TECHNICAL MEMORANDUM

841

Use of ERA5 to Initialize Ensemble Re-forecasts

F. Vitart, G. Balsamo, J. Bidlot, S. Lang, I. Tsonevsky, D. Richardson, M. Balmaseda

Research Department

February 2019

European Centre for Medium-Range Weather Forecasts
Europäisches Zentrum für mittelfristige Wettervorhersage
Centre européen pour les prévisions météorologiques à moyen

Series: ECMWF Technical Memoranda

A full list of ECMWF Publications can be found on our web site www.ecmwf.int.

Contact: library@ecmwf.int

Abstract

This study summarizes the results from re-forecast experiments initialized with ERA5. Results suggest that initializing the re-forecasts from ERA5 instead of ERA Interim improves the extended-range skill scores up to week 4. The use of the ERA5 EDA to perturb the re-forecast initial conditions has been assessed. The ERA5 EDA provides flow-dependent EDA initial perturbations across the re-forecast years instead of the non-flow-dependent perturbations in the current operational set-up. Results show that it is beneficial to use the ERA5 EDA, even if the impact is limited to just the first week of the re-forecasts. Different strategies for land and wave re-forecast initialization have been tested. The consistency of the re-reforecast land surface climatology with real-time forecasts is improved when initializing from ERA5 and, as a consequence, no separate offline land simulation would be needed to initialize the land in the ERA5 reforecasts.

1 Introduction

The ECMWF 11-member operational ensemble re-forecasts are currently run twice a week over the past 20 years. They are used to calibrate the extended-range re-forecasts (e.g. weekly mean anomalies) as well as to produce the Extreme Forecast Index (EFI) from medium-range forecasts. They are also used to assess the extended range forecast skill and its evolution from year to year. Therefore, a large number of years is needed to provide an accurate evaluation of the extend range forecast skill. Take as an example one of the new headline scores for the extended range, the ranked probability skill score of 2-metre temperature at week 3 (day 15-21) over the northern Extra-Tropics. This score is computed from the re-forecasts rather than from the real-time forecasts. This is justified by the fact that extended-range forecast skill scores display a very strong interannual variability, due to slow variability such as ENSO: the temporal sample of the real-time forecast during the life of a model cycle would be insufficient to estimate the skill of that cycle for the extended range. Here we report on the impact of using of ERA5 as initial conditions for the ensemble re-forecast highlighting the differences with the current method and the benefits supporting this choice. Aside of changing the atmospheric initialization, the use of ERA5 for the reforecast has also involved revisiting the initial perturbation strategy – as to include information from the ERA5 EDA- and the initialization of land and ocean waves.

In the current operational configuration, the ensemble re-forecasts are initialized from ERA-Interim for atmospheric and ocean wave fields, while the land initial conditions (soil and snow) are provided by the so-called ERA-Interim-Land, an offline land surface model simulation driven by ERA-Interim surface fluxes. Ocean initial conditions are provided by ORAS5. The main reason a land surface model simulation is used for soil initialization is the inconsistency between the TESSEL land surface scheme in ERA-Interim, which is more than 12 years old, with the HTESSEL scheme used in operational analysis. Since HTESSEL is used in ERA5 (Hersbach et al., 2018, https://www.ecmwf.int/en/elibrary/18765-operational-global-reanalysis-progress-future-directions-and-synergies-nwp), the use of the offline ERA-Interim-Land initialization needs revisiting.

Several revisions were introduced in the HTESSEL land surface scheme, related to treatment of snow, soil hydrology, vegetation, lakes, as well as changes in static fields such as land-sea-mask and orography. These differences between the land surface versions used in a reanalysis and in operations effectively introduce inconsistencies that, due to the long memory in the land surface energy and water reservoirs, can generate systematic differences between the re-forecast climate and the real-time forecasts. They can produce spurious anomalies that persist for weeks into the forecasts. Several examples of spurious signals associated with this inconsistency in the land initialization were noticed in the extended range weekly charts, including a persistent wet anomaly over North Egypt, a persistent cold anomaly over Switzerland in Spring due to a lack of snow in ERA-Interim before 2004 (due to a change in the observing system). To alleviate these problems, an offline land simulation driven by ERA-Interim, using the same surface model as the operational IFS cycle and with the same resolution as the ENS configuration, has been used in every IFS cycle, since June 2012. These land surface model simulations make use of ERA-Interim as meteorological forcing for the land model. Attempt at using GPCP for constraining precipitation accumulation were tried but ERA-Interim precipitation was a preferred forcing, due to issues with Himalayan snow cover when forcing the land surface model with GPCP precipitation (Wegmann et al. 2016, https://www.the-cryosphere.net/11/923/2017/).

