



Environment and
Climate Change Canada

Environnement et
Changement climatique Canada

ADVANCING GLOBAL WEATHER PREDICTION AT ECCC USING SPECTRAL-NUDGING-BASED HYBRID NWP-MLWP MODELLING

Leo Separovic, Syed Zahid Husain, Christopher Subich, Jean-François Caron, and Jing Yang

Numerical Modelling and Prediction Research Division
Environment and Climate Change Canada (ECCC)

5th ECMWF-ESA Machine Learning Workshop
Bologna, Italy, 13-17 April 2026



Canada 

MOTIVATION

- **Current hybrid models from ECCO suffer from two limitations.**
 - Negligible improvements at short lead times (0-72 h) compared to medium range.
 - Relatively small improvements in near-surface meteorology at all lead times.
- **What is needed to improve short lead?**
 - **Primary factor:** increased nudging cutoff wavenumber (incompatible with MSE/MAE-trained AI models that predict the distribution mean/median and lack sharpness).
 - **Secondary factor:** Increased AI model output frequency.
- **What is needed to improve the near-surface prediction?**
 - More AI model levels in the BL to allow for nudging.
- **Focus of today's presentation is short lead accuracy.**

MESOSCALE SMOOTHING ISSUE WITH DETERMINISTIC MSE-BASED TRAINING

- For a global field, defined over a Gaussian grid, spectrally decomposed *MSE*, using spherical harmonics-based expansion, can be expressed as

$$MSE = \sum_n \sigma_{y,y}(n) [(1 - \Gamma(n))^2 + 2\Gamma(n)(1 - \rho(n))].$$

- Here:

- $\Gamma(n) = \sqrt{\sigma_{x,x}(n)/\sigma_{y,y}(n)}$ is the **spectral amplitude ratio**,
 - $\rho(n) = \sigma_{x,y}(n)/\sqrt{\sigma_{x,x}(n)\sigma_{y,y}(n)}$ is the **spectral coherence**,
 - $\sigma_{x,x}(n)$ and $\sigma_{y,y}(n)$ are the forecast and analysis PSDs at total spherical wavenumber n , and $\sigma_{x,y}(n)$ is the corresponding cross-spectral density.
- For a given predictability limit defined by $\rho(n)$, the MSE is strictly minimized when $\Gamma(n) = \rho(n)$, i.e., when the spectral amplitude ratio becomes equal to spectral coherence for every n .
 - As **fine scales** quickly lose predictability with increasing lead time, their coherence drops ($\rho(n) \rightarrow 0$). To avoid the "double penalty" of placing a sharp feature in the wrong location, MLWP models trained to minimize MSE, tend to reduce their own variance ($\Gamma(n) \rightarrow 0$) for these unpredictable scales. The result is a **smoothed forecast** that matches only the predictable component of the variance, given a wavenumber n .

ADDRESSING SMOOTHING WITH AMSE

- To address the smoothing, we proposed $AMSE^*$ (**Adjusted MSE**), where we break the interaction between the spectral amplitudes and coherence contained in the second term in the MSE :

$$AMSE = \sum_n \sigma_{y,y}(n) [(1 - \Gamma(n))^2 + 2 \max(\Gamma^2(n), 1) (1 - \rho(n))],$$

where $\max(\Gamma^2(n), 1)$ is used to make the metric symmetric.

- $AMSE$ is zero if and only if $x = y$ (as in the case of MSE).
- Minimizing $AMSE$ will always point to the direction of increased coherence ($\rho(n) \rightarrow 1$) as well as a correct spectral amplitude, i.e., $\Gamma(n) \rightarrow 1$.
- Spectrally decomposing $AMSE$ permits gradient descent to work for each wavenumber to ensure that the negative interaction between the spectral amplitudes and coherence be avoided at all spatial scales.

