Hybrid NWP-Machine Learning or End-to-End ML?

Massimo Bonavita

Principal Scientist

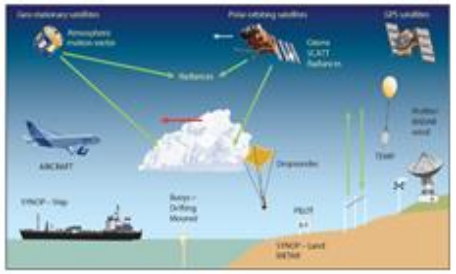
Data Assimilation Team Leader

ECMWF

massimo.bonavita@ecmwf.int

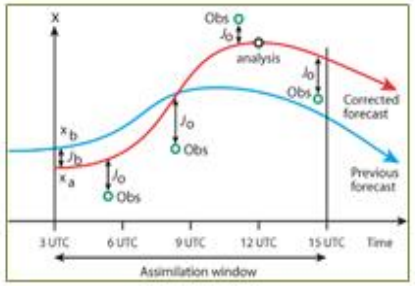
ECMWF

The age of innocence…



Observations

- quality control
- adaptive thinning
- bias correction
- Fast observation models

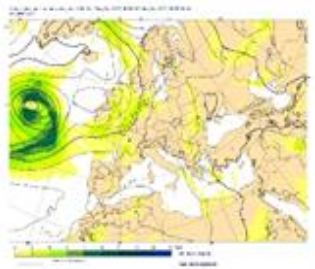


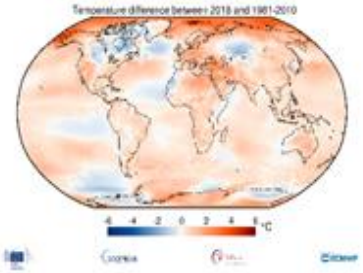
Data assimilation

- Model error estimation
- Model parameter estimation
- Error covariance statistics

Forecast Models

- Surrogate models
- Fast emulator models
- Linearised models for Var





Product generation

- information extraction (resolution, ensembles, weather features)
- data compression

Bonavita et al., 2021

A brave New World: Machine Learning Weather Prediction

- ML models for medium/extended-range weather prediction, mostly trained on ERA5 reanalysis
- Field started with Dueben and Bauer, 2018, attempt at forecasting low-resolution Z500 field as an image-to-image problem, results not too exciting
- Turning point: Keisler, 2022, multiple vertical levels (13), higher resolution (1deg), Graph NN. Results comparable to GFS, ECMWF
- Floodgates open: FourCastNet (NVIDIA, Pathak et al., 2022), Pangu-weather (Huawei, Bi et al., 2022), GraphCast (Google-DeepMind, Lam et al., 2022), FengWu (Academic, Chen et al., 2023)
- Each claims to outperform all previous MLWP model and all traditional physics-based NWP systems, and at a fraction of the cost!!

NVIDIA

HUAWEI

Microsoft

A look under the hood of MLWP models

- Some of these DLWP models are accessible on public repositories
- This allows to test some of the claims in the literature and explore the characteristics of their NWP output
- ECMWF runs Pangu-Weather, FourCastNet and GraphCast on a daily basis
- Today we will look inside Pangu-Weather…

A look under the hood of DLWP models: Pangu-Weather

- Pangu-Weather is based on a variation of Transformer ML Model (e.g., ChatGPT) adapted to computer vision applications
- Trained on 43 years of ERA5 re-analyses, 13 pressure-levels (z/t/q/u/v) plus surface variables (T2m,10u/v,mslp)
- L1-type loss function on forecast errors
- *"Hierarchical temporal aggregation"*: four individual models for 1-hour, 3-hour, 6-hour, and 24-hour prediction, combined to provide forecasts at any hourly fcst range

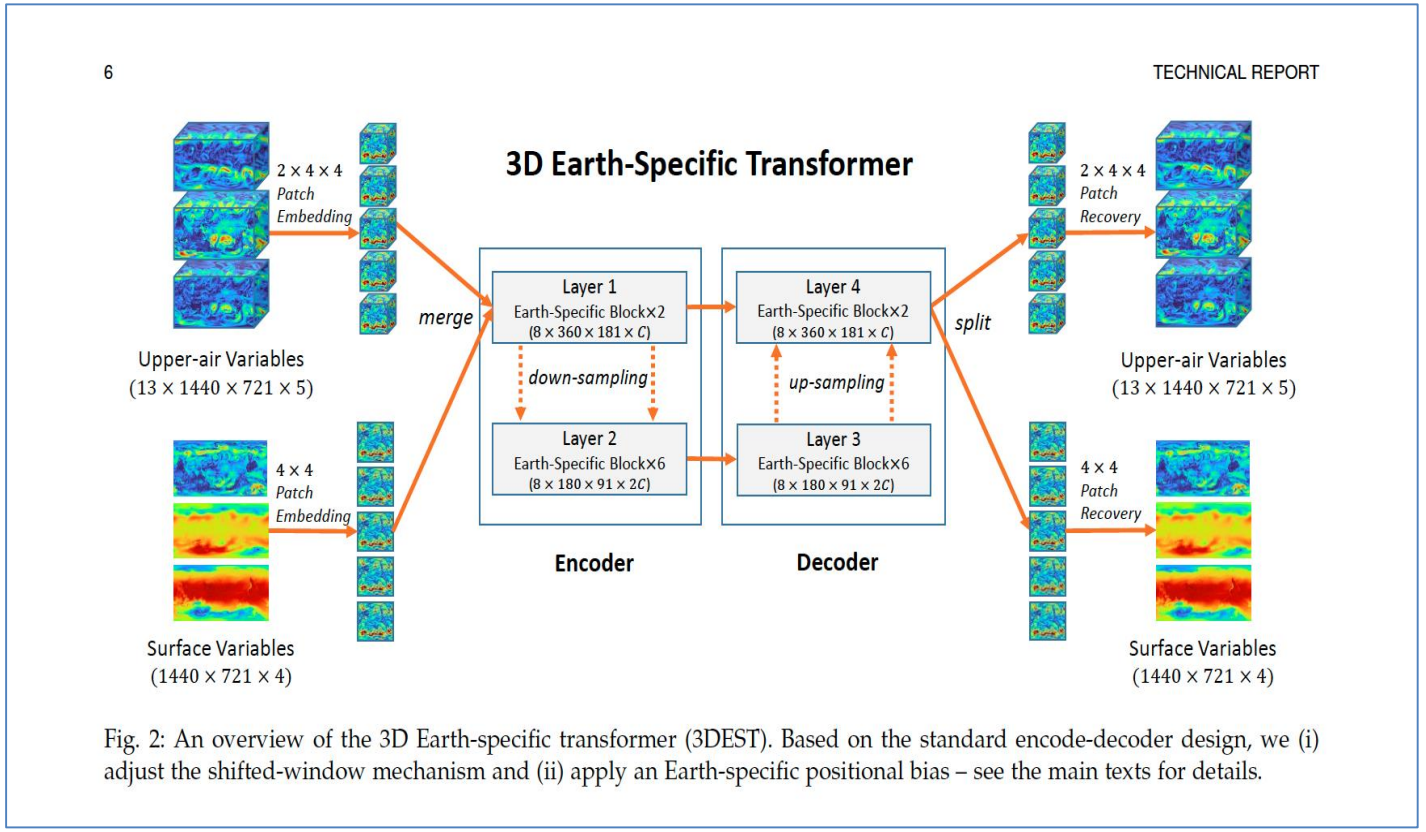


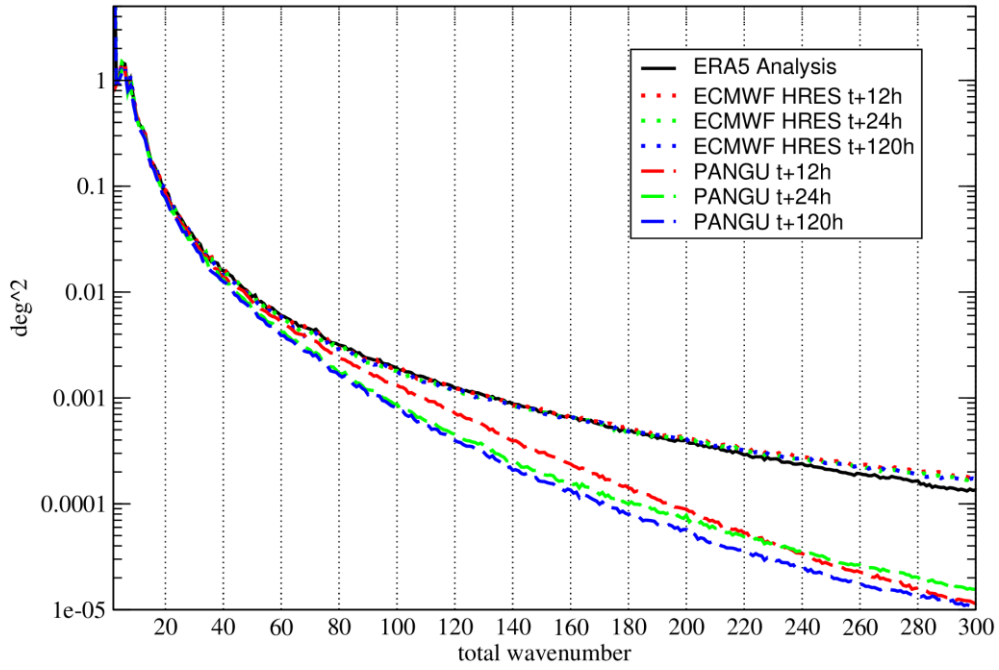
Fig. 2: An overview of the 3D Earth-specific transformer (3DEST). Based on the standard encode-decoder design, we (i) adjust the shifted-window mechanism and (ii) apply an Earth-specific positional bias – see the main texts for details.

