# Evaluating ECMWF AIFS Performance for Severe Weather Events in Japan: A Case Study Analysis

16th Sep 2025, UEF2025

Jumpei Fujino and Kohei Sakamoto
Weathernews Inc.

# Agenda

- **Introduction**
- Case Study 1: Frontal Heavy Rain Across Japan in August
- Case Study 2: Highest Temperature Ever Recorded in Japan
- Accuracy of ERA5
- Statistical Evaluation of Heavy Rain and High Temperature Forecast in Summer in Japan
- Summary and Future Work

# Self Introduction

Jumpei Fujino

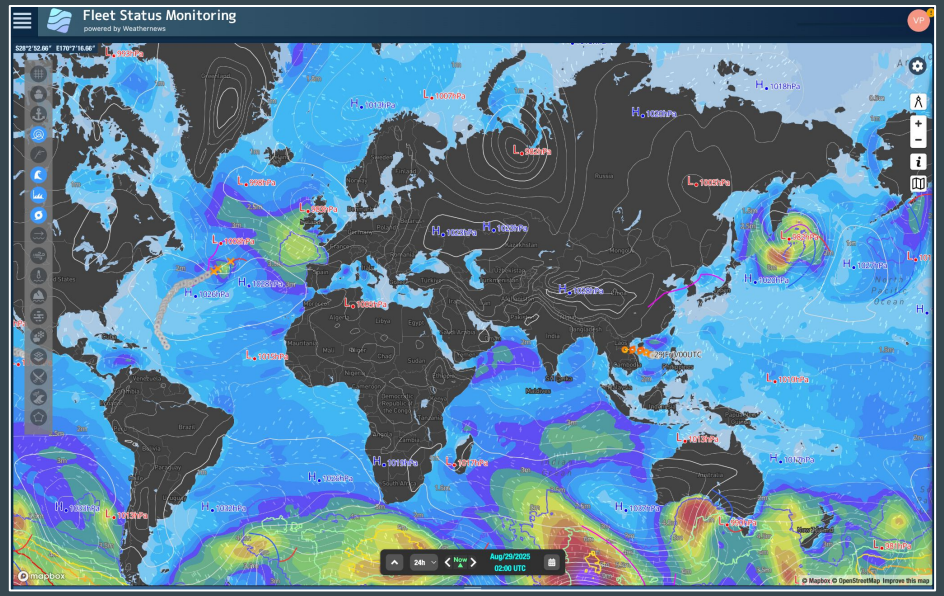
Developer specialized in statistics
Development of probabilistic forecast
and AI weather prediction model
WNI Forecast Center
Weathernews Inc.

Kohei Sakamoto

Forecast system development manager
Development of ocean weather forecast
WNI Forecast Center
Weathernews Inc.

# Weathernews Inc.

- Headquarters in Chiba, Japan
- Founded in 1986
- Consolidated sales : about 160 M USD as of May 31st, 2025
- Number of employees: 1,120 as of May 31st, 2025
- **Our clients: shipping, railway, aviation, road management, energy, broadcasting, general consumers, etc.**

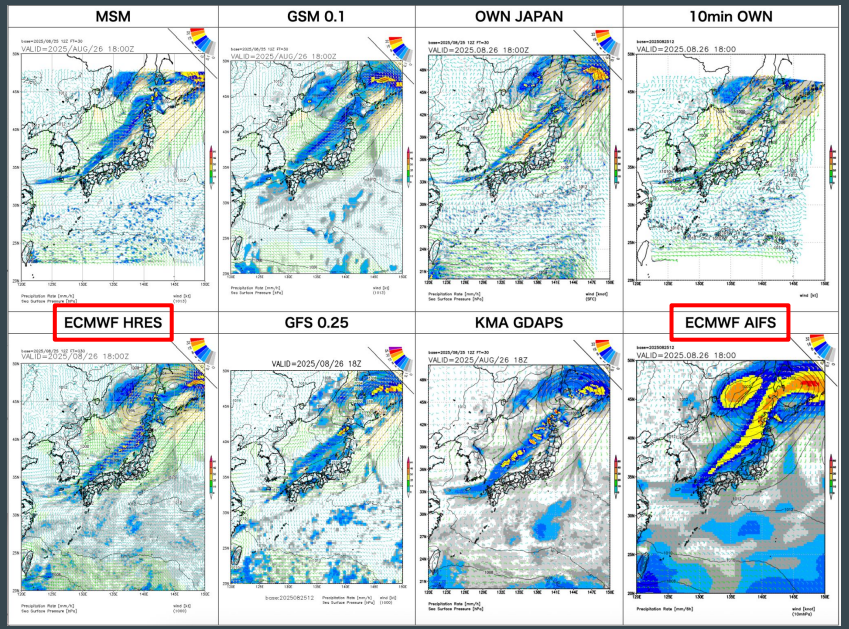



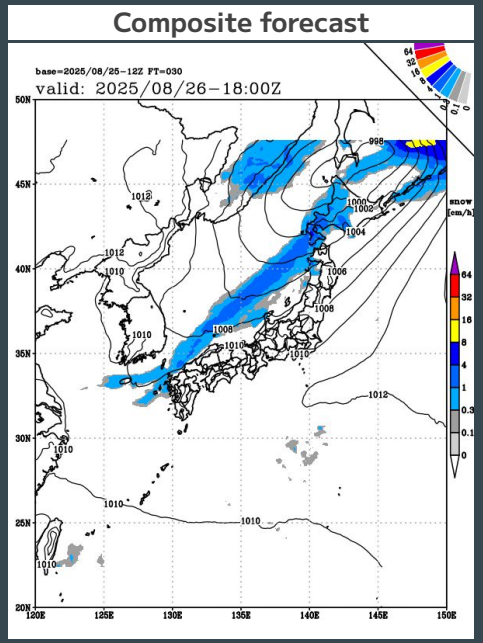

Ocean weather forecast, voyage route recommendation, vessel performance monitoring and forecast for shipping industry

Mobile weather app in Japan
42 M downloads

# Usage of ECMWF data in Weathernews

Our core forecasting system: creating accurate forecast by combining weather predictions gathered from around the world




AIFS is a candidate for future use in this system.

