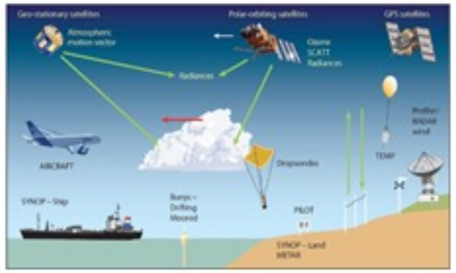
Machine Learning for Model Error Estimation and Prediction

Massimo Bonavita

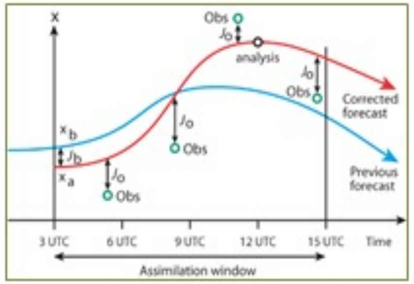
Principal Scientist
Data Assimilation Team Leader
massimo.bonavita@ecmwf.int

ECMWF



Observations

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

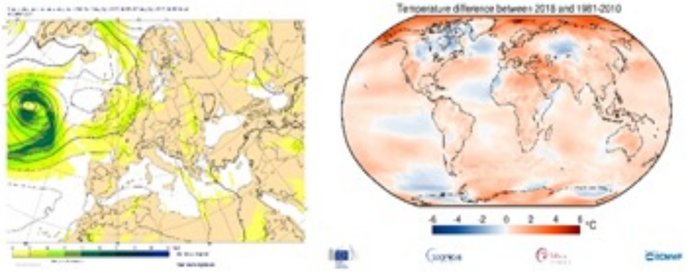


Data assimilation

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

Forecast Models

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



Product generation

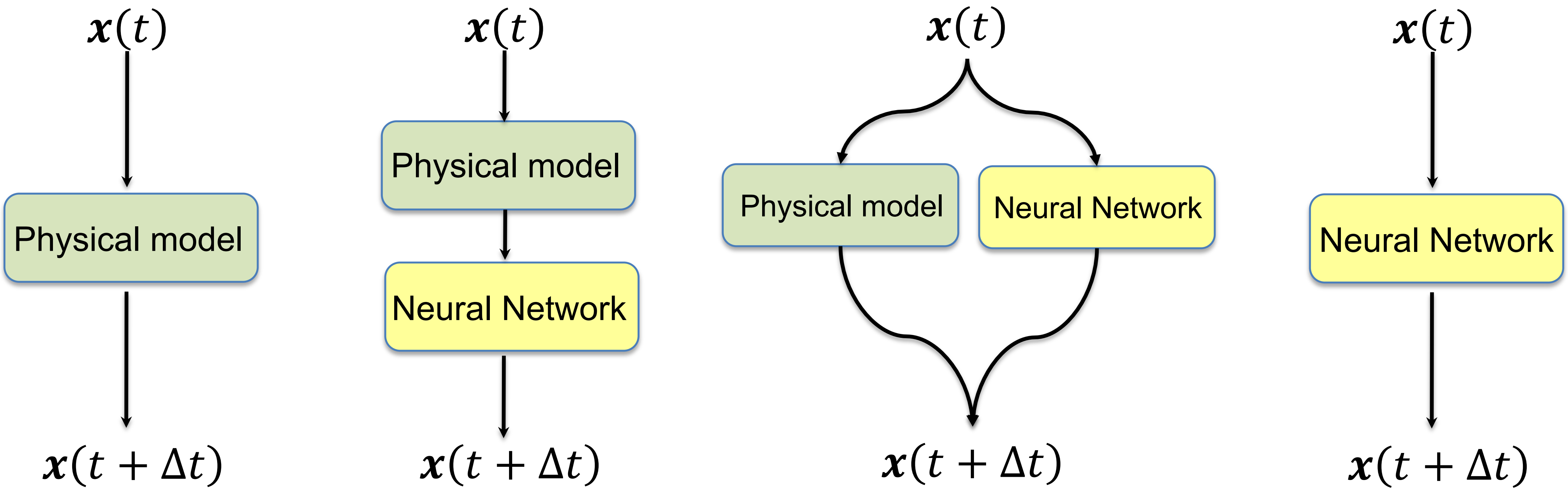
- information extraction (resolution, ensembles, weather features)
- data compression

Bonavita et al., 2021

A taxonomy of forecast models

$$\boldsymbol{x}(t + \Delta t) = M(\boldsymbol{x}(t))$$

Physics-based → **Data-driven**



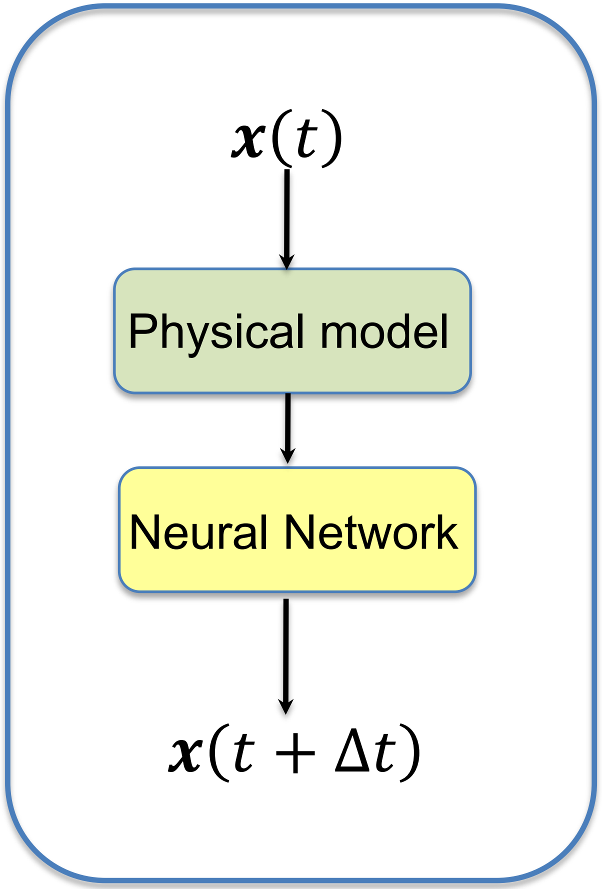
Standard NWP models
(IFS, ICON, FVM,…)

Model error correction
(e.g. Bonavita and Laloyaux, 2020; Farchi et al., 2022, 2023)

Partial model emulation
(e.g. Chantry et al., 2021 Kochkov et al., 2023, NeuroGCM)

Full model emulation
(e.g. Keisler, 2021, PanguWeather, 2022, GraphCast, 2023, …)

Data-driven model error correction



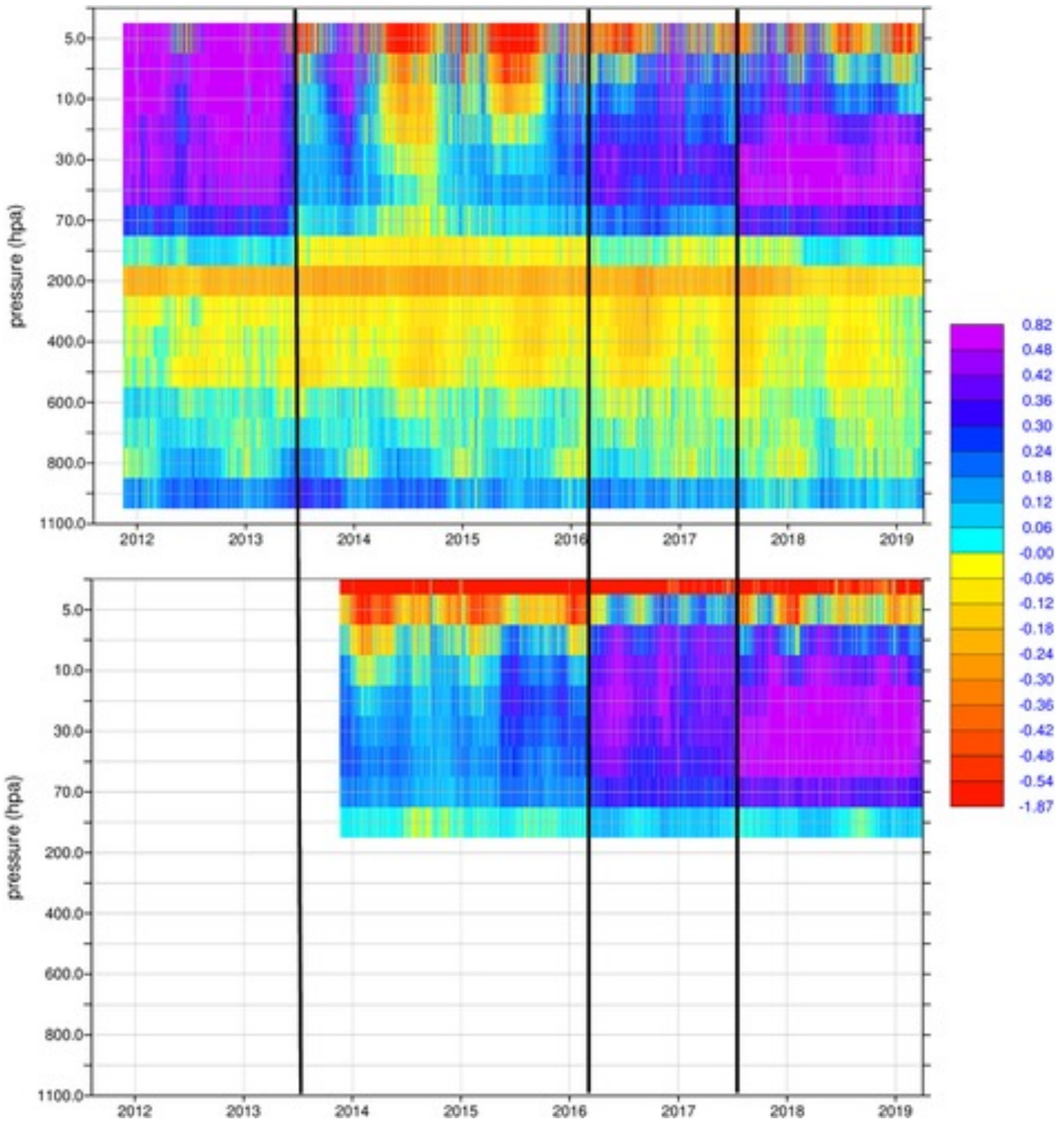
- The physical model will always be affected by systematic, <u>predictable</u> errors due to a variety of causes (unresolved, unrepresented physics, discretisation errors, wrong model forcings, etc.)
- The NN is used as a <u>data-driven (statistical) online correction</u> to the time tendencies forecasted by the physical model:

$$\frac{d\boldsymbol{x}}{dt} = \Phi_{adv}\big(\boldsymbol{x}(t)\big) + \Phi_{phys}\big(\boldsymbol{x}(t)\big) + \Phi_{ML}\big(\boldsymbol{x}(t)\big)$$

- Is model error a big deal?

Model Error: Atmosphere

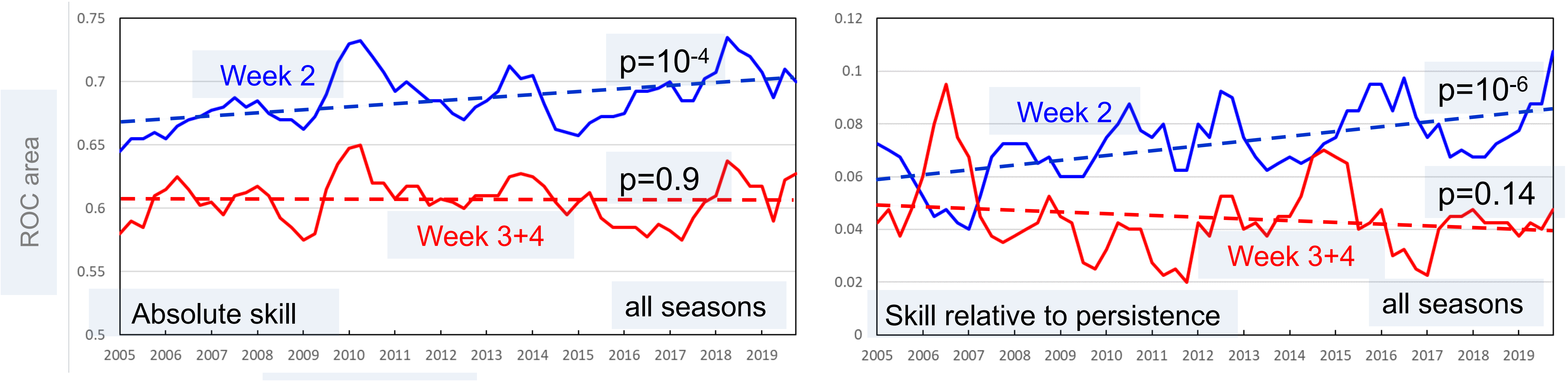
Time series of the difference between observations and temperature first-guess trajectory from the ECMWF operational analysis cycle for radiosonde (top) and Radio Occultations (bottom)



Laloyaux et al., 2020

Model Error: Atmosphere

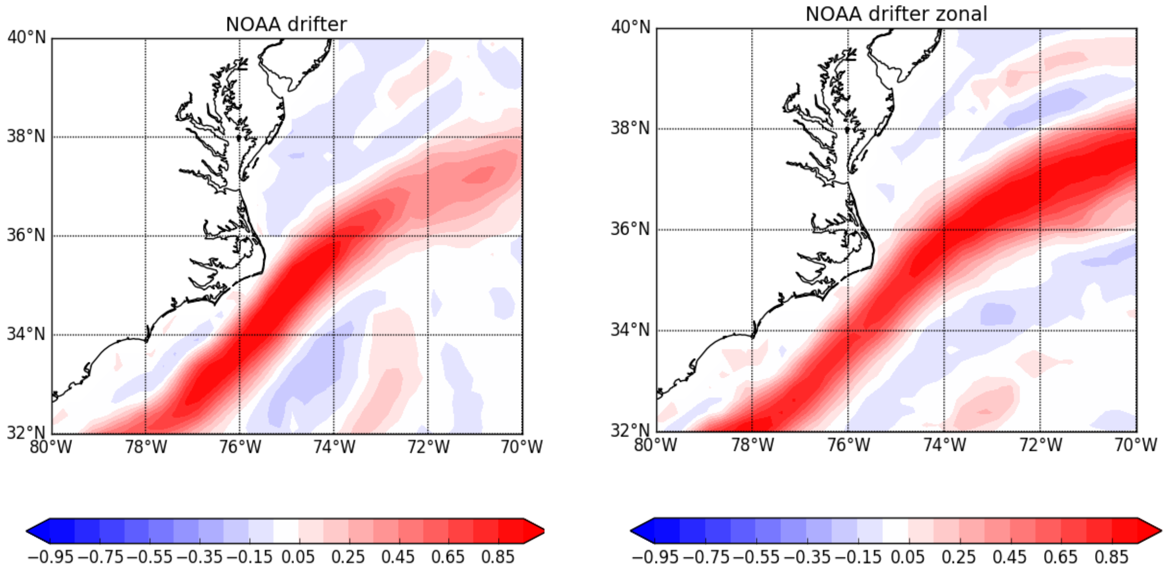
- Traditionally, DA algorithms have used a perfect model assumption and aimed at reducing random initial conditions errors: this approach brings steady progress up to ~2 weeks, but not beyond
- Can Machine Learning help break the 2-week Lorenz predictability barrier?



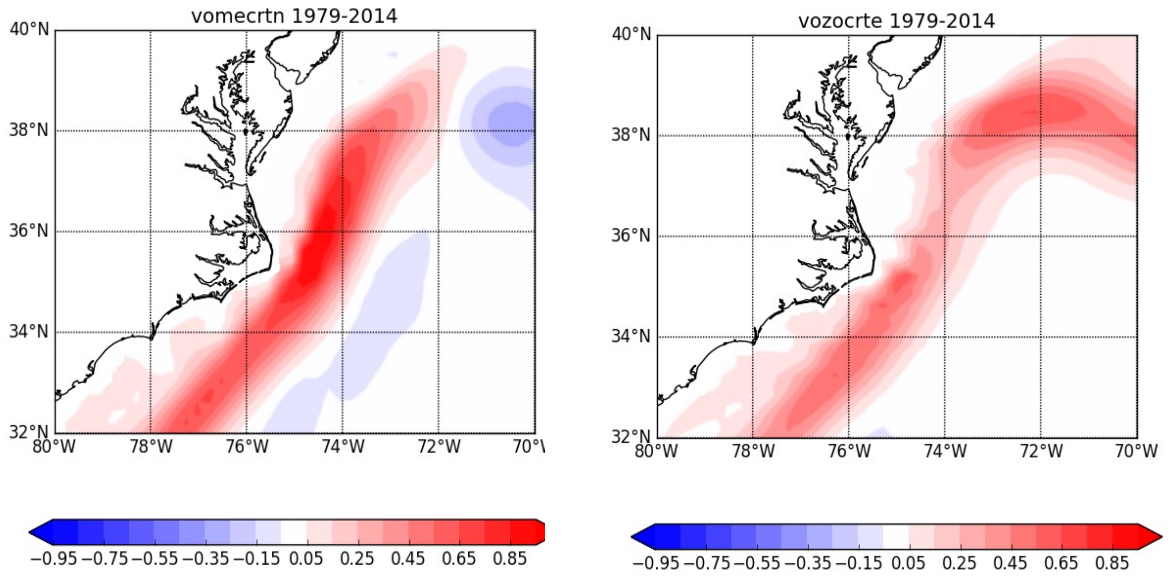
From Frederic Vitart and Thomas Haiden

Model Error: Ocean

NOAA drifter climatology

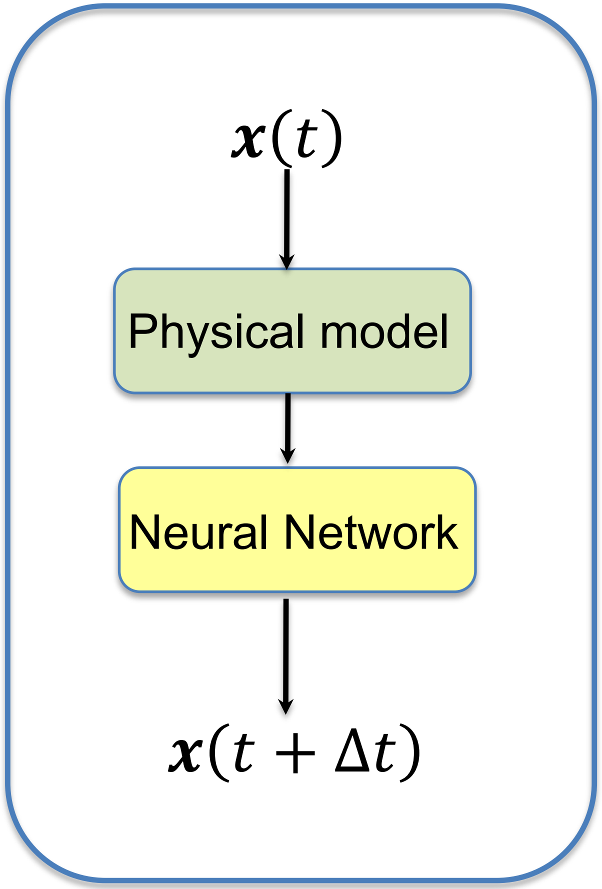


ORAS5 1985-2012.



