,0055 menjadi $9,74 \times 10^{-4}$. Tren yang sama terlihat pada loss validasi yang tetap stabil dan cenderung menurun selama pelatihan. Hasil tersebut menunjukkan bahwa model GRU mampu mempelajari pola kecepatan angin secara efektif. Keselarasan penurunan loss pelatihan dan validasi juga mengindikasikan tidak adanya overfitting yang signifikan, sehingga model layak dilanjutkan ke tahap prediksi dan evaluasi.

8. Prediksi Kecepatan Angin

Tahap prediksi dilakukan dengan menggunakan model Gated Recurrent Unit (GRU) yang telah dilatih untuk memprediksi kecepatan angin pada data uji. Hasil prediksi dikonversi kembali ke skala aslinya melalui transformasi inverse sehingga dapat dilakukan perbandingan langsung dengan data aktual. Gambar 6 di bawah ini menunjukkan perbandingan antara nilai aktual dan hasil prediksi model.

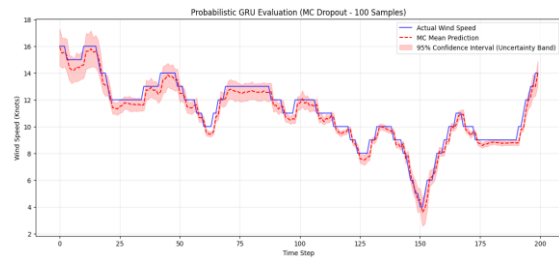


Gambar 6. Perbandingan Aktual dan Prediksi

Visualisasi pada Gambar 6 menunjukkan bahwa kurva prediksi dari model GRU sangat sesuai dengan kurva data aktual. Hal ini menunjukkan bahwa model tersebut efektif dalam menangkap pola perubahan kecepatan angin, baik saat nilainya meningkat maupun menurun. Model juga cukup mampu menangkap perubahan ekstrem, seperti penurunan tajam di sekitar titik pengamatan ke-150 dan kenaikan kembali pada bagian akhir data. Kesesuaian antara data aktual dan hasil prediksi menunjukkan bahwa model *GRU* efektif dalam mengidentifikasi pola temporal pada kecepatan angin.

9. Simulasi Monte Carlo

Metode simulasi Monte Carlo digunakan untuk menghasilkan sejumlah skenario prediktif berdasarkan hasil model GRU. Dalam penelitian ini, simulasi dilakukan dengan menghasilkan 100 skenario prediksi. Pendekatan ini digunakan untuk memberikan gambaran ketidakpastian hasil prediksi. Dengan demikian, hasil yang diperoleh tidak terbatas satu nilai prediksi saja, tetapi juga dapat menggambarkan kemungkinan variasi kecepatan angin. Visualisasi ini ditampilkan pada Gambar 7 dibawah ini.



Gambar 7. Hasil Simulasi Monte Carlo

Berdasarkan Gambar 7, rata - rata hasil prediksi Monte Carlo (garis merah putus - putus) berada sangat dekat dengan data aktual (garis biru), menunjukkan bahwa model GRU mampu merepresentasikan pola perubahan kecepatan angin dengan baik. Selain itu, pita ketidakpastian yang relatif sempit mengindikasikan bahwa variasi antar skenario prediksi rendah dan tingkat kepercayaan model terhadap hasil prediksi cukup tinggi.

Meskipun pita ketidakpastian sedikit melebar pada beberapa titik, terutama saat terjadi perubahan nilai yang tajam di sekitar titik pengamatan ke - 150, data aktual masih berada dalam rentang interval kepercayaan yang dihasilkan. Hal ini menunjukkan bahwa integrasi GRU dan Monte Carlo Dropout mampu menangkap pola temporal kecepatan angin sekaligus mengestimasi ketidakpastian secara efektif.