# Motivations and Contents of our Analysis

- Motivations:
  - As far as we know, there are only a few studies evaluating the accuracy of AIFS for extreme events.
  - For our services, the ability of the models to predict extreme events is most important, and we need to verify whether AI-based models can do it.
- Contents:
  - Case Study 1: Frontal Heavy Rain Across Japan in August
  - Case Study 2: Highest Temperature Ever Recorded in Japan
  - Accuracy of ERA5
  - Statistical Evaluation of Heavy Rain and High Temperature Forecast in Summer in Japan

# Models to be Compared

- **AIFS Single**

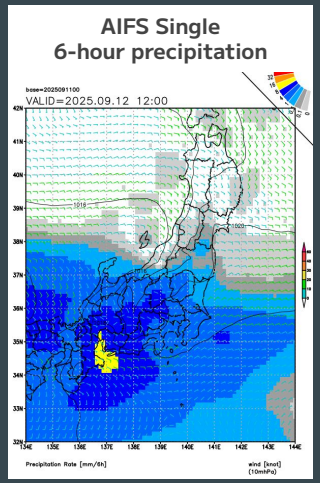
  ECMWF AI-based deterministic model which was trained to minimize Weighted MSE using ERA5 and operational analysis
- **AIFS ENS**

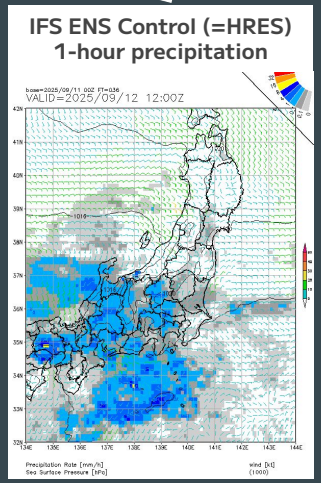
  ECMWF AI-based ensemble model which was trained to minimize CRPS using ERA5 and operational analysis
- **IFS ENS**

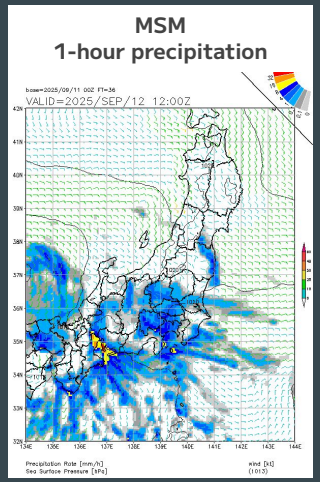
  ECMWF physical ensemble model whose spatial resolution is about 10 km. (Note: in IFS Cycle 49r1, IFS ENS Control and HRES forecast became identical.)
- **MSM**

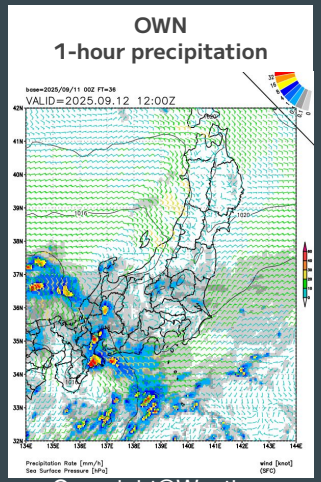
  Japan Meteorological Agency (=JMA) 5km mesoscale physical model.
- **OWN**

  Weathernews 5km mesoscale physical model made with WRF.