In the current configuration, the wave model initial conditions for the re-forecasts are provided by ERA Interim. However, significant contributions in CY46R1 (new wave model parameterisation for wind input and open ocean dissipation) change certain aspects of the wave model climatology. These wave model differences may generate inconsistencies between real-time forecasts and re-forecasts and the use of an offline ERA-Interim-Wave initialization, in a similar way as for land initialization, needs to be considered.

The ensemble generation for the re-forecasts is similar to the one used for real-time forecasts. Singular vectors and ensemble data assimilation (EDA) are used to perturb the re-forecast initial conditions. For singular vectors, the EPSGAMMA had to be rescaled for ERA-Interim initialization. Because ERA-Interim does not include an ensemble of data assimilation, the re-forecast initial conditions are perturbed using the latest operational EDA available at the time of production of the re-forecasts s (namely 2 weeks prior to issuing the real-time forecasts, as to provide the 5-week window needed for EFI calculation). Hence, the EDA initial perturbations are identical from one re-forecast year to another and are not flow-dependent.

2 Use of ERA5 to initialize re-forecasts

2.1 Same perturbations as in current operational re-forecasts

Most of the experiments described in this article were performed in 2018 when the period 1980-1999 was not consolidated yet in the ERA5 MARS archive. Therefore, all the re-forecast experiments cover the period 2000-2016 (17 years) only. The experimental set up is the following:

- 5-member ensemble starting the first of each month
- Re-forecast period: 2000-2016
- Resolution: Tco319L91 with ¼ degree ocean (same resolution as LegB of ENS)
- Cycle 45R1

A control experiment (gtl5) was run with ERA-Interim as atmospheric initial conditions and ERA-Interim-land at Tco319 resolution (experiment gjm2) for land surface initialization. This is the same set up as in CY45R1 operational re-forecasts, except for the resolution of LegA. A second experiment, gw0n, was run with ERA5 as initial conditions for both atmospheric fields and land surface (no use of a land surface model simulation), but with the same initial ensemble perturbation methodology as in the control experiment. Therefore, gw0n does not use the EDA ensemble from ERA5, but uses the operational EDA from 2018. Figure 1 shows a scorecard of the difference of continuous ranked probabilistic skill scores (CRPSS) between both experiments. Figure 1 shows that the skill scores are significantly improved when using ERA5 as initial conditions up to week 3 in the Extratopics and week 4 in the Tropics, except for zonal wind and temperature at 50 hPa in the Tropics, which is slightly degraded, although the difference is not statistically significant. These results suggest that the impact of

ERA5 on extended range forecasts is large and extends well beyond the first few days of the re-forecasts. This can be interpreted as the impact of 10 years of data assimilation system and observing system improvements on extended range forecasts. This result is not trivial, showing that the impact of atmospheric initial conditions extends to week 3 and week 4 in the Tropics, highlighting the importance of atmospheric initial conditions for obtaining good quality extended range forecast.

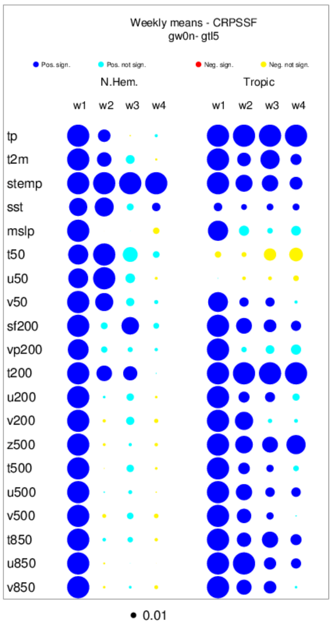

Figure 1: Scorecard of the difference of continuous ranked probabilistic skill scores (CRPSS) between experiment initialized with ERA5 (gw0n) and control (gtl5) over the northern Extratropics (left columns) and the tropics (right columns) for weeks 1 to 4. The size of the dots is proportional to the amplitude of the difference of skill score. The blue (red) colour indicates higher (lower) CRPSS when initializing from ERA5 than from ERA-Interim. Dark blue and dark red colours indicate that the difference is statistically significant at the 1% level of confidence, using a 10,000 resampling bootstrap procedure. The forecasts have been verified against their own re-analysis (ERA5 for gw0n et ERA-Interim for gtl5).

Figure 1 shows the difference of skill scores between the two experiments which are verified against their own re-analysis. Figure 2 shows the verification against ERA-Interim, which in principle should be disadvantageous to the experiment initialized from ERA5 (gw0n). However, Figure 2 shows that when we verify against ERA-Interim, the re-forecasts initialized from ERA5 generally outperform the control experiment, except for the zonal wind at 50 hPa in the tropics and northern extratropical SSTs in week 1. This confirms that the increased skill shown in Figure 1 is not simply due to a change of verification data.