From Hao Zuo

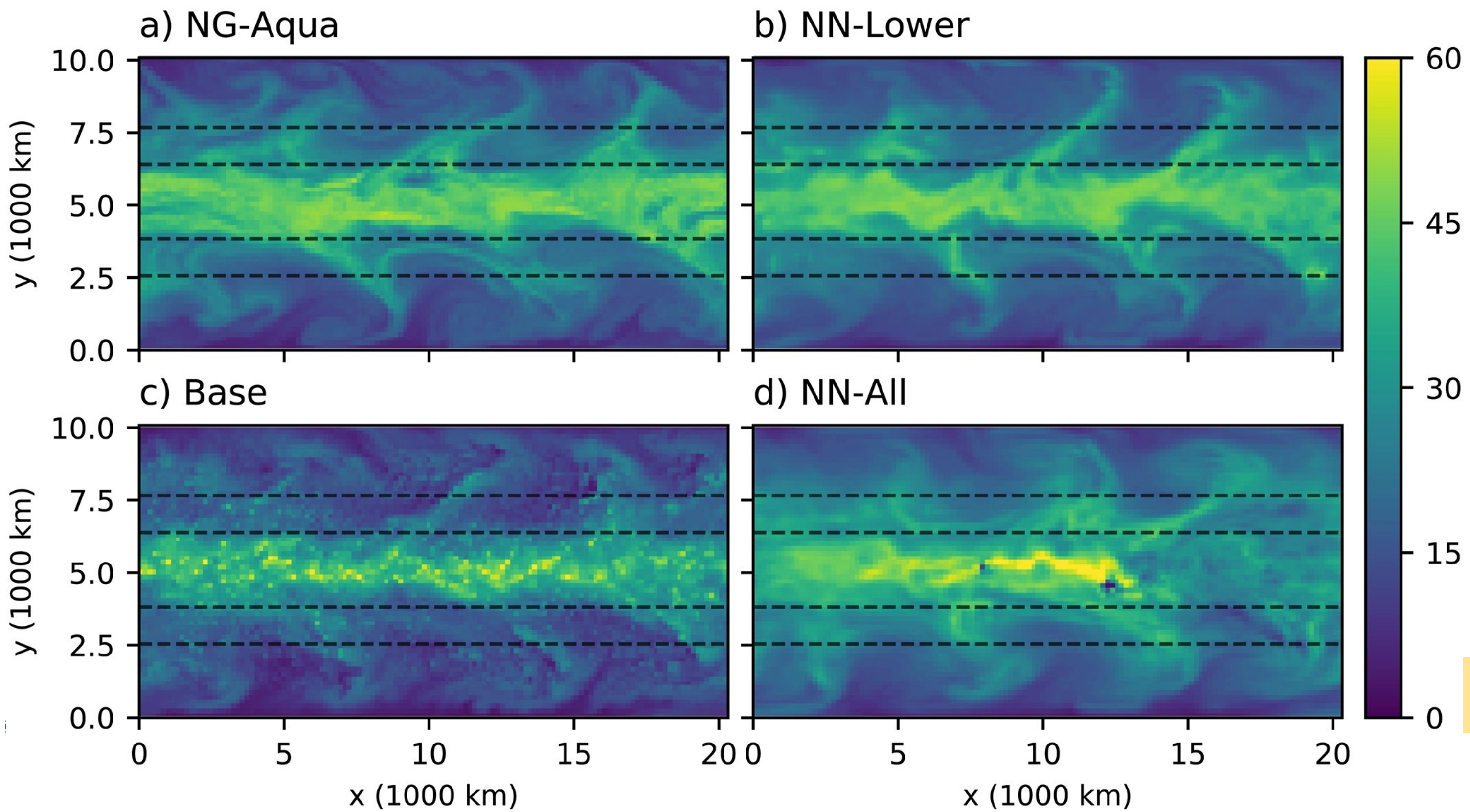
Model error estimation



- The first challenge is the estimation of model error
- Using a statistical model like a NN requires choosing a ground truth to build the training dataset
- We have different options here…

Model error estimation

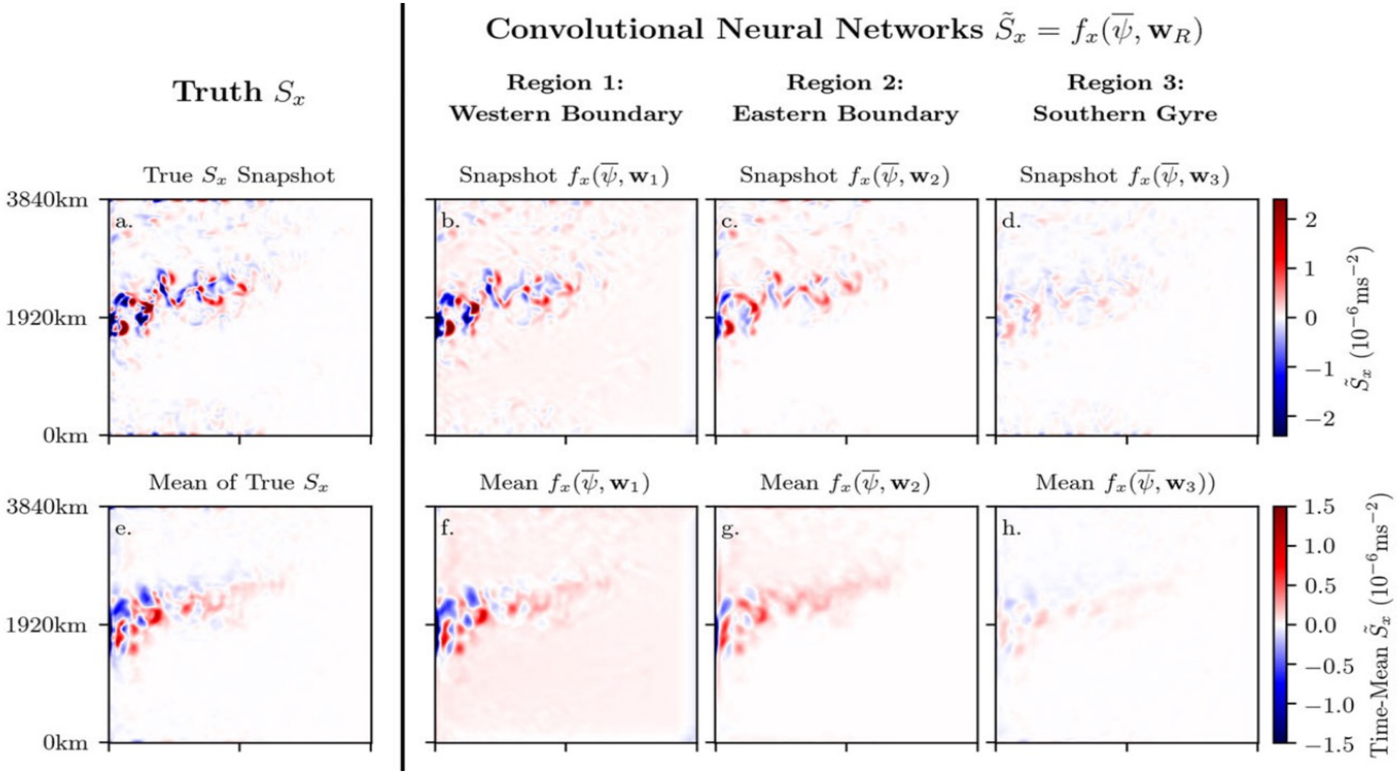
- One option is to train the NN using a more complex, more accurate, higher resolution version of the model as a proxy for the truth
- The goal is to train the NN to learn the effect of the parameterised and unresolved dynamical and physical processes onto the resolved dynamics (e.g., Brenowitz and Bretherton, 2018, 2019; Rasp et al., 2018)



Brenowitz and Bretherton, 2019

Model error estimation

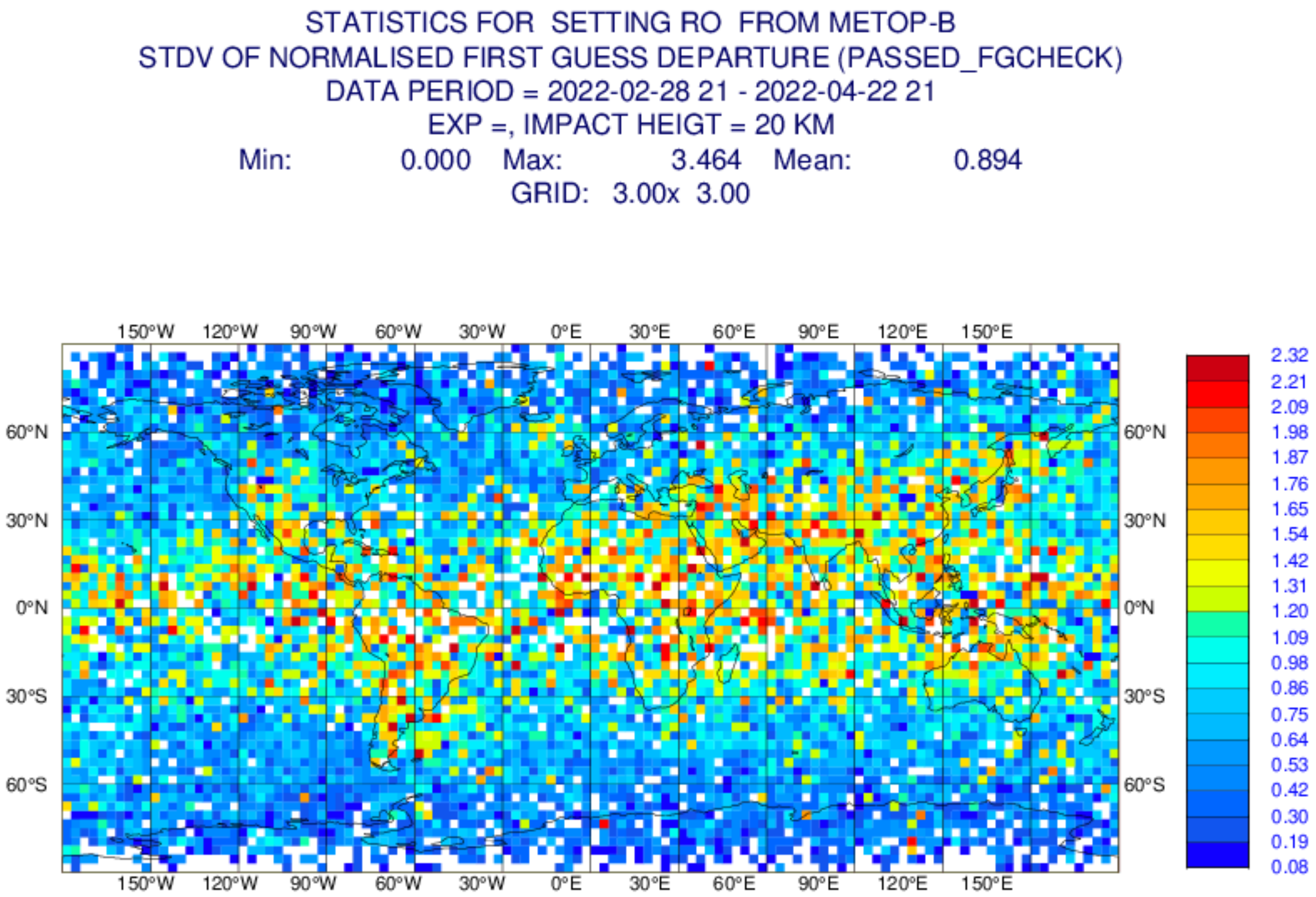
- In Ocean Modelling one of the main sources of errors come from insufficient resolution -> need to account for unresolved physical processes
- This problem has been tackled with ML tools (CNN, DNN, RVM) to build offline data-driven parameterizations of the unresolved dynamics for a fully observed system (Bolton and Zanna, 2019, Zanna and Bolton, 2020, Kutz, 2017, Ling et al., 2016)



Bolton and Zanna, 2019

Model error estimation

- Another option is to train the NN using <u>observations</u> of the system as a proxy for the truth
- The training dataset is built using observations minus (short range) forecasts (O-B departures)
- The NN is trained to learn the O-B departures from a set of available predictors (state, location, time of day, season, etc.)

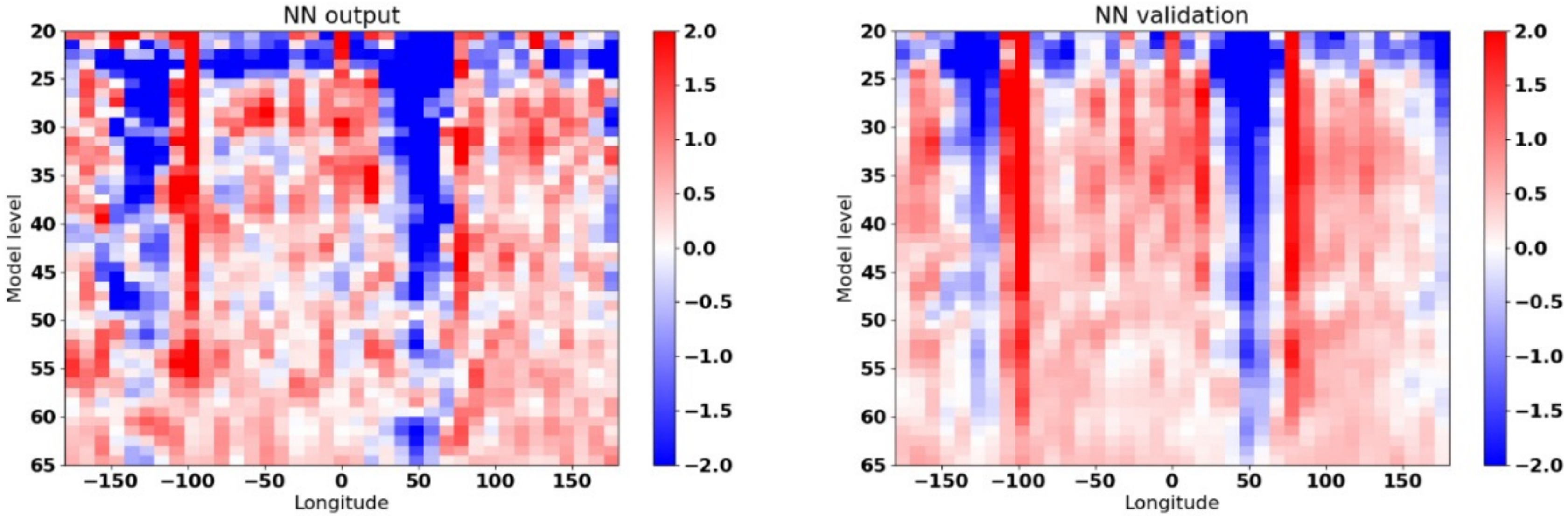


ECMWF Observation Monitoring

Model error estimation

- The idea of using observations as the truth is attractive because it directly introduces independent information on the model errors we want to correct
- It also comes with its own caveats: observation errors, observation coverage, how to extrapolate the corrections in space, time and to other variables in a physically consistent manner?

