Geoscientific Model Development Discussions

Geosci. Model Dev. Discuss., referee comment RC1
https://doi.org/10.5194/gmd-2022-146-RC1, 2022

Comment on gmd-2022-146

Chiem van Straaten (Referee)

Referee comment on "Data-driven Global Subseasonal Forecast Model (GSFM v1.0) for intraseasonal oscillation components" by Chuhan Lu et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2022-146-RC1, 2022

General impression

This paper presents a deep learning framework for sub-seasonal forecasting. The novelty relative to Rasp et al (2020) is the self-attention mechanism. An additional innovation is the extraction of intra-seasonal oscillation components. This extraction eases the learning task, as the only predictable atmospheric motions at the sub-seasonal lead time are the low-frequency components. The authors compare their framework to a version without self-attention and to numerical forecasts. The comparison is detailed, with multiple scores and a case study. Possibilities for clarification do however exist. This could increase the trustworthiness of the results.

Overall comment:

A big finding of the study is that the deep learning models outperform the CFSv2 model beyond 10 days. But currently the comparison lacks information, rendering it untrustworthy. The authors say they use the WeatherBench dataset but CFS is not part of the benchmark dataset. How was it obtained, and especially, how was it processed?
In the verification you make use of the ISO components (e.g. fig 4a), which requires the removal of the seasonal cycle. Was the seasonal cycle for CFS estimated from model reforecasts or something else? Choices like that influence results (see for instance Manrique et al 2020). Also, how was lead time treated in your CFS-processing? At a lead time of 1 days for instance, right after initialization, it is impossible to subtract a 15 day rolling average of the previous days, which is step 2 in your ISO procedure (see section 2.1). Because you often verify the ISO components of circulation, it seems unfair if CFS is the only unfiltered model in that comparison. Please clarify.
Your figures contain a hint that the processing is unfair. Performance of the CFS model is heavily curved: it dips below that of the climatology benchmark and only later levels off (fig, 4a, 6a, 8a). Such behavior is unexpected for a numerical model. The numerical skill in z500 should directly level off at the climatology (Figure 4a, Buizza & Leutbecher 2015).

Specific comments

L42: Is Mayer et al 2021 the best reference for the chaotic nature of the atmosphere? Why not original work like that of Ed Lorenz?

L62: Unclear in what way predictability relates to spatio-temporal scale. I know myself that larger / low-frequency is more predictable, but perhaps good to explicitly state this. Also you can refer to Buizza and Leutbecher (2015).

L89-90: Explain why residual connections give the network this ability.

L115 Vague: "the contributions of evolution among different factors to the forecast may be different." Please clarify.

Section 2.1: Even with the reference to Hsu et al (2015) this is too short a description: "Remove other ISO signals". What is 'other' in this context? Also it would be good to explicitly mention the data you are filtering. My suggestion is to move the data-description at the end of section 2.2. (weatherbench) to here. Then you can also clarify if you do the filtering on a gridpoint basis or not. Also please mention the train/validation/test splits that you make.

L138-145: Introduce ResNet and what a residual block is and refer to the original paper (He et al., 2015), just like you do for the self-attention mechanism.

L162-165: I miss a mention of the domain.

L 203: here you start the discussion of forecasts. You mention that there are two types of predictions, one driven by the original data and one driven by the ISO components. This seems to concern the types of input, not the target against which it is trained (for both forecast models the task seems to be to predict unfiltered Z500/T850). Do I understand this correctly? Please outline the two variations already in the methods.

L 215: Statement: "Furthermore, in this case, the Z500 values predicted by the ISO components are closer to the ERA5 ISO components, with a mean RMSE of 575.96 (m 2 s -2)". It is obvious that the one driven by filtered information will indeed be closer to that information than the one without access to the filtered information. But it is confusing because in the sentences above this one, you discuss scores against unfiltered ERA5, and this seems also be the content of Figure 2. Perhaps make clear that here you discuss scoring against ISO components, and that that is not shown in Figure 2. Add something like “(not shown)”.

Figure 3: From the title of panel a and b and the y labels of c and d I understand the figure presents both scores against ISO components and scores against unfiltered ERA5. Is that correct? Perhaps introduce these two variations of scoring in section 2.3, formula 1, where you would say that $t_{i,j,k}$ is either the filtered or unfiltered component. Or… if you always score against the ISO component (which is not clear to me), then say that $t_{i,j,k}$ is the ISO component.

L307-316: You present the scores stratified against latitude. The conclusion that differences are small in the tropics is not surprising. Z500 in the tropics is a nearly constant field with hardly any variability.

Textual comments

L14-15 “Forecast results tend to become intraseasonal low-frequency components”. Unclear what is meant with this. At large forecast times, model output can of course still be high frequency. Do you mean that “as forecast time increases, the only skillful forecasts are those of low-frequency components”?

L24 CFSv2 acronym mentioned without definition.

L27-28 Planetary wave numbers 3-8.

L144-145 But “with” zero?

L177: More commonly known as “anomaly correlation coefficient”

Figure 7: mention in figure caption that z500 is in contours, and t2m in shading.

Additional literature referred to:

Manrique-Suñén, A.; Gonzalez-Reviriego, N.; Torralba, V.; Cortesi, N. & Doblas-Reyes, F. J. Choices in the Verification of S2S Forecasts and Their Implications for Climate Services Monthly Weather Review, American Meteorological Society, 2020, 148, 3995 - 4008

Buizza, R. & Leutbecher, M. The forecast skill horizon Quarterly Journal of the Royal Meteorological Society, Wiley Online Library, 2015, 141, 3366-3382