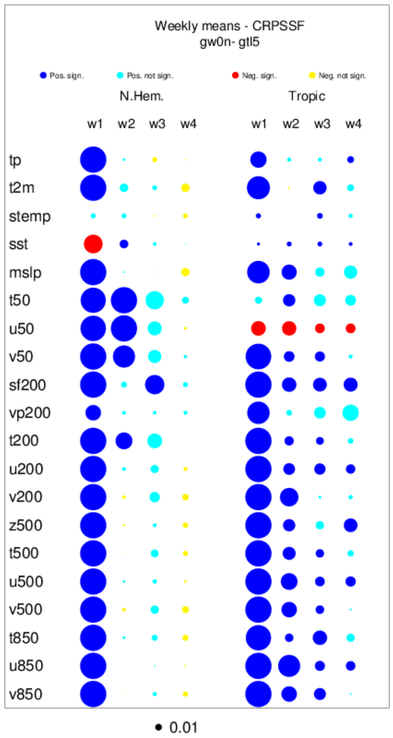

Figure 2: Same as Figure 1 but the verification is against ERA-Interim instead of own re-analysis for both experiments.

The Madden Julian oscillation (MJO), a major source of sub-seasonal predictability, has been diagnosed in both experiments using the Wheeler and Hendon bivariate index, which consists in projecting the forecasts on precomputed combined EOFs of zonal wind at 850 hPa and 200 hPa and outgoing longwave radiation. The forecast skill scores have been computed using a bivariate correlation, as described in Rashid et al. (2011), between the ensemble mean forecast and each experiment's own re-analysis. According to Figure 3, the MJO skill scores are statistically significantly improved when initializing from ERA5 instead of ERA-Interim during the first 20 days of the re-forecasts. The amplitude of the MJO is also stronger with ERA5 during the first few forecast days by 3 to 5%, compared to Control (Fig.4). After 6 days, the difference in MJO amplitude is no longer statistically significant.

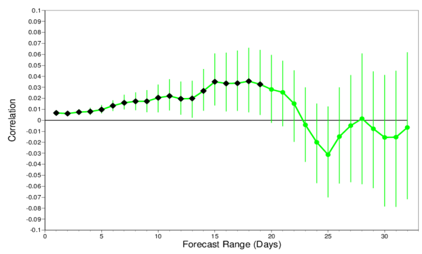

Figure 3: Difference of MJO bivariate correlation as a function of forecast lead time between the experiment initialized from ERA5 and control. The black diamonds indicate statistical significance within the 1% level of confidence according to a 10,000 bootstrapped resampling.

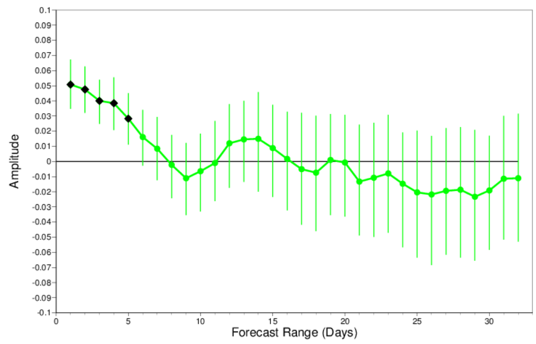

Figure 4: Difference of MJO amplitude error (relative to the MJO amplitude in ERA-Interim). The black diamonds indicate statistical significance within the 1% level of confidence according to a 10,000 bootstrapped resampling.

2.2 Use of ERA5 EDA and rescaled EPSGAMMA

Here, we run an additional set of re-forecasts (experiment gws0) using ERA5 for initialization, but now also to generate the initial perturbations: the ERA5 EDA has been used instead of the operational EDA from recent years. An important advantage of this change is that the ERA5 EDA provides flow-dependent EDA initial perturbations across the re-forecast years instead of the non-flow-dependent perturbations in the current operational set-up. The amplitude of the singular vector initial perturbations is flow dependent because it is linked to the EDA analysis uncertainty estimates of the day. The scaling of the singular vector initial perturbations is controlled by the EDA ensemble standard deviation and a scaling factor EPSGAMMA. The scaling factor is chosen such that on average there is a good match between ensemble standard deviation and ensemble mean RMSE and has to be adapted in case the average EDA ensemble standard deviation changes. In this experiment, EPSGAMMA had been rescaled for ERA5 (value of 0.013 instead of 0.010 in the current operational re-forecast configuration). The change in initial perturbations has a statistically significant positive impact in week 1 in the Tropics and Extra-tropics (Fig. 5), but no statistically significant impact is detected after week 1. These results show that it is beneficial to use the ERA5 EDA and the new value of EPSGAMMA, even if the impact is limited to just the first week of the re-forecasts. The impact on the MJO skill scores is neutral.

