Response to Anonymous Referee #1

First and foremost, we would like to thank the referee for their time devoted to reviewing our work and providing their comments and suggestions that are helpful for improving our manuscript. We will use these comments and suggestions in the revision of the manuscript. In particular, we will add a brief section to elucidate the key limitations of the modelling framework and provide recommendations on the applications of the framework for simulations in different spatiotemporal contexts.

Here, we will respond (in blue font) to the comments made by the referee (in black) point by point.

General Comment

The study focuses on the effects of climate change on green water availability and water-limited attainable yields (AY) for major cereal crops in Ethiopia. It develops an agro hydrological Modelling framework to simulate climatic-hydrological-crop interactions for the reference year (1981-2010) and future periods (2020-2099) under different greenhouse gas emission scenarios. The study discuss the importance of green water management practices to improve crop yields, especially in rain fed agricultural systems. It identifies a significant gap between actual and water-limited yields, emphasizing the need for integrated management strategies to overcome yield-reducing factors. The findings suggest that future changes in AY vary across regions, emission scenarios, and growing seasons, with temperature increase playing a key role in driving changes. The study concludes by recommending the adoption of climate-smart agricultural practices and the use of agro hydrological Modelling for informed decision-making in agricultural management planning. It is well written and have a significant contribution in green water management beyond the study area.

We thank the referee very much for the positive feedback on our manuscript.

The modelling approach used in this study provides valuable insights about the future without using process-based modelling. However, the approach is subjected to uncertainties and assumptions that may impact the accuracy of the results. Can the authors clarify as limitation or future research considerations in the following issues?

We thank the referee for the comment. Before we go into the specific comments below, we would like to clarify the context of this work. In our modelling approach, we emphasized capturing cascades of climatic, hydrological, and crop yield (in relative terms) information at *a climatological* time scale (average conditions over a 30-year period) at a spatial scale of *5 km x 5 km*. We aimed to provide a bigger picture that could be valuable for long-term agricultural water management planning and policymaking at national and sub-national scales, supporting effective climate adaptation, resilience, and rural economic development in Ethiopia. In doing so, it was necessary to make compromises between the *data* used, *models* employed, *geographical area* covered, and the targeted *information*. It is expected that the uncertainties in data, parameters, as well as modelling concepts, unavoidably propagate to the outcomes we presented in the paper. We will briefly elaborate on these uncertainties in the revision.

- The partitioning of rainfall to runoff is based on the CN approach which is completely empirical and needs locally adjusted CN based on soil, land use and hydrologic conditions. How is this affecting the competition where there are no locally contextualized CN values?

  We acknowledge the uncertainties that could arise from the empirical nature of the CN-based rainfall partitioning method implemented in our modelling framework. As the referee rightly stated, the best way to reduce such uncertainty is to adjust the CN value based on the local context. This is feasible for assessments in specific locations (e.g., fields, farms, landscape scales) and small catchments. However, in our case, it is hardly practicable as the model was implemented at a grid-scale covering a significantly large geographical area (the whole agricultural region of the country, ~667,000 km²), which ideally requires adjusting the CN at every grid, preferably based on observed local soil and surface conditions.

  Instead, we adjusted the CN at every grid, considering: *land use* (we used CN corresponding to agricultural land from the USDA lookup table), *soil conditions* (we assigned the CN to each grid based on the Hydrologic Soil Group defining the soil infiltration characterisitics from Ross et al. (2018)), and *soil mositure conditions* (we updated the CN values daily in the soil water balance model using Equation 6 – on page 7). Another most common practice, especially in catchment hydrological modelling is to optimize the CN value through model calibration with observed surface runoff or stream flow data measured at a designated outlet point (e.g., Arnold et al., 2012; Qi et al., 2020). In our case, we focused on the vertical soil water balance without routing the surface as well as subsurface flow, thus this option does not apply in this context.
  Several hydrological models such as SWAT (Arnold et al., 2012), AGNPS (Young et al., 1989), HEC-HMS (USACE, 2000), etc also use CN methods as one option for infiltration estimation. Additionally, there is scientific interest in deriving high resolution CN estimates that can be combined with state-of-the-art landcover maps from remote sensing (e.g., Jaafar et al., 2019), which increases applicability of this method over a large geographical domain where calibration of this parameter is a challenge.