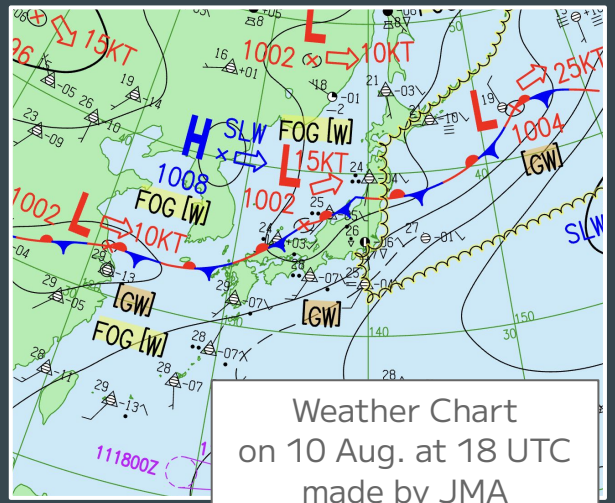



# Agenda

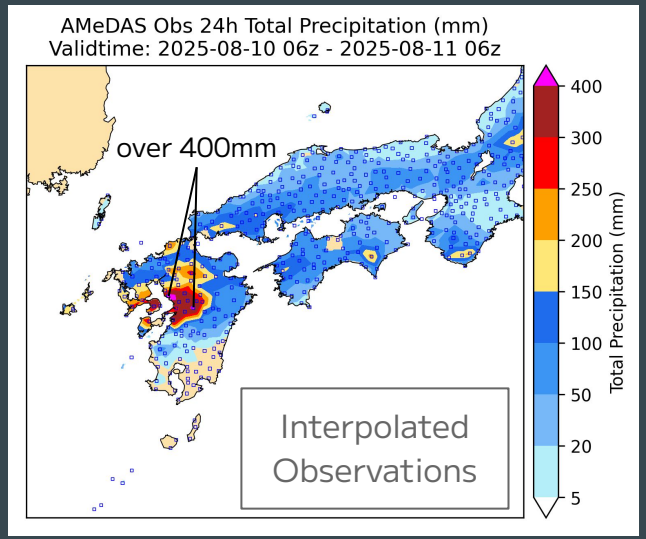
- Introduction
- **Case Study 1: Frontal Heavy Rain Across Japan in August**
- Case Study 2: Highest Temperature Ever Recorded in Japan
- Accuracy of ERA5
- Statistical Evaluation of Heavy Rain and High Temperature Forecast in Summer in Japan
- Summary and Future Work

# Case Study 1: Frontal Heavy Rain Across Japan in August

- What happened: Japan was hit by frontal heavy rain and 24-hour total precipitation exceeded 400 mm at some stations
- Date: August 6 to August 12, 2025
- Causes:
  - Stagnation of the front and low-pressure systems along the front
  - An atmospheric river from southeastern China to the Pacific was formed




Weather Chart
on 10 Aug. at 18 UTC
made by JMA

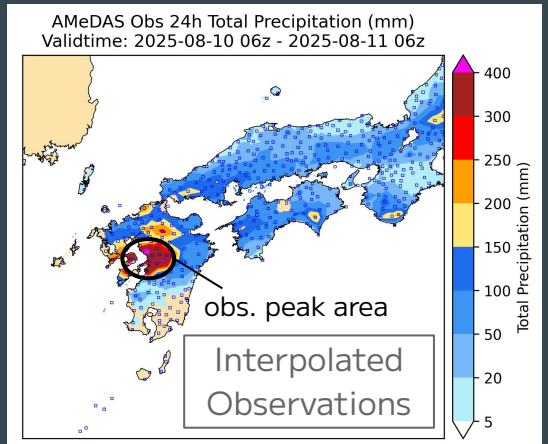


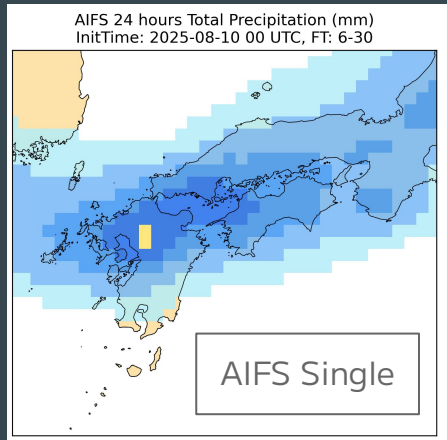

Interpolated
Observations

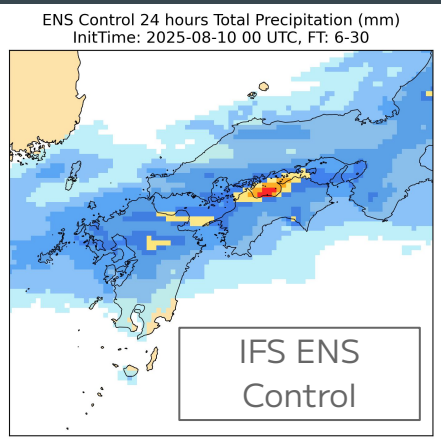
A picture taken by an our app. user

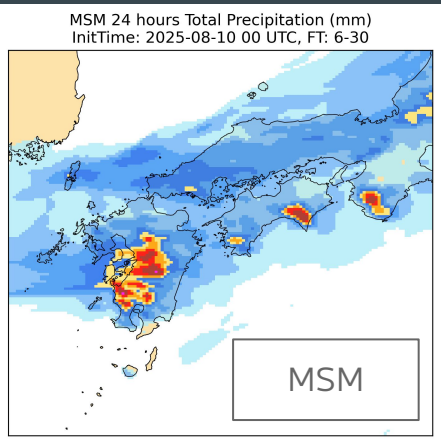
# Deterministic Models Comparison

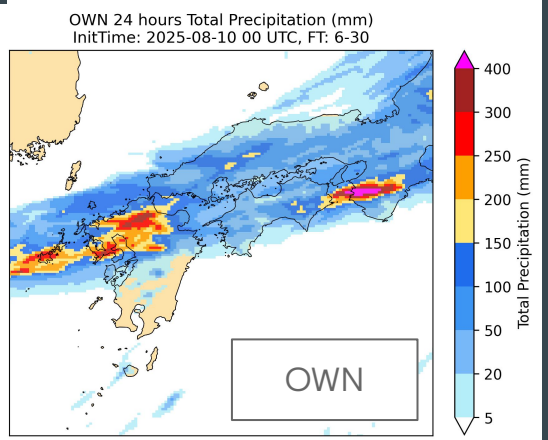
- Spatial peak location of AIFS Single forecast is accurate, but the rainfall intensity of the forecast is too low.
- Physical models are predicting heavier rainfall. Especially the mesoscale models show significant precipitation. However, it seems difficult for them to accurately predict the location of the peak of the rainfall.