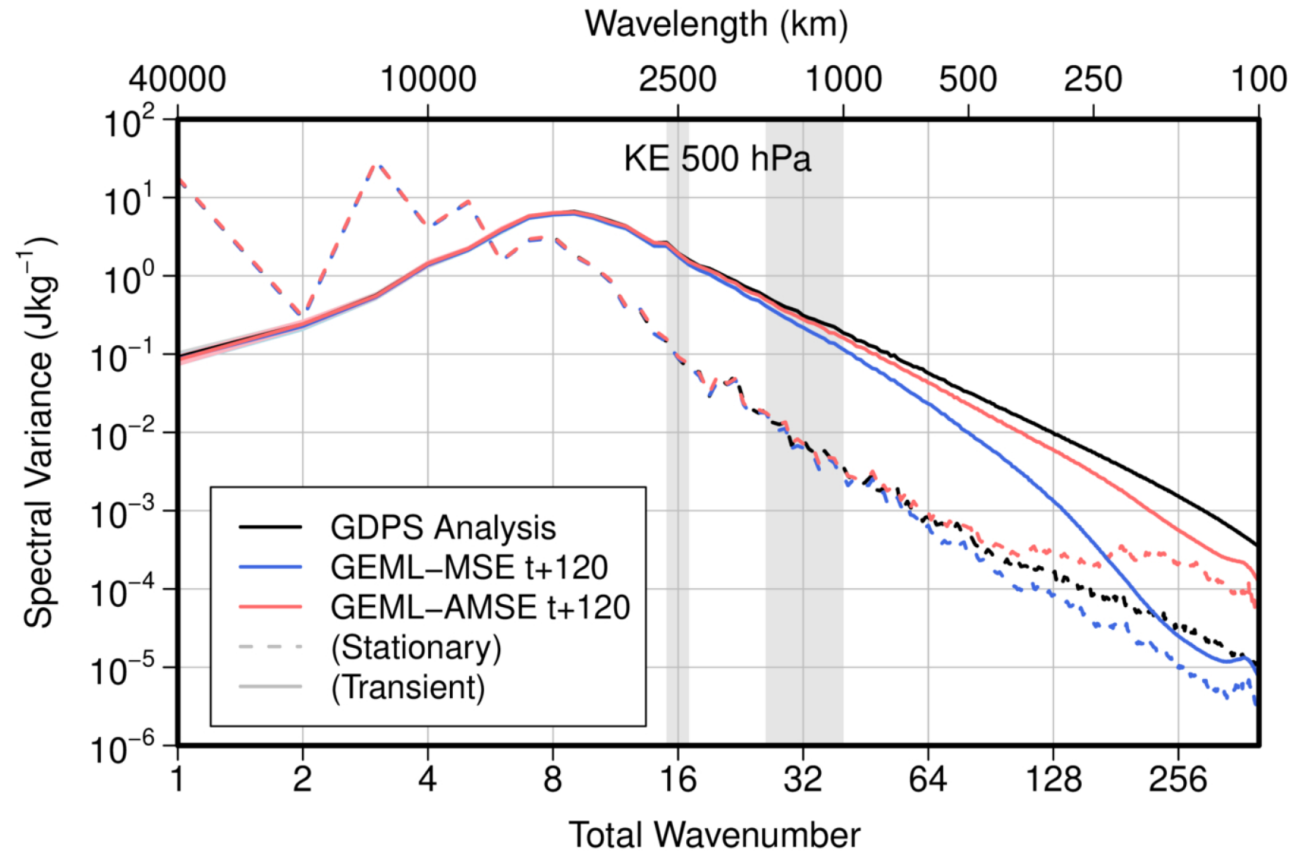
AMSE-FINETUNED GEML

- A new 1/4° **GEML** version (Global Environmental eMuLator – based on Google DM’s *GraphCast* open-source code) has been produced, aimed at reducing the mesoscale smoothing:
 - Trained with ERA5-1979-2015 data (**MSE loss**);
 - Fine-tuned with:
 1. HRES-2016-2018 (**MSE loss**),
 2. GDPS-2019-2023 (**MSE loss**),
 3. GDPS-2019-2023 (**AMSE loss**).
- This gave us two GEML versions, both of which have seen ECCC’s operational GDPS analyses:
 - **GEML-MSE** corresponding to **step (2)** above – to become operational in May 2026* as a part of **medium-range** Global Deterministic Prediction System with Spectral Nudging (**GDPS-SN**);
 - **GEML-AMSE** corresponding to **step (3)** above – to be explored for **84-h** global prediction system that provides the LBCs (and could serve as a nudging reference) for **short-range limited-area km-scale** prediction systems.

* Husain *et al.* (2025) [Leveraging Data-Driven Weather Models for Improving Numerical Weather Prediction Skill through Large-Scale Spectral Nudging in: Weather and Forecasting Volume 40 Issue 9 \(2025\)](#)