- What about using other partitioning approaches such as Thornthwaite and Mather soil moisture water balance model widely applied in the area?

  We believe different models are intended primarily for a specific purpose (although they can also be adaptable for other related purposes), and their parameterization, simulation time steps, and other features are defined based on that main purpose. In this regard, Thornthwaite and Mather is a monthly soil moisture and groundwater recharge model (Sishu et al., 2024; Steenhuis and Molen, 1986). It does not account for soil properties and land surface conditions that determine the rainfall-runoff processes, it is informed only by rainfall and potential evapotranspiration conditions (Sishu et al., 2024). In the version modified by the US

Geological Survey (Westenbroek et al., 2010), the surface runoff routine of the model is based on the CN method. We did not compare this model with our modelling framework, but we can say that a monthly model is too coarse to capture the agro-hydrological fluxes that are pertinent to plant growth response. Our agrohydrological modelling framework determines the soil water balance at a daily time step. This is particularly important to better capture the effects of weather-driven daily variations in antecedent soil moisture conditions that determine the partitioning of rainfall into infiltration and surface runoff. In fact, the soil water balance modules of many crop growth and agroecosystem models, such as DSSAT (Jones et al., 2003), AquaCrop (Hsiao et al., 2009; Raes et al., 2009; Steduto et al., 2009), SIMPLE (Zhao et al., 2019), and SWAT+ (Čerkasova et al., 2023), are built on CN-based approaches.

- Can authors clarify how the lateral and subsurface flow likely affect the uncertainty? This kind of flow is very important in humid and sub-humid parts of the landscape of Ethiopia where there are various research outputs highlighting this issue? How can these be considered in future similar research work as this one?
  We thank the referee for the comments. We believe that the uncertainty arising from lateral and subsurface flow is negligible in our results. This is because i) we applied the model only to the root zone (upper 60 cm) where unsaturated flow dominates the agro-hydrological fluxes, and ii) lateral flow is too slow to account for in a daily model at a coarse spatial resolution like 5 km x 5 km, thus change top soil layer water content from grid to grid. It would have an effect only close to the stream network where soil is more saturated and baseflow is produced, which again we do not resolve well with this course grid. Lateral soil water redistribution is important at spatial scales where topography provides sufficient gradients for subsurface flow, e.g., on the order of 1-100 m. Even at this finer scale where the surface slope is flat, in the top 60 cm soil root zone that we considered, the effect of lateral flow can still be low under a rainfed system. This is because the spatial distribution of rainfall and soil are nearly homogeneous, and thus the soil wets homogeneously unless there are conditions that lead to preferential flow (e.g., soil cracks and localized hardpans, as the referee mentioned). Lateral unsaturated flow in the root zone is more common and should be accounted for in simulations applied to irrigated plots, particularly under furrow and drip irrigation methods.

While authors have tried their best to validate the annual flow with literature data, the authors does not compare their result with other studies for yield.

We agree with the referee that we did not compare our yield results with those from other studies. This is simply because the attainable yield we computed is a relative value—the percentage of the potential yield (water-unlimited) that can be attained under the prevailing climate and soil conditions. We did not find any suitable similar study in the context of Ethiopia for comparison. However, we have tried to validate the attainable yield by correlating it to the actual total cereal production derived from the annual agricultural sample survey reports by the Central Statistical

Agency (CSA) of Ethiopia (Figure 4c). We believe that this provides a clue as to the reliability of the modelling framework we implemented.