10. Evaluasi Model

Tahap evaluasi melibatkan perbandingan antara hasil prediksi model dengan data aktual untuk menilai seberapa akurat model tersebut memprediksi kecepatan angin. Proses evaluasi ini menggunakan beberapa metrik, antara lain Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), dan nilai R^2 . Hasil evaluasi model ditampilkan pada Tabel 2

Tabel 2. Metrik Evaluasi Deterministik

METRIC	MODEL		
	GRU	LSTM	RNN
MAE	0.3472	0.3706	0.400
MAPE	0.4680	0.5016	0.518
RMSE	5.37%	6.03%	6.14%
R ²	0.9879	0.9861	0.9852

Berdasarkan hasil pengujian, model Gated Recurrent Unit (GRU) menunjukkan performa prediksi terbaik dibandingkan LSTM dan RNN. Model GRU memperoleh nilai R² sebesar 0,9878 dan MAE sebesar 0,3472. Selain itu, GRU juga menghasilkan nilai MAPE dan RMSE terendah, sedangkan RNN memiliki tingkat kesalahan tertinggi pada seluruh metrik evaluasi. Hasil ini menunjukkan bahwa mekanisme gating pada GRU lebih efektif dalam menangkap pola temporal kecepatan angin dibandingkan model pembanding. Mengingat arsitektur GRU menunjukkan performa terbaik dibandingkan model lainnya, arsitektur ini dipilih untuk diintegrasikan dengan Monte Carlo Dropout. Oleh karena itu, pengujian lanjutan difokuskan pada evaluasi kualitas estimasi probabilistik berbasis GRU tanpa melibatkan model LSTM maupun RNN.

Metrik Probabilistik	Nilai
Rata-rata Standar Deviasi	0.3533 Knots
Rata-rata Variabilitas Relatif (CV %)	6.94%

Gambar 8. Metrik Evaluasi Probabilistik

Berdasarkan hasil pengujian yang disajikan pada Gambar 8, simulasi probabilistik menghasilkan rata - rata standar deviasi sebesar 0,3533 knots dan rata - rata koefisien variasi (CV) sebesar 6,94%. Nilai tersebut menunjukkan bahwa interval prediksi yang dihasilkan cukup presisi dengan tingkat ketidakpastian yang stabil dan terkontrol. Temuan ini mengindikasikan bahwa model GRU tidak hanya mampu menghasilkan prediksi yang akurat, tetapi juga andal dalam

mengestimasi interval kepercayaan untuk mendukung pengambilan keputusan pada sektor maritim.

KESIMPULAN

Berdasarkan hasil penelitian, integrasi model Gated Recurrent Unit (GRU) dan Simulasi Monte Carlo mampu menghasilkan prediksi kecepatan angin yang sangat baik dengan nilai MAE 0,3718, RMSE 0,5014, MAPE 5,78%, dan R² Score 0,9861, yang menunjukkan kemampuan model dalam menangkap pola temporal serta fluktuasi kecepatan angin secara efektif. Selain menghasilkan prediksi titik (*point forecast*), Simulasi Monte Carlo juga dapat mengestimasi ketidakpastian prediksi melalui berbagai skenario sehingga memberikan informasi rentang kemungkinan nilai yang dapat meningkatkan interpretabilitas hasil dan mendukung pengambilan keputusan pada sektor maritim. Penelitian ini berhasil dengan menggabungkan GRU dan Simulasi Monte Carlo pada prediksi kecepatan angin, utamanya pada aspek ketidakpastian sehingga menghasilkan informasi yang lebih komprehensif. Namun, penelitian ini masih memiliki keterbatasan karena menggunakan data dari satu lokasi dan periode pengamatan tertentu, serta variabel yang masih terbatas sehingga belum sepenuhnya mewakili seluruh faktor meteorologi yang memengaruhi kecepatan angin.

