Response to RC2 on nhess-2022-162

NOTE: Reviewer's comments are in black, our responses to the comments are given in blue below.

The manuscript describes an evaluation study of EURO-CORDEX hindcast simulations and additional WRF simulations with focus on the SPEI drought index. The methods are clearly described and mostly sound. In particular, the authors discuss the potential benefits of higher spatial resolution and regionally-tuned model setup, which may help to design efficient yet suitable model setups in future studies. The manuscript is suggested for publication after minor revision as described below.

Minor Revisions suggested:

Please provide your motivation for the model selection you applied (only 6 simulations). E.g, I'm pretty sure that EURO-11 simulations with several WRF configurations are available since many years. Why are they not used in this study?

> The models selected are all those that were available at that time that met two requirements: containing the necessary variables (precipitation, Tmax and Tmin) and coverage of the period 1980 – 2009. To make this clearer in the text, we will change it in L131 to: "At the time of selection, these were all available model runs that cover the study period 1980 – 2009 and contain the relevant meteorological variables needed for the analysis."

In most practical applications, RCM simulations are only used after bias correction. Did you analyze the effect of bias correction (of temperature and precipitation) on your results? Please comment.

> Bias correction is usually based on observational data sets and especially meant for purposes of future projections, which are out of scope for this study. We aim at evaluating RCM performances regarding their capabilities to simulate droughts, not at bias correction (which could be the next step for a projections study). Here, the RCMs were forced by reanalysis data, where observations are already included. Since also requested by reviewer 1, to give the reader an idea of the biases involved, we have calculated the spatially and temporally averaged bias values of the monthly values of the relevant variables and will add them along with some description. No relation could be detected between the bias values and the respective RCM SPEI performances. This suggests that a bias correction would probably not have much effect. This finding we will add to the conclusions as well.

I would like to see a critical discussion on the practical relevance of the SPEI in climate change (i.e. trend) studies. Since NHESS has a focus on hazards and the hazard is rather the agricultural drought than the meteorological drought, the question arises to which degree the SPEI is able to describe drought hazard and what the of the SPEI to serve as proxy for agricultural droughts are. E.g., SPEI does not cover the effect of increasing surface runoff during heavy precipitation events, which, however, is a loss in the soil moisture budget. Similarly, SPEI does not regard increased transpiration due to longer vegetation periods in a warmer climate. I respect that this is not the main topic of the manuscript and therefore cannot be treated in a quantitative manner, but it is a very important boundary information and should be discussed in the conclusions and/or introduction sections.

> We can understand the points listed here. The SPEI describes the rainfall deficit in a certain time period based on a reference period. It does not describe any further impacts except that a lack of soil moisture (especially important for agriculture) or discharge, water levels etc. is assumed after a certain number of months with dry anomaly. As stated, the points mentioned are especially relevant in the climate change context, but climate change was not an issue in

the strict sense of this study. Nevertheless, we see the importance of this point and will add some words to the methods section 3.2: “In this context it should be emphasized that the SPEI (and also SPI) has limitations regarding the practical relevance for climate change, when the focus is primarily on impacts. Apart from an implied lack of soil moisture (agricultural drought) and decline of streamflow, groundwater, reservoir and lake levels (hydrological drought), which completely rely on the degree of dry anomaly over a certain time period, impacts going beyond this are not addressed. Due to the complete reliance on dry anomaly, effects of a warming world (e.g., longer vegetation period and thus modified transpiration behavior), cannot be included either or would be considered only indirectly.”.

L295: “Therefore, it can be assumed that if existent, the benefits of a resolution increase from 5 to 12.5 km are less distinct,” A single simulation does not allow drawing such general conclusions. This statement refers also to other conclusions (e.g. line 321, 322).

True, a single simulation is not enough for such clear and strict general conclusion. Thus, the sentence will be adjusted to “From our results, we obtain that, if existent, the benefits of a resolution increase from 12.5 to 5 km are less distinct.”. The sentence in line 321 will be adjusted to: “Our findings indicate that the WRF benefits can be attributed to the WRF model settings and not to the increased resolution.”. Moreover, in the sections 4.3 and 4.4 some sentences will be changed to make a clearer connection to our study.

L329: “Because of the results in the previous section, here we focus on the values from the WRF@5 km run in direct comparison to the reference values from OBS (Figure 3).“ In the previous section, you argue that the 15km WRF simulation outperforms the 5km simulation, at least for temporal correlation. How does this lead to the decision to use only the 5km simulation in the subsequent section?

This is a valid objection. Therefore, we will add the WRF@15 km map to Figure 3 and the relevant scores to Table 5. That implies adding a bit more text to the section as well as to the conclusions.

More generally speaking: By (partly) removing the 15km WRF simulation from the analysis, you lose the option to directly compare it to the WRF 5km simulation and therefore you lose the most direct indicator for the effect of the spatial model resolution. On the other hand, major results of the study are statements like “computation resources could therefore be saved, since a coarser resolution can provide similar results” (Abstract) or “WRF’s increased resolution and setup is turned out to be beneficial in the analysis of the monthly values of the meteorological variables and the correlations of the SPEI time series”. Therefore, I suggest keeping both the 5km and 15km WRF simulations in each part of the analysis in order to better support your conclusions. (Side benefit: The additional Fig.5 could be avoided, if you had the 15km simulation in the general analysis (Figure 4).)

We agree that more information of the 15 km simulation would be desirable. Therefore, we will add information from the WRF@15 km run in every section, at least in the tables and give further descriptions in the text, e.g., to compare the two WRF runs with each other. Adding one more domain map to the figures would increase the figure dimensions, since there would be nine domain maps in total per figure. Especially in section *4.5 Drought Characteristics Analysis* we consider the information from the tables as more meaningful. Thus, Figure 5 needs to be kept this way.