GEML-AMSE VS. GEML-MSE: POWER SPECTRA

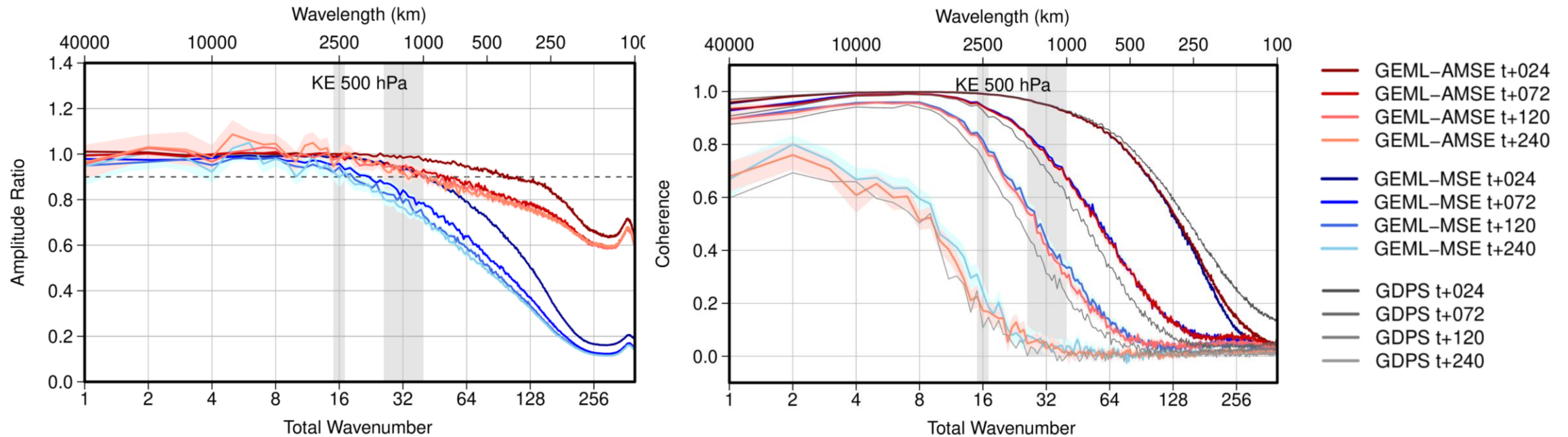
- AMSE training significantly improves GEML effective resolution.
 - In a strict sense, resolution of transient eddies increases from T16 to T32.
- There is some variance inflation of the stationary eddies (seasonal climatology) with respect to GDPS analyses.
 - Relative contribution of this noise to the total spectral variance is, however, very small.



GEML-AMSE VS. GEML-MSE: SPECTRAL AMPLITUDE RATIO AND COHERENCE

- Despite better resolving fine-scale details, GEML-AMSE tightly matches the coherence of GEML-MSE across the entire range of scales, giving a clear advantage to GEML-AMSE.
- At around forecast day 5, large-scale GEML-AMSE's coherence begins to fall behind GEML-MSE.

Verification against the GDPS analysis
Transient anomaly KE500 spectral coherence



EXPERIMENTAL SETUP – SPECTRAL NUDGING

- Configuration **GDPS-SN-MSE**:

- Aimed for medium-range
- Nudging towards GEML-MSE
- Standard spectral nudging configuration:

- Spn_nudge_HU_L = **.false.**
- Spn_cutoff_scale_large = **2750.**
- Spn_cutoff_scale_small = **2200.**
- Spn_start_lev = **0.85**
- Spn_const_lev_bot = **0.70**

