Anonymous Referee #1

a) The manuscript entitled "Resolution-dependence of future European soil moisture droughts" by Eveline C. van der Linden presents differences between low (standard) resolution and high-resolution runs of EC-Earth with respect to drought conditions over central Europe. They find that droughts happen to be more severe and durable in the high-resolution experiments and explore potential causes leading to the differences between the model runs.
The manuscript is generally in a good shape, mostly well structured and well written. The overall presentation of the results is good with mostly concise and high quality figures. The methodological approach is well explained and technically sound, but some clarifications are needed in this context. However, the considered ensemble is relatively small and some of the conclusions might not be robust. The authors thus need to discuss some limitations of their approach before final publication.

We wish to thank the reviewer for this positive evaluation, and for the constructive comments and insightful suggestions on our paper. They have helped us to substantially improve the quality of the manuscript. Our detailed responses to the comments are presented below.

Major comment

b) The ensemble used in this study consists of 6 ensemble members, each containing 5 years of simulation, resulting in a sample size of 30 years. This is mostly fine for assessing the averaging land surface water and energy balance components. However, I doubt the sample size is large enough for assessing droughts. Droughts are extreme events. An extreme drought event is thus defined as a 1/30 year event, which does not correspond to the 1st quantile, simply because the sample size is too small. It would be good if the authors could discuss the robustness of their results and provide a concise reasoning why the ensemble is not larger. Please also provide some insights into why you choose the years 2002-2006 for present day climate conditions. Maybe because these are the last yers of the CMIP5 historical runs? I was just wondering because in 2003 central Europe experienced a major drought and heatwave.

Thank you for these questions. We agree that, ideally, extreme drought events as discussed in this paper should be studied with longer model runs. Unfortunately, there are currently no longer runs at this high resolution available for EC-Earth and we currently do not have the resources to increase the ensemble size. This experiment with exceptionally high spatial resolution for a global climate model is computationally very expensive and was therefore performed for multiple research questions. The larger ensemble approach with shorter runs was motivated by a study on the impact of climate change on teleconnection responses to specific tropical SST patterns (Haarsma et al.

2013). As you suggest, the focus of the current paper should therefore be on average changes in the land surface water and energy balance components and on the underlying physical processes rather than on extreme drought events. Therefore we will shift the focus from soil moisture droughts, which are extreme events by definition, to mean changes in soil drying and the associated physical processes, for which 30 years are sufficient. In the revised paper, we will change the title and text accordingly. We will discuss the impact of soil drying on droughts shortly as an impact in the final section of the current paper and we will add a note in which we clarify that a larger sample size is required to obtain robust answers.
Years 2002-2006 are indeed chosen for present-day conditions since these are the last years of the CMIP5 historical runs. This choice was made for multiple research questions and therefore did not consider the major drought and heatwave of 2003. Since this is a model ensemble with perturbed initial conditions, there will not be a major heatwave in each 2003 ensemble member though.

More comments

c) Introduction: It might be good to add a few more references in the first part of the introduction. Especially regarding the uncertainty in European drought trends, such as e.g. Samaniego et al., 2018 (DOI: 10.1038/s41558-018-0138-5)

We will add more references in the first part of the introduction, including your suggestion.

d) p. 1, l. 22-25: You write that potential evaporation is enhanced through larger atmospheric moisture demand due to the increasing temperatures. You also write that humidity and wind speed might affect evapotranspiration. This is all a bit confusing, since the atmospheric moisture demand is also defined through humidity and wind speed. Maybe consider to rephrase this part.

Thank you for this comment. We understand that this part could be a bit confusing. In the first part of this paragraph we focus on global mean changes (higher temperatures and the associated increase in saturation vapour pressure), whereas the second part focuses on regional changes such as wind speed and associated moisture transports. We will rephrase this paragraph to make the clear distinction between global mean and regional effects.

e) p. 2, l. 3: What do you mean by hydroclimatic components?

Precipitation and evaporation. We will clarify this in the text.

f) P. 2, l. 12-14: This is also a bit confusing. You write about quantifying drought severity, and later drought characteristics (such as e.g., severity). Seems redundant.

Thank you for pointing this out. We will remove the redundancy.

g) p. 2, l. 18: Please outline what variables are needed to compute PDSI.