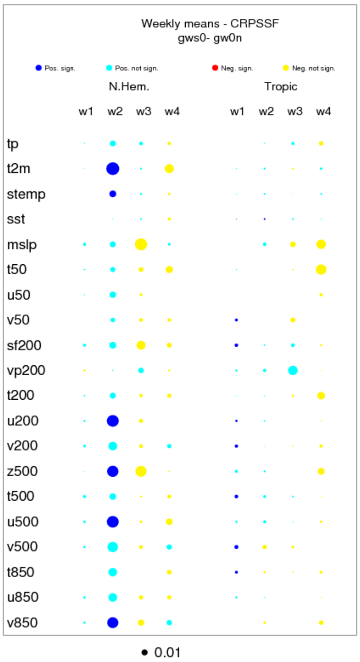

Figure 5: Same as Figure 1, but this time the difference is between two experiments initialized from ERA5 but one using ERA5 EDA and new value of EPSGAMMA while the second experiment uses the same initial perturbations as in CY45R1 operational suite.

The results presented so far focussed on the impact on the forecast skill scores. This is important for the skill assessment of the extended range forecasts, particularly since one of these scores is now an ECMWF headline score. However, as mentioned in the introduction, the re-forecasts are also used for the calibration of real-time forecasts and for the calculation of EFI. Therefore, it is important to check that the ERA5 climate does not introduce new inconsistencies which would translate into spurious signal in the week 3 or 4 mean anomalies and in the EFI calculation. For the atmospheric fields, the model climate generated by the ERA5 initialized experiment is overall very consistent with the model climate of the control experiment. For surface parameters, the ERA5 model climate seems also consistent with the control experiment climate but over some regions, the 2-metre temperature climate from the ERA5 experiment is slightly warmer than the climate of the experiment initialized by ERA-Interim and the offline land simulation. This is particularly the case over Central Asia, South Africa and significantly warmer over the Central plains of North America (Fig 6). According to Figure 6, the 2-metre temperature biases are significantly reduced over North America. The large biases in the current system over this region led to systematic spurious cold anomalies during summer and are probably due to the lack of data assimilation in the land surface model simulations used to initialize the current re-forecasts. This was mentioned in the 17 July 2018 daily report, which showed by comparing with station data that the climate of the current reforecast is too warm over Central North America, generating systematic cold anomalies. Using the re-forecasts initialized from ERA5 helped remove these spurious warm biases. This suggests that, compared to the current approach, the reforecast initialized by ERA5 should result in a climate more consistent with the real-time forecast. Section 5 will show an example of weekly mean anomaly chart. As a consequence, no separate offline land simulation would be needed to initialize the land in the ERA5 reforecast. This is further explored below in section 3.

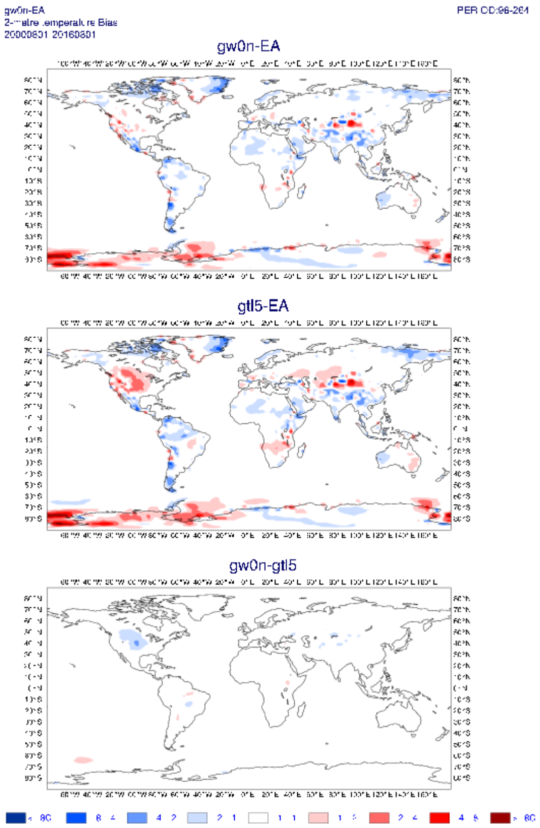

Figure 6: 2 metre temperature mean biases computed for day 5-11 of the re-forecasts starting on 1st August 2000 to 2016 relative to ERA5. The top panel shows the biases obtained with the ERA5 initialized experiment, the middle panel shows the biases with the current configuration and the bottom panel shows the difference between top and middle panels.