The water limited attainable yield was estimated based on FAO equation. How certain is the result from the equation? Was it not possible to compare with experimental plots yield under various treatments? It is essential to support their findings with relevant literature and analysis of the result by comparing the increment or decrement of SMD or GWA with other studies too.

We thank the referee for the comments. The FAO productivity function is a basic and established relationship between climate and crop yield, explaining the relative yield as a function of the evapotranspiration ratio, the ratio of the actual to potential evapotranspiration. Crop coefficient-based models are built on this function and are used to assess crop yield responses to water-limited climate conditions (Foster and Brozović, 2018). For our stated purpose—assessing water-limited attainable yield to serve as a basis for reducing crop yield gaps in the rainfed farming systems in Ethiopia—we believe that the FAO water production function is a fair choice over more complex and data-intensive process-based crop models. We recognize the uncertainties arising from the simplicity of this model, but we believe that the level of uncertainty is acceptable, as we tried to demonstrate in our validations. We strongly suggest using process-based crop models if one aims to simulate absolute crop yields, especially at field and farm scales. We will stress these points in the revision of our manuscript. We agree with the referee that such analyses could benefit from validation with field experiments. However, the scale of this particular study makes it challenging to set up experimental plots that sufficiently cover the entire agricultural region. The time frame and resources allocated for this work did not allow for such efforts. Instead, we decided to rely on existing experimental data.

What was the problem collaborating with people and institutions in Ethiopia? There are institutions such as Ethiopia Agricultural Research Institute who do experiments under various agro-ecology for various crops. There was a possibility validating some of the results.

This study was a first level analysis of agroclimatic and hydrological effects on water-limited crop yield potential across Ethiopia. In a higher level analysis, one could focus on farm and landscape scales, at which point local experimental data from the Ethiopian Agricultural Research Institute and others would be interesting. We will pursue this in the future. However, it remains true that our methodology is not directly applicable to small scales which are represented by experimental crop yield data, as our approach is designed for large (national) scale assessment to estimate the 'relative yield'. For the purpose of validation of our results, we have used the best available crop data from the CSA, which we disaggregated from the zonal scale to the grid scale (Wakjira et al., 2021). For surface runoff, we collected published runoff plot measurement data at 17 locations, which we believe is fairly sufficient to test the performance of our model.

Specific comments

Page 4 line 110: why CHRIPS and ERA-5 were used? Why not other products?

In preliminary analyses we did explore a range of other rainfall products, and we concluded that CHIRPS is one of the most suitable daily rainfall data for Ethiopia and a large part of East Africa, offering one of the highest spatial resolutions, temporal continuity, and record lengths (e.g., Ahmed et al., 2024; Bayissa et al., 2017; Dinku et al., 2018; Gebrechorkos et al., 2018; Musie et al., 2019). For temperature and other climate variables, ERA5-Land provides better spatial resolution (9 km x 9 km) compared to other products. We performed bias-correction to ERA5-Land temperature and downscaled it to 5 km x 5 km grid resolution over Ethiopia (Wakjira et al., 2023).

Page 9 section 2.4 Assessment of Green water availability and its yield potential seems like result rather than methodology. So I recommend to put this section in the result part

Thank you. We will consider this in the revision.

Page 10, lines 27–29: The authors evaluated the simulations of AY [%] in terms of their correlation to variations in TCP [tonne y -1], however, the computed values of AY from the model are not mentioned in the manuscript. So, what are the order of magnitudes of these total crop production by showing the range under different agroclimate?

The total cereal production (TCP) is the sum of all cereal crops produced in Ethiopia. The 16-year (1995-2010) average TCP across the study region ranges from 28 tonnes to 24,000 tonnes.

Page 10 section 3.1: the wording of "observation" used in the section is misleading.

Thank you. We will revise our use of the terminology.

Page 12 Figure 4a and b. The performance evaluation of the model for runoff and actual evapotranspiration, the study used R2 and NSE for runoff and R2 for actual evapotranspiration. Why NSE is not used for actual evapotranspiration.