Following the suggestions of the second reviewer, we have decided to limit the discussion on the PDSI and other off-line drought metrics and we will concentrate more on studies of actual soil moisture changes. The introduction paragraph focusing on PDSI will therefore be replaced.

h) p. 2, l. 26-27: The north-south wetting vs. drying pattern in Europe is actually a well-known feature which was assessed in many studies.

We agree that this large-scale pattern has been shown in many studies. However, our point is that the magnitude of these changes is highly uncertain and that over many land areas in the transition regions between north and south it is even uncertain whether there will be a wetting or drying trend. Regionally, there are still large uncerainties. We will rephrase this part to avoid confusion.

i) Sec. 2.1 and 2.2: Why do you choose the years 2002-2006? How do the model runs differ? Why dont you use more recent SST data? What version of HTESSEL do you use? Does HTESSEL include river routing or where does the runoff go? Are there open water bodies in HTESSEL?

We choose years 2002-2006 because these are the final years of the CMIP5 historical runs. Since the individual ensemble members have perturbed initial conditions, the exact years are not of major relevance. We will explain this more extensively in the text.
The model experiments were performed about five years ago, therefore the SST data are not the most recent ones. However, our purpose is to demonstrate the effect of model resolution on future soil moisture changes. To demonstrate this effect these experiments were suitable.
Open water bodies in HTESSEL are represented by open (or frozen) water tiles in the land surface scheme, as described in the data and methods section. HTESSEL does not include river routing. For each grid cell, runoff is transferred to a designated region in the ocean.
We will add these explanations to the 'data and methods' section.

j) p. 5, eq. 1: Do you really mean dθ within the integral? Shouldnt it just be θ?

Thank you for noticing this mistake. We will correct it in the revised manuscript.

k) p. 8, l. 15: The validation period is 1982-2011, right? Might be worth to mention this here as well. What happens if you choose just 2002-2006 as validation period? How is an event like 2003 represented in the model?

Thank you for these questions. We understand that this period could be confusing. We will add the section on model validation to the observations description, since this is what the observations are used for. Please note that we use an ensemble of model simulations with perturbed initial conditions (we will explain this in the experimental setup). This indicates a 'present-day'-like forcing similar to 2002-2006, but due to the perturbed initial conditions these years are (and should) not be exactly the same as the observed 2002-2006 conditions. The initial conditions create climate variability, so the simulated 2002-2006 years are not supposed to be exactly the same as the observed 2002-2006 state. Moreover, using only 2002-2006 as a validation period is not statistically robust, since the period only represents five years. The impact of internal climate variability on such a time scale is large, both in the model and in observations. To compare the climatology, you need to compare about 30 years of data. That is why we use a validation period of 30 years for the observations as well. We will explain this in the text.

l) p. 9, eq. 4: Well, dS/dt is not necessarily just soil moisture. This might include also snow/ice water storage and water in open water bodies. How is this represented in HTESSEL?

In HTESSEL, a grid box is either 100% land or 100% sea. Each non-land point (grid point with less than 0.5 land cover) can have two fractions, open and frozen water. Open water bodies are thus not included in dS/dt, which only takes land grid points into account. Each land point (grid point with 0.5 or more land cover) can have six fractions of which two include snow (snow on low vegetation/bare ground and high vegetation with snow underneath). Snow is treated as an additional surface layer on top of the upper soil layer. You are therefore right that snow and ice (permafrost) are part of the soil water S in this context and a melt term should be added. However, since we focus on the warm season months and choose our study region outside mountainous regions, ice and snow can be neglected in the land surface water balance. We will add this explanation in the text.

m) p. 9, l.10-11: Does the soil water content determines runoff?

We understand that this sentence implies a causal relation between soil water content and runoff, which is only partly correct (as described under the next comment) since it is rather a negative feedback. We will rephrase this paragraph.

n) p. 9, l. 23-27: Here it would help if you could provide more information on how runoff is treated in HTESSEL.

Runoff consists of two parts: surface runoff and subsurface runoff. Whenever precipitation or snow melt occurs a fraction of the water is removed as surface runoff. The ratio runoff/precipitation scales with the standard deviation of orography, and depends on the complexity represented in the gridbox, as well as on soil texture and soil water content. Subsurface runoff is associated with free drainage at the bottom. We will add this information in the description of the land surface scheme and will rephrase these sentences accordingly.

o) p. 12, l. 19: remove "is"

Thank you, we will correct this sentence.