Review Zhang et al., “The climate of the Eastern Mediterranean and the Nile River basin 2000 years ago using the fully forced COSMO-CLM simulation”.

General comments

The paper by Zhang et al. presents results from a transient climate simulation covering 500 BCE to 2018 CE. This simulation has been performed with the paleoclimate version of the COSMO-CLM regional climate model (RCM), which was forced with a global simulation carried out with the MPI-ESM-LR model. Such a transient simulation with an RCM is novel and provides a spatially detailed view of the climate. In this paper, the RCM is set up for a domain spanning the Nile River basin and the eastern part of the Mediterranean. The results section consists of two parts. In the first part, the model’s performance is evaluated by comparing the simulated climate for the period 1980-2018 with observations and reanalysis data for the same time frame. This comparison indicates that the model is doing a reasonable job, except for the precipitation in the Nile River domain. The second part provides a comparison of the results for two periods (the Early Roman Period (ERP, 400-362 BCE) and the preindustrial period (PI, 1800-1850 CE)), and presents also an analysis of the link between temperature and precipitation on the one hand, and the atmospheric circulation on the other hand.

Although the performed simulation is innovative and represents a technological improvement that could play an important role in studies of past climates, there are several issues with this paper that need to be resolved before it can be published. These issues are discussed in detail below, and relate to the purpose of the paper, the set-up of the simulation and the analysis of the results.

Thank you very much for taking the time to review out manuscript. We will modify the manuscript accordingly, please see the following responses to the different comments. The primary aim of the paper is to provide a comprehensive characterization of the ERP climate. To achieve this, we, in the first part of the manuscript, focusing on the evaluation of the CCLM output. Additionally, we acknowledge the necessity of refining the employed methodology to offer a more accurate description of the applied analysis. It's essential to clarify that the primary focus of the paper is to compare the BCE climate to the pre-industrial climate and further characterize the climate over EM and NR during the BCE period.

Main comments

Purpose of the paper. It is not clear to me what the purpose of this paper really is. One goal is to provide an evaluation of the performance of the model for the present-day climate, but this evaluation is not presented as being the main aim. The results for the ERP and PI periods are compared, and it may be expected that a characterization of the ERP climate is the main objective of this paper, as the PI period is usually taken as the reference period in paleoclimate modelling. However, the introduction section of this paper does not even mention the Early Roman Period once, and thus does not explain at all why there is a focus on the ERP. My advice is therefore to streamline the introduction section and make a case why it is important to obtain a more detailed climatological understanding of the ERP. This should include an explanation of why the 400-362 BCE period is chosen. Why not another period, and why these specific 38 years? To my knowledge, there is no established definition of the Early Roman Period that falls between 400 and 362 BCE, so this requires a thorough discussion.

Thank you very much for the suggestions regarding on how to more concisely present the purpose of the manuscript. The purpose of this paper includes two parts, part one: an evaluation of the performance of the model for the present-day climate; part two: compare the ERP and PI periods, and a characterization of the ERP climate. We agree with your

opinion of adding a paragraph in the "Introduction" section to thoroughly discuss the reason of why we are choosing the ERP period specifically. For choosing these two periods, we have analyzed the time series from 500 BCE to 1850 CE of the MPI-ESM output and the forcing data. In the study of (McCormick et al., 2012a), one have tried to explore the relationship between climate change during the Roman Empire period. Nevertheless, this study only goes to 100 BCE while the entire roman empire starts back to around 500 BCE. Thus, together with comparing the long time series of precipitation and temperature from the MPI-ESM output, we choose this period: 400-362 BCE as this period goes early back to the time when the roman empire is rising from a kingdom. In addition to this, the volcanic activities (Toohey and Sigl, 2017) are also taken into consideration for choosing these two periods.

Given the substantial computational and time resources required for this process, the Regional Climate Model (RCM) paleo transient simulation results spanning the entire 2500-year period is not available now, therefore comparing the entire period output with proxy records is in our future study scope.

Set-up of the simulation - CCLM. The results are obtained with the newly developed paleoclimate version of the CCLM model. However, it is not explained what has been modified compared to the "normal" version. More details on the model should thus be provided. What is meant by the phrase "by implementing CCLM in model version 5.0 with CLM version 16"? this phrase suggests that CCLM is a different model from CLM, but CCLM just stands for COSMO-CLM, doesn't it? And what is actually the spatial resolution of the version of CCLM applied here?

We have implemented the external forcings (solar, orbital, volcanic, GHG and land use change) to the normal version of CCLM. The reason for this is that in paleoclimate applications of CCLM, those forcings have never been fully implemented while they are in the driving models. For example in some papers, when it is concerning the climate of LGM (Ludwig et al., 2016), one has implemented the orbital forcing specifically, and in some other research regarding the GHG, one have implemented the GHG forcing into CCLM. But until now, there is not such a simulation using CCLM with all the forcing implemented. Thus, we have decided to run a fully forced CCLM simulation with all the external forcings.

Regarding the confusion of the phase: "by implementing CCLM in model version 5.0 with CLM version 16", here we intend to say by implementing are the external forcing into the CCLM (COSMO model version 5.0 with CLM version 16). We are going to correct that also in the manuscript.

For choosing the spatial resolution for this simulation (0.44 °), we did test simulations with 0.11° and 0.44° horizontal resolution. By comparing the results of simulation output for these two different spatial resolutions, we did not find much added value of higher resolution at 0.11 in the regard of presenting the regional precipitation and temperature. But instead, by improving the horizontal resolution to 0.11, the computation time and storage resources consumed is much higher (16 times) than 0.44 °. Thus, we have chosen 0.44 ° for this specific simulation.