We thank the referee for raising this comment. The corresponding NSE for evapotranspiration is 0.8. We will add this to Figure 12b in the revision.

Figure 4: The 17 surface runoff data points obtained from the literature should be located on the map to see their spatial distributions.

Thank you. We will add the locations of the surface runoff plot data to Figure 12a in the revision.

Page 13 line 3-4: The reference climatology of growing season GWA and water-limited yield across the RFA region based on the computed SMD and AY values is presented, and considering alfalfa reference grass (Ky = 1.1), what is the source for this value?

We thank the referee for the comment. The yield response factor (Ky) values for alfalfa grass were taken from the FAO Irrigation and Drainage Paper 56 (Allen et al., 1998). We will indicate this in the revision.

Page 26 line 485: There are various research works that show the infiltration of the soil is very high compared to rainfall intensity. The problem is the hardpan formation limiting the infiltration through the root zone. The hardpan formation at lower depth facilitates the later subsurface flow with the terrain high slope. Authors need to mention this as part of their recommendation.

Thank you. We believe that this is true for specific locations that are subjected to the conditions that result in hardpan formation such as high heavy machinery traffic like in highly mechanized agricultural lands, flat areas with heavy clay soils exposed to inundation followed by drought conditions, etc. Also limiting infiltration are very steep terrain gradients, which as mentioned above we cannot capture with the coarse spatial resolution of our framework. We will mention these limitations in the revision.

References

Ahmed, J. S., Buizza, R., Dell'Acqua, M., Demissie, T. and Pè, M. E.: Evaluation of ERA5 and CHIRPS rainfall estimates against observations across Ethiopia, Meteorol. Atmos. Phys., 136(3), doi:10.1007/s00703-024-01008-0, 2024.

Allen, R. G., Pereira, L. S., Raes, D. and Smith, M.: Crop Evapotranspiration, Rome., 1998.

Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., Van Griensven, A., Van Liew, M. W., Kannan, N. and Jha, M. K.: SWAT: Model use, calibration, and validation, Trans. ASABE, 55(4), 1491–1508, 2012.

Bayissa, Y., Tadesse, T., Demisse, G. and Shiferaw, A.: Evaluation of satellite-based rainfall estimates and application to monitor meteorological drought for the Upper Blue Nile Basin, Ethiopia, Remote Sens., 9(7), 1–17, doi:10.3390/rs9070669, 2017.

Čerkasova, N., White, M., Arnold, J., Bieger, K., Allen, P., Gao, J., Gambone, M., Meki, M., Kiniry, J. and Gassman, P. W.: Field scale SWAT+ modeling of corn and soybean yields for the contiguous United States: National Agroecosystem Model Development, Agric. Syst., 210(June), doi:10.1016/j.agsy.2023.103695, 2023.

Dinku, T., Funk, C., Peterson, P., Maidment, R., Tadesse, T., Gadain, H. and Ceccato, P.: Validation of the CHIRPS satellite rainfall estimates over eastern Africa, Q. J. R. Meteorol. Soc., 144, 292–312, doi:10.1002/qj.3244, 2018.

Foster, T. and Brozović, N.: Simulating Crop-Water Production Functions Using Crop Growth Models to Support Water Policy Assessments, Ecol. Econ., 152(March), 9–21, doi:10.1016/j.ecolecon.2018.05.019, 2018.

Gebrechorkos, S. H., Hülsmann, S. and Bernhofer, C.: Evaluation of multiple climate data sources for managing environmental resources in East Africa, Hydrol. Earth Syst. Sci, 22, 4547–4564, doi:10.5194/hess-22-4547-2018, 2018.

Hsiao, T. C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D. and Fereres, E.: Aquacrop-The FAO crop model to simulate yield response to water: III. Parameterization and testing for maize, Agron. J., 101(3), 448–459, doi:10.2134/agronj2008.0218s, 2009.