O-B departures on 01-05-2019 averaged between 10S and 20N (left) and their prediction with a Convolutional Neural Network (right)



From P. Laloyaux, 2021

Model error estimation

- Option #3 is to train the NN using analysed states of the system as a proxy for the truth
- The training dataset is built using analysis minus (short range) forecasts (A-B increments)
- The NN is trained to learn the A-B increments from a set of available predictors (state, location, time of day, season, local solar zenith angle, SST, etc)
- Advantages:
  1. The model error estimates are directly available in model space and globally;
  2. Analyses are more accurate than any individual observation type
- Disadvantages:
  1. The analyses will be affected to some extent by the model error we want to estimate
  2. The (A-B) increments will also be affected by ”errors of the day” (i.e., initial condition errors)

Model error estimation: weak constraint 4D-Var

- We already do model error estimation inside the 4D-Var analysis cycle: it is called weak constraint 4D-Var!

$$J_{wc4DVar}(\boldsymbol{x}_0, \boldsymbol{\eta}) = J_B + J_O + J_Q = \frac{1}{2}\left(x_0 - x_0^b\right)^T B^{-1}\left(x_0 - x_0^b\right)$$

$$+ \frac{1}{2}\sum_{i=0}^{N}(H_i(x_i) - y_i)^T R_i^{-1}(H_i(x_i) - y_i)$$

$$+ \frac{1}{2}\sum_{i=1}^{N}\left(x_i - M_i(x_{i-1}, \boldsymbol{\eta})\right)^T Q_i^{-1}(x_i - M_i(x_{i-1}, \boldsymbol{\eta}))$$

$\boldsymbol{x}_0 = state\ estimate$
$\boldsymbol{\eta} = model\ error\ estimate$

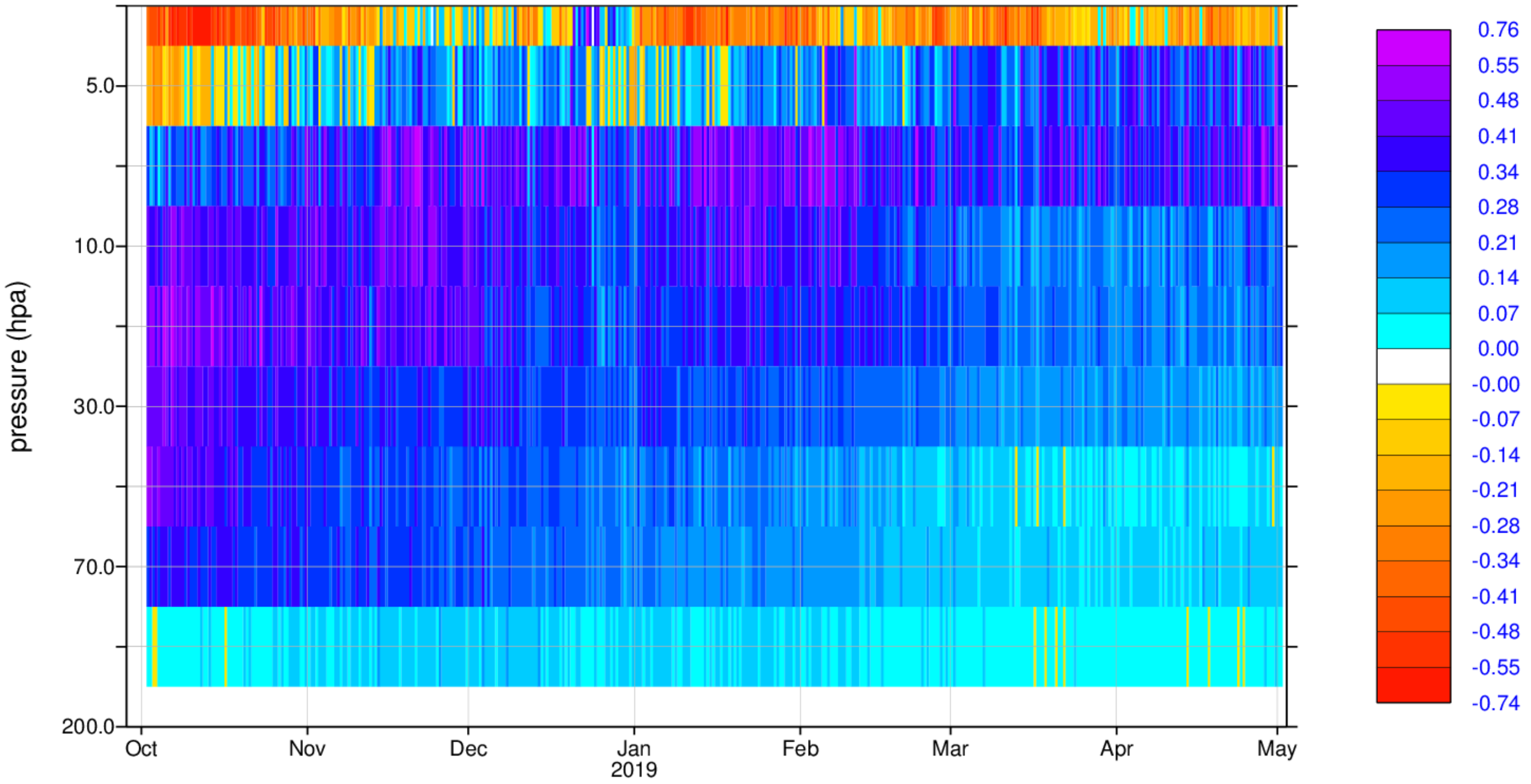
Loss functions common in ML/DL can be obtained as particularisations of wc-4DVar (Brajard et al., 2020; Bocquet et al., 2020; Farchi et al., 2020):

1. ML/DL models are not used for state estimation $\rightarrow J_B = 0$
2. ML/DL loss functions typically assume full, noiseless observations $(H_i = I, R_i \rightarrow 0) \rightarrow J_O = 0$
3. ML/DL models can optionally have a regularization term function of the NN model parameters $\mathrm{L}(\boldsymbol{\eta}) = \mathrm{L}(\boldsymbol{W}, \boldsymbol{b})$, e.g. Tikhonov regularisation, drop-out, etc.

Model error estimation: weak constraint 4D-Var

- wc-4DVar progressively learns a model error tendency correction and applies it to subsequent background forecasts in the DA cycle
- wc-4DVar is an online machine learning algorithm for model error estimation and correction!

Mean first-guess departure with respect to GPS-RO temperature retrievals



From Laloyaux et al, 2020

Model error estimation: weak constraint 4D-Var

- What can Machine Learning bring to table?

- wc-4DVar produces an estimate of model error valid over the length of the assimilation window: $J_{wc4DVar}(\boldsymbol{x}_0, \boldsymbol{\eta})$

- The NN will produce a model of model error, which can be applied and used at any point in time:

$$\mathbf{M_W}(\boldsymbol{predictors}) = \mathbf{M_W}(\boldsymbol{x}, time, lat, lon, SST, \dots)$$

Hybrid ML-DA

- Option #3 is to train the NN using <u>analysed</u> states of the system as a proxy for the truth
- In this approach we have a number of possibilities on how to fuse the Data Assimilation system with the ML estimate of model error:

  1. Offline training of NN and use it in a data assimilation cycle (e.g, Bonavita and Laloyaux, 2020, Watson, 2019);
  2. Online training of NN in a data assimilation cycle with a coordinate descent approach (e.g. Brajard et al., 2020, Bocquet et al., 2020),
  3. Full online training of NN inside data assimilation system (e.g., Farchi et al., 2021b)

Hybrid ML-DA in low-order systems

- In all generality we would like to minimise the following cost function of the state ($\mathbf{x}_{0:N}$) and the NN parameters ($\mathbf{p}$) (Farchi et al., 2021):

$$
\begin{aligned}
\mathcal{J}\left(\mathbf{p}, \mathbf{x}_{0:N_t}\right) \triangleq\ & \frac{1}{2}\left\|\mathbf{x}_0 - \mathbf{x}_0^{\mathrm{b}}\right\|^2_{\mathbf{B}^{-1}} + \frac{1}{2}\sum_{k=0}^{N_t} \left\|\mathbf{y}_k - \mathcal{H}_k(\mathbf{x}_k)\right\|^2_{\mathbf{R}_k^{-1}} \\
& + \frac{1}{2}\sum_{k=0}^{N_t-1} \left\|\mathbf{x}_{k+1} - \mathcal{M}_k(\mathbf{p}, \mathbf{x}_k)\right\|^2_{\mathbf{Q}_k^{-1}} + \mathcal{L}(\mathbf{p}),
\end{aligned}
\tag{4}
$$

- Here the background error ($\mathbf{x}_0 - \mathbf{x}_0^{b}$)$\sim \mathrm{N}(\mathbf{0},\mathbf{B})$, the observations are sparse and are related to the system state by $\mathbf{y}_k = \mathbf{H}_k(\mathbf{x}_k) + \mathbf{v}_k$, ($\mathbf{v}_k \sim \mathrm{N}(\mathbf{0},\mathbf{R})$)
- Model errors ($\mathbf{x}_{k+1}$-$\mathrm{M}(\mathbf{p},\mathbf{x}_k)$) follow a Gaussian distribution with covariance $\mathbf{Q}_k$, and are assumed uncorrelated with other error sources

Hybrid ML-DA in low-order systems

- How to optimise this cost function of $(\mathbf{p}, \mathbf{x}_{0:\mathrm{N}})$?
- One way is to alternate DA steps (to estimate x0:N) and ML steps (to estimate the model error parameters p) in a coordinate-descent framework:

FARCHI ET AL. | 5

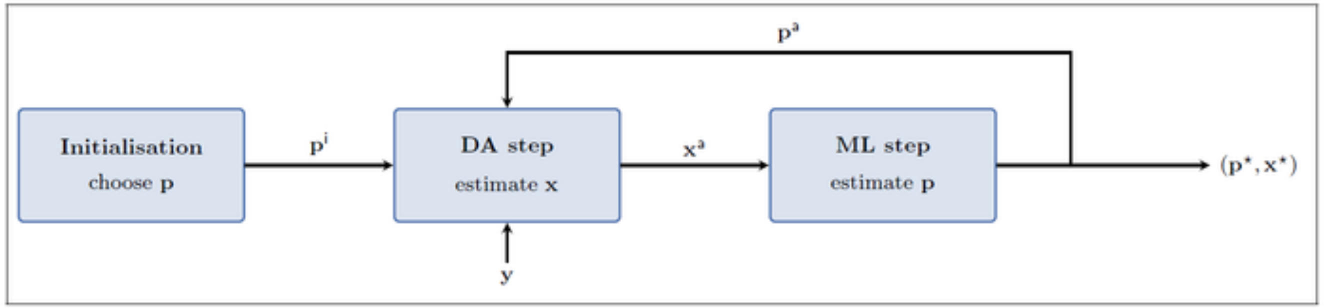


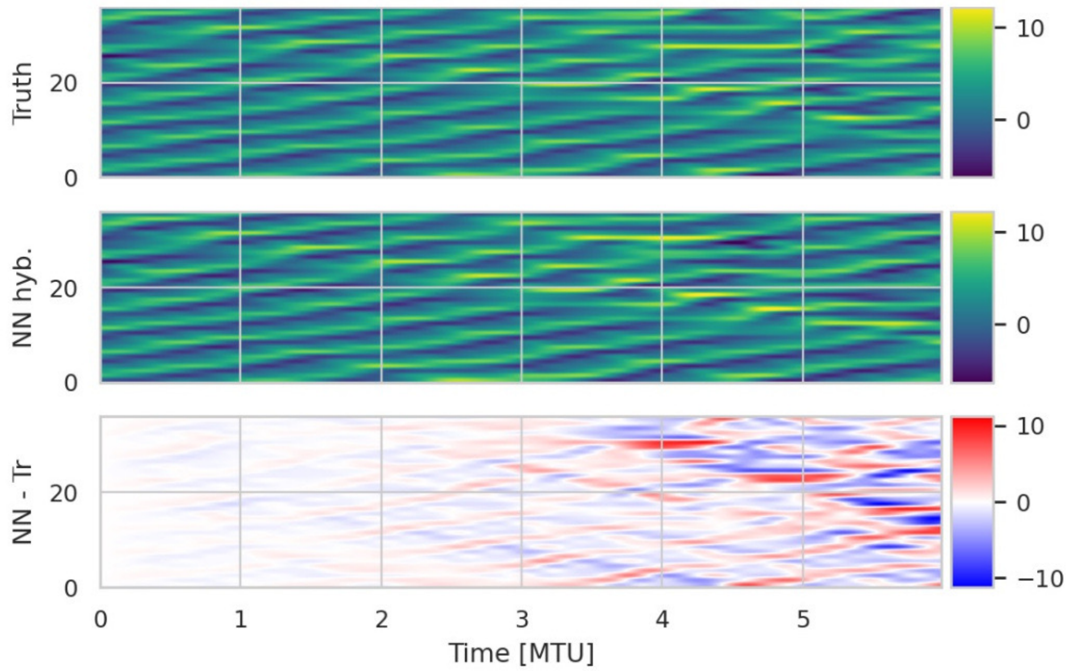
FIGURE 1 Minimisation strategy for Equation (4): alternate DA steps with ML steps to estimate the parameters $\mathbf{p}$ and the trajectory $\mathbf{x}_{0:N_t}$ with an increasing accuracy.

Coordinate descent approach for the online training of a NN

From Farchi et al., 2020

Hybrid ML-DA in low-order systems

- This coordinate-descent idea was used with good results, e.g. in Brajard et al., 2021 to reconstruct the state and the dynamics of Lorenz-96 model using convolutional neural networks



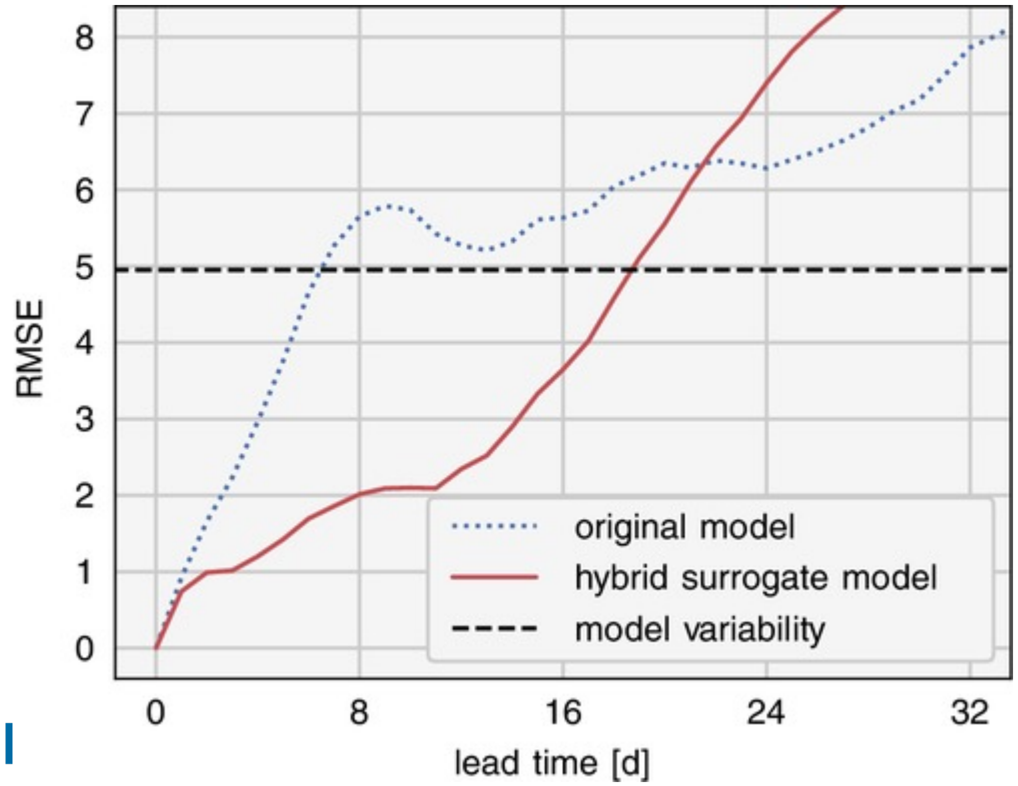
State estimation with the coordinate descent approach for the online training of a NN

from Brajard et al., 2021

- However: 1) the initialization of $\mathbf{p}$ is critical and cold-starting can easily lead to divergence, and 2) the number of DA-ML cycles required to reach convergence can be high, problematic for a realistic application

Hybrid ML-DA in low-order systems

- Farchi et al., 2021a, have used an offline training approach to construct a hybrid ML-physics model for a quasi-geostrophic system
- Their surrogate model based on Convolutional Neural Networks was shown to be able to significantly improve on the original perturbed model both in forecast mode and in the assimilation cycle

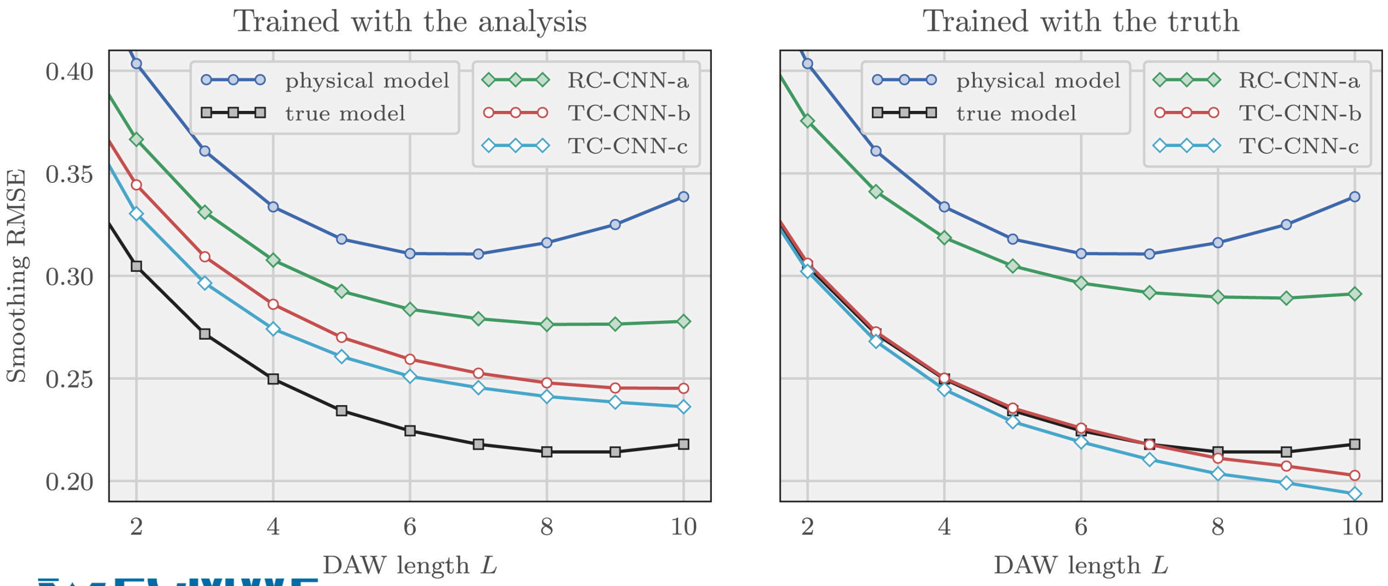


Forecast skill of the original model (dotted line) and the hybrid surrogate model (continuous line) as a function of the lead time in days