As mentioned earlier, the Ensemble re-forecasts are also used for the calculation of EFI. Inconsistencies between model climate and real-time forecasts are likely to be captured by the EFI. In order to test the impact of ERA5 on EFIs, a CY45R1 test suite has been run in parallel to the operation re-forecast suite during the period June to September 2018. The only difference between the two suites is the use of ERA5 for initialization and initial perturbations. To reduce the cost of this experiment, the test suite has been run with a re-forecast ensemble size of 5, instead of 11 in operations and once a week only, instead of twice a week. Figure 7 shows the results for the EFI calculated for Summer 2018 using the same re-forecast sample from operations as in the test-suite (period 2000-2016, 5 members, once a week). The summer 2018 real-time data used for the EFI calculations is of course the same in both cases. Figure 7 suggest a neutral impact on the EFIs for total precipitation and small but statistically significant positive impact on the 2-metre temperature EFI globally.

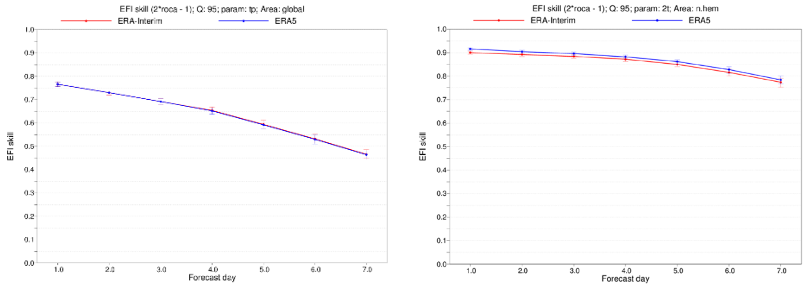

Figure 7: EFI skill as a function of the forecast lead time (up to day 7) for global precipitation (left panel) and global 2-metre temperature (right panel). The blue (red) curve corresponds to the skill in the test suite (oper).

More EFI scores can be found at
https://confluence.ecmwf.int/display/EVAL/EFI+verification+-+ERA-Interim+versus+ERA5

All these results suggest that it is safe to use the new configuration of re-forecasts with ERA5.

3 Land Surface Initialization

There are several possible land surface initialization strategies for the re-forecasts:

1. Use ERA5 to initial land surface (as in the experiments described above)
2. Use a land surface model simulation for land surface initial conditions at LegA resolution
3. Use a land surface model simulation for land surface initial conditions at same resolution as operational analysis
4. Use a stand-alone land reanalysis

Each option has its advantage and inconvenience. All the results presented so far were produced with option 1 (use ERA5 to initialize Land Surface). This option assumes that the land surface data assimilation in ERA5 and operational analysis are close enough so that we don't need to use a different dataset to initialize reforecast land surface. It also has the advantage of ensuring the consistency between initial land surface and upper level fields. 2) is what is currently done in operations and could be an option for CY46R1. In addition to ensuring the consistency between the land surface modelling used for the re-forecasts and real-time forecasts, an advantage of this option is that the land surface initial condition being at the same resolution as legA do not need any interpolation. However, this is not such a big advantage since the real-time forecast land surface initialization needs anyway to go through fullpos to interpolate from Tco1279 to Tco639, which means that error in fullpos interpolation may still be present in the forecast charts. 3) would have the advantage over 2) of having a perfect consistency between the land surface interpolation for real-time and re-forecasts. However, this is technically difficult to realize since ERA5 is not at Tco1279, which means that we cannot do fullpos interpolation from Tco1279 for the re-forecasts. Option 4) would be a better option than 2) or 3) since it includes data assimilation, but this is not available yet. In summary, of these 4 options, only two could be tested, namely: 1) and 2).

In order to compare these two options, two additional Cycle 46r1 experiments have been run with the same experimentation setup as the previous section:

h2g8 Use of a land surface model simulation (INILANDEXPVER=h286) h2gf Control (Use of ERA5 land surface)

The scorecard (Fig 8) shows a significant degradation in the surface temperature skill scores when using offline experiment land simulation, but this is expected because the verification is against ERA5. For the upper level fields, there is no statistically significant differences in forecast skill scores. The experiment using a land surface model simulation (h2g8) has larger biases inT850 relative to ERA5 in winter over North India (Fig. 9). This difference in biases is robust and consistent across all winter's months. h2g8 also displays larger dry biases over Central US in Summer (not shown). This latest problem is the same as the one present in our current system and discussed above (see also Figure 6).

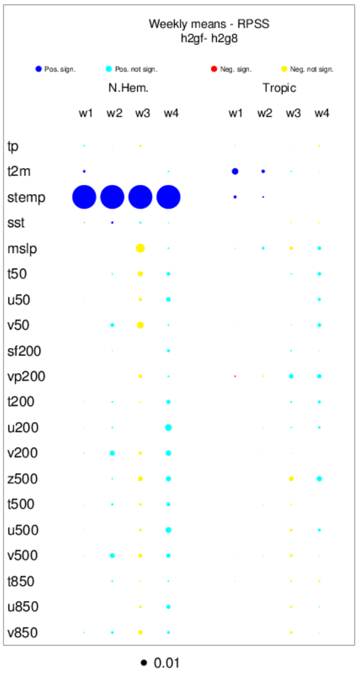

Figure 8: Scorecard of the difference of continuous ranked probabilistic skill scores (CRPSS) between experiment with land surface initialized with ERA5 (h2gf) and land surface model simulation (h2g8) over the northern Extratropics (left columns) and the tropics (right columns) for weeks 1 to 4.