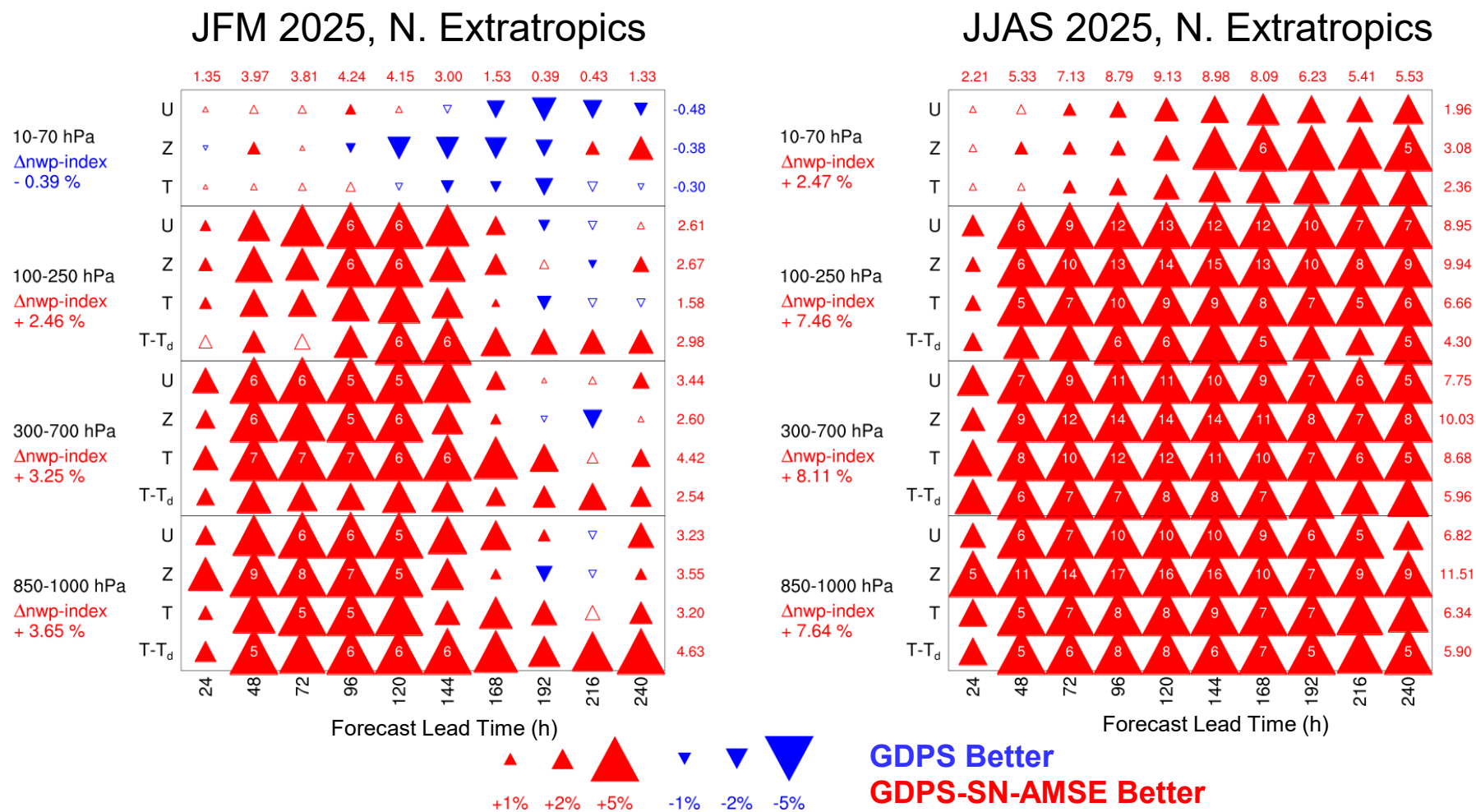
- Configuration **GDPS-SN-AMSE**:

- Aimed for short-range
- Nudging towards GEML-AMSE
- Nudging moisture, reduced cutoff wavelength and extension into BL:

- Spn_nudge_HU_L = **.true.**
- Spn_cutoff_scale_large = **1000. -> 1500.**
- Spn_cutoff_scale_small = **500 -> 1000.**
- Spn_start_lev = **0.95**
- Spn_const_lev_bot = **0.85**