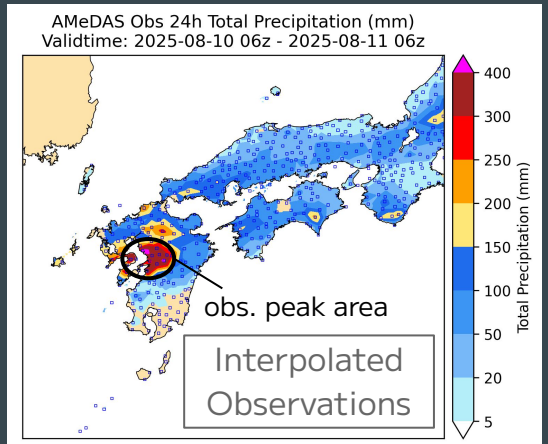



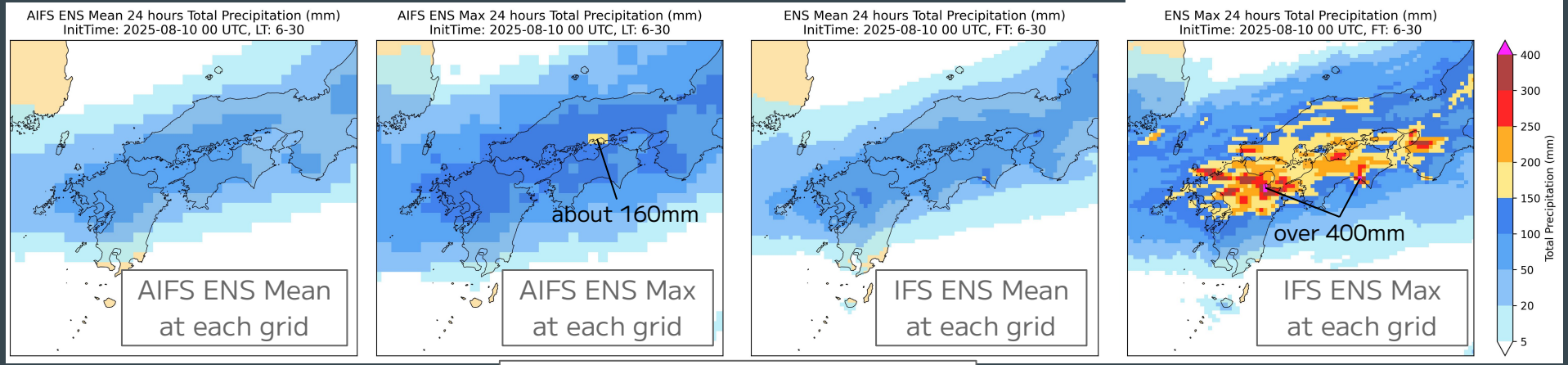




6 to 30 hours from initialized time

# Ensemble Models Comparison

- The rainfall predicted by AIFS ENS is too low. Even at its maximum, it's only about 160 mm/24h.
- Looking at IFS ENS Max, it can be seen that some members predict over 400 mm/24h at some grids, although the locations of the grids differ from the observations' peak.



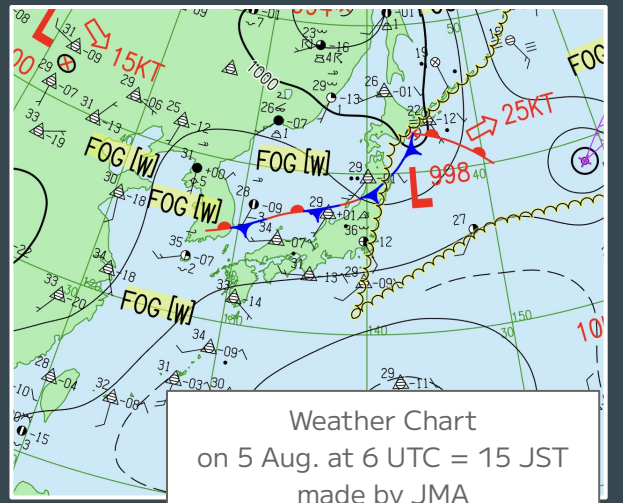

6 to 30 hours from initialized time

# Agenda

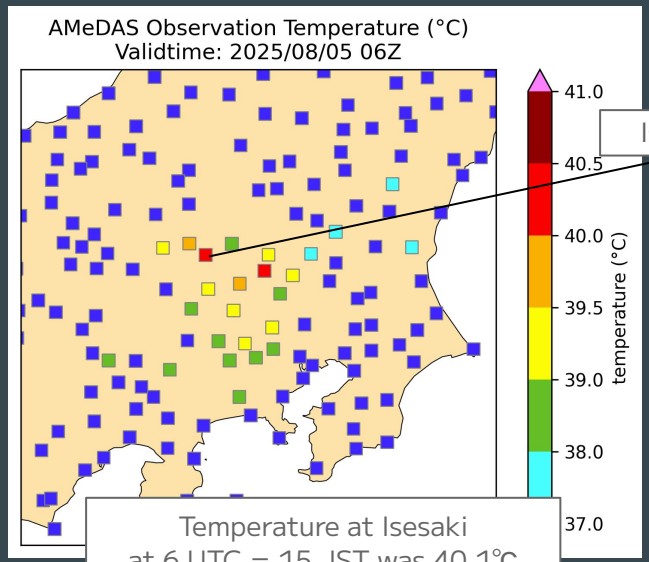
- Introduction
- Case Study 1: Frontal Heavy Rain Across Japan in August
- **Case Study 2: Highest Temperature Ever Recorded in Japan**
- Accuracy of ERA5
- Statistical Evaluation of Heavy Rain and High Temperature Forecast in Summer in Japan
- Summary and Future Work

