Earth Syst. Dynam. Discuss.,
doi:10.5194/esd-2016-14-AC1, 2016

Earth System Dynamics Discussions

Interactive comment on "The use of regression for assessing a seasonal forecast model experiment" *by* R. E. Benestad et al.

R. E. Benestad et al.

rasmus.benestad@met.no

Received and published: 24 August 2016

We are grateful for the comments raised by reviewer 1, which reveals the need for more careful explanation about the model outcome.

We will revise the paper to make the message and description clearer, with a better connection between figures, tables, and the main text. There may also have been some confusion arising from the question of sensitivity and the identification of a systematic effects associated with different model options, which we will try to explain more carefully in the revised version.

We argue that the sensitivity to various model set-up options are key in terms of predictability, however, the systematic effect can be used for bias correction but may be more difficult to predict. Our model experiment involved an ensemble too small for

answering the latter, but we can deduce a high sensitivity from our runs.

1. The information conveyed through Fig 2 is the raw results of the factorial regression analysis applied to the model experiments - both regression coefficients (intercept and anomalies in terms of this based on different model set-up option) and error estimates. This will be explained more carefully in the revisions. It shows the results of the factorial regression analysis applied to the results from the model experiments. The sensitivity to different sea-ice model options is slightly less than the others, but not by a whole lot. The differences are mainly in the regional anomalies, and the response-to-noise ratio is affected by whether the error estimates are higher in the same region. The small ensemble size used in this experiment precludes high precision when it comes to details.

2. The results are clear - subjective choices about model settings such as choice of atmosphere top (vertical levels) and representation of sea-ice has an effect on the predictions. However, the ensemble we used was too small to detect a robust effect (p-values) in the sense of a systematic bias associated with the settings. In our study of predictability, we limited the test for differences due to model set-up options (this, however, does not apply to different initial conditions) to the final month, which is expected to indicate the largest sensitivity to the choice. Operational seasonal forecasts are usually made for a three-month period, and first and second months are expected to show smaller differences and the effect is not as visible. We also looked at 70N, which doesn't change much, however, it's more interesting to look at a latitude over e.g. Oslo in terms of seasonal predictability. We will expand on this in our revision. The Walker test is applied to traditional assessments such as the the p-value from regression (which is pretty standard), and is not a replacement for each individual test. A chi-squared could also provide a similar base for a Walker test, but was not done as regression analysis was considered to be the best choice and sufficient for these purposes.

3. This is a good comment, and the paper needs to explain more carefully that the op-

tions considered in our experiments were expected to have strongest effects in the high latitude regions. There are other factors too which are expected to affect the tropics, however, the scope of this study was limited to the mid- to high latitudes and the search for causes for poor seasonal predictability. Furthermore, such factors are included in the model simulations, but we did not check their effects by including more experiments with changing their set-up options. The response to different model set-ups is nonlinear, but by considering additional snow-cover, the picture could potentially change: it could give a net response that looked more linear, or it could be that changing sea-ice but not snow (or other model aspects) is not really physically consistent. However, the results still indicate that it is easy to get nonlinear biases in model predictions depending on the model settings.

4. Thanks for asking this: C(.) is the change due to option setting, and is the result from the regression coefficient (one number per grid box). There is no need for adding residuals as the experiments were strictly controlled whereby one factor was changed whereas the other unconsidered factors remained the same.

5. Table 1 was missing (an unfortunate glitch), and will be inserted in the revised manuscript.

6. The reference to Fig 2 was left over from an early version with more figures. The reference will be corrected in the revised manuscript.

7. Only 200hPa - 50hPa is not shown here. The reference to 50hPa was for an additional figure that is not longer shown due to similarities.

8. Now. It should be 200hPa. It will be corrected in the revisions.

9. Yes. Thanks for pointing this out!

10. Yes.

11. Both matter for predictability if one does not know which option is best. High sensitivity (large response) and a robust response both indicate that the option setting

has an effect, but the latter indicates that the effect is more the same for all situations.

12. Yes it's correct as the last term represents the difference between L91 and L62.

13. It has been dropped. The text will be fixed in the revised version.

14. We will try to improve the graphics in the revised version.

Minor: thanks for pointing out. Regarding the nonlinear aspects, all other factors are included although have not been changed here.

Interactive comment on Earth Syst. Dynam. Discuss., doi:10.5194/esd-2016-14, 2016.