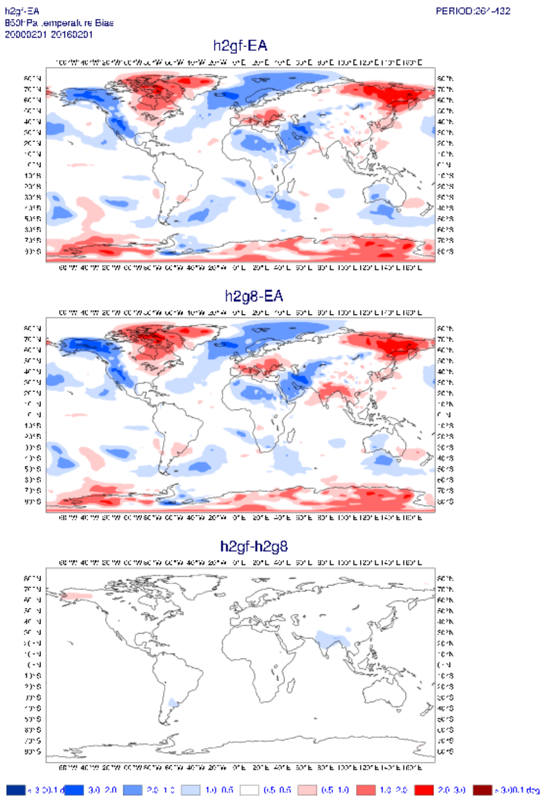

Figure 9: 850 hPa temperature biases relative to ERA5 for week 2 (day 12-18) in re-forecasts starting on 1st February 2000 to 2016. The top panel shows the biases of the experiment initialized from ERA5, the middle panel shows the biases from the experiment initialized from a land surface model simulation and the bottom panel shows the difference between the two.

Based on these results, there is no clear reason for using a land surface model simulation with ERA5, at least for the IFS cycle 46R1 when ERA5 and operational land assimilations are still sufficiently consistent. Therefore, it is proposed to implement ERA5 in Cycle 46R1 with land surface initialized directly from ERA5, which will result in a simpler setup. The negligible differences shown in Figure 9 have also a reassuring message for the ERA5-driven land surface simulations as plausible initial conditions to avoid inconsistencies and spurious anomalies. The option of a stand alone land simulation or reanalyses will still be useful when new changes to the land surface model (e.g. 5-layer snow, 9-layer soil, new lake mapping, ...) are introduced operationally.

4 Wave Model Initialization

With CY46R1, there is a new wave model parameterisation for wind input and open ocean dissipation. This yields a systematic change in certain aspects of the wave model climatology. In a similar way as for the land surface, a large change in the forcing or in the wave parameterisations could introduce spurious anomalies in the forecasts. To produce initial conditions that are both consistent with the wave model climate of CY46R1 and ERA5 winds, an offline wave model hindcast has been carried out with ERA5 wind forcing (experiment h1pe). A new ERA5 initialized experiment has been run with wave model initialized from this wave model simulation along with a control experiment with a cheap configuration of the coupled system (atmosphere resolution at Tco199 and 1-degree ocean). Results suggest that there is no statistically significant impact on the forecast skill scores (Fig. 10). This is to be expected since the memory of the wave model initial conditions is of the order of a week to 10 days. As shown in Fig. 11, both simulations yield similar significant wave height mean climate after a week, however, from Fig. 11 is also clear that in the early part of the forecasts, there is a difference between forecast initialized from ERA5 and those from the offline wave initialization with the new physics, that slowly disappears with forecast lead time. Whether these differences are due entirely to the new wave model physics or have a component of the wave data assimilation used in ERA5 is still to be determined, but the difference pattern has the characteristic signature of the model changes (more slowly in tropical areas as expected from the dominance of swell conditions in those areas). When EFI products for waves will be computed with CY46R1, this signature may appear as a spurious anomaly rather than a true signal.

Since we have currently no proof that the wave model initialization from the offline experiment has a significant positive impact, it is planned to have the 46R1 re-forecast e-suite configured with the wave model initialized from ERA5. If the wave EFI shows a spurious signal in the first days of the forecasts, then the use of the experience, as described above, will be considered.