From Farchi et al., 2021a

Hybrid ML-DA in low-order systems

- In Farchi et al., 2021b, two crucial aspects of the training process have been studied using the two-scale Lorenz model
- One aspect was whether it is more efficient to correct the integrated-in-time model error (i.e., the model resolvent), or the model error tendencies
- Results show that both techniques perform well in forecast mode, but tendency correction is preferable when the hybrid model is used in a DA cycle

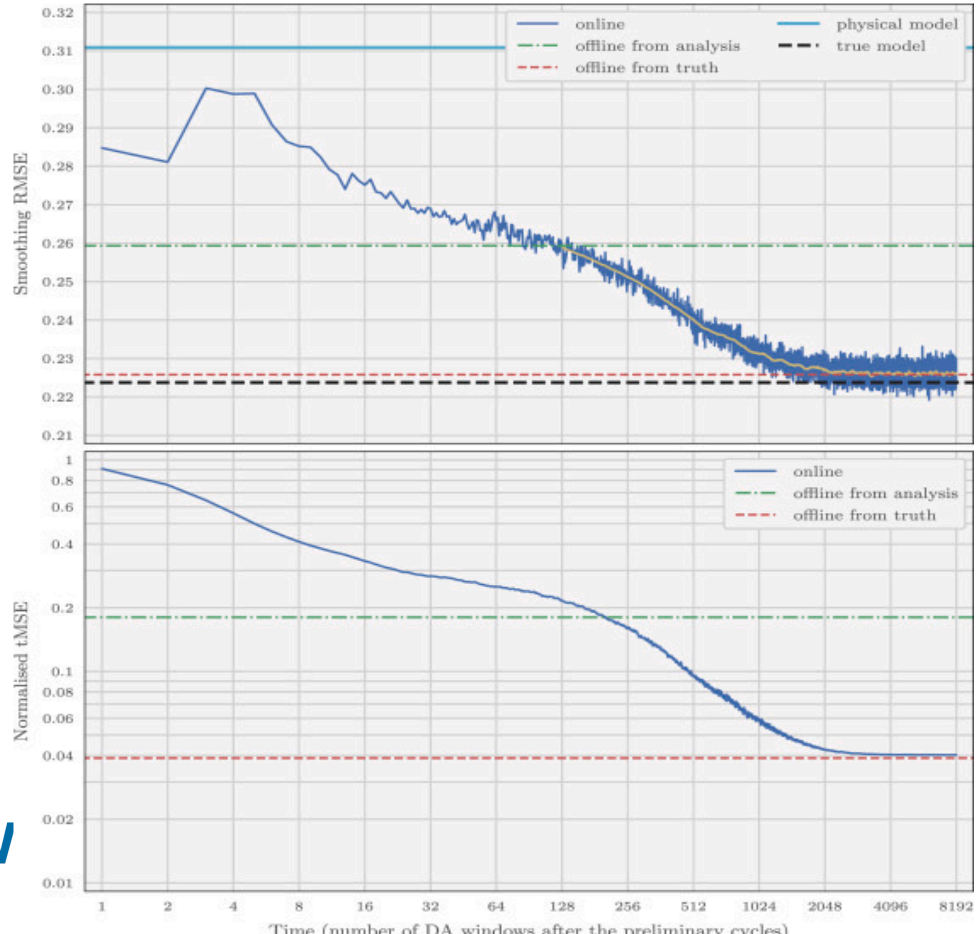


Analysis RMSE for the physical model (in blue), the true model (in black), and the trained surrogate models: RC-CNN-a (in green), TC-CNN-b (in red), and TC-CNN-c (in cyan). The surrogate models are trained either with the analysis (left panel) or with the truth (right panel).

From Farchi et al., 2021b

Hybrid ML-DA in low-order systems

- The other aspect that was investigated was the effectiveness of offline vs online (i.e., inside 4D-Var) model error learning
- Online learning appears to give best results in this experimental setup



Time series of sRMSE (top panel) and tMSE (bottom panel) for the online experiment with TC-CNN-b (in blue). For comparison, the horizontal lines show the scores for the physical model (in cyan), the true model (in black), TC-CNN-b trained offline with the analysis (in green) and trained offline with the truth (in red).

From Farchi et al., 2021b

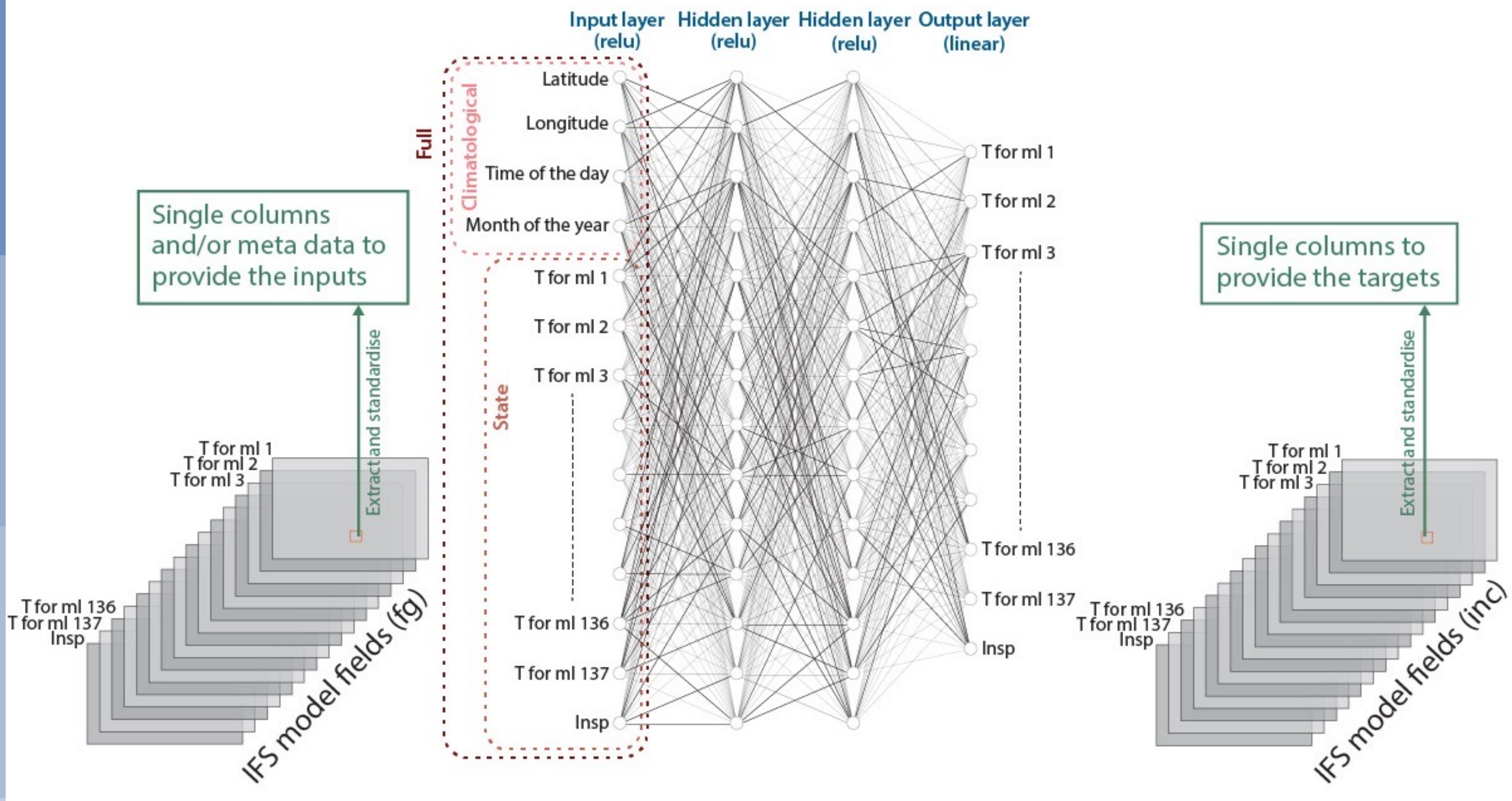
Hybrid ML-DA in operational systems

- How can we adapt these ideas to operational NWP and Climate prediction? There are some issues:
  1. Typically there is not enough time to iterate to convergence the DA and ML steps in operational NWP;
  2. We have a much more complex model, but a very good one!
  3. Most importantly, the size of the model error space is orders of magnitude bigger than in low-order models

Hybrid ML-DA in operational systems

- What kind of prior knowledge do we have on the atmospheric model error generating distribution?
- Prior assumptions:
    1. *We can consider the atmospheric flow to be subject to homogeneous dynamics and heterogeneous forcings*;
    2. *Physical parameterisations of unresolved motions and radiation plus surface forcings are the dominant sources of model error*
    3. *Physical parameterisations are computed and applied over model columns*.
- This led us to define a set of predictors made up of the concatenation of <u>climatological predictors</u> (<u>time of day, month, lat, lon</u>) and the <u>vertical columns</u> (137 levels) of the <u>model first guess</u> prognostic variables of the model (t, lnsp, vo, div, q).
- This choice amounts to splitting <u>the full 3d regression problem into a 1d x 2d problem</u> and is conceptually similar to having a separable representation of a 3D covariance matrix

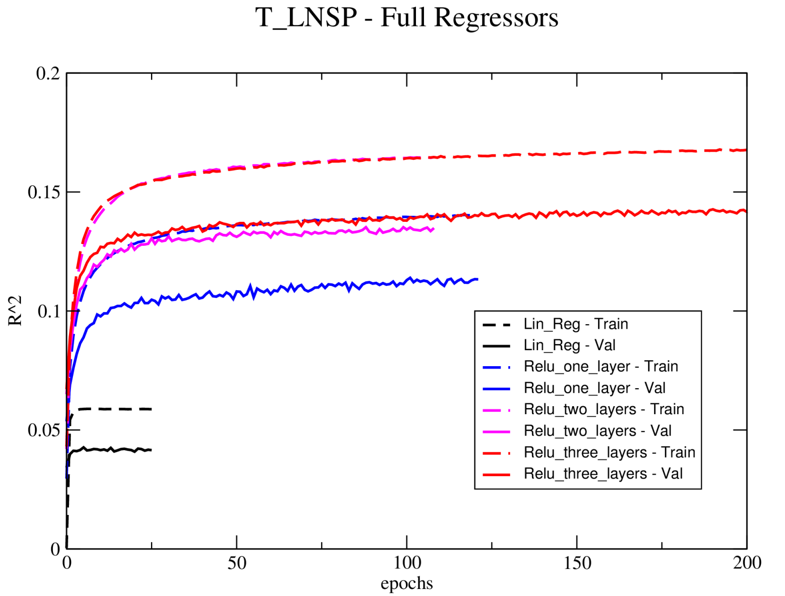
Hybrid ML-DA in operational systems

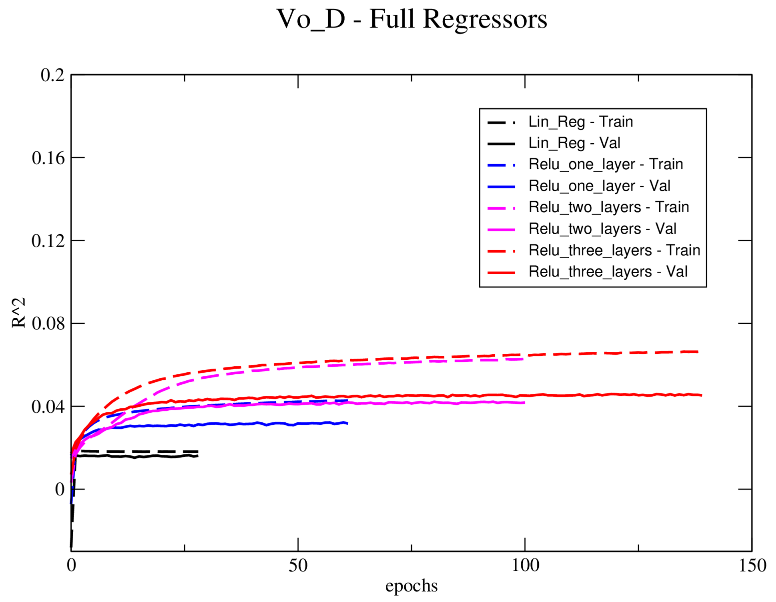


- Dense Neural Network with Relu activations
- Three layers with nonlinear activations give best results: problem with only moderate nonlinearities
- Dropout layers used to control overfitting, input/outputs pre-normalised for training, Adam minimiser
- Number of trainable parameters ~6*10^4, size of training dataset ~10^6