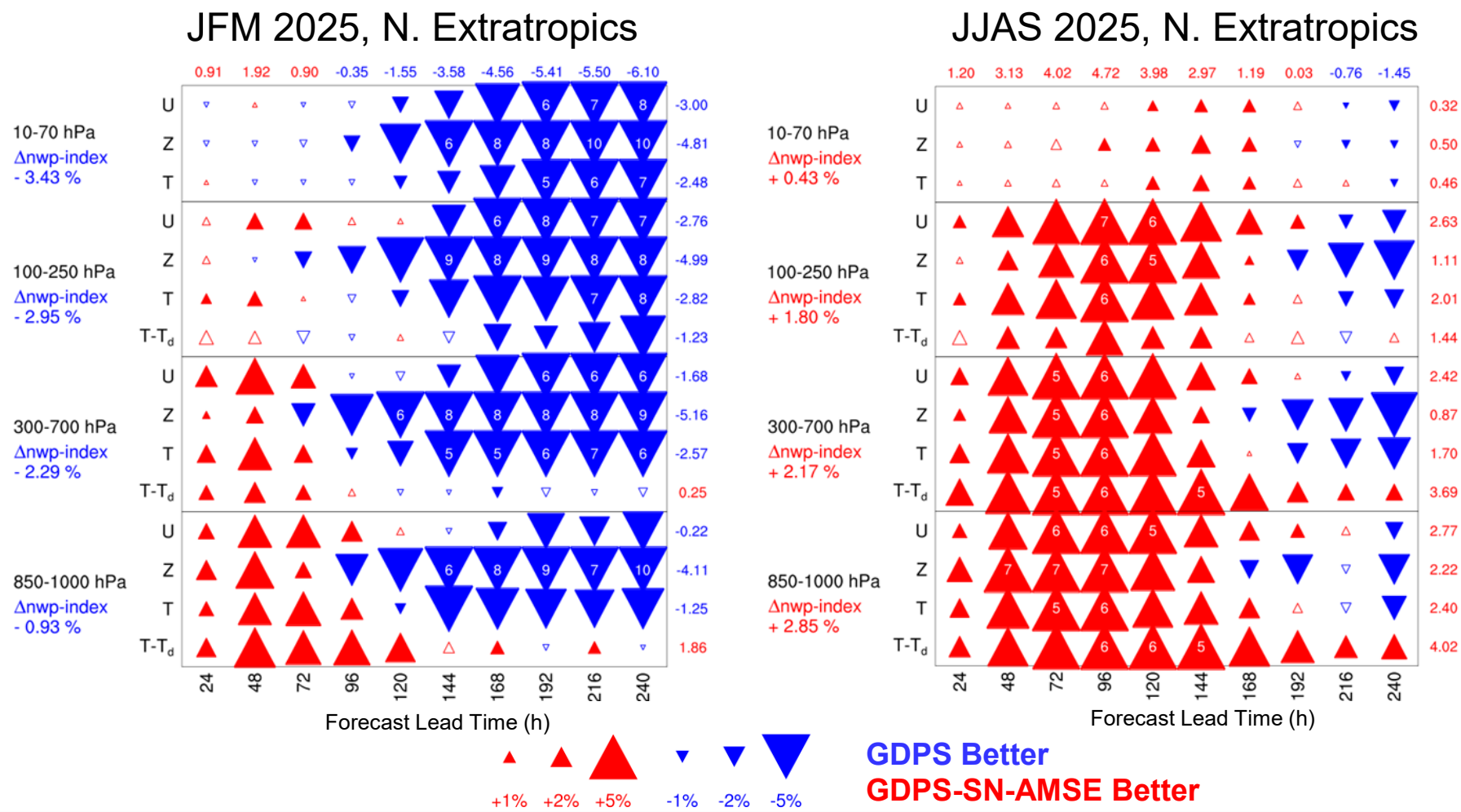
GDPS-SN-AMSE VS. GDPS: VERIFICATION AGAINST RADIOSONDES

- Compared to the operational **GDPS**, the **GDPS-SN-AMSE** configuration brings 5-15 % RSME reductions throughout the troposphere at days 2-5.



GDPS-SN-AMSE VS. GDPS-SN-MSE: VERIFICATION AGAINST RADIOSONDES

- Compared to **GDPS-SN-MSE**, the **GDPS-SN-AMSE** configuration brings up to 10 % RSME deteriorations beyond day 3 (6), during boreal winter (summer).
- It is still superior at short range.



TOWARDS HYBRIDIZATION OF KM-SCALE SYSTEMS

- The improvements at short lead times are genuine: they are not associated with excessive smoothing and double penalty reduction (not shown).
- This strongly motivates us to test spectral nudging in the 10-km global system that provides LBCs for kilometer-scale **limited-area** systems:
 - **HRDPS** (High-Resolution Deterministic Prediction System) at **2.5-km** resolution, focused over Canada.
 - **CAPS** (Canadian Arctic Prediction System) at **3-km** resolution, coupled with the ocean and sea-ice model, large domain covering the entire Arctic and high-latitude regions.
- Conditional on global results, internal spectral nudging towards global predictions can easily be activated in HRDPS/CAPS to try to leverage these improvements in the two systems.
- We are still experimenting with optimization of GEML finetuning with AMSE, as well as spectral nudging configurations for short-range predictions.