# Case Study 2: Highest Temperature Ever Recorded in Japan

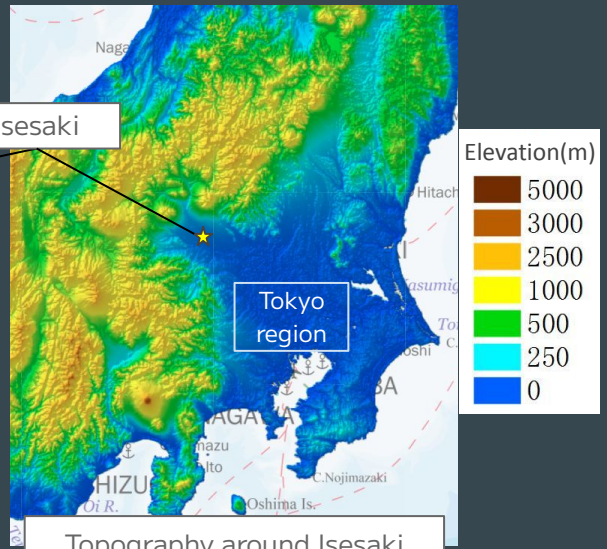
- What happened: A record high temperature of 41.8℃ was recorded
- Location: AMeDAS (=JMA's observation network) Isesaki
- Date: August 5, 2025
- Causes: Foehn-type downslope westerly winds parallel to the front and solar heating




Weather Chart
on 5 Aug. at 6 UTC = 15 JST
made by JMA



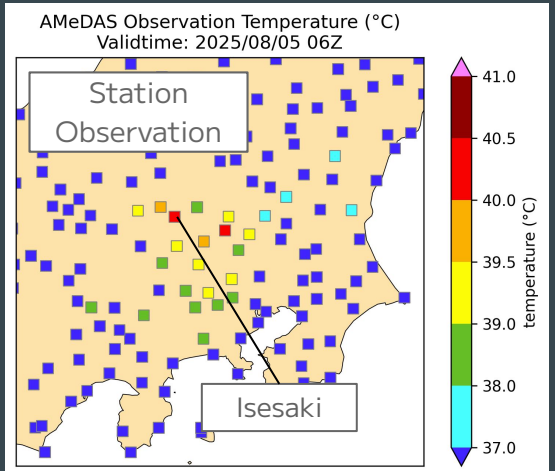

Temperature at Isesaki
at 6 UTC = 15 JST was 40.1℃

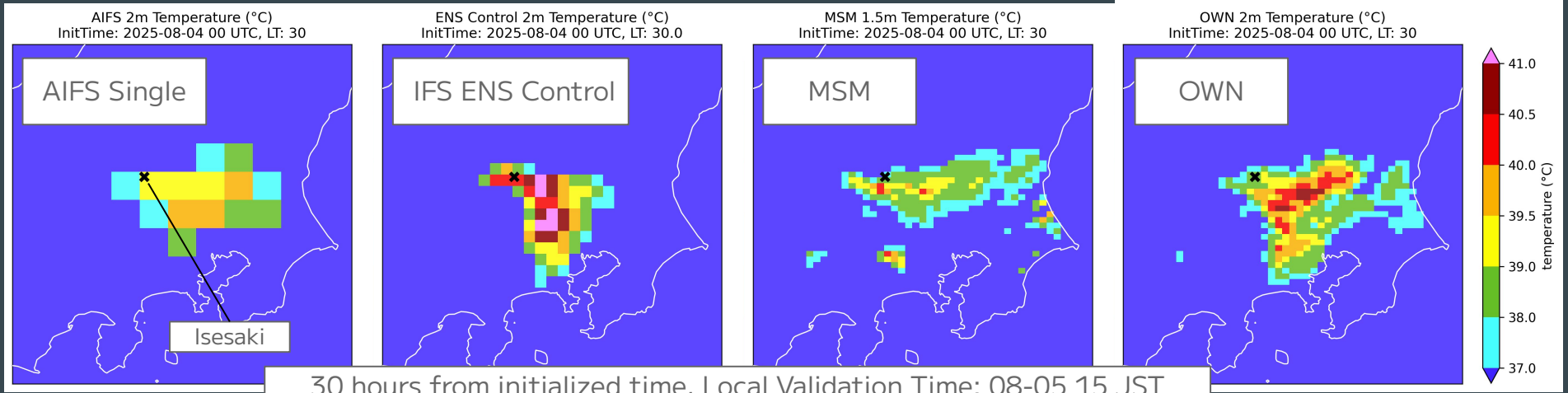



Topography around Isesaki

# Deterministic Models Comparison

- The accuracy of AIFS Single forecast seems high, although the highest temperature of the forecast is slightly lower than that of observations.
- The IFS ENS Control temperature forecast is high enough, but it appears to be slightly overestimating temperatures in some areas.
- No clear advantage of the mesoscale models can be seen.