Bonavita & Laloyaux, 2020

Hybrid ML-DA in operational systems



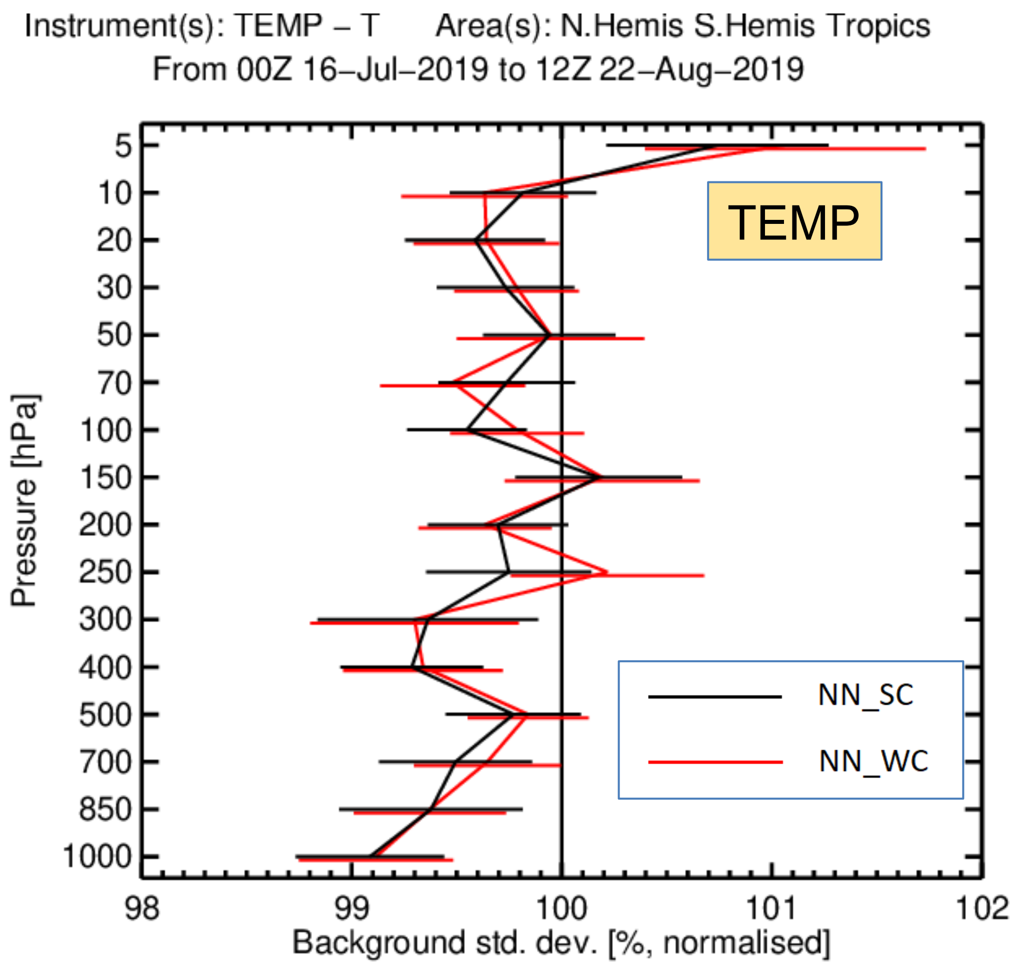


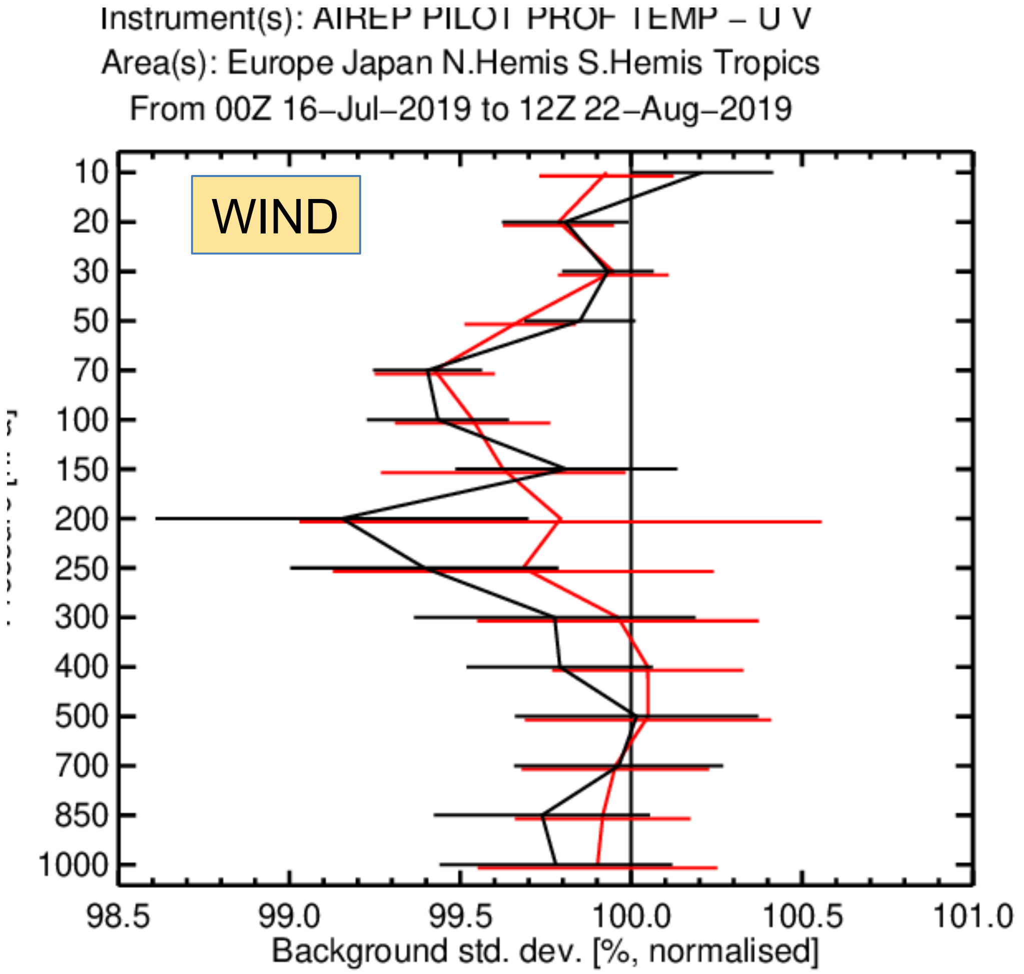
- Training/Testing curves are shown in terms of explained variance (R^2)
- Saturation of explained variance is used as stopping criterion during training
- Mass (T, lnsp) errors can be better predicted (~14-15% explained variance) than wind (~4-5% explained variance) and humidity (~0%) errors.
- State-dependent predictors (first guess values) have more predictive power than climatological predictors
- The NN model provides a state-dependent correction beyond climatological bias correction.

Bonavita & Laloyaux, 2020

Model Error Estimation and Correction in the IFS: Random errors

Globally-averaged Observation-First Guess StDev norm. diff.





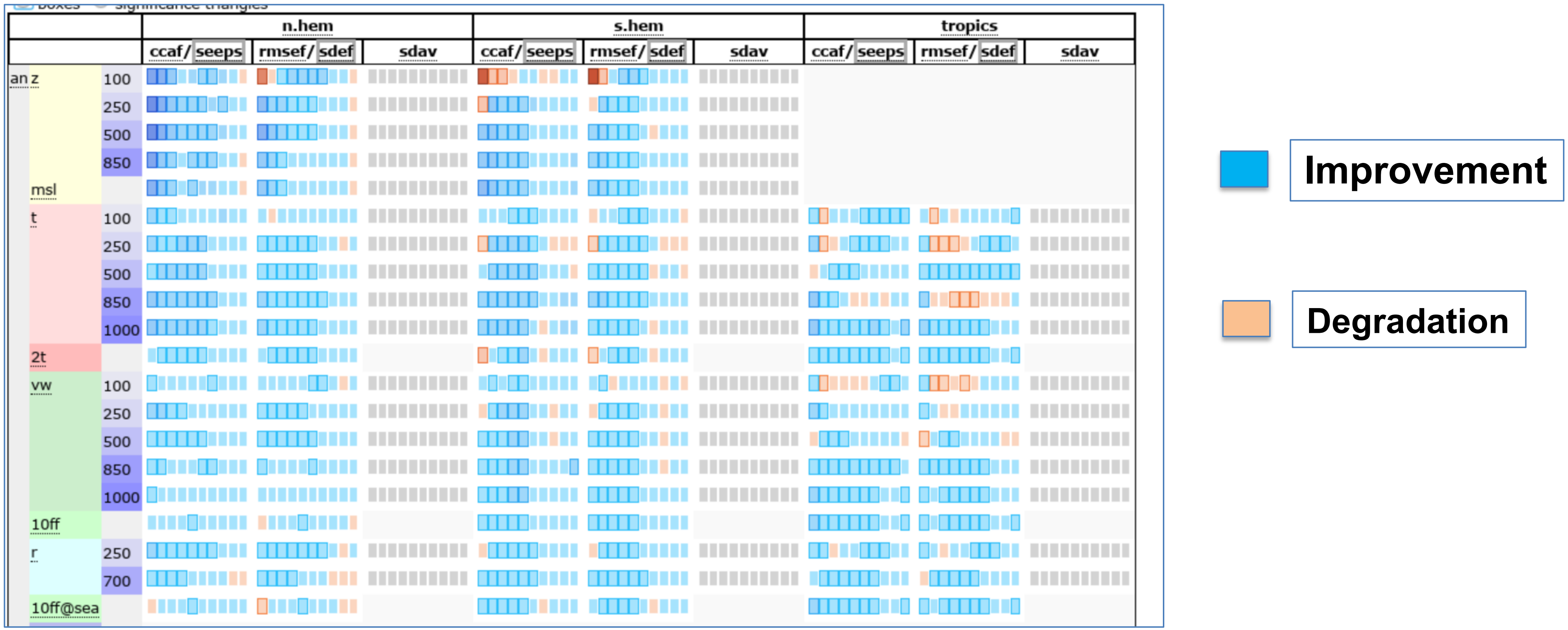
Note: 100% Baseline is current operational Weak Constraint 4D-Var

ECMWF

Bonavita & Laloyaux, 2020

Model Error Estimation and Correction in the IFS: Forecast Skill

- ANN in combination with Weak Constraint 4DVar improves the fit of observations to the model, both in the mean and in the random component.
- What can the ANN bring to forecast skill?



Model Error Estimation and Correction in the IFS

- The original idea of Bonavita and Laloyaux, 2020 has been further developed
- Specifically, the NN training can now be done inside the IFS 4DVar (NN 4D-Var, Farchi et al., 2023)
- Where are we with this line of development?

Model Error Estimation and Correction in the IFS

Training the NN parameters inside 4D-Var results in further forecast skill improvements for most variables.

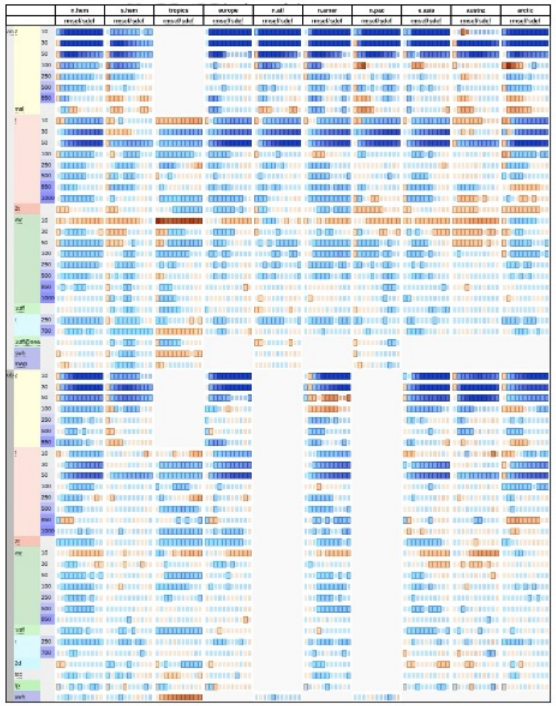


Figure: Score card 2022/06/03 to 2022/08/31. 12H assimilation window with NN model error correction trained **online**.

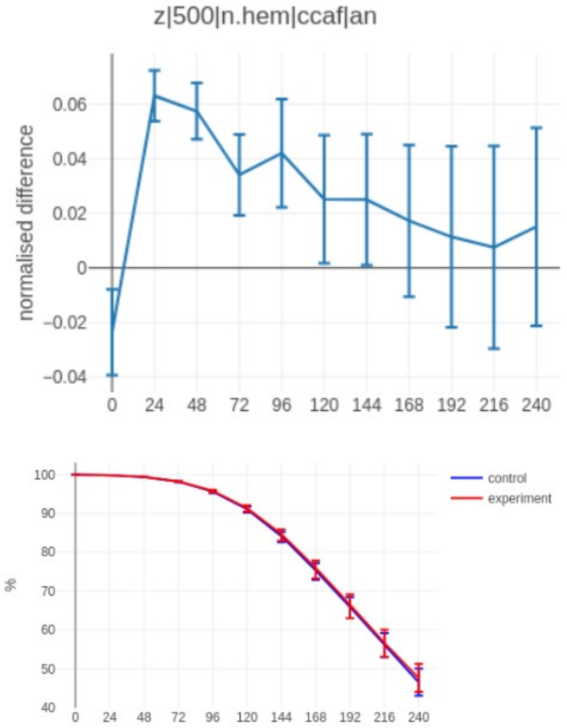


Figure: Z500 NH anomaly correlation. 2022/06/03 to 2022/08/31. 12H assimilation window with NN model error correction trained **online**.

From Marcin Chrust (ECMWF)

Model Error Estimation and Correction in the IFS

- Up to now we have been learning model errors optimising NN 4D-Var over the standard 12-hour assimilation window
- The problem with this approach is that some of the errors of the initial conditions (analysis) will get aliased into the model error estimates
- Initial conditions errors will be less of a factor if we do the NN 4D-Var optimisation over a longer assimilation window

Model Error Estimation and Correction in the IFS

Impact of online model error training with a 24h DA window.

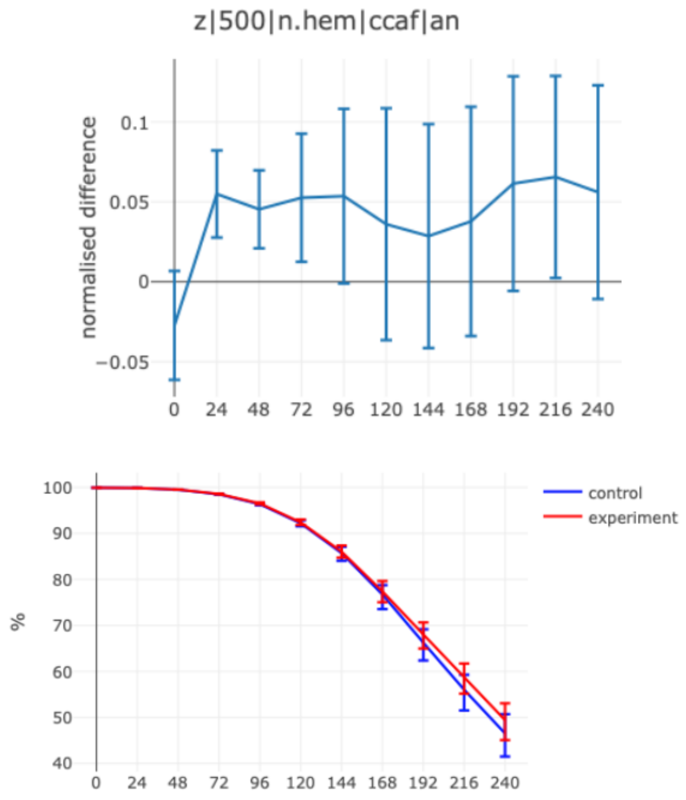


Figure: Z500 NH anomaly correlation. 2022/06/03 to 2022/07/28. 24H assimilation window with **offline** NN model error correction.

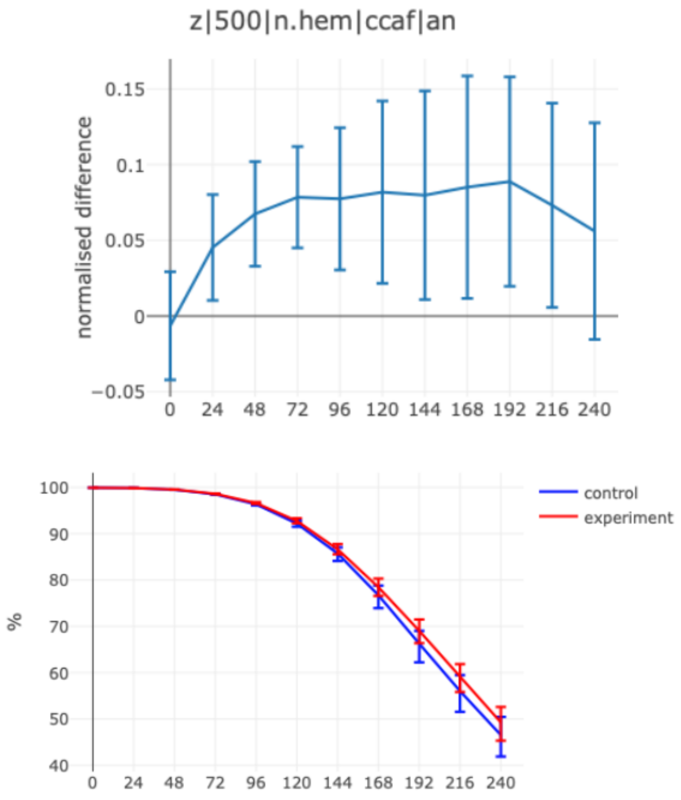


Figure: Z500 NH anomaly correlation. 2022/06/03 to 2022/07/28. 24H assimilation window with **online** NN model error correction.

From Marcin Chrust (ECMWF)

Model Error Estimation and Correction in the IFS

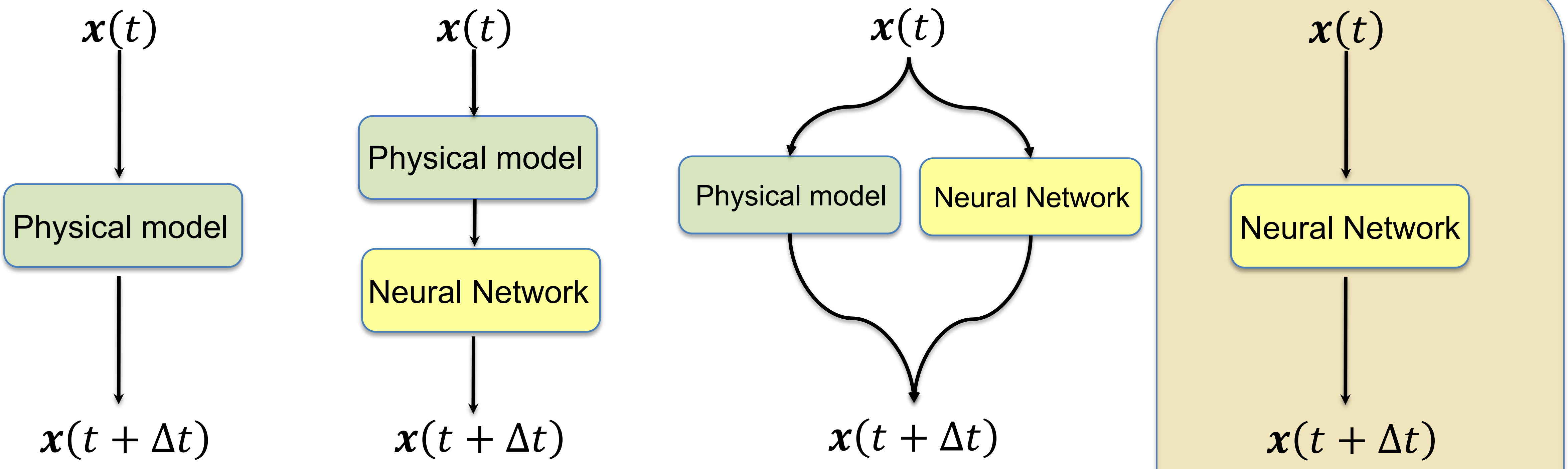
- Optimising over a 24-hour assimilation window shows clear improvement over 12-hour window (it recovers ~ half the gap in performance between the IFS and GraphCast/AIFS)
- Can we do better with an even longer assimilation window?
- Up to now we have used the original fully connected NN column model in NN 4D-Var. What if we use more expressive 3D NN architectures* (CNN, Graph networks, Transformers, etc.)?
- To be continued…

* Note that this is more complex than putting together NN blocks in Tensorflow/Pytorch, as it requires coding the NN machinery inside 4D-Var