Jaafar, H. H., Ahmad, F. A. and El Beyrouthy, N.: GCN250, new global gridded curve numbers for hydrologic modeling and design, Sci. Data, 6(1), 1–9, doi:10.1038/s41597-019-0155-x, 2019.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., Wilkens, P. W., Singh, U., Gijsman, A. J. and Ritchie, J. T.: The DSSAT cropping system model., 2003.

Musie, M., Sen, S. and Srivastava, P.: Comparison and evaluation of gridded precipitation datasets for streamflow simulation in data scarce watersheds of Ethiopia, J. Hydrol., 579(September), 124168, doi:10.1016/j.jhydrol.2019.124168, 2019.

Qi, J., Lee, S., Zhang, X., Yang, Q., McCarty, G. W. and Moglen, G. E.: Effects of surface runoff and infiltration partition methods on hydrological modeling: A comparison of four schemes in two watersheds in the Northeastern US, J. Hydrol., 581(November 2019), 124415, doi:10.1016/j.jhydrol.2019.124415, 2020.

Raes, D., Steduto, P., Hsiao, T. C. and Fereres, E.: Aquacrop-The FAO crop model to simulate yield response to water: II. main algorithms and software description, Agron. J., 101(3), 438–447, doi:10.2134/agronj2008.0140s, 2009.

Ross, C. W., Prihodko, L., Anchang, J., Kumar, S., Ji, W. and Hanan, N. P.: HYSOGs250m, global gridded hydrologic soil groups for curve-number-based runoff modeling, Sci. data, 5, 180091, doi:10.1038/sdata.2018.91, 2018.

Sishu, F. K., Tilahun, S. A., Schmitter, P. and Steenhuis, T. S.: Revisiting the Thornthwaite Mather procedure for baseflow and groundwater storage predictions in sloping and mountainous regions, J. Hydrol. X, 24(March), 100179, doi:10.1016/j.hydroa.2024.100179, 2024.

Steduto, P., Hsiao, T. C., Raes, D. and Fereres, E.: Aquacrop-the FAO crop model to simulate yield response to water: I. concepts and underlying principles, Agron. J., 101(3), 426–437, doi:10.2134/agronj2008.0139s, 2009.

Steenhuis, T. and Molen, W.: The TM procedure as a simple engineering method to predict recharge, J. Hydrol., 84, 221–229, 1986.

USACE: Hydrologic Modeling System Technical Reference Manual., 2000.

Wakjira, M. T., Peleg, N., Anghileri, D., Molnar, D., Alamirew, T., Six, J. and Molnar, P.: Rainfall seasonality and timing: implications for cereal crop production in Ethiopia, Agric. For. Meteorol., 310, 108633, doi:10.1016/J.AGRFORMET.2021.108633, 2021.

Wakjira, M. T., Peleg, N., Burlando, P. and Molnar, P.: Gridded daily 2-m air temperature dataset for Ethiopia derived by debiasing and downscaling ERA5-Land for the period 1981–2010, Data Br., 46, 108844, doi:10.1016/j.dib.2022.108844, 2023.

Westenbroek, M. S., Kelson, V. a., Dripps, W. R., Hunt, R. J. and Bradbury, K. R.: SWB — A Modified Thornthwaite-Mather Soil-Water- Balance Code for Estimating Groundwater Recharge, U.S. Geol. Surv. Tech. Methods 6-A31, 60, 2010.

Young, R. A., Onstad, C. A., Bosch, D. D. and Anderson, W. P.: AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds, J. Soil Water Conserv., 44(2), 168–173, 1989.

Zhao, C., Liu, B., Xiao, L., Hoogenboom, G., Boote, K. J., Kassie, B. T., Pavan, W., Shelia, V., Kim, K. S., Hernandez-Ochoa, I. M., Wallach, D., Porter, C. H., Stockle, C. O., Zhu, Y. and Asseng, S.: A SIMPLE crop model, Eur. J. Agron., 104(February), 97–106, doi:10.1016/j.eja.2019.01.009, 2019.