Set-up of the simulation – applied forcings. Several forcings are applied in the simulation, but it is not clear what the main differences between these forcings are. The authors mention orbital, solar, GHG, volcanic and land-use changes. I suggest explaining what the main differences in radiative forcings are between the main periods of interest in this paper: present period, PI and ERP. How do each of these radiative forcings change between the three periods? What forcings are most important, and how large are the differences when expressed in Wm-2? This information is important for interpreting the results and should be included.

We are currently preparing a paper which mainly discussing the set up and the changes in the different external forcing (Hartmann, et al. 2024). Due to the different research focus, regarding the detailed information of the forcing are presented in Hartmann, et al. 2024. We can share the manuscript confidentially to further illustrate the above-mentioned information. The external forcings are based on the recommendation for the PMIP4 past1000 contribution to CMIP6 (Jungclaus et al., 2017). Here is some more detailed information:

The orbital forcing is represented by the eccentricity, the obliquity and the longitude of perihelion. Those values constantly change with time. The effect of the orbital forcing is further explained in the next response.

The total solar irradiance represents the solar forcing. The ERP is in a solar minimum phase where TSI values maximum as high as the lowest values of PI. Today's values are similar to PI.

The volcanic by the aerosol optical depth. Both periods cover silent periods and larger volcanic eruptions.

The changes of greenhouse gas concentrations are given in equivalent CO2 concentrations which consists of CO2, CH4 and N2O. The PI has higher values as the ERP. For the present-day simulation we use only CO2 concentrations at the level of 1950, which is lower than both equivalent CO2 concentrations.

The land use changes are introduced as LAI, maximum and minimum plant coverage derived from the MPI-ESM transient simulation.

Analysis of the results. The differences between the ERP and PI climates should be discussed in terms of the differences in radiative forcings between the two periods (see point 3 above). In addition, it should become clear how the results for the ERP compare to proxy-based climate reconstructions for the same period. On line 497 it is noted that "comparing modeled data with proxy records is essential for a comprehensive understanding of climate variability", but surprisingly the authors do not make such a comparison for the results presented in this paper. This is clearly a missed opportunity, and in my opinion a model-proxy comparison should be included. Moreover, the model results should also be discussed more extensively relative to previous modelling studies.

We will include the information of radiative forcing of the two periods as well as in the discussion. Nevertheless, as we also plan to submit a paper dedicated to explaining the setup and external forcings, we will only bring up the difference of external forcing such as radiative and volcanic over the two periods. For details regarding the forcings please refer to the response to the "set up of the simulation".

For the further information about the radiative forcings between the two studied period, we will include a new plot here which shows the differences of the insolation of the two periods (see Figure R1). For the volcanic forcing, we have used the reconstruction from Toohey and Sigl, 2017.

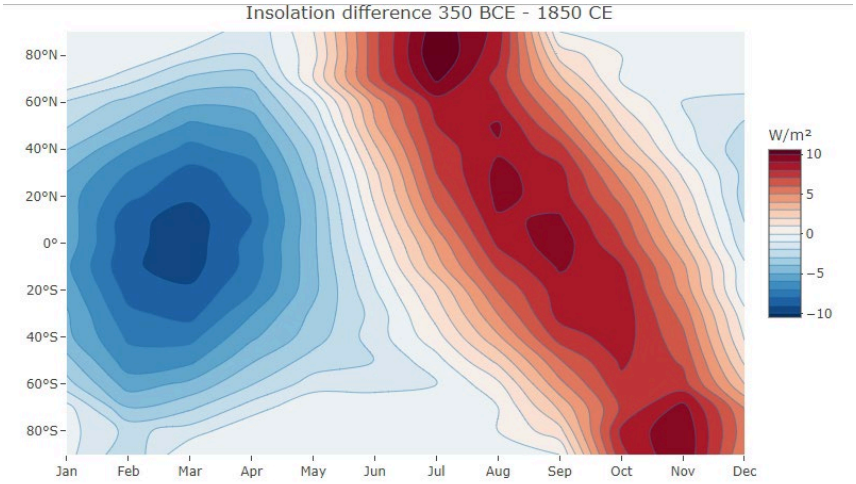

Figure R1. The insolation difference between the centuries of ERP and PI caused by different orbital forcing values.

As comparing to the proxy records. We completely agree that it will be largely important to the whole study, as we are currently analyzing the data. Nevertheless, given the substantial computational and time resources required for this process, the Regional Climate Model (RCM) paleo transient simulation results spanning the entire 2500-year period is not available now, therefore comparing the entire period output with proxy records is in our future study scope.

Additionally, this paper, has a different focus in evaluating the RCM simulation results and characterized the ERP climate. We will include in the discussion how this study can benefit the future study such as comparing the simulation output to proxy records. Nevertheless, we will modify the introduction and give a more consistent presentation of the

overall subject of the paper. We are working specifically to and preparing a manuscript which is dedicated to compare our simulation output with the proxy records.

Discussion. The present paper presents the first transient RCM simulation of the past 2500 years, so one question to answer is if there is a clear advantage of making the effort of running an RCM in transient mode. If the authors would run three snapshot RCM experiments for present-day, PI and ERP, to what extent would the results be different? In other words, what have we learned about the ERP climate from the transient simulation that could not be obtained from a snapshot experiment with mean forcings for the ERP?

The added value of running the RCM in a transient mode will be discussed in detail another manuscript we are preparing currently, as mentioned in the response to comment: set up of the simulation. In that paper, the impacts of running a transient RCM with all the forcings are presented. If necessary, we can share the manuscript with the reviewer confidentially in order to resolve the confusion mentioned in the comment.

Other comments

All the other comments will be addressed as suggested. Especially the plots with significance dot have been modified by only plotting the significant grid box.