from Bi et al., 2022

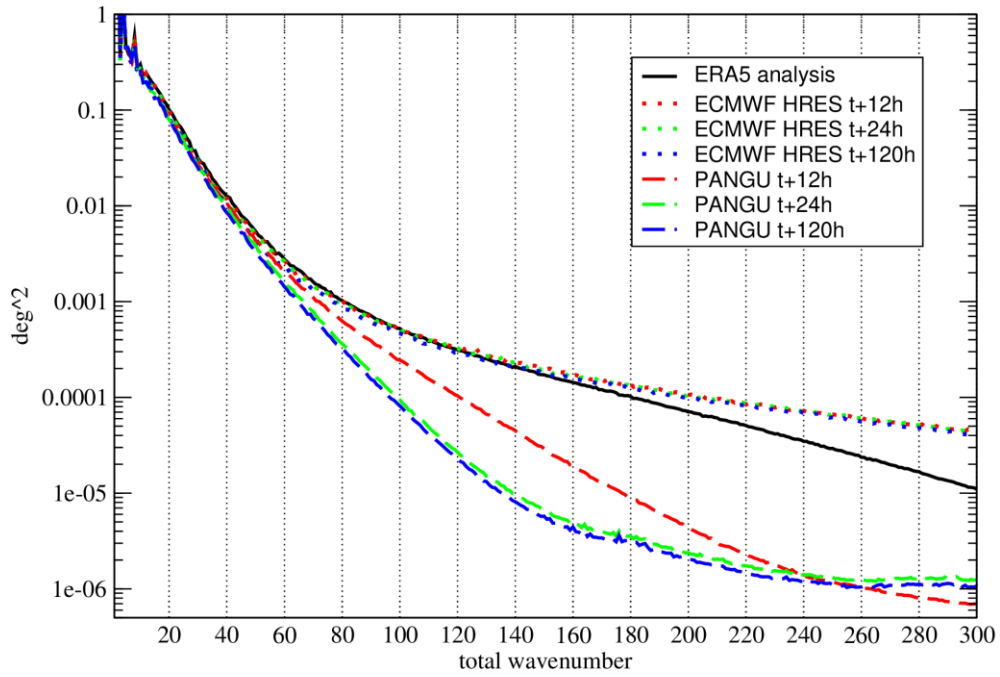
A look under the hood of DLWP models: Pangu-Weather

- Is Pangu-Weather a realistic emulator of the atmosphere?
- Let us start with a look at forecast energy spectra:

temperature @850hPa



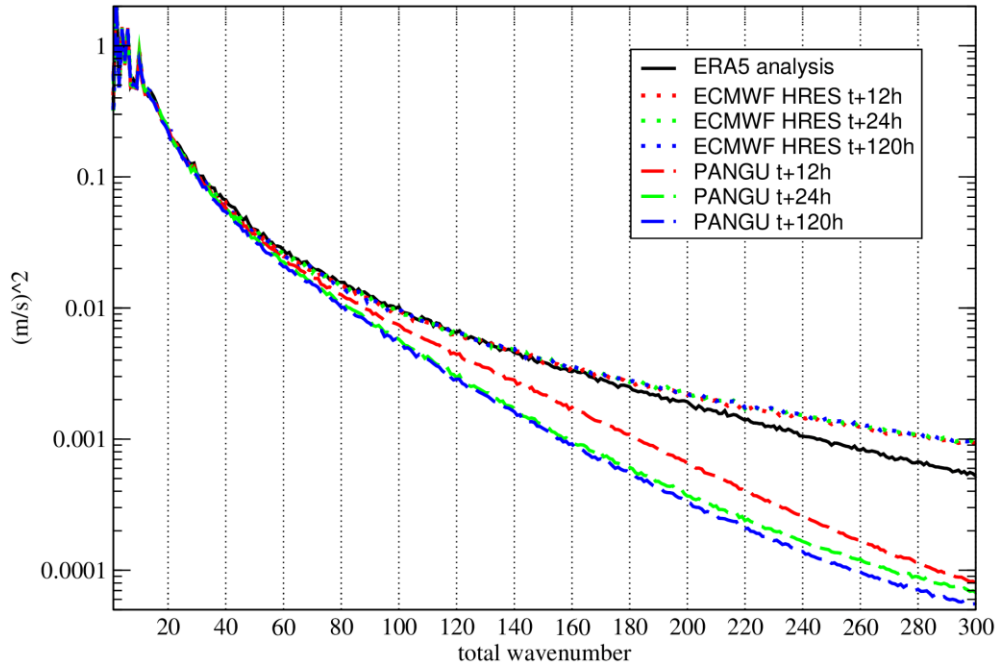
temperature @250hPa



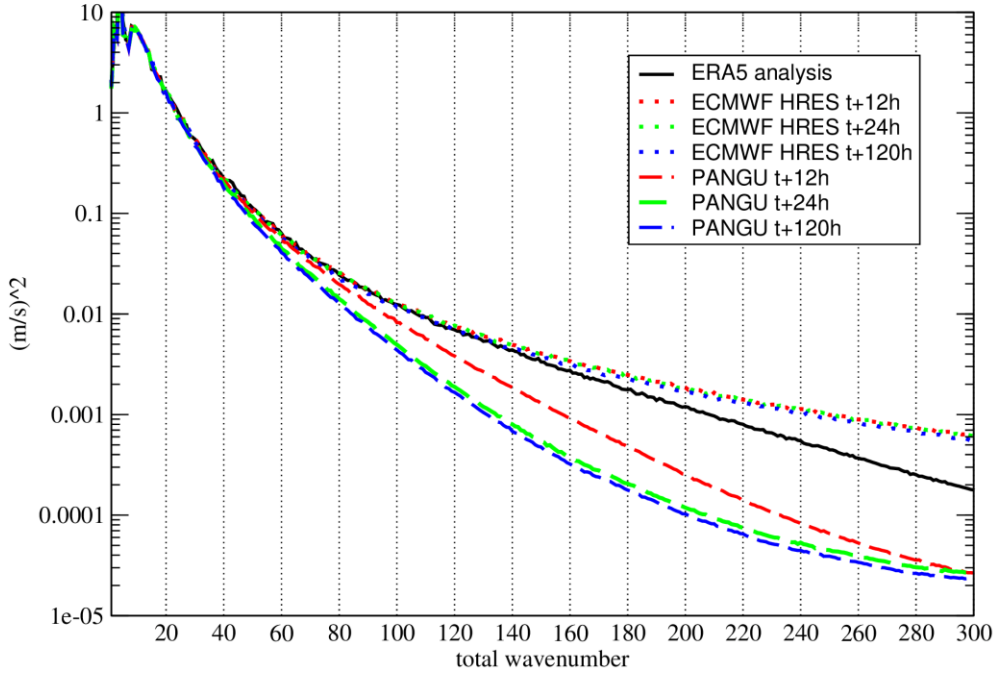
A look under the hood of DLWP models: Pangu-Weather

- Is Pangu-Weather a realistic emulator of the atmosphere?
- Let us start with a look at forecast energy spectra:

Wind speed @850hPa



Wind speed @250hPa



A look under the hood of DLWP models: Pangu-Weather

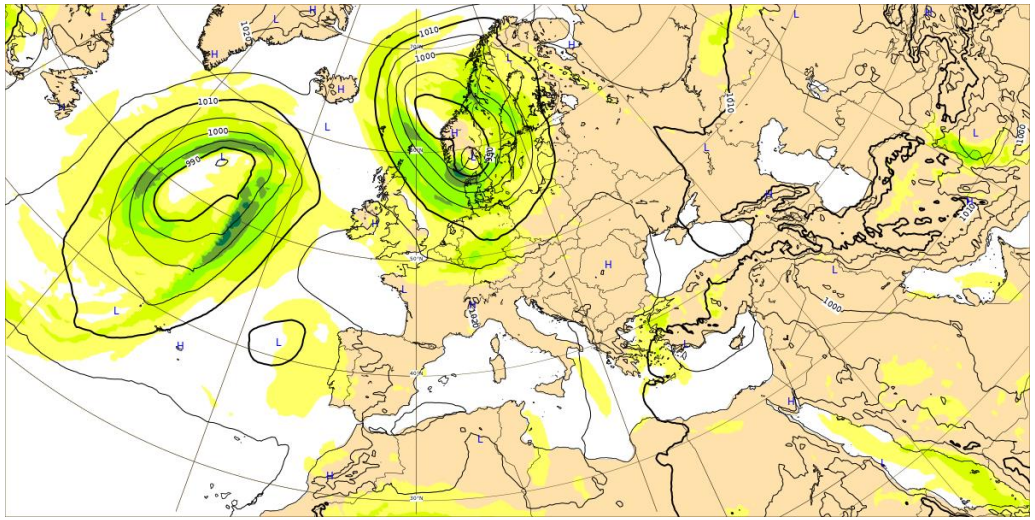
- Pangu-Weather forecast energy spectra are significantly less active than ERA5 analyses (and ECMWF operational forecasts)
- Pangu-Weather is under-active at all wavenumbers, but spatial scales smaller than ~400-500 km are heavily suppressed
- The activity of the Pangu-Weather forecasts decreases progressively with forecast lead time

- Does it matter?