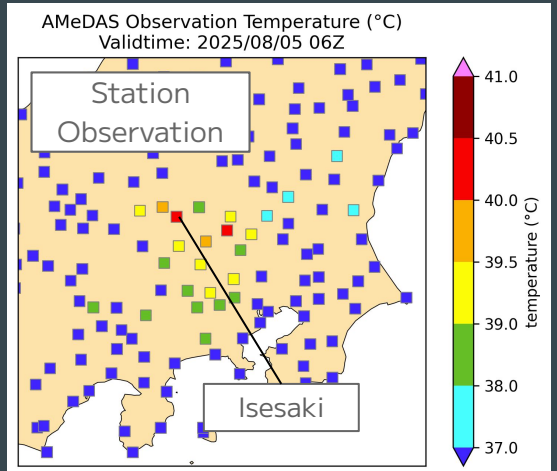


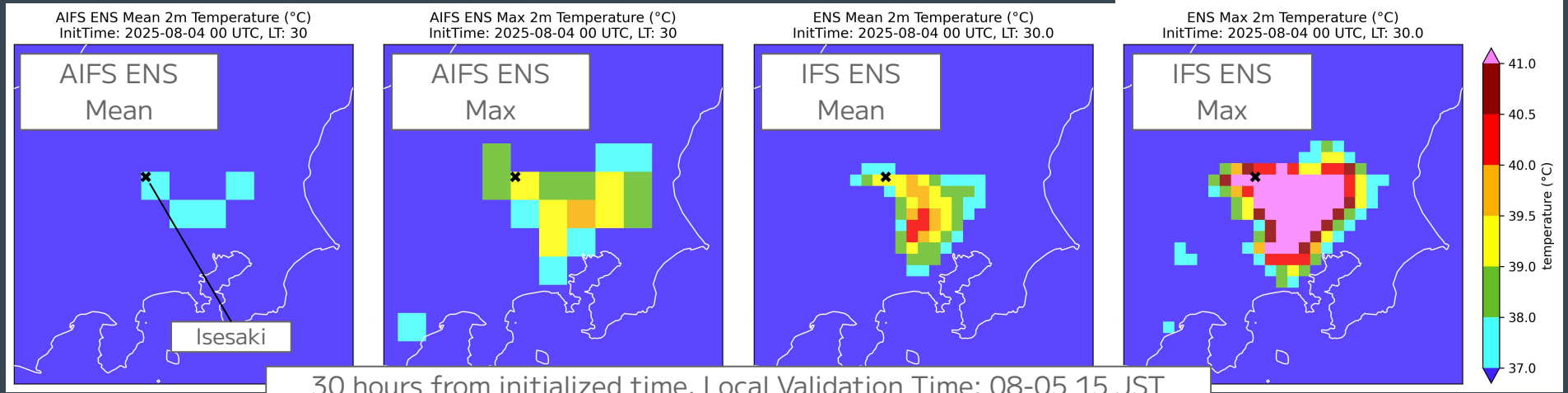




30 hours from initialized time, Local Validation Time: 08-05 15 JST

# Ensemble Models Comparison

- The temperatures predicted by AIFS ENS is too low. For example, in AIFS ENS Mean, there are no grids exceeding 38℃.
- IFS ENS Mean forecast seems to be accurate. Looking at IFS ENS Max, it can be seen that IFS ENS can output sufficiently high temperatures.



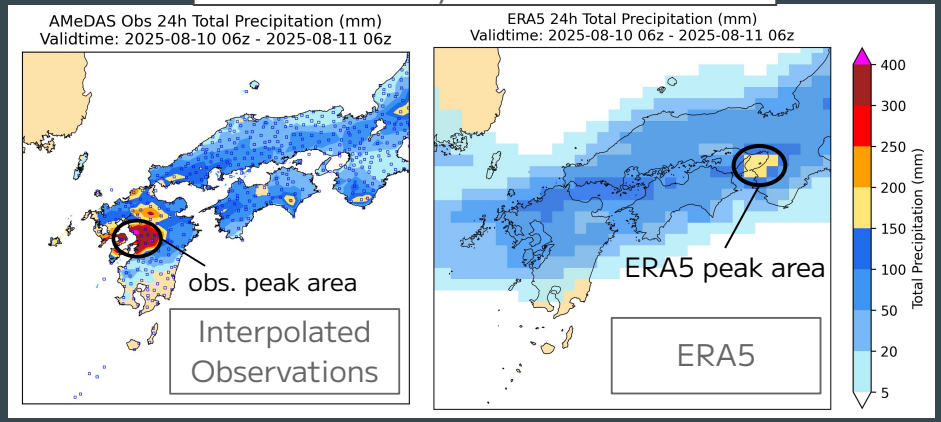
# Agenda

- Introduction
- Case Study 1: Frontal Heavy Rain Across Japan in August
- Case Study 2: Highest Temperature Ever Recorded in Japan
- **Accuracy of ERA5**
- Statistical Evaluation of Heavy Rain and High Temperature Forecast in Summer in Japan
- Summary and Future Work

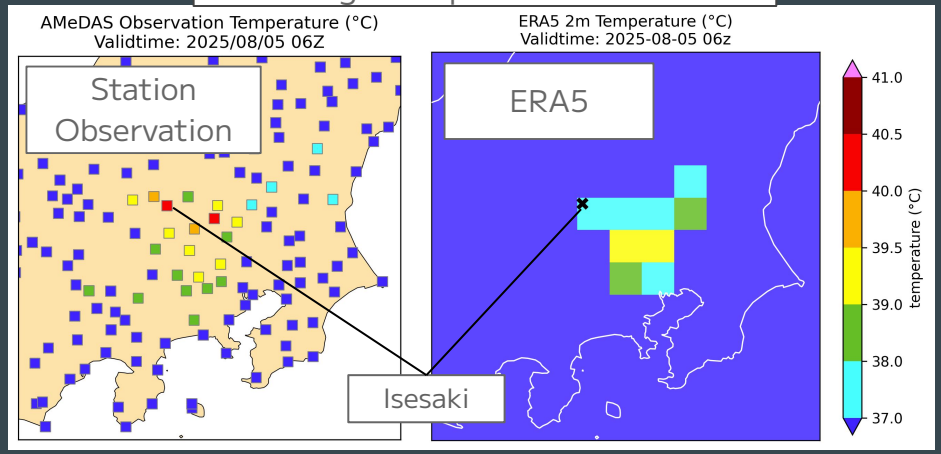
# Accuracy of ERA5

- ERA5 = AIFS's pre-training data
- For the heavy rain case, the rainfall peak location of ERA5 is not accurate and the rainfall intensity of ERA5 is too low.
- For the high temperature case, EAR5's temperature appears to be a bit lower than observed values.
- ERA5's accuracy for extreme events in Japan doesn't seem to be very accurate.