Saran

Berikut adalah beberapa saran yang dapat digunakan untuk penelitian berikutnya:

1. Penelitian selanjutnya dapat membandingkan performa GRU dengan metode *deep learning* lain,

- seperti LSTM, BiLSTM, dan Transformer, untuk memperoleh model prediksi kecepatan angin yang paling optimal.
- Estimasi ketidakpastian dapat dikembangkan dengan membandingkan Simulasi Monte Carlo dengan pendekatan probabilistik lain, seperti Monte Carlo Dropout, Bayesian Neural Network, dan Quantile Regression, guna mengevaluasi kualitas interval prediksi secara lebih komprehensif.
 - Penelitian selanjutnya disarankan menggunakan data meteorologi multivariat, seperti suhu udara, tekanan udara, kelembapan, curah hujan, arah angin, dan radiasi matahari, untuk meningkatkan akurasi serta kemampuan model dalam menangkap dinamika atmosfer.
 - Pengujian pada lokasi berbeda, periode pengamatan yang lebih panjang, dan data multistasiun perlu dilakukan untuk mengevaluasi kemampuan generalisasi model sehingga dapat diterapkan lebih luas pada sistem peramalan cuaca dan keselamatan maritim.

DAFTAR PUSTAKA

- Arifuddin, D., Kusnawati, & Kusnawati. (2025). *Perbandingan Performansi Algoritma Multiple Linear Regression dan Multi Layer Perceptron Neural Network dalam Memprediksi Penjualan Obat*. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 5 (2), 722 - 737. <https://doi.org/10.57152/malcom.v5i2.1952>
- Arwansyah, Suryani, Hasyrif SY, Usman, Ahyuna, & Alam, S. (2024). *Time Series Forecasting Menggunakan Deep Gated Recurrent Units*. Digital Transformation Technology (Digitech), 4 (1), 410 - 416.
- Islam, C. C., & Wahabi, A. (2025). *Perbandingan Model LSTM dan GRU untuk Peramalan Angin*. Sistemasi: Jurnal Sistem Informasi, 14 (6), 2871 - 2831.
- Jaya, W. P., Wintoro, P. B., Septiana, T., & Mulyani, Y. (2025). *Peramalan Multivariate Time Series Harga Aspal Menggunakan Algoritma Gated Recurrent Unit*. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 5 (4), 1517 - 1530. <https://doi.org/10.57152/malcom.v5i4.2198>
- Jordi, R., & Narullita, L. F. (2026). *Rancang Bangun Sistem Informasi Prediksi Risiko Penyeberangan Laut Berbasis Data Cuaca Menggunakan Metode Decision Tree*. JATI (Jurnal Mahasiswa Teknik Informatika), 10 (1).
- Liwijaya, A., Risma, P., Dewi, T., & Oktarina, Y. (2025). *Implementasi Deep Learning Dalam Prediksi Real time Irradian Surya*. Journal Of Applied Smart Electrical Network And Systems (Jasens), 6(2), 90 -96. <https://doi.org/10.52158/pvdpsr36>
- Matondang, I. J., Budianita, E., Syafrina, F., & Afrianty, I. (2026). *Sistem Prediksi Produksi Kelapa Sawit Berbasis Gradio Menggunakan Algoritma Regresi Linear Berganda*. Bulletin Of Computer Science Research, 6 (2), 664 - 672.
- Minarto, D., & Santoso, K. T. (2023). *Pengembangan Sistem Monitoring Dan Prediksi Cuaca Maritim Untuk Peningkatan Keselamatan Navigasi*. SAMMAJIVA: Jurnal Penelitian Bisnis dan Manajemen, 1 (4), 231 - 138.
- Murase, K., Nakamoto, A., & Tomimaya, N. (2024). *Performance of recurrent neural networks with Monte Carlo dropout for predicting pharmacokinetic parameters from*

- dynamic contrast-enhanced magnetic resonance imaging data*. Journal of Applied Clinical Medical Physics, 2, 1-19.
- Rais, E. R., Sovia, R., & Sumijan. (2025). *Analisis Prediksi Penjualan Suku Cadang Motor dengan Metode Monte Carlo*. Bit-Tech (Binary Digital - Technology), 8 (1), 41 - 48.
- Ramadila, H., Susandri, S., & Zamsuri, A. (2026). *Optimization of LSTM and GRU Deep Learning Models for Tidal Sea Level Time Series Prediction to Support Early Warning Systems*. JAISEN: Journal of Advanced Information Systems and Engineering, 2 (1).
- Tabas, S., & Samadi. (2022). *Variational Bayesian dropout with a Gaussian prior for recurrent neural networks application in rainfall-runoff modeling*. Environmental Research Letters, 17, 1-17.