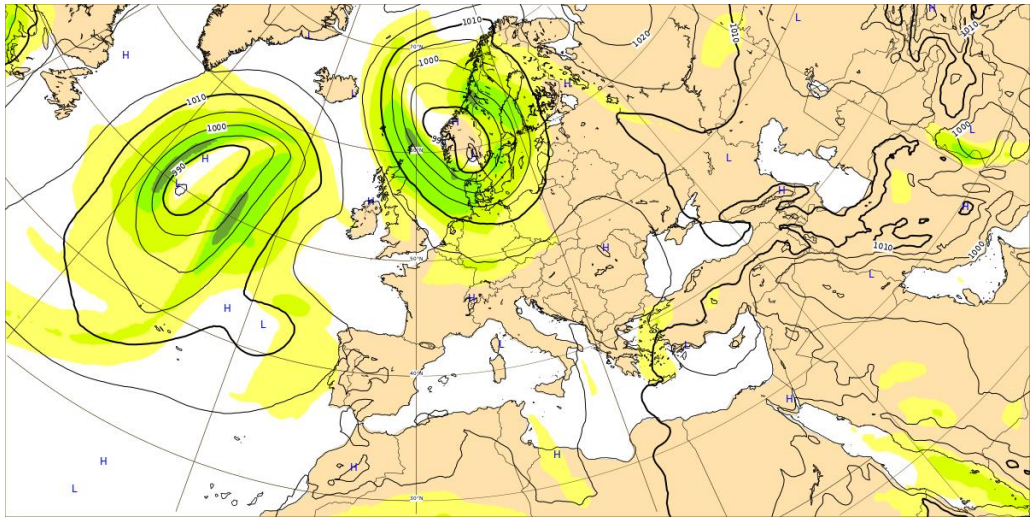
A look under the hood of DLWP models: Pangu-Weather

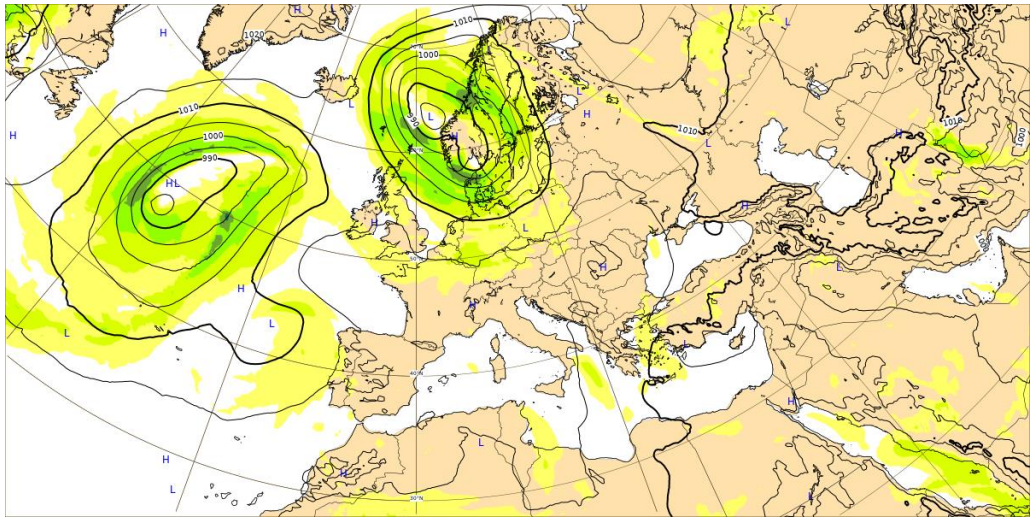
- If one looks at standard synoptic maps at continental scales, (almost) all seems fine!

ECMWF HRES mslp & windspeed @850 t+72h



Pangu-w mslp & windspeed @850 t+72h



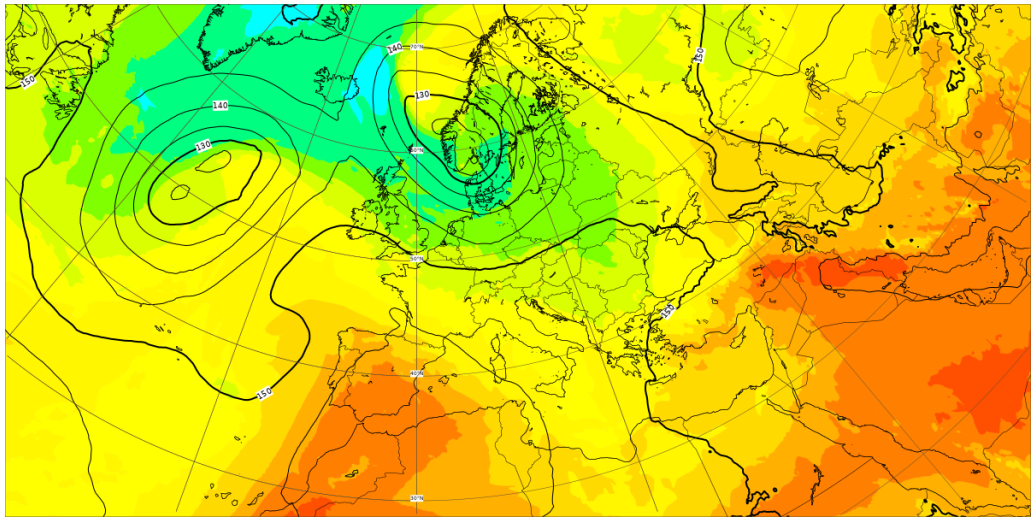


ECMWF analysis 09 Aug 2023 00Z

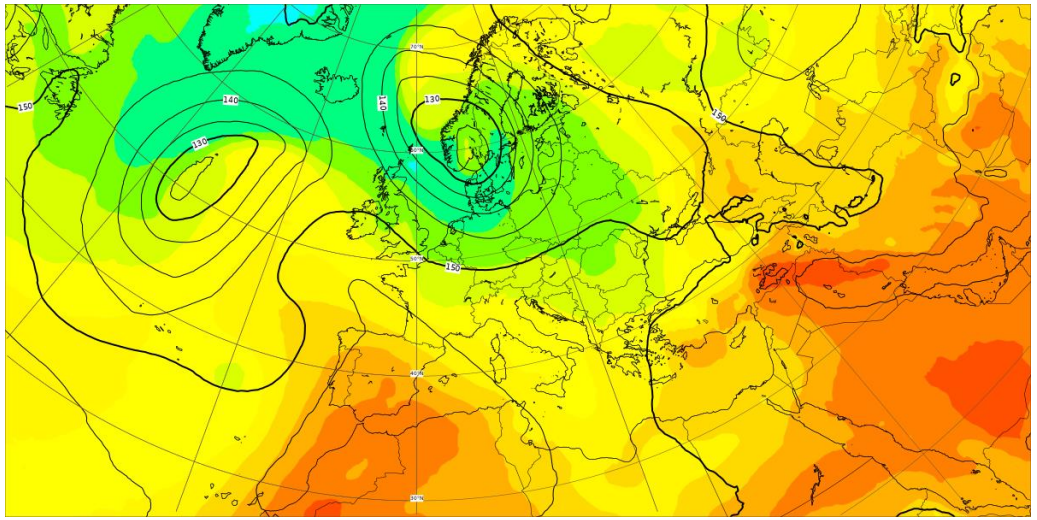
A look under the hood of DLWP models: Pangu-Weather

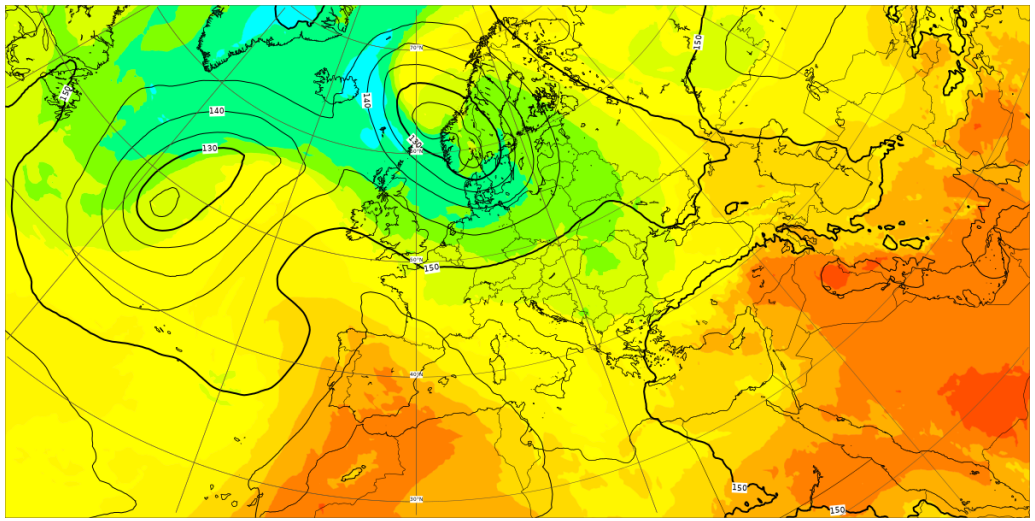
- If one looks at standard synoptic maps at continental scales, (almost) all seems fine!