A taxonomy of forecast models

$$\boldsymbol{x}(t+\Delta t) = M\big(\boldsymbol{x}(t)\big)$$

Physics-based → **Data-driven**



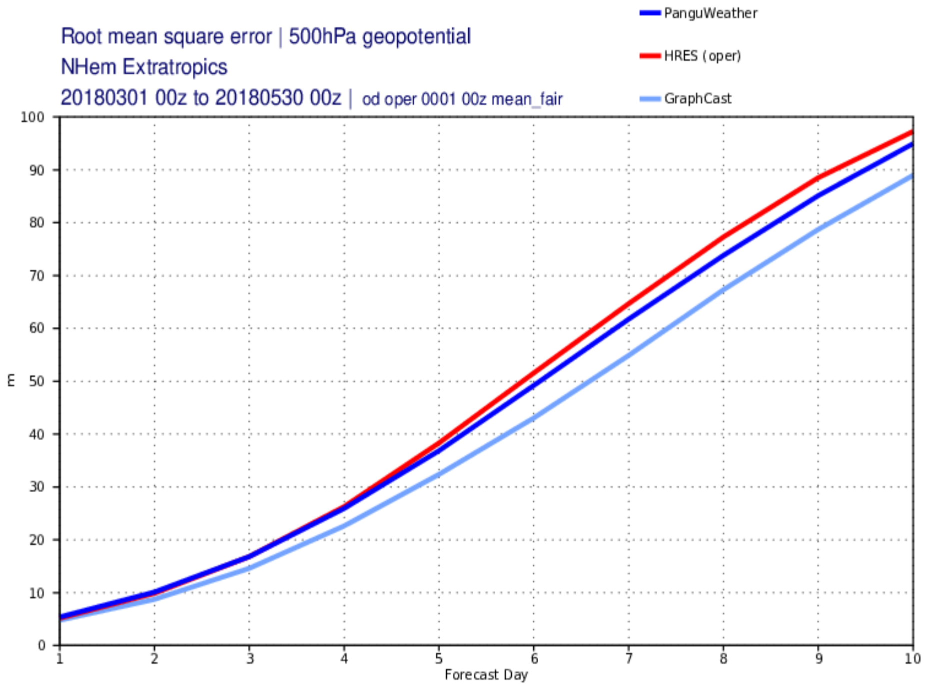
Standard NWP models
(IFS, ICON, FVM,…)

Model error correction
(e.g. Bonavita and Laloyaux, 2020; Farchi et al., 2022, 2023)

Partial model emulation
(e.g. Chantry et al., 2021 Kochkov et al., 2023, aka NeuroGCM)

Full model emulation
(e.g. Keisler, 2021, PanguWeather, 2022, GraphCast, 2023, …)

ML forecast models



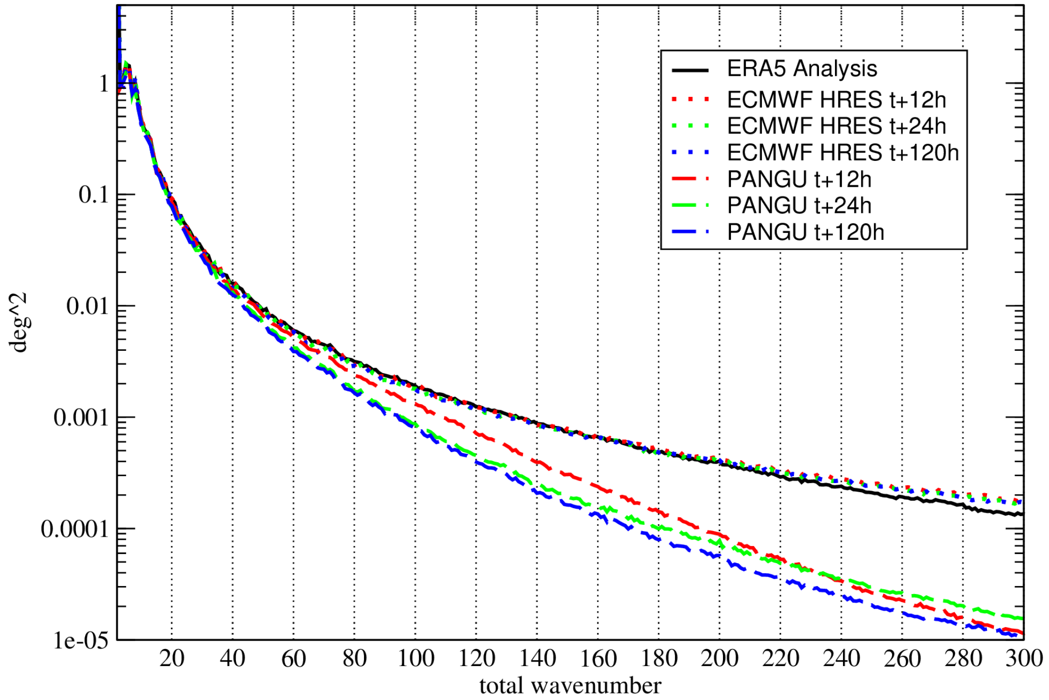
From Zied Ben Bouallegue (ECMWF)

- Field took off with Keisler, 2022, then in rapid succession 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)…
- All trained on ERA5 re-analysis (deterministic + EDA)
- Superior forecast scores, 3-4 order of magnitude cheaper to run (not to train!)
- What's not to like?

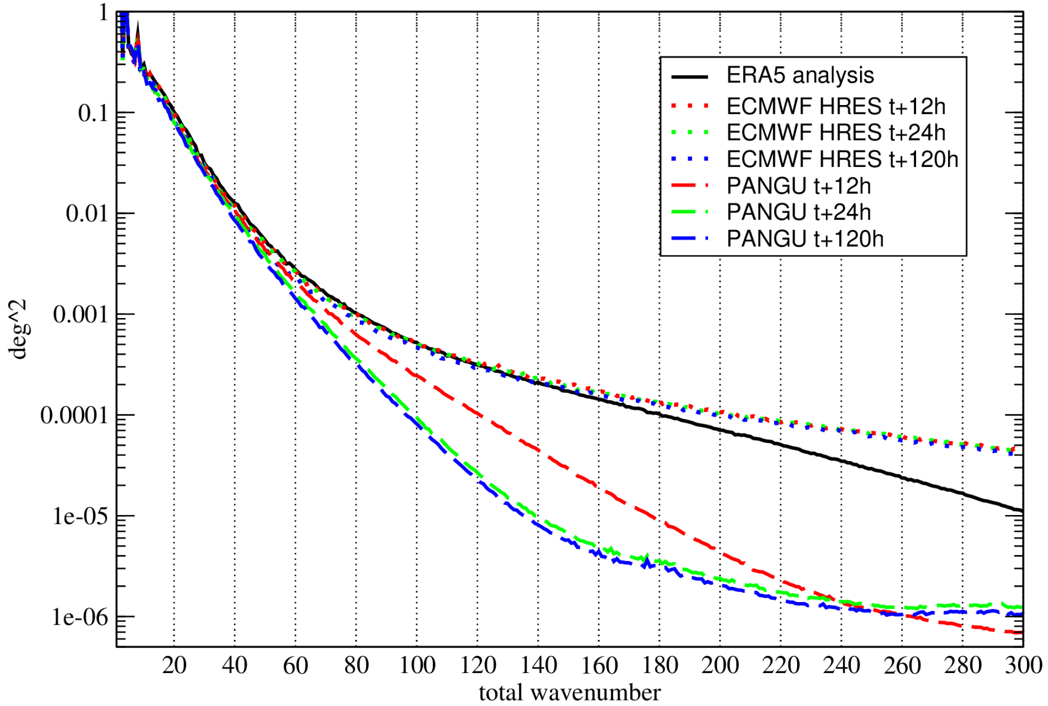
A look under the hood of MLWP models: Pangu-weather

- Are Pangu-weather (and other DLWP models) realistic atmosphere emulators?

Temperature 850hPA



Temperature 250hPA



Bonavita, M. (2023) On some limitations of data-driven weather forecasting models. GRL, under review. (arXiv: 10.48550/arXiv.2309.08473)

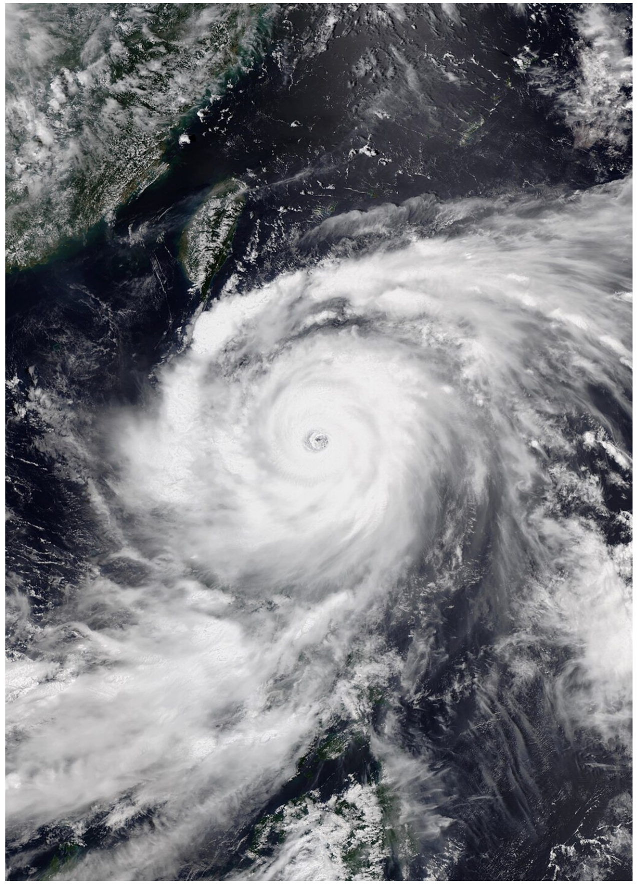
A look under the hood of MLWP models: Pangu-weather

- Are Pangu-Weather (and other DLWP models) realistic atmosphere emulators?

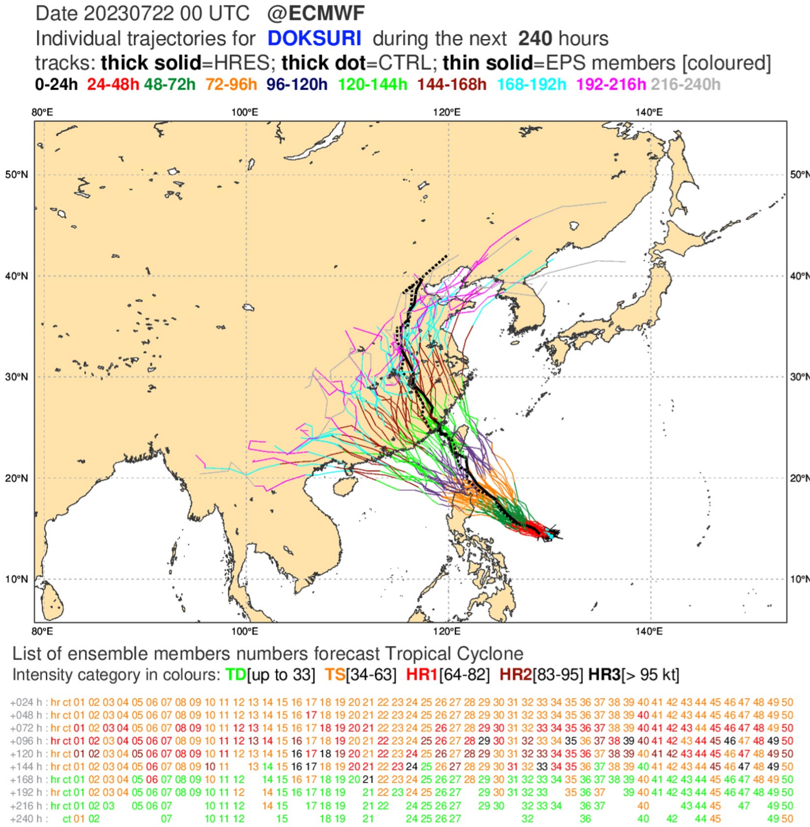
- Pangu-Weather forecasts show sharply decreased levels of variability wrto ERA5 analyses beyond ~wavenumber 50 (~400 km) from the start of the forecast
- Differently from IFS forecasts, which show consistent variability at all forecast ranges, PW forecast variability decreases with forecast range, noticeable jump at t+24h and beyond -> increasing "blurriness" of predictions

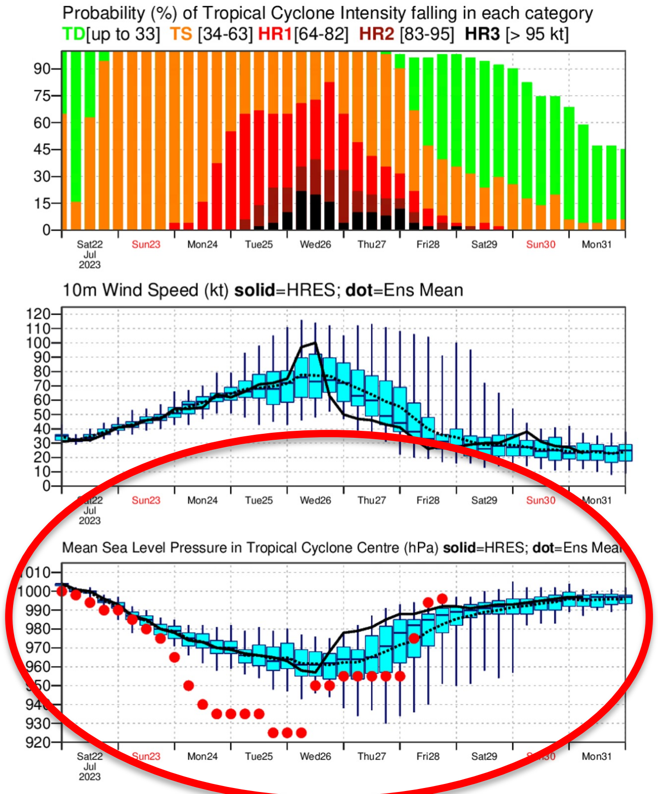
- Does it matter?

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



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.

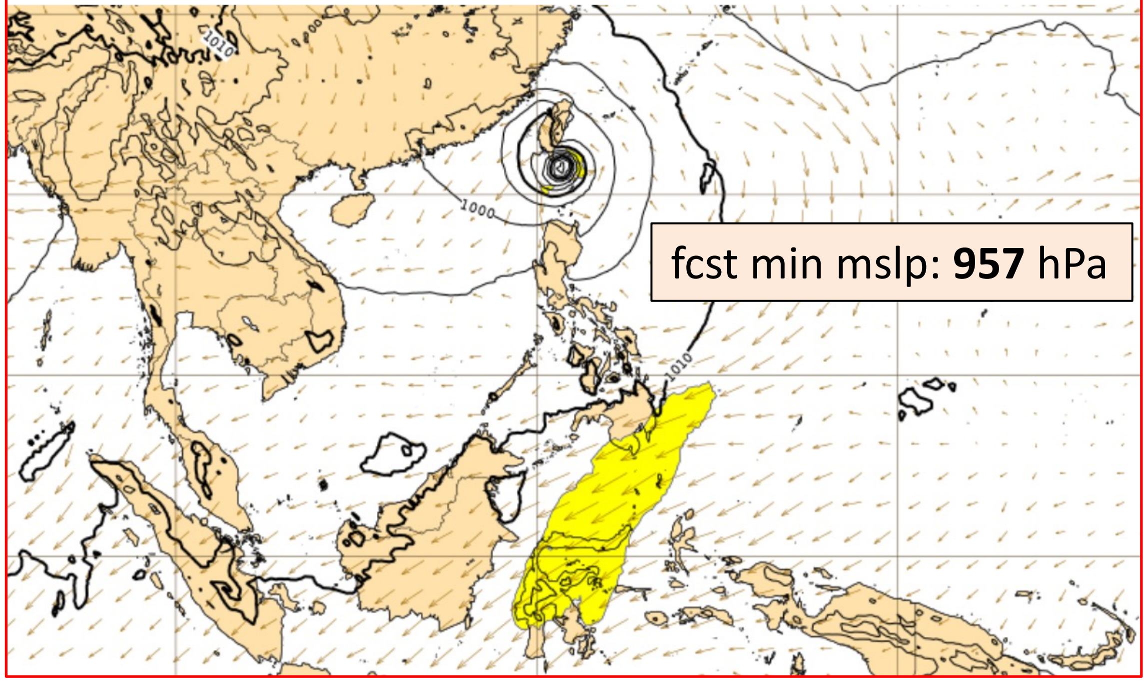




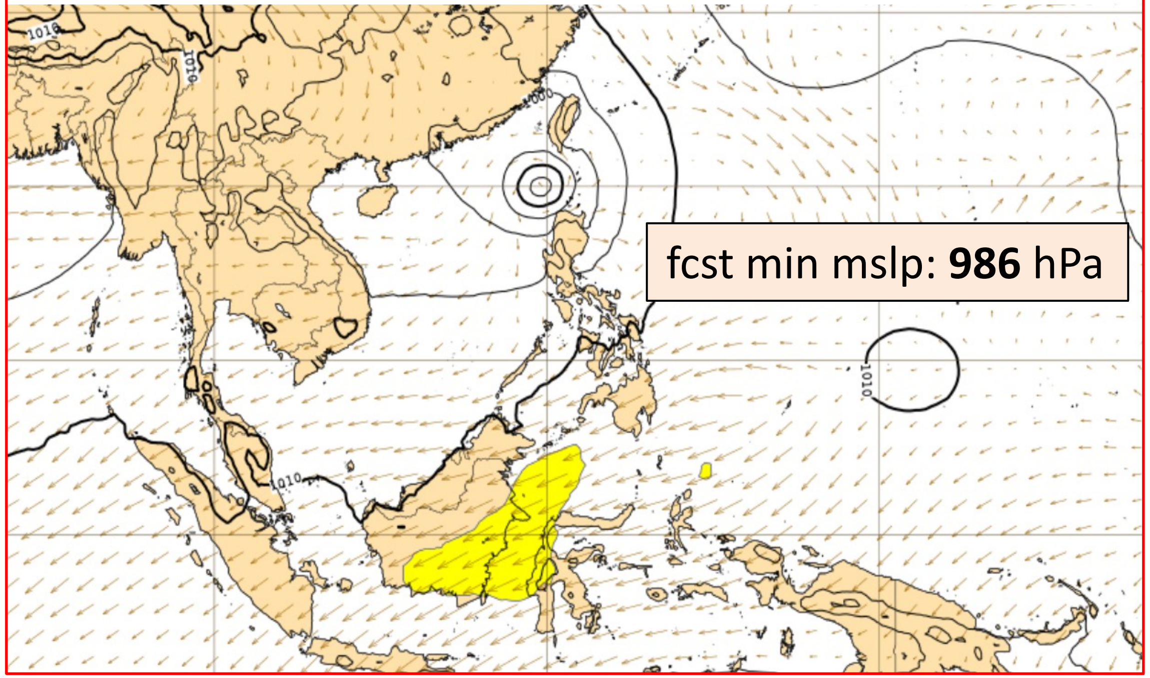
A look under the hood of MLWP models: Pangu-Weather

