

Response to Reviewer 2

With the revised version, the authors have sufficiently addressed all of my concerns, except of my main concern. It is good to see the authors, at least, agree with the premise of my main concern that conditional verification may introduce artificial biases. And I agree with the authors that their focus on 12-h forecast **minimizes** this statistical artefact. The authors, however, have not eliminated the artefact completely! The relevant verbatim citation from my initial review is “Forecasts with less than 100% skill hence exhibit the tendency to underestimate anomaly amplitude in conditional verifications.”, and not “when considering forecasts that have lost all skill”. The important question that this study thus still needs to address is: In the presence of an artificial bias (due to conditional verification), how can we identify the forecast biases that are introduced by deficiencies inherent in the parameterizations (of moist processes)?

With respect to this important question, I did not find the authors’ responses to be satisfactory. In particular, I do not see that the authors have modified their manuscript in response to my comment.

We thank the reviewer for reading and commenting on the revised version of our manuscript. We agree that conditioning the verification on an analysis-based case definition can introduce a selection bias. As the reviewer points out and seems to agree, our choice of verifying against a 12-forecasts minimises this effect, though there, of course, remains a part that is difficult to disentangle. We have added an explicit acknowledgement of this artificial bias and the limitations of our analysis to the ‘data and methods’ section and ‘conclusions’.

The authors response ‘1.’ is not to the point, because my initial comment did not mean to say that conditional verification is not useful in general. The authors response ‘2.’ is not to the point, at least if the authors do not mean to claim that 12h forecasts are 100% perfect. The authors response ‘3.’ is not (fully) to the point, because my initial comment did not question the usefulness of evaluating parameterized physics using short-range NWP forecasts.

My initial comment, inter alia, specifically referred to the cyclones’ intensity bias. My initial comment may thereby have been misread as if the problem of “anomaly underestimation” is with cyclone intensity only. I would here like to re-iterate that any anomalous process or feature tends to be underestimated on average in less-than-perfect forecasts: anomalous translation speeds, rain rate, moisture fluxes, latent heat release I would certainly expect this artificial bias to be small in 12-h forecasts, but so are the signals found by the authors! The au-

thors claim (in their response ‘3.’) that “our study reveals that cyclones with strong heating exhibit distinct biases tied to the representation of specific physical processes.” (my emphasis) is true only to the extent that the authors can disentangle their signal from the artificial biases introduced by conditional verification of less-than-perfect forecasts.

In a published version of this manuscript, the authors need to provide an answer to the question at the end of the first paragraph. At least, there needs to be an acknowledgement of the artificial bias, and possibly the acknowledgement that disentangling biases may in fact be quite difficult.

While we agree that we cannot fully disentangle the origin of the biases, the biases shown in our study cannot be solely attributed to the statistical artefact that the reviewer brought forward. Our results contain biases of both signs across different fields, featuring strong asymmetries and physically coherent spatial patterns. Neither the sign reversal nor the structured patterns can be solely explained by this statistical artefact associated with the conditional analysis, which shows at least part of the presented biases are due to inherent model biases.

As indicated above, we agree that conditioning the verification on an analysis-based case definition can introduce a selection bias and have added an explicit acknowledgement of this artificial bias to the ‘data and methods’ section and the ‘conclusions’.