Replies to comments by Anonymous Referee #1, 06 April 2023

In their article Boas et al. explore crop yield forecasts using seasonal and sub seasonal forecast models coupled to a land surface model that simulates crop yields. They explore the fidelity of their forecasts with respect to crop yields, and explore the source of muted year-to-year differences in yield levels in simulations relative to observations by evaluating the simulated LAI, ET, and soil moisture.

General comments:

Overall this research was well conducted and evaluated, but it could be communicated more effectively. The introduction contains information that belongs in the methods and the methods includes information that belongs in the results. The introduction in particular could be clarified to make the paper more accessible to the reader.

> Thank you very much for your constructive comments and suggestions. We have restructured large parts of our introduction and moved certain paragraphs from the introduction to methods, and from methods to the results section as indicated. Please find our detailed replies below.

Replies to the list of specific comments by reviewer #1:

Abstract:

L 23: please quantify what you mean by "very close to reference simulations results"

> We have clarified our results discussion throughout the manuscript and added several paragraphs in the results and discussion sections accordingly. Please also see our reply to the comments on the discussion below.
>
> (Line 22-25): We found that after preprocessing of the forecast products (i.e. temporal downscaling of precipitation and incoming shortwave radiation), the simulations forced with seasonal and sub-seasonal forecasts were able to provide a model output that was very close to the reference simulation results forced by reanalysis data (the mean annual crop yield showed a maximum difference of 0.28 and 0.36 t/ha for AUS-VIC and DE-NRW, respectively).

L26-28: I don't think "inter-annual variability" is the right phrase to use in the abstract. This would imply that you calculated a variance from a time series. You are instead describing the magnitude of the difference between two years relative to total yield levels. This portion of the abstract and the equivalent discussion in the results may be less confusing if you use the phrase "inter-annual differences" or describe the absolute change in yields between years (e.g. observations showed differences of up to XXXX tons/ha, while the model only simulated YYY) or similar.

> Thanks for pointing this out. We have decided to rephrase this accordingly throughout the manuscript.
>
> (Line 27-30): In addition, they also reproduced the generally higher inter-annual differences in crop yield across the AUS-VIC domain (approximately 50 % inter-annual differences in recorded yields and up to 17 % inter-annual differences in simulated yields) compared to the DE-NRW domain (approximately 15 % inter-annual differences in recorded yields and up to 5 % in simulated yields).

Introduction:

Although the vast majority of the paper is written clearly and effectively, the introduction is difficult to follow. The first and last paragraph in particular could be clarified to make the motivation and goals of the analysis more clear.

Mixing a discussion of the challenges of making climate forecasts with the challenges related to crop yield forecastss in the introduction can be difficult to follow for the reader. I suggest breaking the first paragraph into two paragraphs at around line 70. This would provide enough space to do a more thorough evaluation of crop yield forecasting literature as well. For example, in Australia see Wang et al. (2020) and Potgieter et al. (2022), as well as references therein, both of which are relevant to the present study.

Thanks for the suggestion. We gladly incorporated the suggested literature and added a paragraphs on crop yield modelling as follows:

(Line 108-116): Wang et al. (2020) investigated the impact of pre- and early-season El Niño Southern Oscillation (ENSO) related large scale climate signals on wheat yields in Australia. They found that these ENSO signals can have a significant impact on wheat yields in the Australian wheat belt and could explain up to 21% of the yield variation. In another study by Potgieter et al. (2022), the lead time and skill of Australian wheat yield forecasts using seasonal climate forecasts derived from a statistical ENSO-analogue system were compared with using a dynamic general circulation model (GCM). They found that ENSO-derived forecasts showed higher skills at a longer lead time (6 months), with a higher correlation coefficient of 0.48 compared to 0.37 for GCM forecasts, while GCM forecasts provided higher skill at shorter lead times (1-3 months) with a higher correlation coefficient of 0.44 compared to 0.35 for ENSO-analogue forecasts.

L75-82: discussion of MetSim and VIC belongs in the methods section, not the introduction

We agree and moved this paragraphs to the methods section:

(Line 281-287): MetSim is based on algorithms from the Mountain Microclimate Simulation Model (MTCLIM) (Hungerford et al., 1989; Thornton and Running, 1999; Thornton et al., 2000; Bohn et al., 2013) and the Variable Infiltration Capacity (VIC) macroscale hydrologic model (Liang et al., 1994). MetSim can be used to either generate spatially distributed sub-daily time series of meteorological variables from a smaller number of input variables (daily minimum and maximum temperatures and elevation data), or to disaggregate meteorological data from a coarse temporal resolution to a finer one (Bennett et al., 2020).

L115 - 137: Much of this information again belongs in the methods and makes the introduction difficult to follow. For example, specifying how you forced the MetSim model and how it was evaluated belongs in the methods.

Please see our reply to the previous comment.

L137-141: This is a nice and clear articulation of the study objectives. The paragraph leading up to this point is difficult to follow because these objectives were not outlined clearly. I would suggest moving these lines up to the beginning of the paragraph as it will help the reader to understand the study objectives before you go into detail about which models you use and why.

Thank you for your detailed suggestions for the introduction. We have restructured the text accordingly and moved the description of our study objectives towards the beginning of the section.