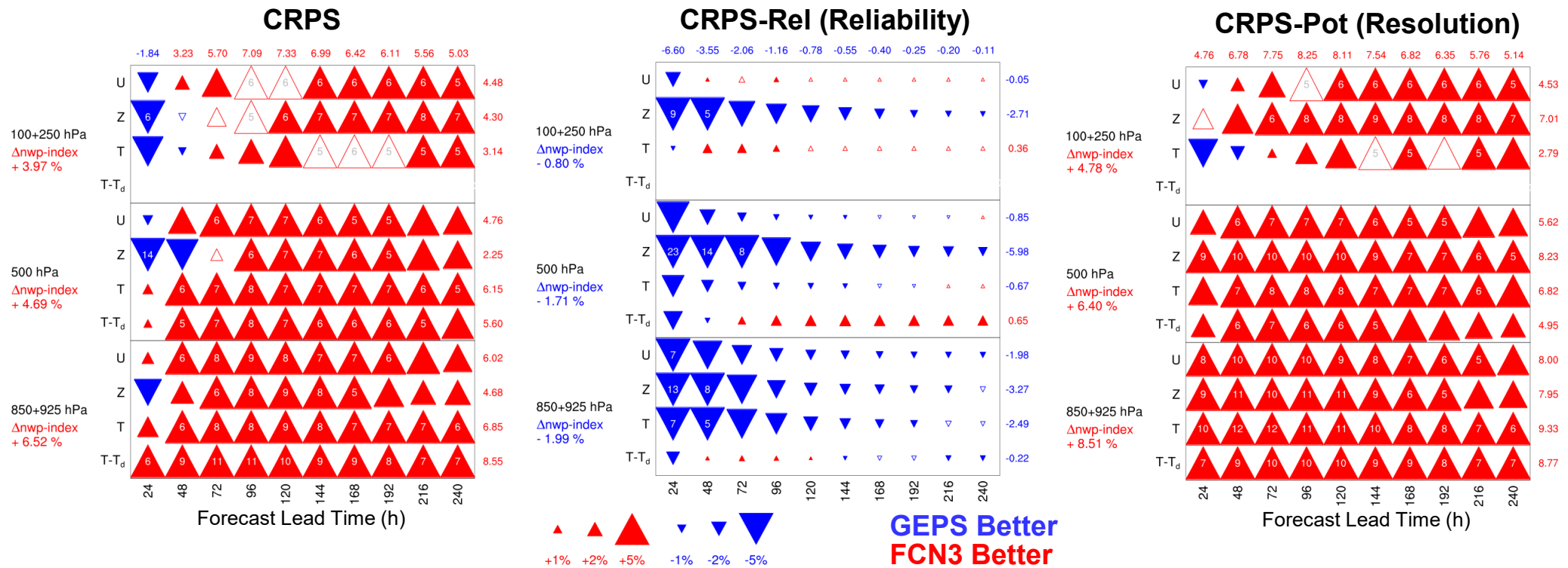
SPECTRAL NUDGING TOWARDS AI IN THE GEPS

- We are still in the phase of development and rigorous validation of MLWP predictions:
 - Nvidia's FourCastNet v.3 (**FCN3**; Bonev *et al.* 2025).
 - Mix of spatial (grid-point) and spectral CRPS loss terms.
 - Ensemble version of own model Paradis* (Pereira *et al.* 2026).
 - CRPS-based training.
- Here we'll show some results from FCN3 validation against GEPS, in their native setups, over JJA 2022 test period.
 - Experiment: $\frac{1}{4}^\circ$ -resolution **FCN3** trained with ERA5 (1980-2016) and **initialized deterministically** with ERA5.
 - Control: 25-km operational Global Ensemble Prediction System (**GEPS**) initialized with ensemble analyses.

FCN3 VS. GEPS:

SCORES AGAINST RADIOSONDE OBSERVATIONS

- Degradation of CRPS reliability but spectacular improvements in CRPS resolution component.
- Reliability is degraded due to a strong overconfidence at shorter lead times.
- Initialization with ensemble of analysis is expected to improve reliability but degrade the resolution.