ECMWF HRES T & GH @850 t+72h



Pangu-w T & GH @850 t+72h

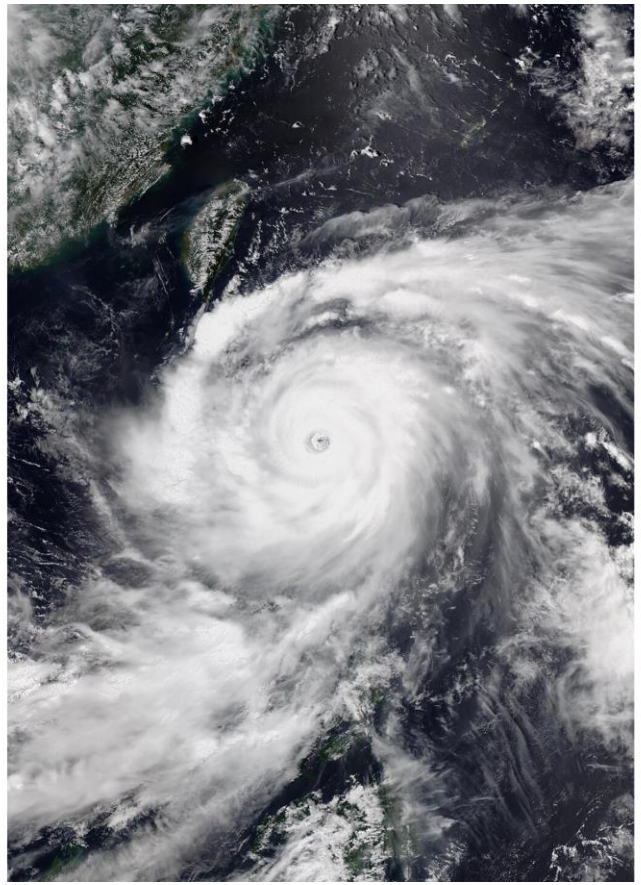




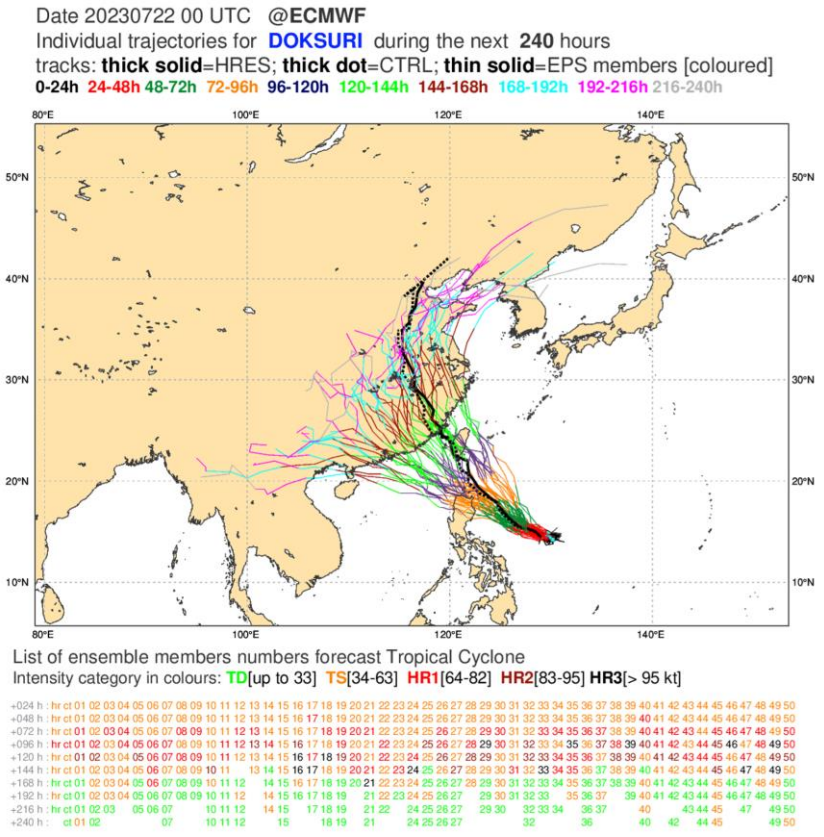
ECMWF analysis 09 Aug 2023 00Z

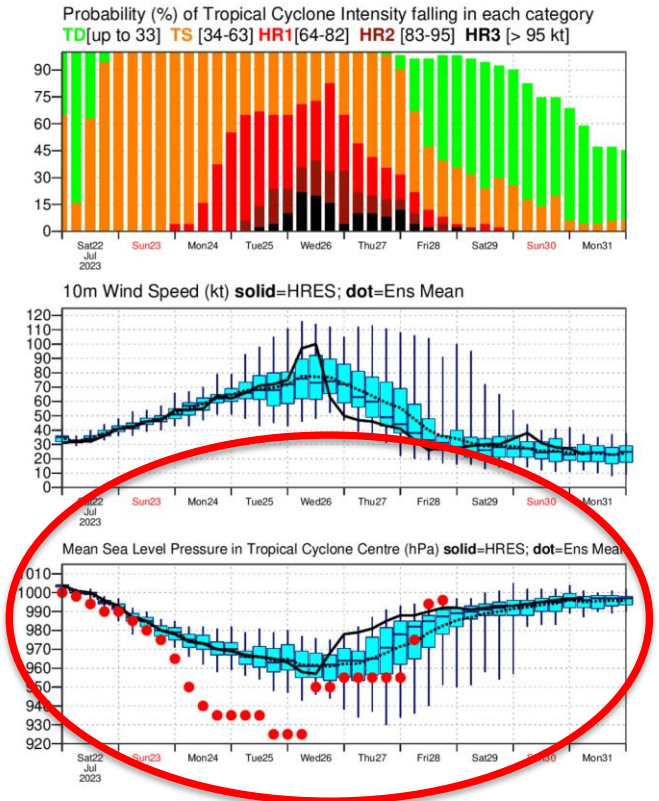
A look under the hood of DLWP models: Pangu-Weather

- It is a different story if one looks at high impact weather events where resolution matters…



Typhoon Doksuri (Egay) of the 2023 Pacific typhoon season near its peak intensity while off the coast of Luzon during the afternoon of July 25, 2023. It had 10-min sustained winds of 175 km/h (110 mph) (JMA) and 1-min sustained winds of 230 km/h (145 mph) (JTWC) and an official minimum central pressure of 935 mbar (27.6 inHg) at the time this image was captured.