## Heavy Rain Case




## High Temperature Case

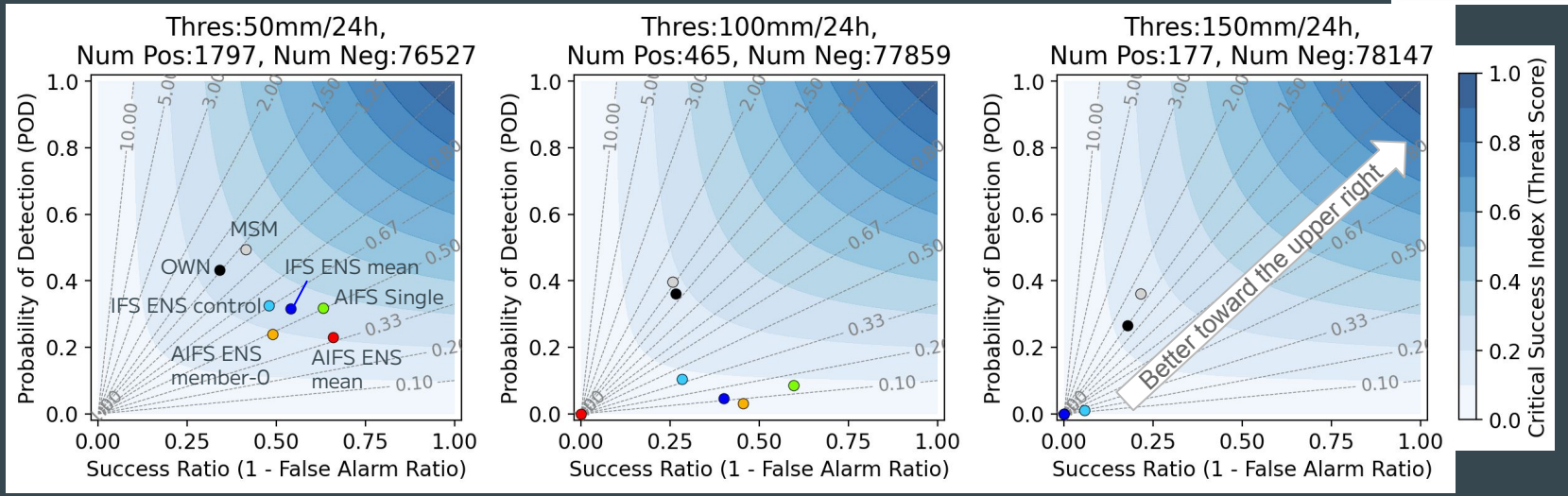


# Agenda

- Introduction
- Case Study 1: Frontal Heavy Rain Across Japan in August
- Case Study 2: Highest Temperature Ever Recorded in Japan
- Accuracy of ERA5
- **Statistical Evaluation of Heavy Rain and High Temperature Forecast in Summer in Japan**
- Summary and Future Work

# Roebber's Performance Diagram for Heavy Rain

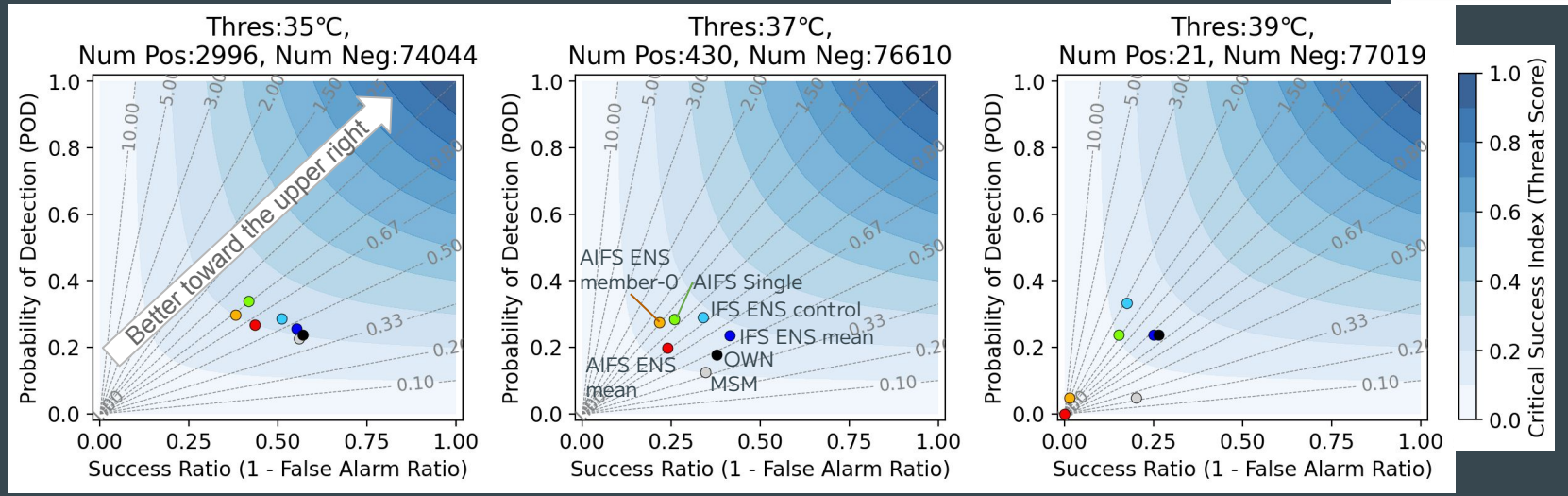
- Performance of 24h precipitation forecast in 2 months of July and Aug 2025
- MSM has the highest CSI for all the three thresholds.
- IFS and AIFS models are almost incapable of predicting heavy rain events of 150mm/24h.
- Because of its higher SR, AIFS Single has slightly higher CSI compared to IFS ENS Control and Mean.
- AIFS ENS has lower CSI than IFS ENS.