FCN3 VS. GEPS: SPECTRAL VERIFICATION

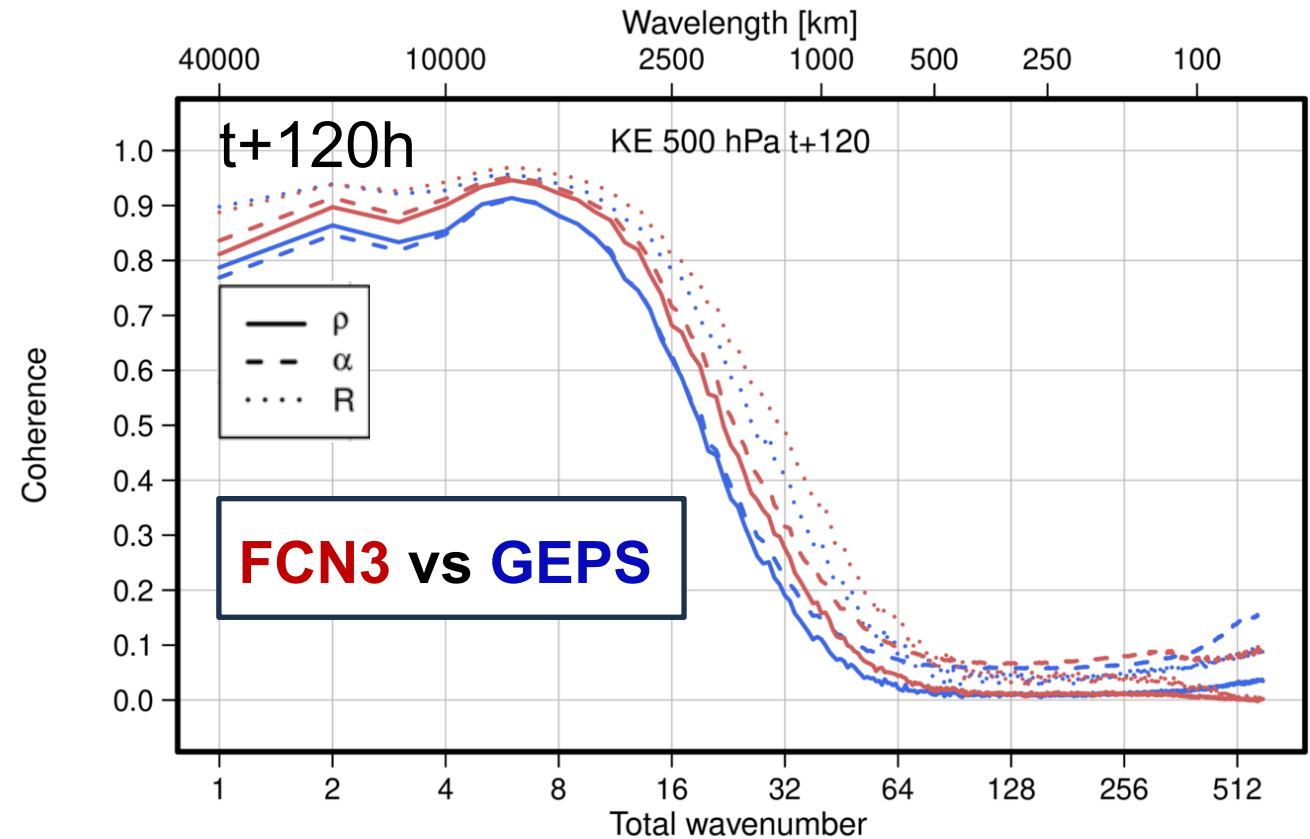
- **Improved True Predictability**

- At day 5, the **FCN3** member-analysis spectral coherence (ρ) is generally larger than that of **GEPS** at all scales where there is at least some predictability. This also holds for ensemble mean coherence (R).
- This explains the large skill gains measured by CRPS resolution.

- **Short-Lead Predictability Bias**

- For **GEPS**, inter-member coherence (α) and member-analysis coherence (ρ) are similar, implying a good balance between true and intrinsic predictability.
- For **FCN3**, α catches up with ρ , implying a much-reduced overconfidence compared to short leads (not shown) due to deterministic initialization.
- This explains the FCN3 degradations in CRPS reliability component for the first 3-4 days.

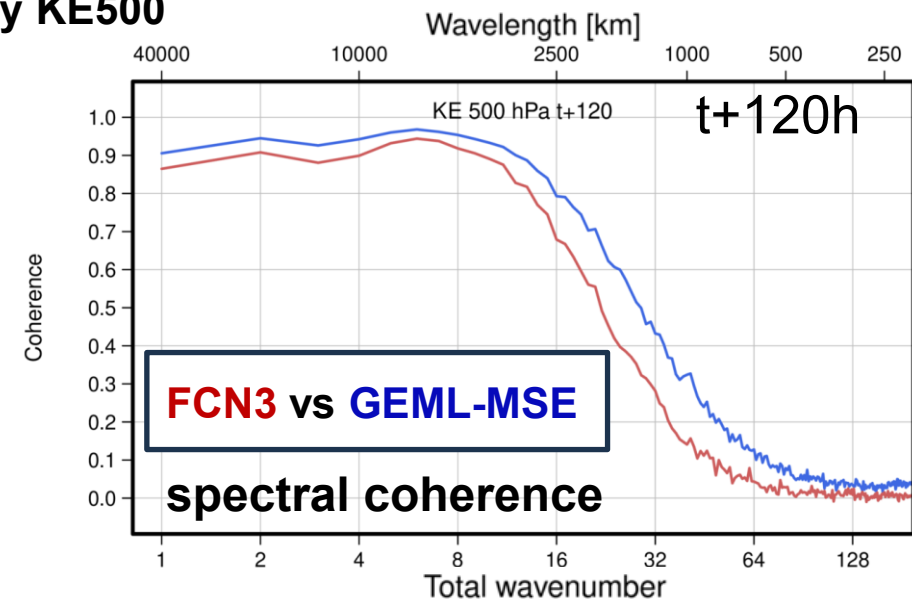
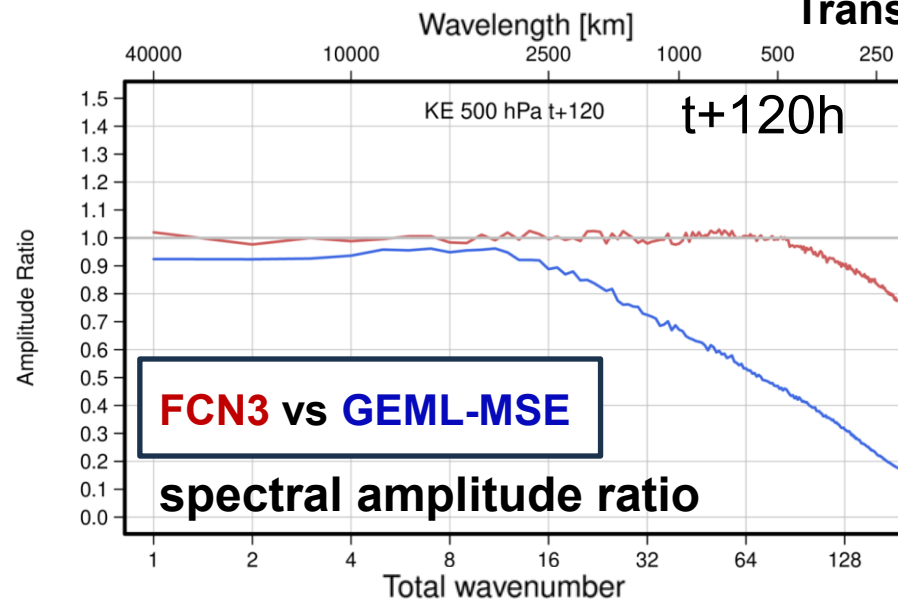
Verification against the GDPS analysis
Transient anomaly KE500 spectral coherence