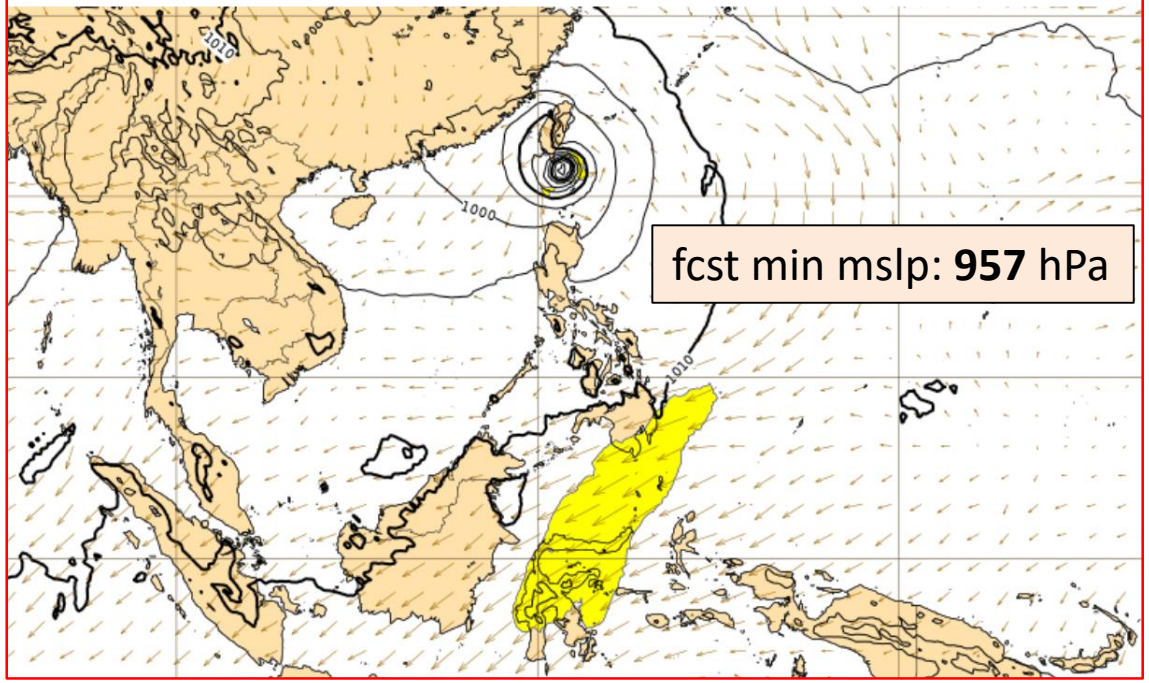




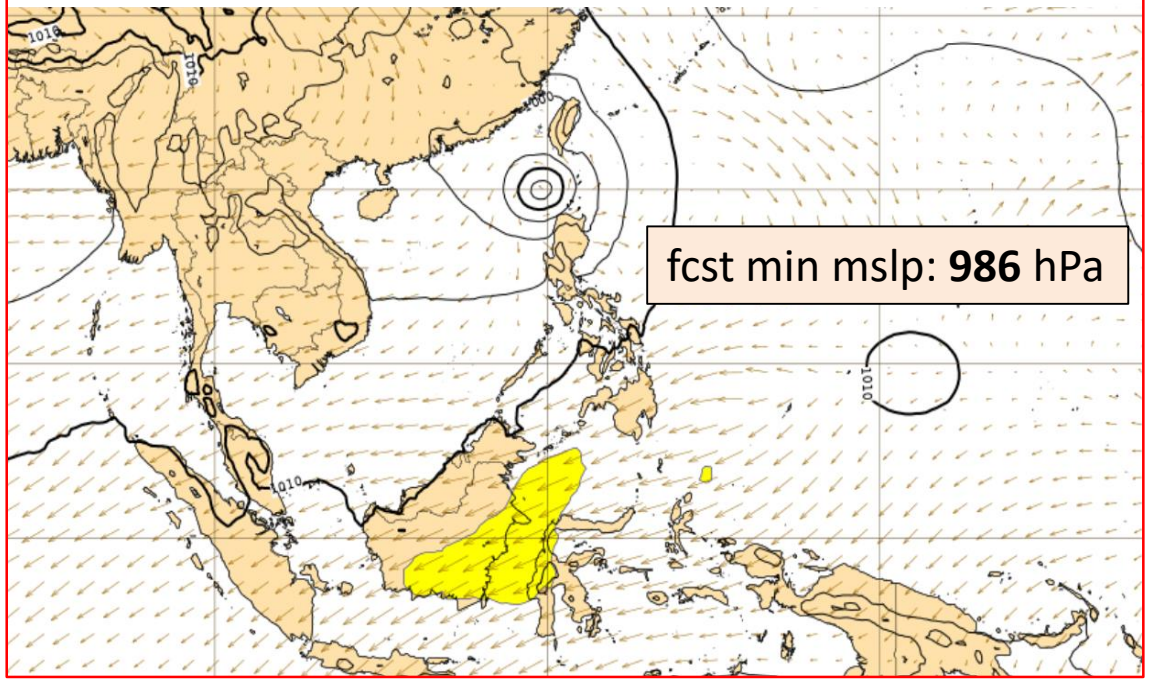
A look under the hood of DLWP models: Pangu-Weather

TC Doksuri 26 Jul 2023 12Z
Estimated Best Track min mslp: **944** hPa

ECMWF HRES t+132h valid 26/07/23 12Z

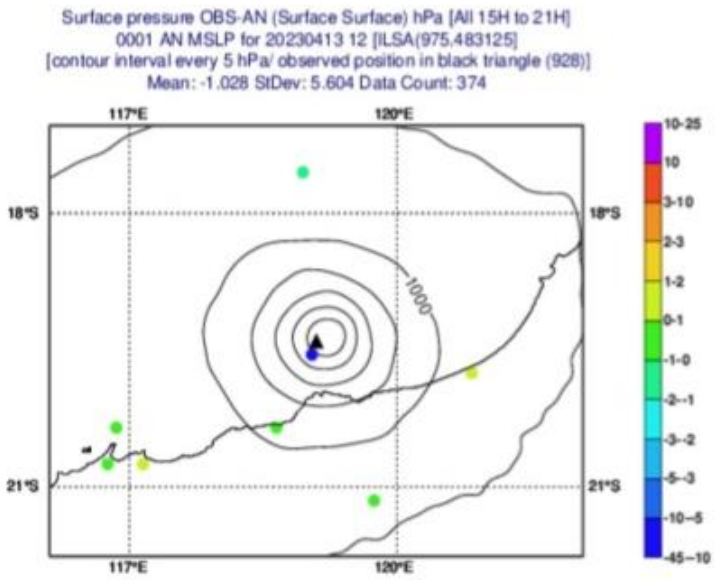


Pangu-weather t+132h valid 26/07/23 12Z

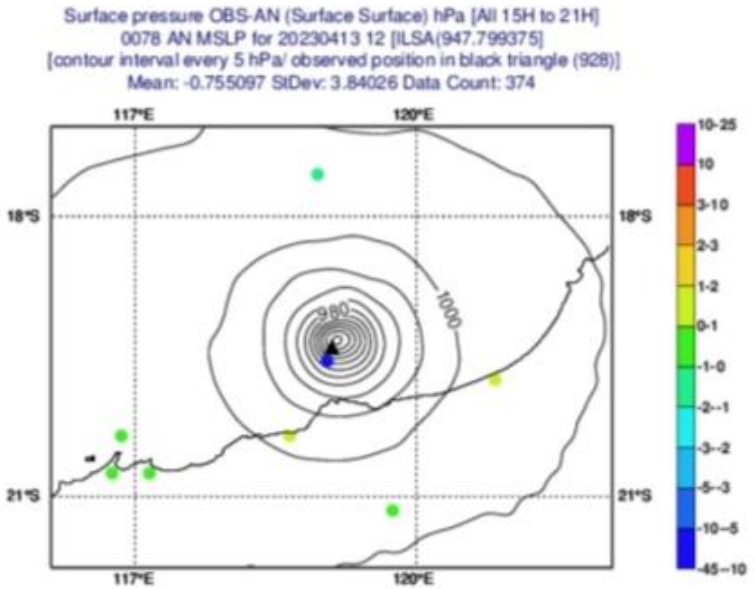


On the importance of resolution…

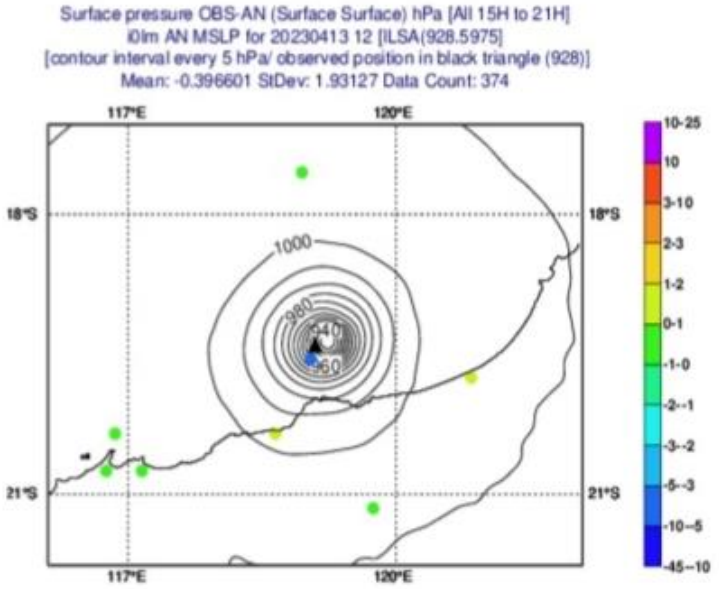
Tropical Cyclone ISLA – 13/04/2023 12UTC
Best Track estimated min MSLP **928** hPa



Oper 47r3 (**TL399**): **975.5** hPa



48r1 e-suite (TL551): **947.8** hPa



4DVar **TCo511**: **928.6** hPa

Zaplotnik, Bonavita and Holm, 2023

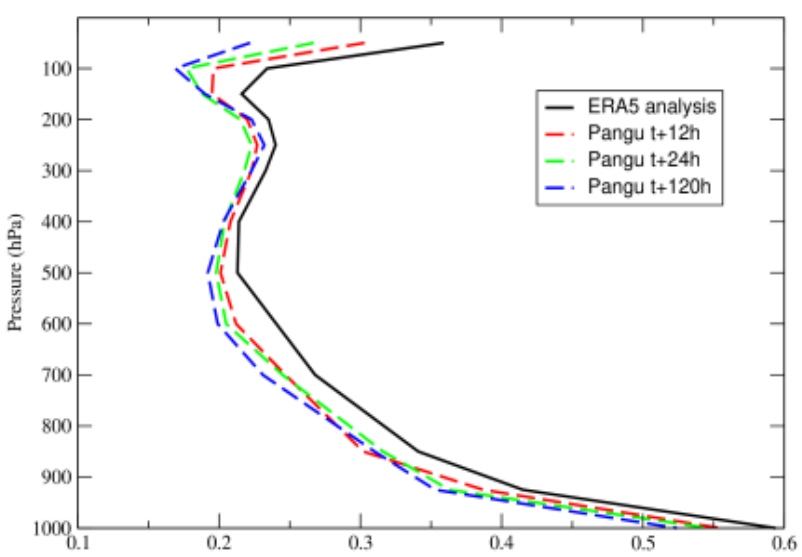
A look under the hood of DLWP models: Pangu-Weather

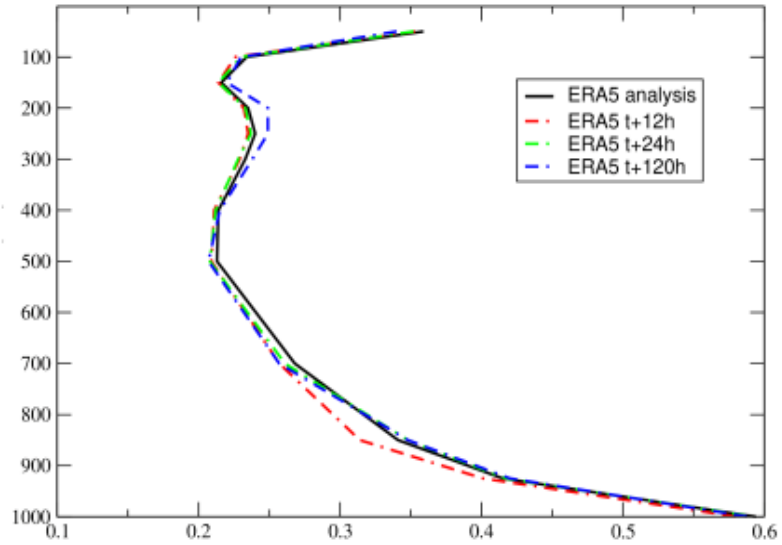
Unrealistic forecast energy spectra imply dynamically inconsistent forecast fields