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

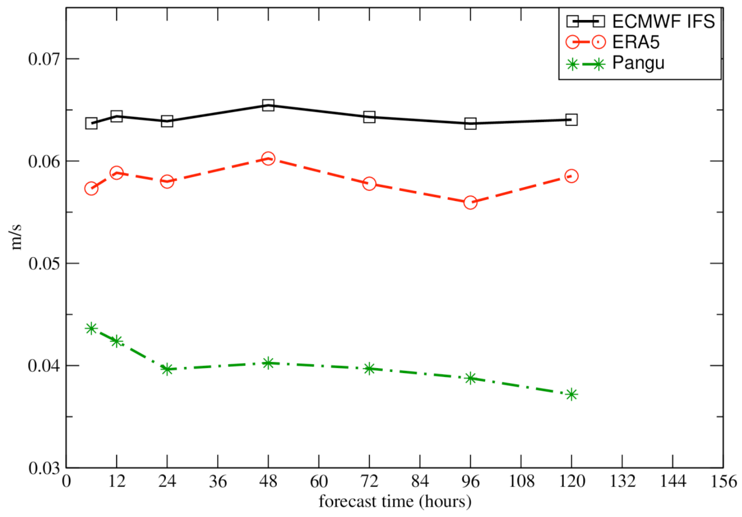


Pangu-Weather dynamical fields

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

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

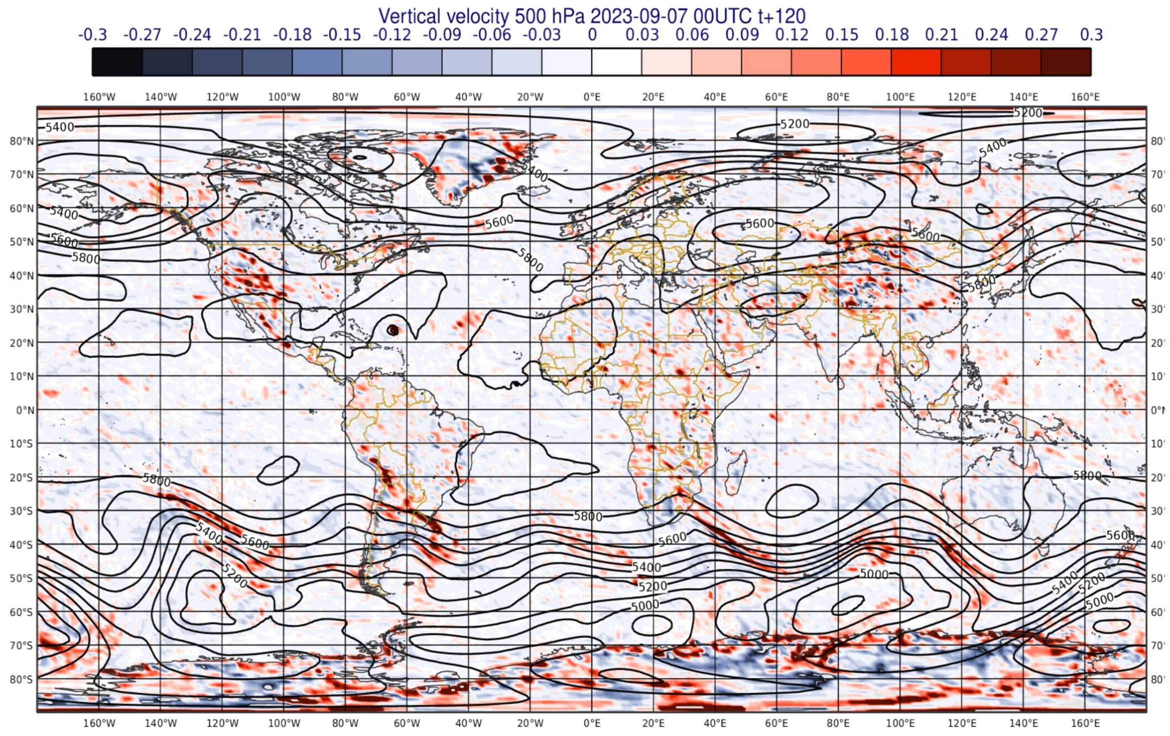
Unsurprisingly, the progressive reduction in the magnitude of the predicted divergence field leads to increasingly weak vertical velocity predictions:



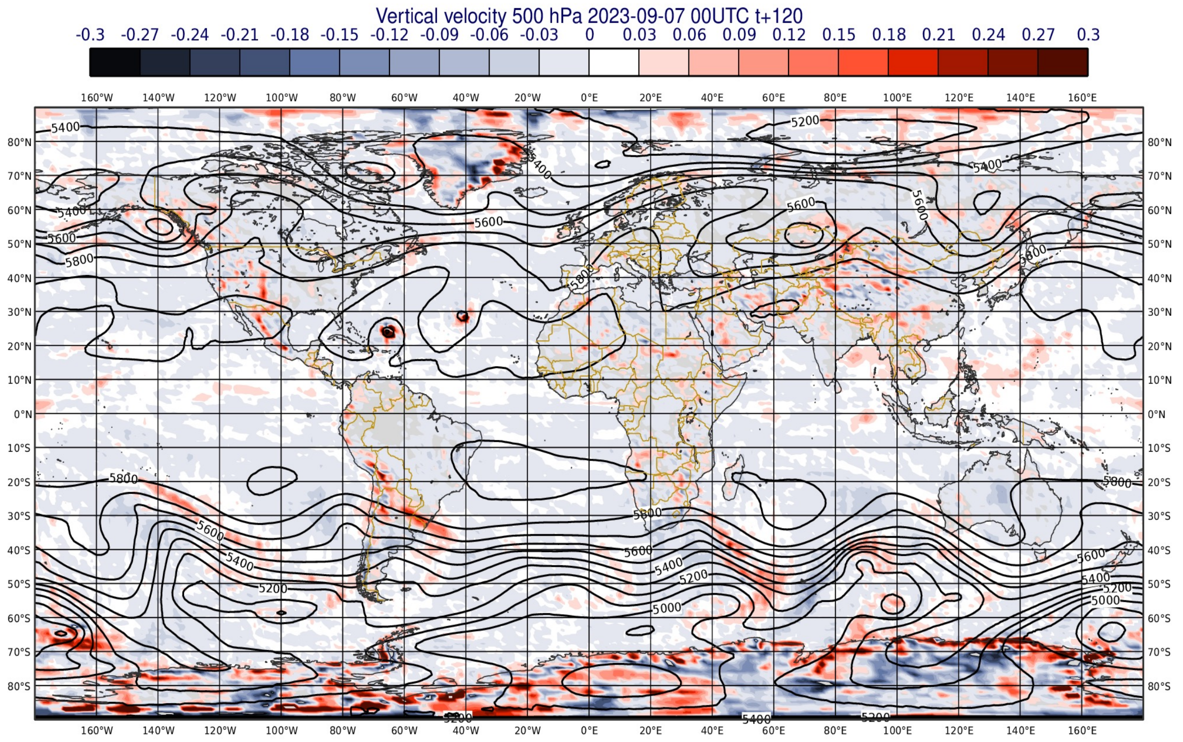
Evolution of stdev of fcst vertical velocity field
IFS, ERA5, Pangu

Pangu-Weather dynamical fields (3)

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



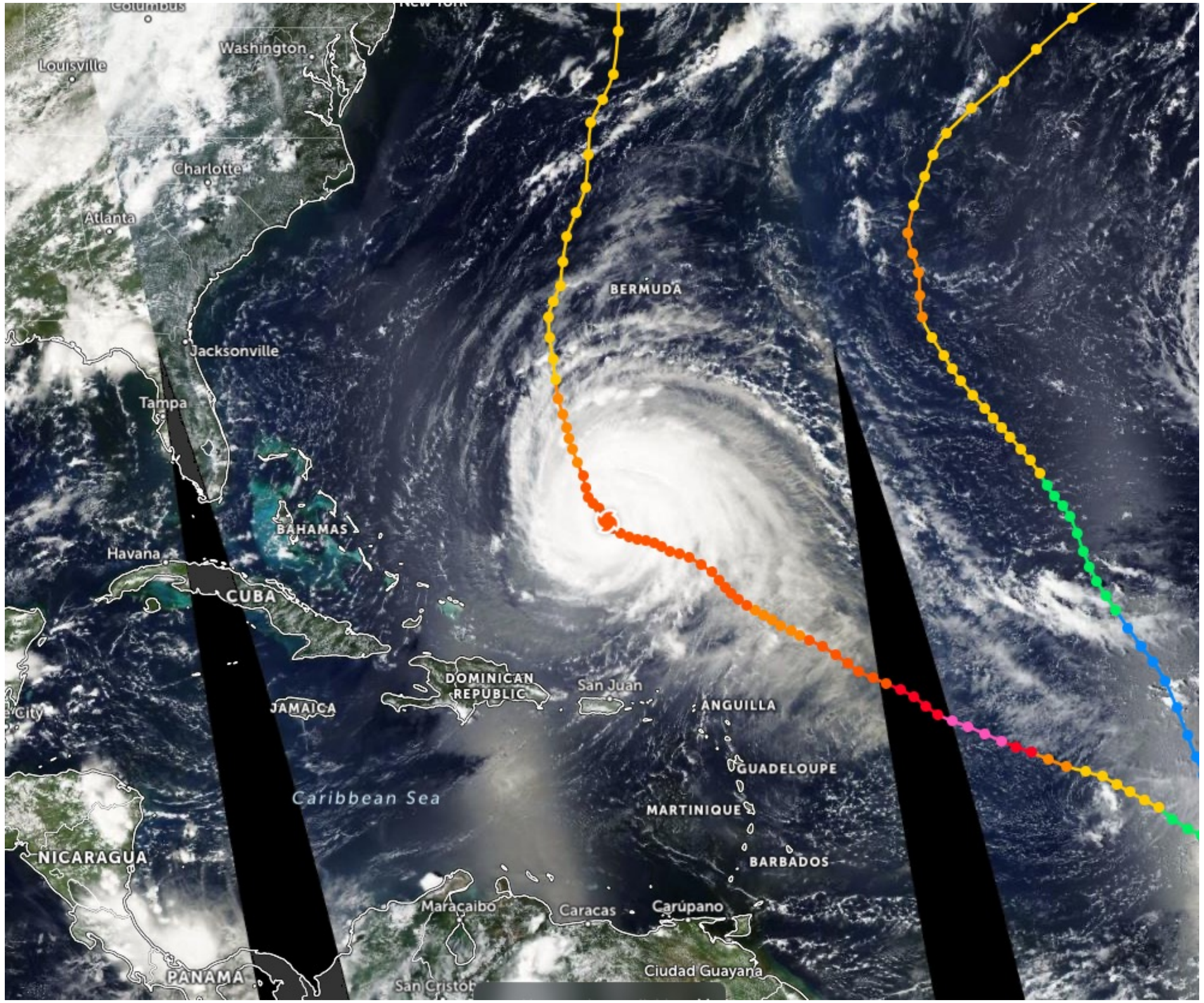
Pangu-Weather fcst vert. vel.
2023-09-07 00UTC t+120h



Pangu-Weather dynamical fields (3)

Hurricane Lee, 12 September 2023 01UTC

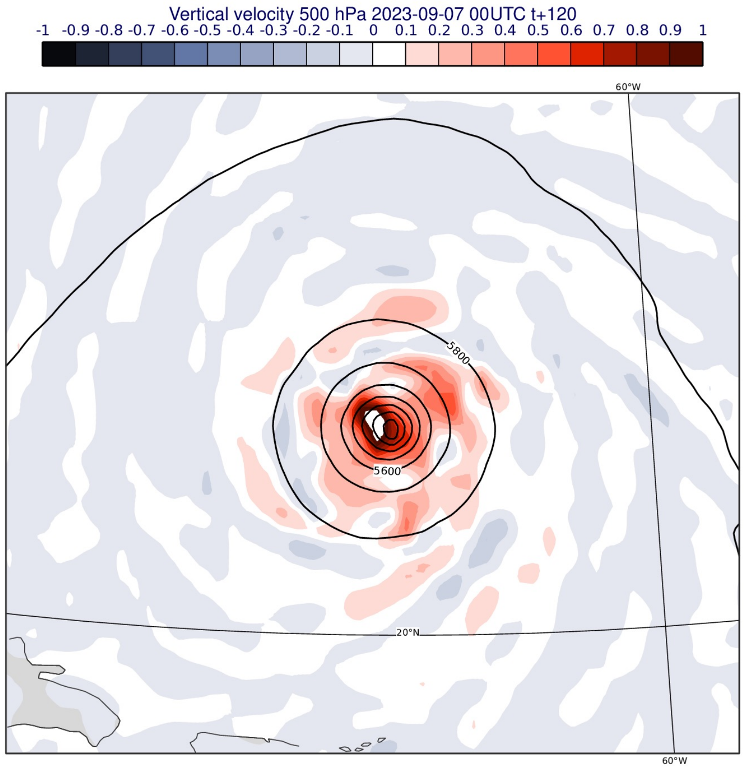
Strongest TC of the 2023 Atlantic Season so far, Category 3 at the time



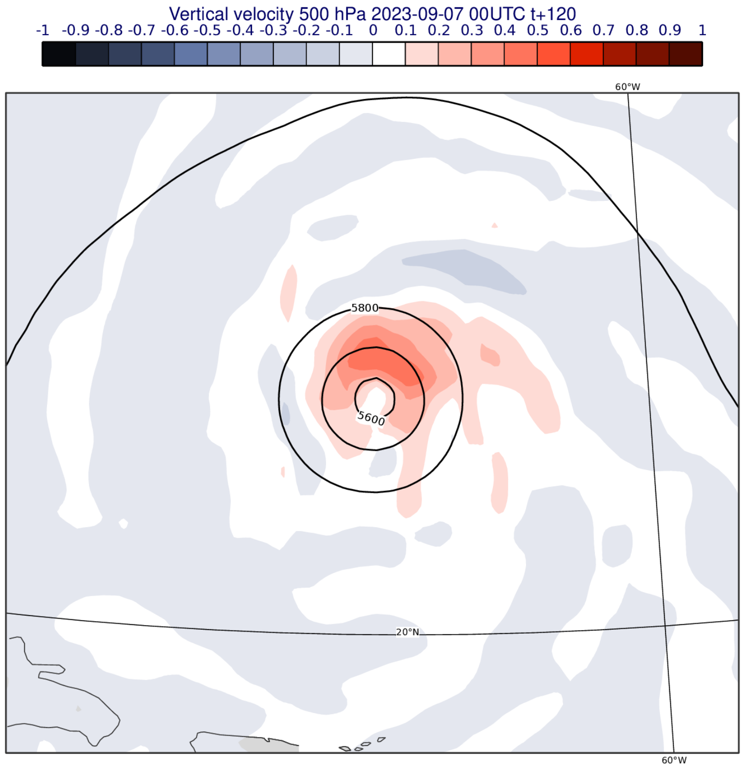
https://zoom.earth/storms/lee-2023/#map=satellite-hd

Pangu-Weather dynamical fields (3)

IFS vert. vel. (m/s) + Z500
2023-09-07 00UTC t+120h



ERA5 vert. vel. (m/s) + Z500
2023-09-07 00UTC t+120h

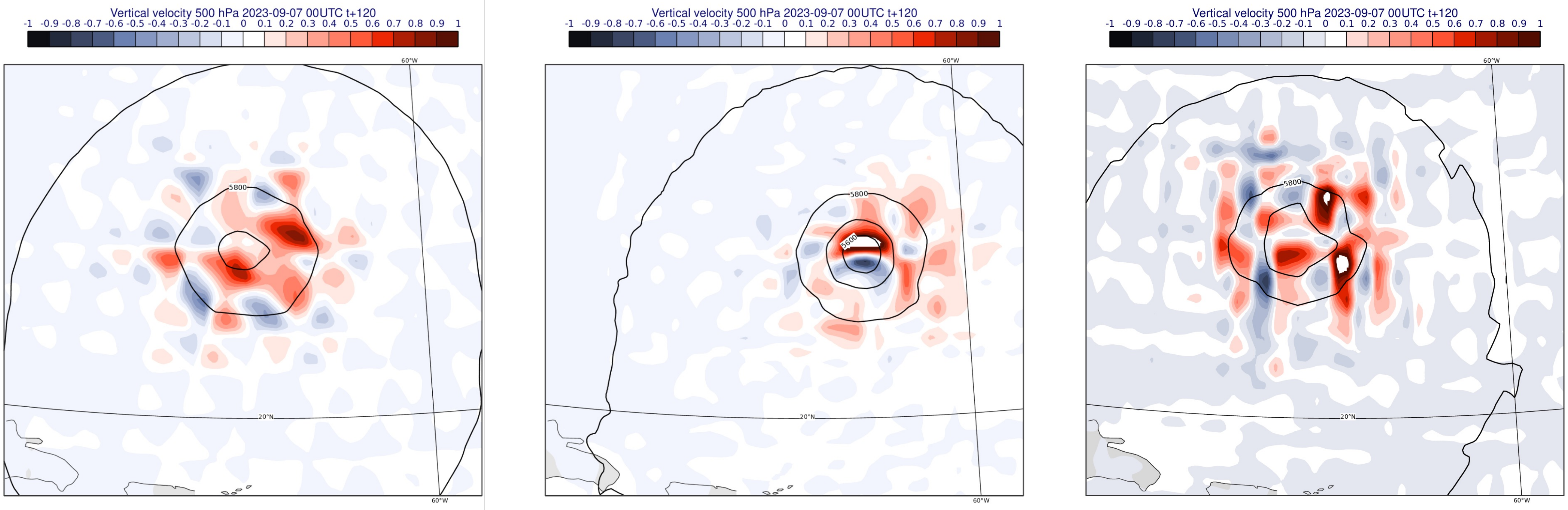


Pangu-Weather dynamical fields

AIFS v0.1 vert. vel. (m/s) + Z500
2023-09-07 00UTC t+120h

GraphCast vert. vel. (m/s) + Z500
2023-09-07 00UTC t+120h

Pangu-Weather vert. vel. (m/s) + Z500
2023-09-07 00UTC t+120h



Take-home Messages

- We are in the Big Data era! There is an unprecedented and rapidly growing amount of geophysical data ready to be used, both online (at ECMWF we use ~80 million obs a day in 4DVar: this is still less than 5% of the total amount of obs that reach the building every day!) and through reanalysis datasets
- This means that we have enough data not only to improve the initial state estimates, but also the models (both forecast model and forward models for observations)
- Machine Learning techniques can play a crucial role in tackling model deficiencies and improving predictive capabilities in both NWP and Climate.
- This field is evolving rapidly, next year's presentation will likely be quite different from this year's! Stay tuned…

Outlook

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

- Can MLWP extend to DA, i.e. from observations to forecasts? (no examples of this yet, contrary to what is sometimes claimed)

- 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: Can we match ML models forecast accuracy and provide physically credible forecasts?