(Line 75-87): The major aim of this study was to evaluate the efficacy and applicability of this state-of-the-art forecasting product for physical and biogeochemical land surface responses and regional crop production in an ecosystem process model approach. To this end, we tested the combination of the Community Land Model version 5 (CLM5) (Lawrence et al., 2018; 2019) and seasonal forecasts from ECMWFs latest seasonal forecasting system SEAS5 (Johnson et al., 2019). Regional simulations were conducted for two domains with different climate regimes and agricultural characteristics, one covering the state of North Rhine-Westphalia in Germany (DE-NRW), and one the state of Victoria in Australia (AUS-VIC), using sub-seasonal and seasonal forecasts with different lead times as input. In our evaluations we focussed on (1) the model's sensitivity to seasonal changes in weather patterns and their effect on regional vegetation properties, e.g., leaf area index (LAI), evapotranspiration (ET), and crop yield; (2) the representation of the surface soil moisture content; and (3) the overall applicability and potential of seasonal weather forecasts for the prediction of regional agricultural production in model applications such as CLM5. In addition, we addressed the pre-processing steps required for the usage of the SEAS5 product in this model application and briefly discuss the importance of temporal downscaling.

Materials and methods

Section 2.4: There are periodically results sprinkled throughout your methods section. Please move these to the results section. For example, in L263-266 you discuss the performance of the forecasts relative to the observations. These belong in the results section as do all similar discussions of model performance and evaluation

We moved the whole paragraph with the comparison of seasonal forecasts and official weather statistics to the results section:

(Line 360-387): 3.1 Comparison of seasonal forecasts to recorded weather statistics

Results:

L383-4: "Only minor differences between the seasonal and sub-seasonal experiments can be observed for both domains". Is this true? It looks to be the case for AUS-VIC, but in the post-July period I see substantial differences In the DE-NRW domain between the sub seasonal and seasonal forecast runs.

Thanks for pointing this out. We added a short comment on this:

(Line 418-420): Only minor differences between the seasonal and sub-seasonal experiments can be observed for AUS-VIC, while for DE-NRW, the sub-seasonal experiment yielded lower mean soil moisture contents compared with the seasonal model runs in the late growing season, especially in August and September of 2017.

L454-456 and 465-467: As in the abstract, I think that the clarity of the results could be improved if you are more clear about what you are measuring here. The phrase "inter-annual differences" that you use is exceptionally clear, but then later you use the phrases "variation" and inter-annual variance, which could be confused with the variance of a time series. Consistently using "inter-annual differences" would clarify this confusion. Alternatively, you could describe differences using the absolute difference in yield levels rather than a percent.

Thanks again for the suggestion. We understand how this wording could be confusing to the reader and thus have rephrased this throughout our manuscript, e.g.;

(Line 570-587): In general, the inter-annual differences in simulated LAI and ET were relatively low in the forecast experiments as well as in the reference simulations. This is also reflected in low inter-annual differences of simulated crop yields. The seasonal experiments were able to reproduce the generally higher inter-annual differences in crop yield throughout the AUS-VIC domain (up to 50 % in records and 17 % in simulated yields) compared to the DE-NRW domain (up to 15 % in records and 5 % in simulated yields). After weather conditions, regional agriculture and crop yield is largely impacted by agricultural management decisions (e.g., on crop varieties, planting dates, irrigation, and fertilizer types and application techniques) as well as other environmental factors such as pests and crop damage from wildlife, which are not sufficiently well represented by CLM5. In addition, the crop module of CLM5 lacks parameterizations for most crop types and varieties and the fertilizer application routine is highly simplified. These deficiencies in the model structure lead to considerable uncertainties in the crop phenology simulated by CLM5.

Thus, the inter-annual variability in crop yield simulated by the CLM5 is primarily influenced by the variability of model forcing data and soil moisture states, as it does not consider further anthropogenic or economic factors affecting crop yield, as discussed above. Consequently, the small inter-annual differences in simulated yield suggest that the CLM5 crop module has limited sensitivity to changes in climate conditions. Uncertainties in the simulated annual crop productivity and its low inter-annual differences can be partly explained by the observed systematic biases of the simulated soil moisture content compared to satellite derived soil moisture products, i.e., ESA-CCI and SMAP L3, and CRNS measurements for both domains.

Discussion:

L492: Perhaps use language that is more precise than saying that the forecasts were "very close" to the results of the reference simulations. The phrase does not give a lot of guidance to the reader and what is considered "very close" is subjective and often application dependent. You could, for example, describe which aspects of the simulations were well simulated (inter-annual yield differences, overall yield levels, etc).

Thanks for pointing this out. We added results in form of total yield differences (t/ha) in the results section and discussion:

(Line 492-494): Inter-annual differences of the mean annual crop yield (averaged for the regarded crops) of up to 1.31 t/ha can be observed in the records, while crop yield simulated by CLM5 showed only differences of up to 0.30 t/ha in the forecast simulations (0.28 t/ha for CLM-SUB) and up to 0.24 t/ha in the reference simulations.

(Line 506-509): While inter-annual differences in crop yield up to 1.23 t/ha were observed in official records, CLM5 simulations resulted in smaller differences of up to 0.45 t/ha in CLM-S, 0.35 t/ha in CLM-SUB and 0.38 t/ha in reference simulations, on average for the regarded crops.

(Line 533-535): Overall, annual crop yield predictions from the forecast experiments were close to results from the reference simulations, with a maximum difference between the mean annual crop yield simulated with forecasts and with reanalyses of 0.28 and 0.36 t/ha for AUS-VIC and DE-NRW, respectively.

L531-533: See above discussion about terminology with respect to “inter-annual yield differences” vs “inter-annual variability”

Please see previous comments.