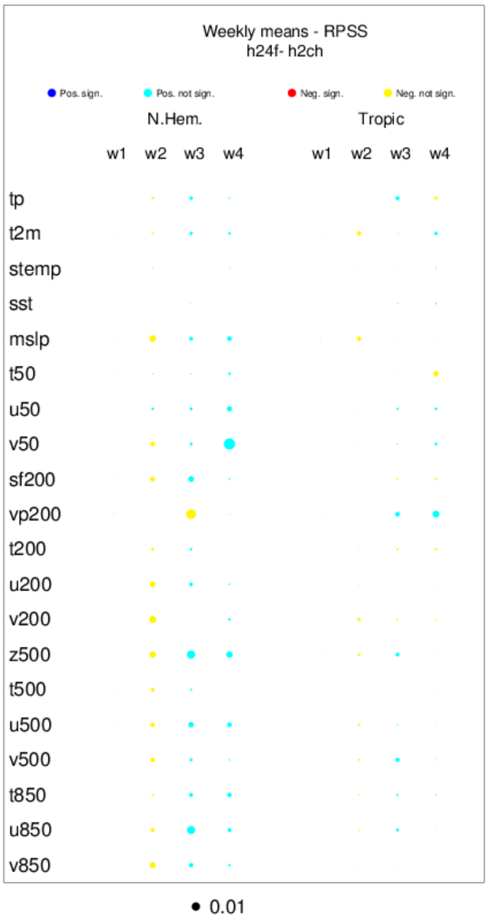

Figure 10: Scorecard of the difference of continuous ranked probabilistic skill scores (CRPSS) between experiment with wave model initialized with offline wave model simulation (h24f) and control experiment (h2ch) over the northern Extratropics (left columns) and the tropics (right columns) for weeks 1 to 4.

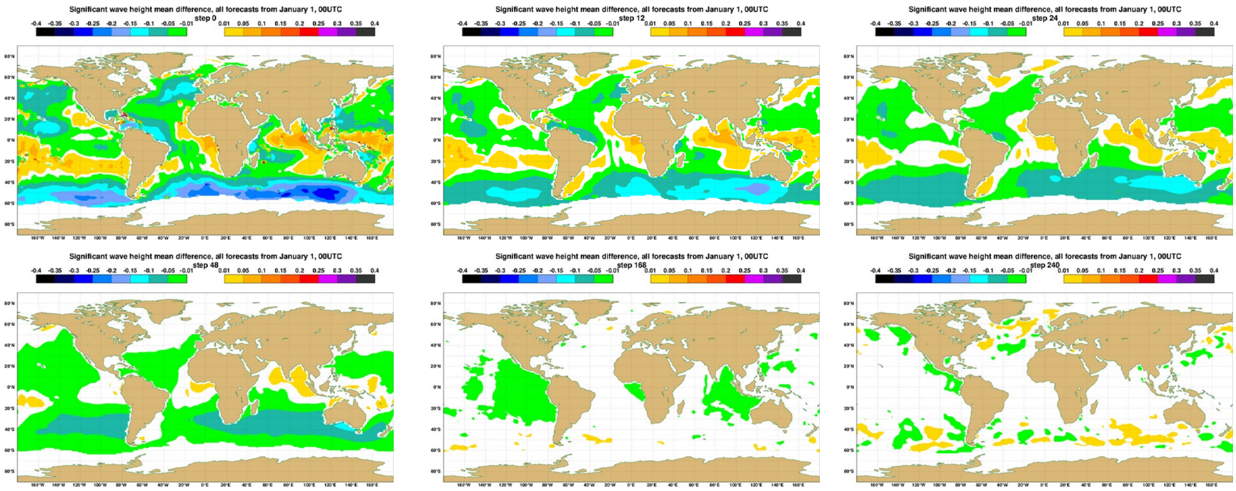

Figure 11: Significant wave height mean difference for all (14+1) forecasts initiated from January 1st 0 UTC (2001 to 2016) for different forecast lead times (step) between runs with wave initial conditions from the CY46R1 ERA5 forced wave hindcast (h1pe) and those with ERA5 wave data.

5 Extended-Range Forecast Charts

As already described in Section 2, a CY45R1 re-forecasts test suite has been run in parallel to the operational re-forecast suite during Summer 2018. Extended-range forecast charts have been produced by using the test-suite to produce the Model Climate used to calibrate the real-time forecasts and compared to charts using the operational re-forecasts with the same frequency, ensemble size and re-forecast period as in the test suite. In this comparison, the real-time forecasts are the same, the only difference being the model climate used for calibration. The forecasts look in general similar to a first order, but the slight difference in model climate can generate some regional differences. Figure 12 shows an example of weekly mean anomaly chart: week 1 (day 5-11) anomaly of 2-metre temperature from the ensemble forecast starting on 26 July 2018. Globally the charts look similar, but in this particular case, the use of the new re-forecasts produces warmer anomalies over Central US, Australia and South Africa. These anomalies produced using the new re-forecasts are more consistent with the verification produced either from ERA5 or ERA interim.