Data Period: 2 July to 31 August, Initialised Time: 00 UTC each day, Stations: AMeDAS (about 1,300 stations all around Japan), Lead Time: 12 to 36 hours from initialized time

# Roebber's Performance Diagram for High Temperature

- Performance of high temperature forecast in 2 months of July and Aug 2025
- The mesoscale models don't show a significant advantage over AIFS and IFS models.
- For higher temperature (37℃ and 39℃) forecasts, AIFS Single has slightly lower CSI than IFS ENS Control and Mean.
- AIFS ENS has lower CSI than IFS ENS.




Data Period: 2 July to 31 August, Initialised Time: 00 UTC each day, Stations: AMeDAS (about 840 stations all around Japan),
Lead Time: 30 hours from initialized time, Local Validation Time: 15 JST each day

# Agenda

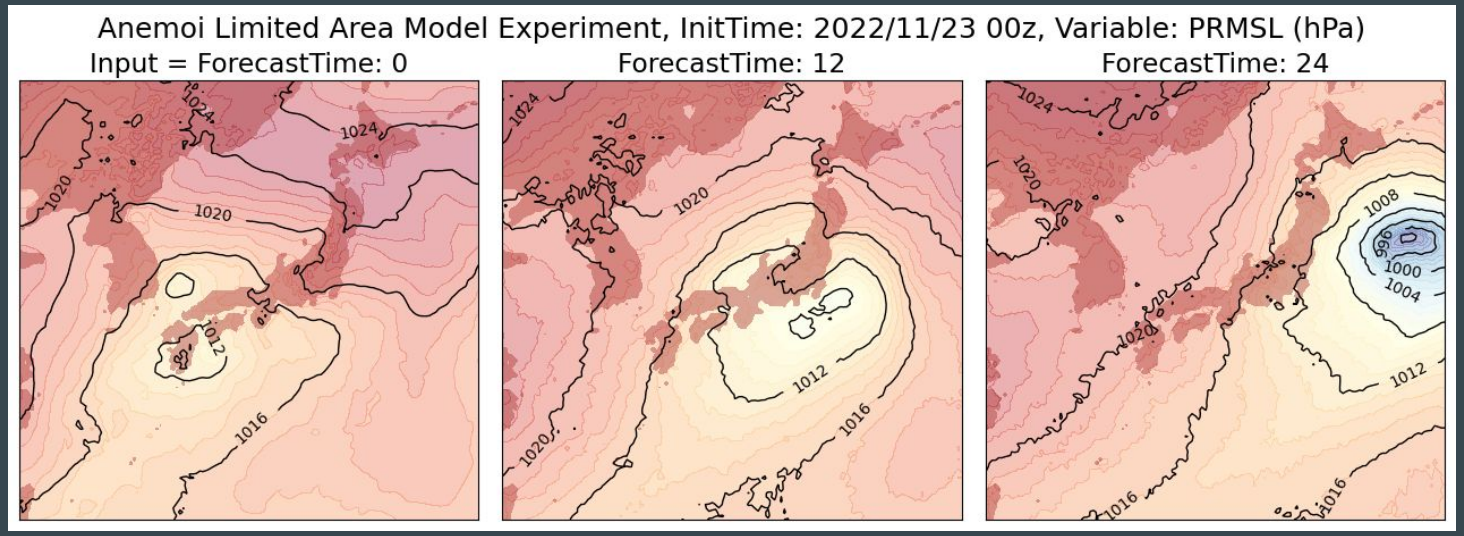
- Introduction
- Case Study 1: Frontal Heavy Rain Across Japan in August
- Case Study 2: Highest Temperature Ever Recorded in Japan
- Accuracy of ERA5
- Statistical Evaluation of Heavy Rain and High Temperature Forecast in Summer in Japan
- **Summary and Future Work**

# Summary

- As for deterministic forecast,
  - AIFS Single shows results that are comparable to IFS ENS Control (=HRES) and IFS ENS Mean for both heavy rain forecasts and high temperature forecasts.
  - However, for events with 24-hour rainfall amounts reaching several hundred millimeters, AIFS and IFS rainfall amounts were too low.
- As for ensemble forecast, AIFS ENS has lower capability for forecasting extreme phenomena than IFS ENS. (Further investigation is needed.)
- To create an AI weather model capable of predicting disaster-level events in Japan, ERA5 may be inadequate; training data with higher resolution and higher accuracy would be required.

# Future Work: AI-based High Resolution Model

- Creating AI weather models with frameworks like Anemoi, using higher-resolution regional analysis data as training data.
- Investigating which loss functions are suitable for which tasks.



Thank you for your attention.

weathernews