EGUsphere, referee comment RC2
https://doi.org/10.5194/egusphere-2022-706-RC2, 2023

Comment on egusphere-2022-706

Anonymous Referee #2

Referee comment on "Nudging allows direct evaluation of coupled climate models with in situ observations: a case study from the MOSAiC expedition" by Felix Pithan et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-706-RC2, 2023

Peer review for manuscript: Nudging allows direct evaluation of coupled climate models with in-situ observations: A case study from the MOSAiC expedition

Climate models are important tools for understanding complex interactions of physics and dynamics in the atmosphere. Descriptions of dynamical and especially physical processes are not complete in the models and simplifications made in parametrizations of physical processes causes uncertainty in model results. Therefore, it is very important to evaluate accuracy of models.

Typically accuracy of climate models is estimated by comparing model results against long time series of observations (or data sets which are strongly constrained by observations as reanalyses). As the state of large circulation, which causes a large part of variability of atmospheric conditions in middle and high latitudes, in models are not connected to state of large circulation in the real atmosphere, long time series is needed for comparisons, so that they are able to present a sufficient part of climatological variability which is caused by variations in large scale circulation. The requirement of long time series rules out a lot of observational data which otherwise could be used for evaluation of climate models. Especially in high latitudes, where the permanent observational network is sparse and a lot of observations have been collected from relatively short measurement campaigns. This strongly limits the evaluation of accuracy on climate models in the high latitudes where the presentation of atmospheric conditions in climate models is often worse than in lower latitudes. Therefore, it is highly important to evaluate the capability of climate models to simulate polar climate, which further allows development of models, so that they can better simulate atmospheric conditions also in the high latitudes. The manuscript presents a method how to overcome the requirement of long time series using nudged model simulations. When nudged simulations are utilised, large scale circulation in the model is strongly constrained by the observed large scale circulation. This kind of methodology allows direct comparison of model results and observations. As the large circulation in nudged simulation is constrained by the observed large scale circulation, model physics are responsible for uncertainty of models. This method does not allow estimate biases which are associated with presentation of large scale circulation in the models, but often a large part of the uncertainty is associated with parameterized physical processes. However, the possibility to use short time series of observation provides large advantages for model evaluations and further model development.

The novelty of the manuscript is in its methodology. The set of models that is evaluated in the manuscript is not comprehensive by any means, but I think that is not the scope of the manuscript. However, even this set of models shows interesting differences in their capability to simulate atmospheric conditions and clearly demonstrate the usefulness of nudged simulation for model evaluation and also shows some deficiencies in models especially associated with treatment of snow surface. Therefore, the scientific value of this manuscript supports the publishing of the manuscript in Geoscientific Model Development.

General comments:

1) Nudging is probably familiar for many readers, but I still suggest adding a short general description of nudging in the beginning of the nudging paragraph in the method section. Does nudging cause artificial effects on time series when model state is nudged towards real atmospheric state?

2) As model biases are related to weather conditions, I would start the result section with short description of weather conditions where, in addition to temperature changes, you could shortly describe e.g. cloud conditions, cloud cover and cloud liquid and ice water content, stratification, longwave and shortwave radiative fluxes, turbulent heat fluxes and how they are related.

3) The differences in surface temperature and skin temperature between the models are well explained in the method section, but it is still sometimes tricky to follow which is meant by surface/skin temperature in some parts in the results section. Therefore, I would suggest paying attention to clarity when surface/skin temperature is discussed in the result section and add remainder about their meanings when it is not very clear which temperature is in question.

4) Longwave radiation often has a remarkable effect on surface and near surface temperatures. However it has not received a lot of attention in the manuscript. If you have observational data of longwave radiative fluxes, I would suggest adding comparison of longwave radiative surface fluxes between models and observations, and how differences in cloud cover and cloud water (liquid and ice) content affect longwave radiative fluxes as well as how differnces in longwave radiative fluxes affect surface temperatures.

Specific comments:

Lines 100 – 101, Are the different values for rhcrit and rhsat used for the whole column below 700hPa or only in the inversion layer if a temperature inversion exists below 700 hPa?

Lines 141 – 144, How representative observations are for the whole grid cells used in comparisons? Were the conditions in the area of grid cell homogeneous enough (e.g. occurrence open water causes inhomogeneity) that point measurement could represent average condition in the grid cell?

Lines 188 – 189, How coarser resolution affects the delay?

Line 195 Do you compare surface temperatures from models against observed skin temperature? The next paragraph maybe gives an answer, but maybe it is good to clarify also here.

Line 270, Which kind of weather conditions are associated with unstable stratification in AWI-CM models?

Lines 286 – 297, Have you calculated the relationship between sensible heat flux and difference between skin temperature and 2m temperature in AWI-CM3/IFS. How would it look? Maybe the relationship between sensible heat flux and difference between skin temperature and 2m temperature looks better than relationship between sensible heat flux and difference between surface temperature and 2m temperature because in the method section it has been mentioned that AWI-CM3/IFS uses skin temperature for turbulent heat fluxes.

Lines 330 – 331, Could stronger cooling in AWI-CM1/ECHAM be associated with clouds?

Lines 344 – 346, Has the overestimation of cloud ice content so large an impact on radiation that it can compensate the underestimation of cloud liquid water content?

Overall, the manuscript is well written conclusion based on evidence of results. Methods are appropriately described allowing readers to understand how study is done. Therefore, I recommend publishing the manuscript in Geoscientific Model Development after minor revision.