Jua Vilhelm: A Novel End-to-End AI System for Global Weather Prediction

Marvin Vincent Gabler
, *J. Wuilloud, H. Taheri Shahraiyni, D. Neupert, A. Grigoryev, R. Almeida, G. Martin Hernandez, J. D. Daubinet, N. Ekhtiari, R. J. Song, P. Dudbridge, and E. Tarakci*
Jua.ai AG, Pfäffikon, Switzerland

32nd Conference on Weather Analysis and Forecasting(WAF)/
28th Conference on Numerical Weather Prediction (NWP)/
20th Conference on Mesoscale Processes
17-21 July 2023, Madison, WI
American Meteorological Society

Abstract

What is Jua Vilhelm system?: It is an innovative end-to-end artificial intelligence (AI)-powered weather forecasting system, trained on an extensive global dataset.
Scale & Resolution: Hourly high-resolution (1x1 km grid) global predictions.
Performance: The Jua Vilhelm system demonstrates superior performance in predicting near-surface parameters, such as air temperature and wind speed, when compared to the Integrated Forecast System (IFS) model. Furthermore, the IFS model occasionally generates inconsistent outputs, while Jua Vilhelm maintains complete consistency.
In addition, the results of benchmarking of jua Vilhelm System using global SYNOP observations proved its excellent performance.
Other specific achievements: Vilhelm achieves rapid forecast run times in a matter of seconds, enabling the execution of hundreds of ensemble runs within minutes, an accomplishment previously unattainable.
Conclusion: The development of Jua Vilhelm system marks a significant advancement in the field of weather forecasting, offering unprecedented accuracy and speed.

Material & Methods

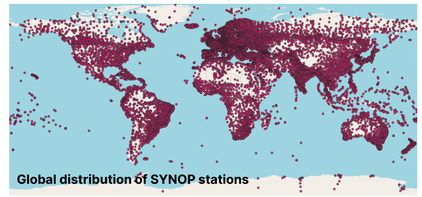
What is the methodology behind Jua Vilhelm system? It's an ensemble of specialized models, each fine-tuned for different aspects of the forecast. These models are underpinned by transformer neural networks, optimized for solving partial differential equations with unparalleled efficiency.
What is the key to multi-step forecasting? To predict multiple timesteps, the model uses an autoregressive method: the output from one step serves as input for the next. To minimize compounding errors over longer forecast horizons, we apply sophisticated fine-tuning techniques.
How does the model handle spatial resolution? We employ a series of models to produce a final output at a native 1x1km grid resolution near surface.
What about the output format and processing time? The model generates global forecasts, written to Zarr files at a resolution of about 0.008 degree (43200x21600), within just 20 minutes from receiving initial conditions.
Which datasets were used? We primarily use ERA5 reanalysis and radar data, complemented by IoT data and observations from SYNOP and METAR weather stations.
How did you train the model? Our model was trained using an in-house GPU cluster of 32 A100 GPUs.

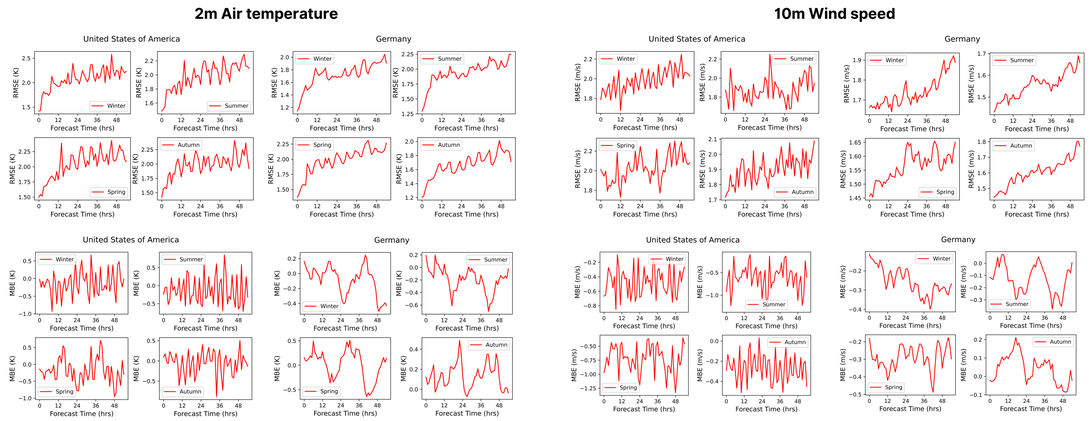
Evaluations:

Variables: 2m air temperature & 10m zonal and meridional wind speed.
Metrics: We utilized several metrics, but there is not enough space to present all of them here, hence we present here only Mean Bias Error (MBE) and Root Mean Square Error (RMSE)
1 - Benchmarking using observations:
Observation data: The SYNOP observations (More than 10000 stations, globally distributed) for 2022 were used for country- and seasonal-based benchmarking of JUA Vilhelm model
Time steps: Hourly benchmarking of JUA Vilhelm up to 54 hours ahead
2 - Comparing Jua Vilhelm against IFS:
Ground truth data for comparative analysis:
JUA Vilhelm: The ERA5 reanalysis global dataset for 2018 with 0.25 degree resolution
IFS: IFS global Initial conditions for 2018 with 0.25 degree resolution.
Time steps: *6-hour intervals for JUA Vilhelm and 12 hours intervals for IFS (global IFS initial conditions for evaluations were available 2 times per day)*

Benchmarking using SYNOP observations 2022

Benchmarking using SYNOP observations has been done for all of the countries and seasons and here only the results for two countries (USA & Germany) are presented

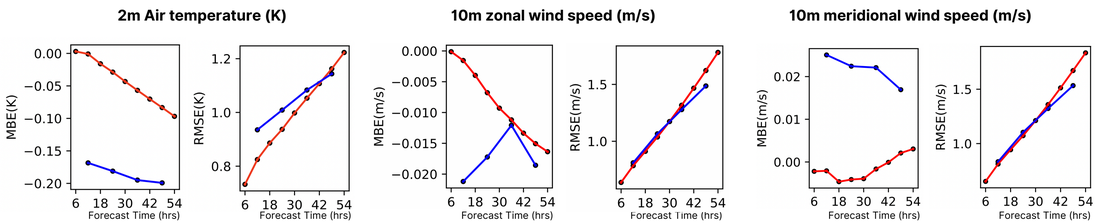




The model shows a good performance for short term prediction of air temperature and wind speed for all of the seasons. We notice significant bias between the initial conditions and observations (Forecast time=0 in the figures). Therefore, the significant part of the error in the predictions is related to the inherent difference between observations and initial conditions.

Comparing Jua Vilhelm against IFS (Year:2018)

Benchmarking using SYNOP observations has been done for all of the countries and seasons and here only the results for two countries (USA & Germany) are presented



In terms of air temperature, Jua Vilhelm not only outperforms IFS in RMSE but also demonstrates reduced bias.

In terms of wind speed (Zonal and meridional), IFS models exhibit comparable RMSE levels for wind speed; however, Jua Vilhelm outperforms IFS significantly in terms of MBE.

www.jua.ai
marvin@jua.ai

Jua