- Too early to say which approach will prevail, but certainly things are moving at speed! 4th ECMWF-ESA Workshop on ML for Earth Observation and Prediction, Frascati, Rome, 7-10 May 2024 (https://www.ml4esop.esa.int)

Thanks for your attention!

Bocquet, M., Brajard, J., Carrassi, A. and Bertino, L. , 2019: Data assimilation as a learning tool to infer ordinary differential equation representations of dynamical models. Nonlinear Processes in Geophysics, 26 (3), 143–162. doi:10.5194/npg-26-143-2019.

Bocquet, M., Brajard, J., Carrassi, A. and Bertino, L. (2020). Bayesian inference of chaotic dynamics by merging data assimilation, machine learning and expectation-maximization. Foundations of Data Science, 2 (1), 55–80. doi:10.3934/fods.2020004

Bolton, T., & Zanna, L. (2019). Applications of deep learning to ocean data inference and subgrid parameterization. Journal of Advances in Modeling Earth Systems, 11, 376– 399. https://doi.org/10.1029/2018MS001472

Bonavita, M. and P. Laloyaux, 2020: Machine Learning for Model Error Inference and Correction, JAMES https://doi.org/10.1002/essoar.10503695.1

Bonavita, M., Arcucci, R., Carrassi, A., Dueben, P., Geer, A. J., Le Saux, B., Longépé, N., Mathieu, P., & Raynaud, L. (2021). Machine Learning for Earth System Observation and Prediction, Bulletin of the American Meteorological Society, 102(4), E710-E716, DOI: 10.1175/BAMS-D-20-0307.1

Bonavita, M.. (2021) Exploring the structure of time-correlated model errors in the ECMWF data assimilation system. Q J R Meteorol Soc, 147(739), 3454– 3471. Available from: https://doi.org/10.1002/qj.4137

Continued…

Bonavita, M. (2023). On some limitations of data-driven weather forecasting models. GRL, under review. https://doi.org/10.48550/arXiv.2309.08473

Brajard, J., Carrassi, A., Bocquet, M. and Bertino, L. (2020). Combining data assimilation and machine learning to emulate a dynamical model from sparse and noisy observations: A case study with the lorenz 96 model. Journal of Computational Science, 44, 101171. doi:10.1016/j.jocs.2020.101171.

Brenowitz, N. D., & Bretherton, C. S. (2018). Prognostic validation of a neural network unified physics parameterization. Geophysical Research Letters, 17, 6289–6298. https://doi.org/10.1029/2018GL078510

Brenowitz, N. D., and Bretherton, C. S. (2019), Spatially Extended Tests of a Neural Network Parametrization Trained by Coarse-Graining, J. Adv. Model. Earth Syst., 11, 2728– 2744. https://doi.org/10.1029/2019MS001711.

Chantry, M., S. Hatfield, P. Dueben, I. Polichtchouk, and T. Palmer. Machine learning emulation of gravity wave drag in numerical weather forecasting. Journal of Advances in Modeling Earth Systems, 13(7):e2021MS002477, 2021. doi: https://doi.org/10.1029/2021MS002477

Farchi, A., Laloyaux, P., Bonavita, M. & Bocquet, M.(2021a) Using machine learning to correct model error in data assimilation and forecast applications. Q J R Meteorol Soc, 147(739), 3067– 3084. Available from: https://doi.org/10.1002/qj.4116

Farchi, A., M. Bocquet, P. Laloyaux, M. Bonavita and Q. Malartic, (2021b)A comparison of combined data assimilation and machine learning methods for offline and online model error correction, Journal of Computational Science, Volume 55, 2021,101468, ISSN 1877-7503, https://doi.org/10.1016/j.jocs.2021.101468.

Keisler, R., (2021) Forecasting Global Weather with Graph Neural Networks, https://arxiv.org/pdf/2202.07575.pdf

Kutz, J. N. (2017). Deep learning in fluid dynamics. Journal of Fluid Mechanics, 814, 1–4.

Laloyaux, P, Bonavita, M, Dahoui, M, et al. 2020: Towards an unbiased stratospheric analysis. Q J R Meteorol Soc. 146: 2392–2409. https://doi.org/10.1002/qj.3798

Ling, J., Kurzawski, A., & Templeton, J. (2016). Reynolds averaged turbulence modelling using deep neural networks with embedded invariance. Journal of Fluid Mechanics, 807, 155–166.

Navon, I.M., 1998: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography, Dynamics of Atmospheres and Oceans, Volume 27, Issues 1–4, 55-79

Rasp, S., Pritchard, M. S., & Gentine, P. (2018). Deep learning to represent subgrid processes in climate models. Proceedings of the National Academy of Sciences of the United States of America, 115(39), 9684–9689. https://doi.org/10.1073/pnas.1810286115

Ruiz, J.J., Pulido, M. and T. Miyoshi, 2013: Estimating Model Parameters with Ensemble-Based Data Assimilation: A Review, Journal of the Meteorological Society of Japan. Ser. II, 91, 2, 79-99. https://doi.org/10.2151/jmsj.2013-201

Additional Material

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

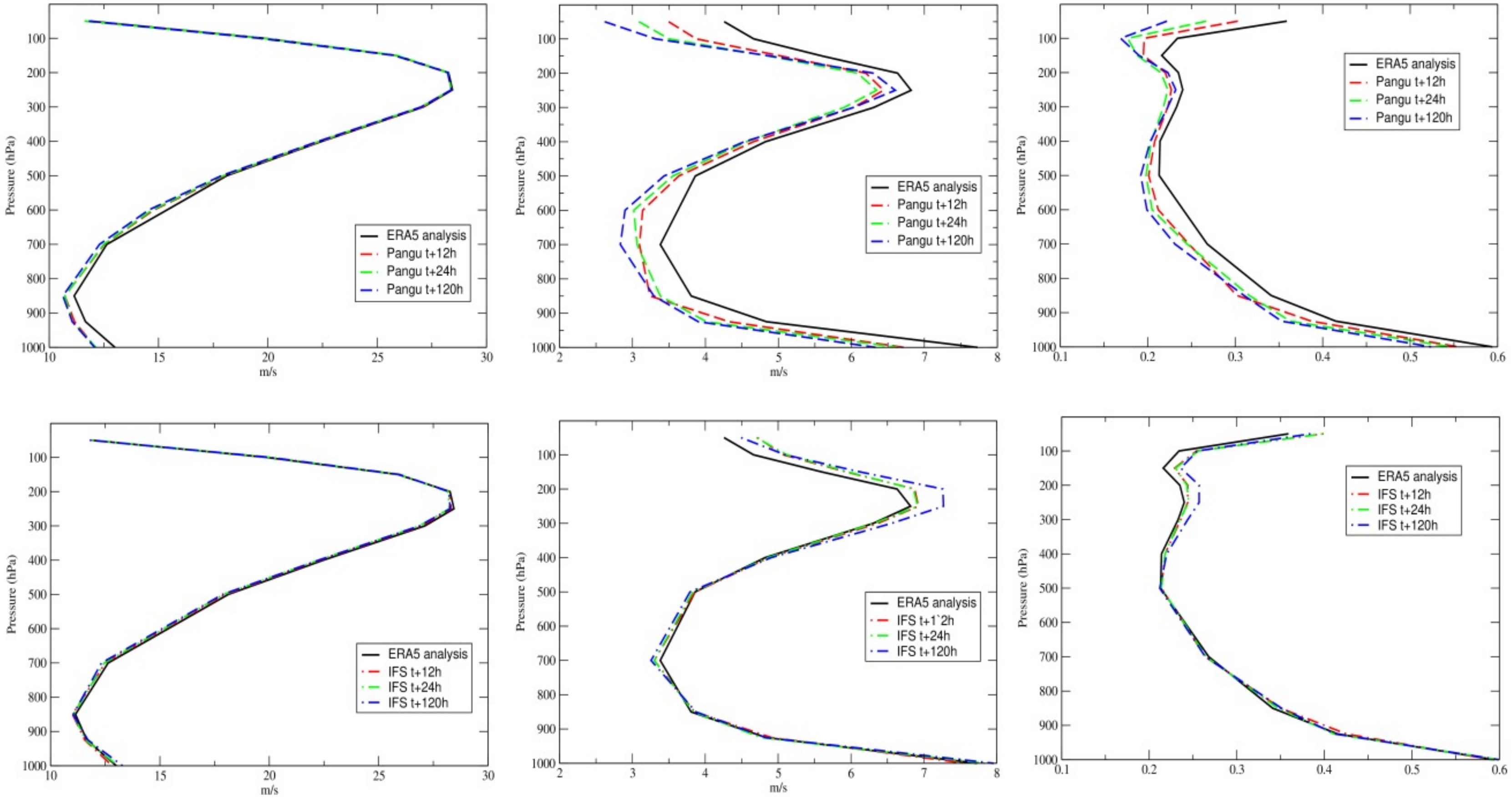
Unrealistic forecast energy spectra imply dynamically inconsistent forecast fields

Pangu-Weather dynamical fields (1)

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

Pangu

IFS

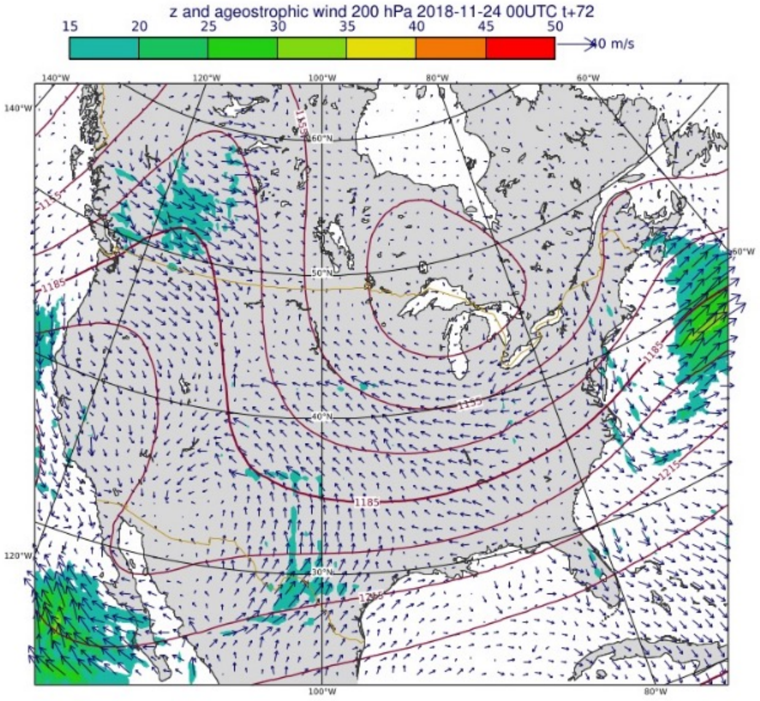


$|\boldsymbol{V}_g|$ $|\mathbf{V}_{ag}|$ $|\mathbf{V}_{ag}|/\,|\mathbf{V}_g|$

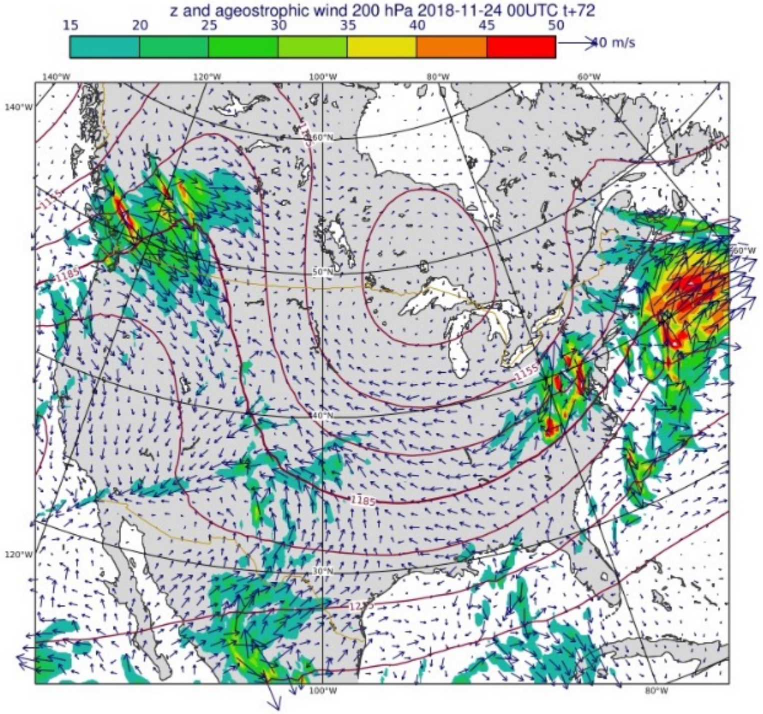
Pangu-Weather dynamical fields (1)

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



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

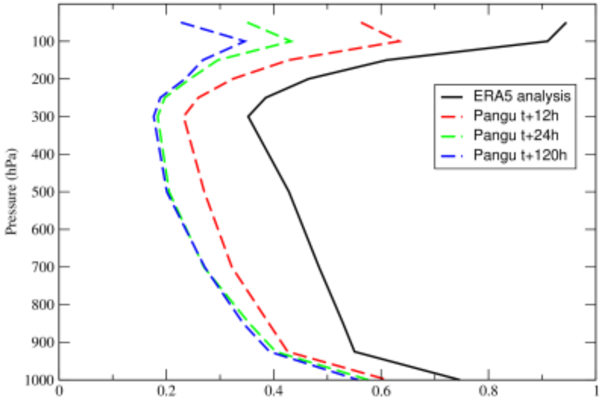
Pangu-Weather dynamical fields (2)

Vorticity and divergence decomposition of the circulation

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

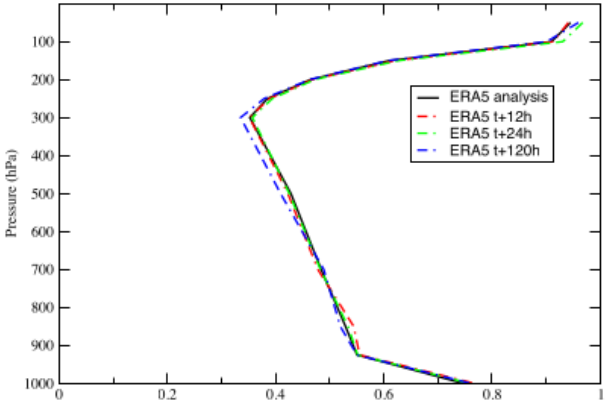
$$\nabla^2\chi = \delta, \ \nabla^2\psi = \zeta$$

Pangu



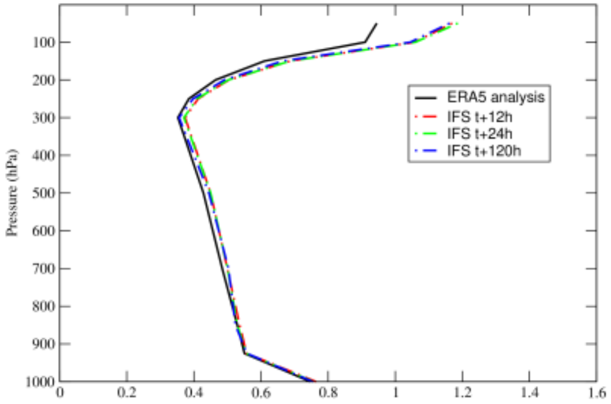
δ/ζ

ERA5-fcst



δ/ζ

IFS



δ/ζ