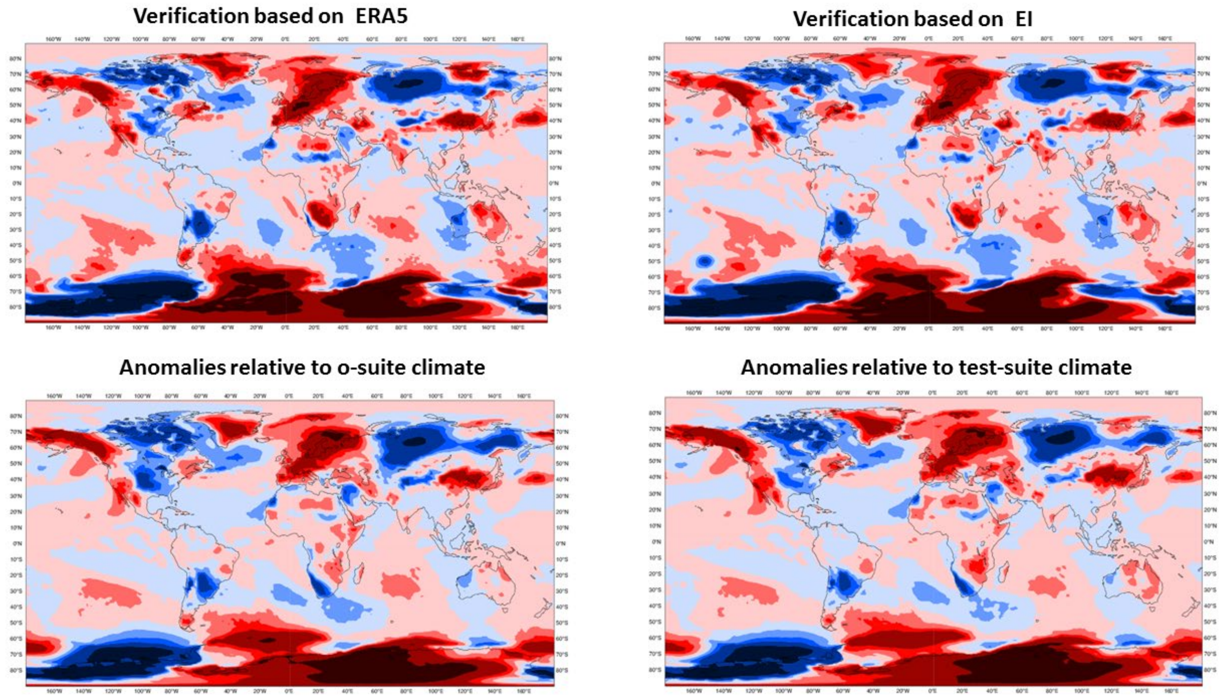

Figure 12: 2-metre temperature anomaly charts using the current operational re-forecasts (bottom left panel) to calibrate the operational real-time forecasts or a set of re-forecasts initialized from ERA5 (bottom right panel). The start date is 26 July 2018 and the lead time is day 5 to 11. The top panel show the verification using ERA5 (top left panel) and ERA Interim (top right panel).

6 Conclusions

This study suggests that using ERA5 instead of ERA-Interim to initialize the operational re-forecasts will not degrade the quality of the ECMWF extended-range forecasts and EFI. On the contrary, the re-forecast skill will be largely improved up to at least week-3 and the model climatology will be more consistent with the real-time forecasts, which will help remove some known issues in the current operational system. The impact on EFI skill score is neutral to positive. Therefore, it is planned to use ERA5 as re-forecast initial conditions in cycle 46R1. To summarize, the differences with the current re-forecast initialization will be:

- ERA5 instead of ERA-Interim over the past 20 years
- No land surface model simulation needed to initialize the reforecast land conditions
- No Wave model simulation used to initialize the reforecast wave field (pending final assessment in the 46R1 e-suite).
- EPSGAMMA=0.013
- ERA5 EDA

Acknowledgements

We would like to thank Martin Leutbecher for useful discussions on the use of ERA5 to initialize re-forecasts.

References

Rashid, H. A., H. H. Hendon, M. C. Wheeler, and O. Alves, 2011: Prediction of the Madden–Julian oscillation with the POAMA dynamical prediction system. Climate Dyn., 36, 649–661, doi:https://doi.org/10.1007/s00382-010-0754-x.

Wegmann, M. Orsolini, Y., Dutra, E., Bulygina, O., Sterin, A., and Bronnimann, S.: Eurasian snow depth in long-term climate reanalysises, The Cryosphere, 11, 923-935, https://doi.org//10.5194/tc-11-923-2017, 2017.