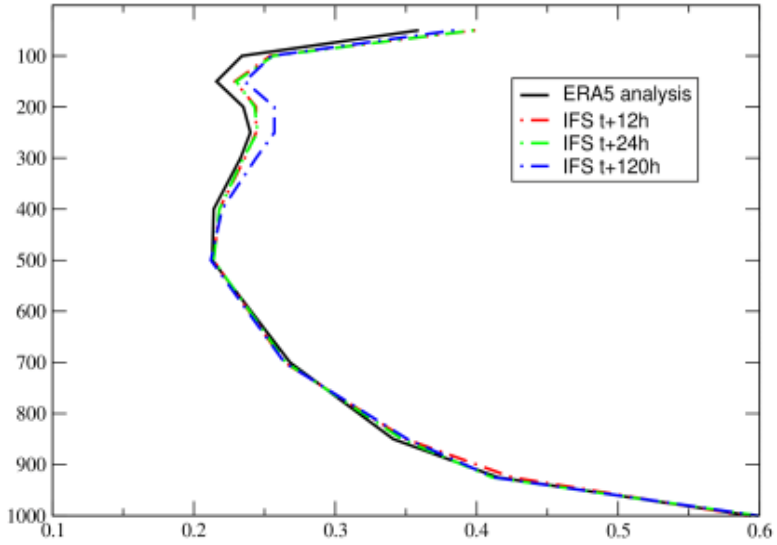
Pangu-Weather dynamical fields (1)

Geostrophic wind ($\mathbf{V}_g = \frac{1}{f}\hat{\mathbf{k}} \times \nabla_p \Phi$) vs ageostrophic wind $\mathbf{V}_{ag} \equiv \mathbf{V} - \mathbf{V}_g$

$|\mathbf{V}_{ag}| / |\mathbf{V}_g|$







Pangu

ERA5

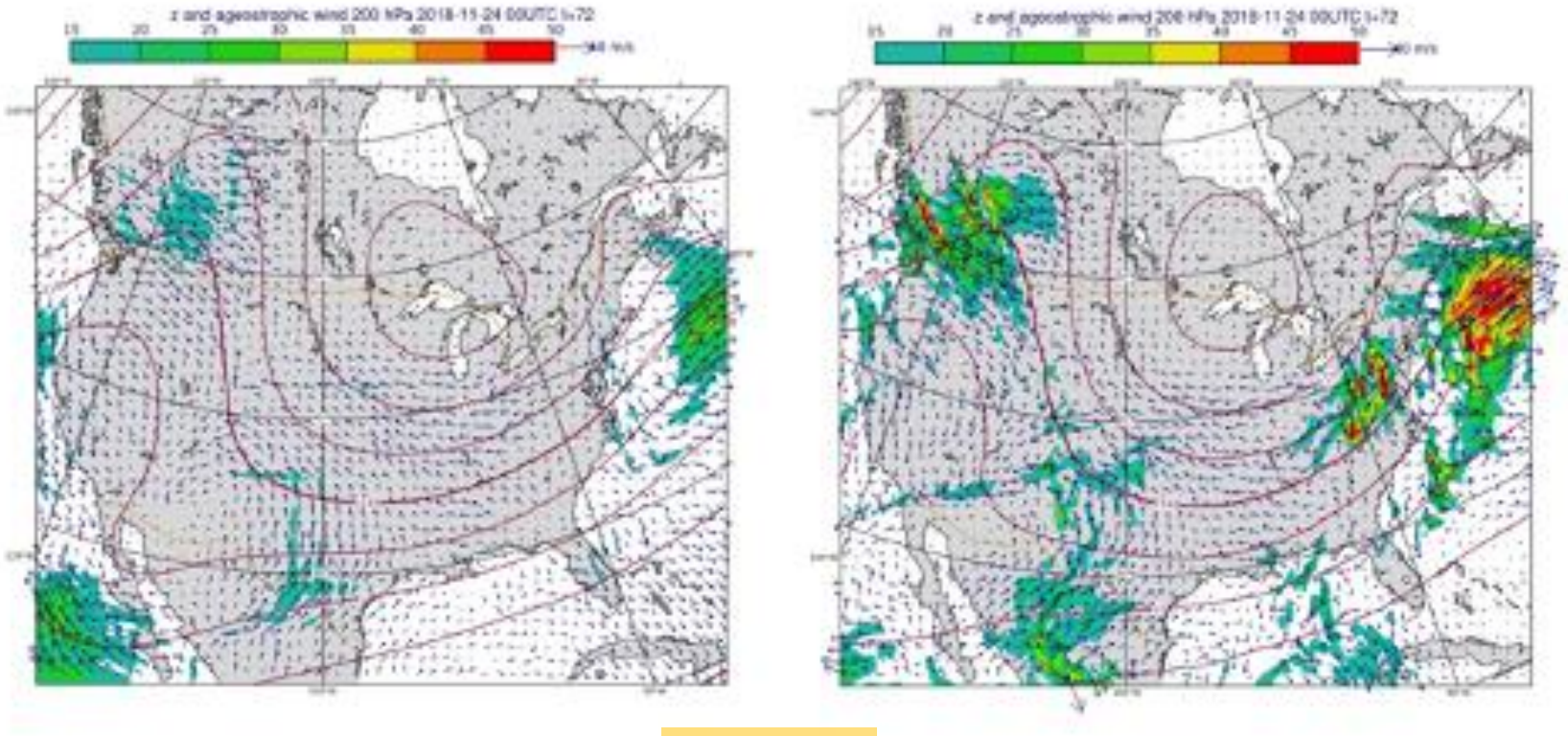
IFS

Pangu-Weather dynamical fields (1)

Geostrophic wind ($\mathbf{V}_g = \frac{1}{f}\hat{\mathbf{k}} \times \nabla_p \Phi$) vs ageostrophic wind $\mathbf{V}_{ag} \equiv \mathbf{V} - \mathbf{V}_g$

Pangu

IFS



$|\mathbf{V}_{ag}|$

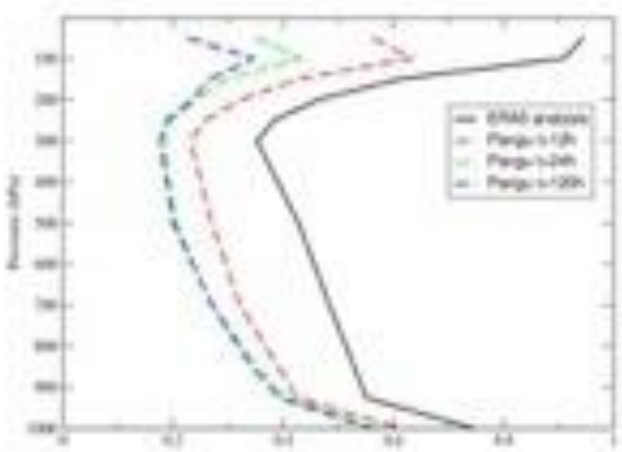
Pangu-Weather dynamical fields (2)

Vorticity and divergence decomposition of the circulation

$$\boldsymbol{u} = \boldsymbol{u}_d + \boldsymbol{u}_v = -\nabla\chi + \mathbf{k} \times \nabla\psi$$

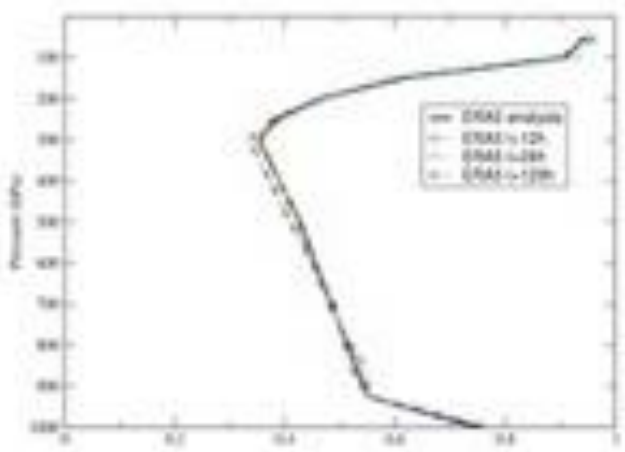
$$\nabla^2\chi = \delta,\ \nabla^2\psi = \zeta$$