FCN3 VS. GEML-MSE: SPECTRAL VERIFICATION

- Unlike the MSE-trained GEML, FCN3 shows no attenuation of amplitudes at scales larger than 400 km. Some amplitude bias beyond this scale is likely associated with the difference between ERA5 and GDPS-Analysis spectral variances.
- FCN3 individual members, however, cannot compete with GEML's coherence at any scale. This shows clear advantage of AMSE- over CRPS-based training if the goal is deterministic prediction with more realistic mesoscales.

Verification against the GDPS analysis
Transient anomaly KE500



FCN3 Member 1 vs GEML

SUMMARY

- Hybrid modelling based on spectral nudging towards MLWP model inferences has a potential to provide large boosts of forecast accuracy in a wide range of NWP systems.
 - The one-way spectral nudging strategy makes it technically very easy to plug-in new model versions or new models as the MLWP component in the hybrid systems.
 - More frequent updates of MLWP components in the operational hybrid systems may become a new norm.
- Work is ongoing at ECCO towards hybridizing **km-scale short-range** NWP systems (HRDPS and CAPS), as well as the **global ensemble** (GEPS), targeted for 2027 and 2028, respectively.
- The Global Deterministic Prediction System with Spectral Nudging (GDPS-SN) is the first hybrid system scheduled to become operational:
 - GDPS-SN is scheduled to replace the operational physics-based GDPS on **May 20, 2026**.

SUPPLEMENTAL MATERIAL

GEML – TRAINING CURRICULUM

Dataset	Loss	Steps	Batches-of-32	Learning Rate
ERA5 (1979-2015)	MSE	1 step (6h)	300,000	1e-3 → 3e-7
ERA5	MSE	2 steps .. 12 steps (12h – 3d)	1,000 each	3e-7
HRES (2016-2018)	MSE	12 steps (3d)	5,000	3e-7
GDPS (2019-2023)	MSE	1 step (6h)	1,000	3e-7
GDPS (2019-2023)	MSE	12 steps (3d)	5,000	3e-7
GDPS (2019-2023)	AMSE	1 step (6h)	25,000	1e-5 → 3e-7
GDPS (2019-2023)	AMSE	2 step (12h)	2,500	1e-4 → 3e-7
GDPS (2019-2023)	AMSE	4 step (1d)	2,500	1e-4 → 3e-7
GDPS (2019-2023)	AMSE	8 step (2d)	1,250	1e-4 → 3e-7
GDPS (2019-2023)	AMSE	12 step (3d)	1,250	1e-4 → 3e-7

- 2016–2023 period split between HRES and GDPS analyses (consistent GDPS unavailable before 2019).
- Short **1-step** GDPS G2 training stage for stability – "*alien analysis problem*" applies in training too!