Pangu



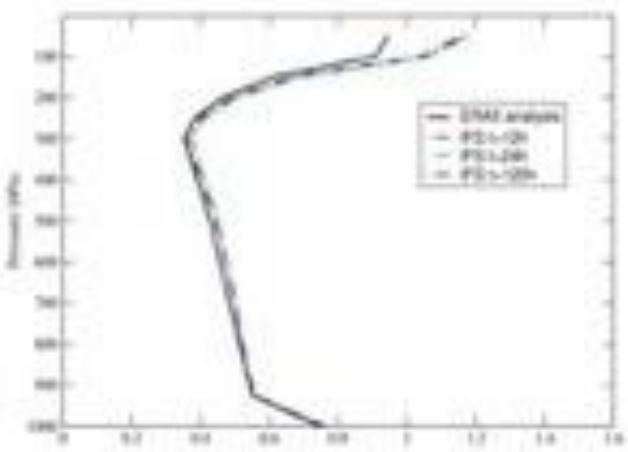
δ/ζ

ERA5-fcst



δ/ζ

IFS



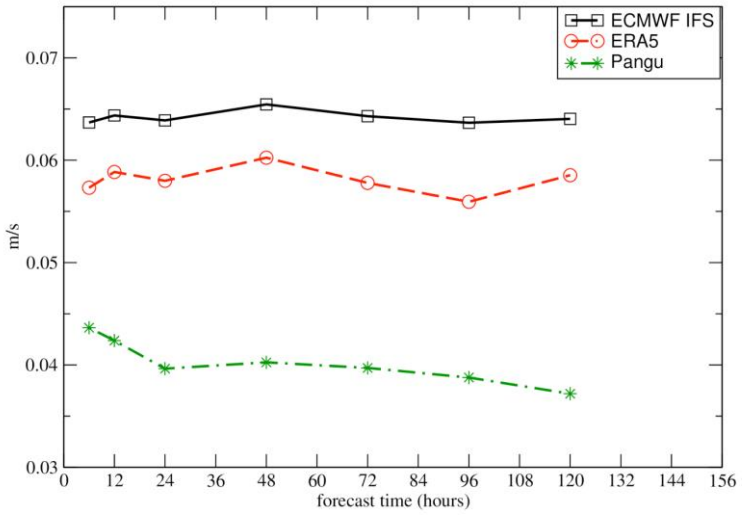
δ/ζ

Pangu-Weather dynamical fields (3)

Vertical velocity is not predicted by Pangu-Weather but can be diagnosed by integrating the continuity equation on forecasted pressure-level fields (Holton and Hakim, 2012):

$$\omega(p) = \omega(p_s) - \int_{p_s}^{p} \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right)_p dp$$

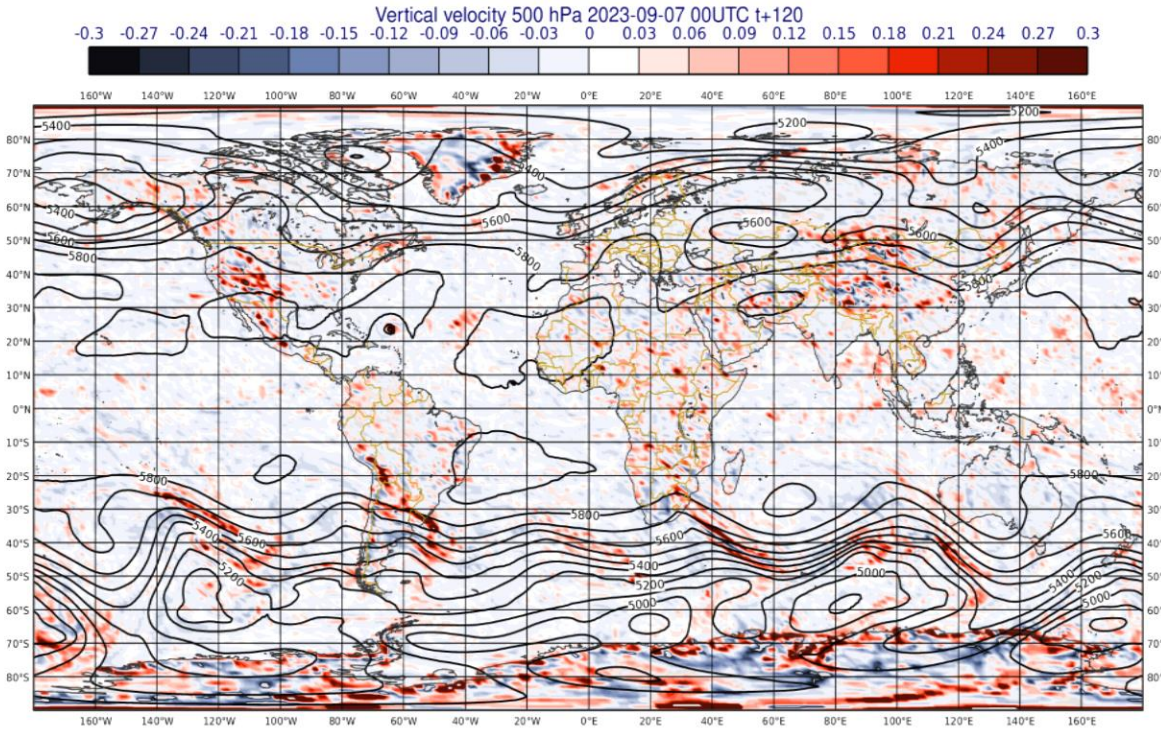
Unsurprisingly, the progressive reduction in the magnitude of the predicted divergence field leads to increasingly weak vertical velocity predictions:



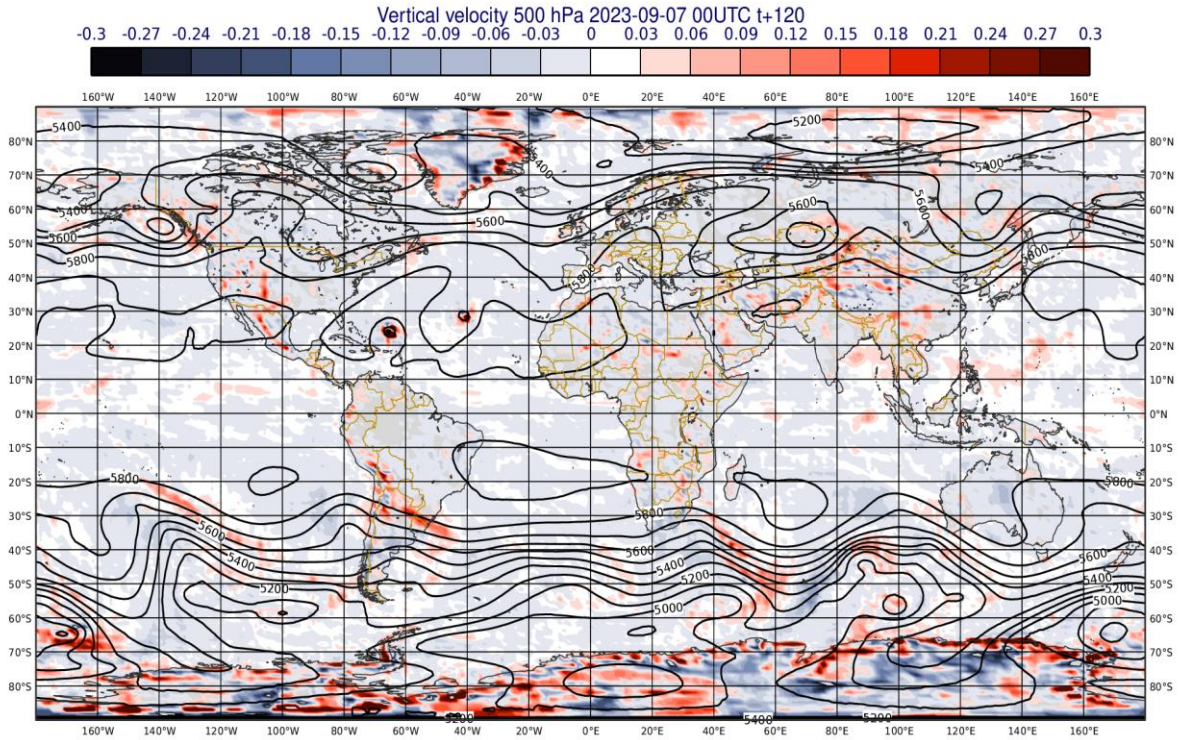
Evolution of stdev of fcst vertical velocity field at 500 hPa IFS, ERA5, Pangu

Pangu-Weather dynamical fields (3)

ERA5 fcst vert. vel. 500 hPa
2023-09-07 00UTC t+120h



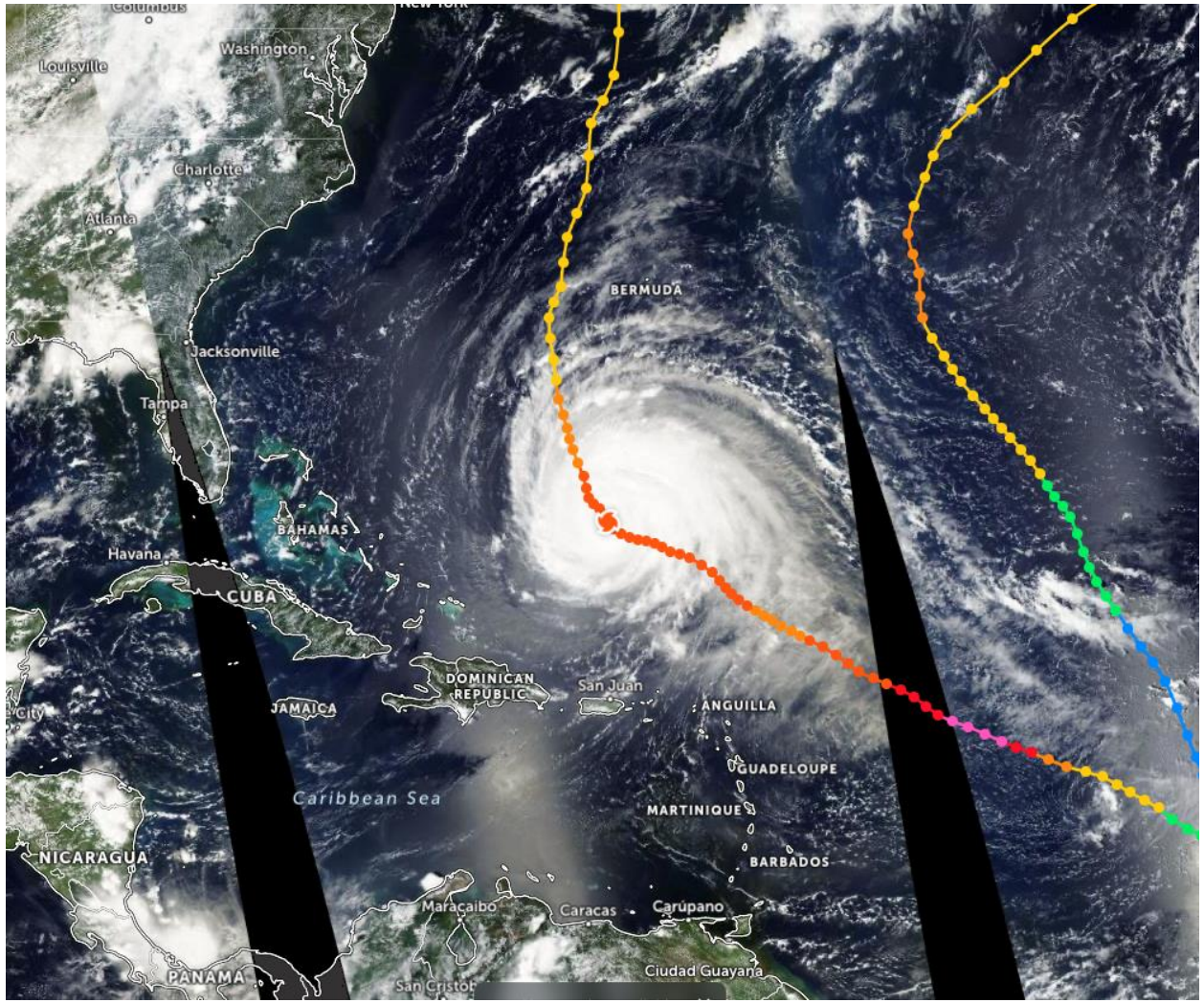
Pangu-Weather fcst vert. vel. 500 hPa
2023-09-07 00UTC t+120h



Pangu-Weather dynamical fields (3)

Hurricane Lee, 12 September 2023 01UTC

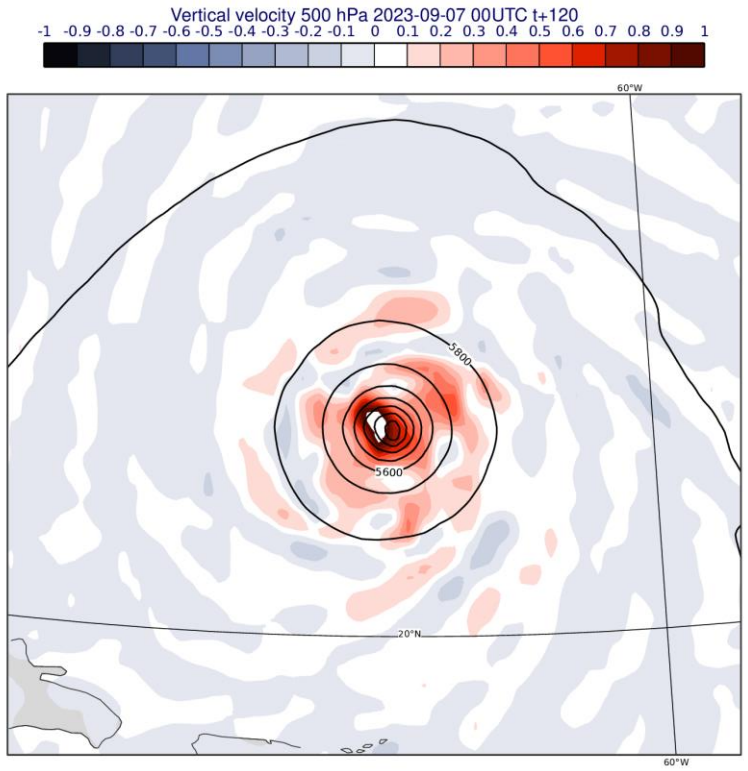
Strongest TC of the 2023 Atlantic Season so far, Category 3 at the time



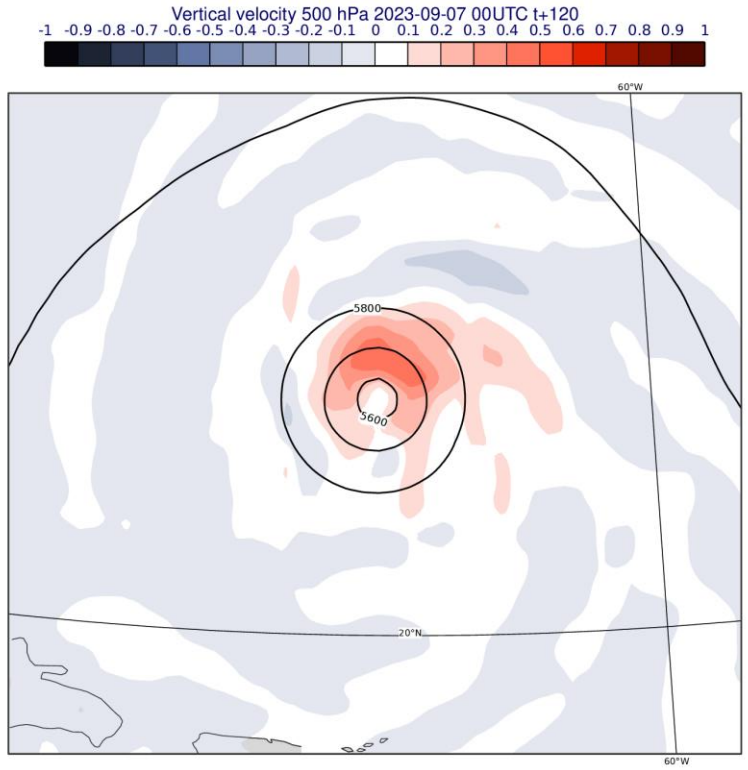
https://zoom.earth/storms/lee-2023/#map=satellite-hd

Pangu-Weather dynamical fields (3)

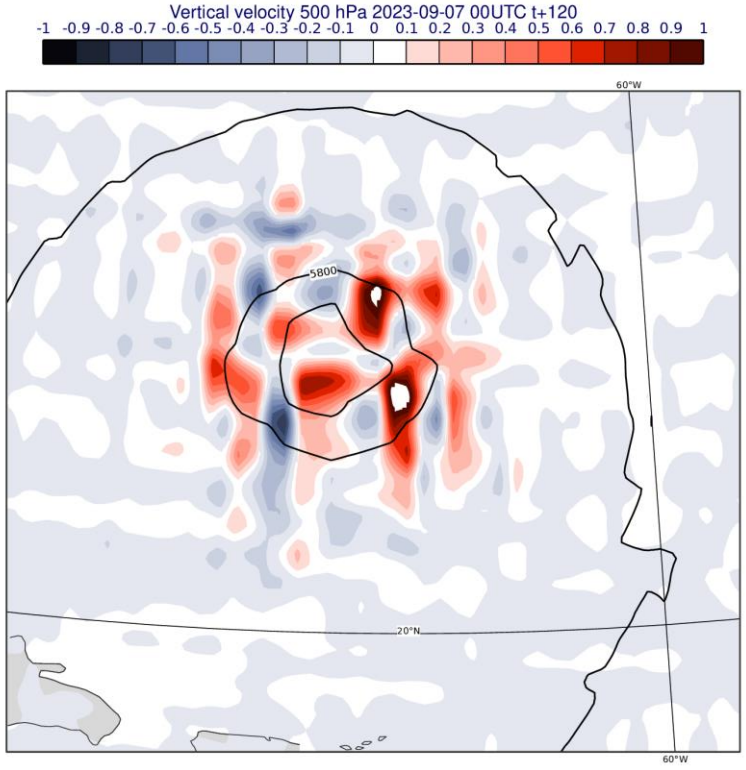
IFS fcst vert. vel. 500 hPa
2023-09-07 00UTC t+120h



ERA5 fcst vert. vel. hPa
2023-09-07 00UTC t+120h



Pangu-Weather fcst vert. vel hPa.
2023-09-07 00UTC t+120h



Conclusions

"*Pangu-Weather not only ends the debate on whether AI-based methods can surpass conventional NWP methods, but also reveals novel directions for improving deep learning weather forecast systems*."

Bi et al., 2022

- Pangu-Weather and co. are trained on ERA5 re-analysis dataset by minimising a L2/L1 loss function of forecast errors. By construction, this leads to increasingly blurry forecast fields and unrealistic forecast activity spectra.
- It is sometimes claimed that MLWP models produce forecasts which are closer to traditional ensemble forecasts mean. Closer inspection reveals that this is not the case (Bonavita, 2023).
- While synoptic-type maps look OK, on closer inspection Pangu-Weather is shown to produce physically inconsistent forecast fields

Outlook

- For the next generation of MLWP models the challenge will be to produce physically consistent forecasts with realistic activity and maintain forecast skill

- For the traditional DA-NWP community the challenge is to speed up adoption of ML techniques to make traditional DA and NWP processes significantly more effective and efficient, e.g. talks by Alban and Marcin

- Too early to say which approach will prevail, but certainly things are moving at unprecedented speed! 4th ECMWF-ESA Workshop on ML for Earth Observation and Prediction, Frascati, Rome, 7-10 May 24

Bonavita, 2023: https://doi.org/10.48550/arXiv.2309.08473

ECMWF