

## General Remarks for *Greciano-Zamorano et al. 2026*

In this study, the authors present a new high-resolution simulation focusing on the Sierra de Guadarrama, located in the center of the Iberian Peninsula for the period 1991-2020. The dataset was created with the WRF model, and ERA5 reanalysis data provided the initial and boundary conditions. To test the accuracy of the simulated precipitation, many indices were compared against observations from the Spanish Meteorological Agency and the Guadarrama Monitoring Network. This manuscript is valuable for the climate research community working on high mountain environments, but also for broader topics such as regional climate modelling or hydrology.

The manuscript follows a logical structure and fits into the scope of Geoscientific Model Development. However, the authors need to address some comments before it is ready for publication.

### Major comments:

1. The authors used a remarkable monitoring network installed and maintained in such a complex environment. However, Table 1 shows that some stations are only available for 3 or 4 years instead of the long coverage provided by most of the stations. For example, stations 3 (Hoyas), 8 (Cabeza Mediana), 18 (Alameda) and 26 (Herrera). Since the simulations are run for 30 years, I would not include these stations in the analyses, as they can produce fictitious biases when compared against the entire period as done in the first part of the analyses (Figs 2 and 3). In fact, the authors already mention this in lines 184-186. Thus, what is the reason to include these potentially problematic stations in the analysis?
2. The authors employ the New Tiedtke scheme as cumulus parameterization even in the 1 km resolution domain. However, at the end of section 2, they mention some sensitivity experiments done for year 2009 to evaluate the performance of the model when the parameterization is switch-off in that domain, or when using alternative parameterization options. However, why was 2009 considered to run those experiments? Did these alternative runs also include a 1 year spin-up as the main simulation?
3. Also related to the sensitivity experiments. The authors stated in the discussion that they obtain better results with these simulations than with the configuration of the main simulation (406-408). Thus, why was that configuration used in the first place and not the one from the sensitivity experiments?
4. The authors investigate the precipitation also during the days in which it rains in reality. The authors justify this stating that “the assessment focuses on the realism of the simulations in representing precipitation when it rains in the reality” (lines 156-157). However, using such metric, I think that the authors are not evaluating the full performance of the model and that they are biasing their results. For example, such metric is only applicable in this particular case in which some stations are available, but it could not be valuable for alternative regions where the stations

are limited. Thus, I would consider removing this metric from the final version and using only the one based on the *temporal masking*, which includes also the days in which the model simulates rain, but it is not observed in reality.

5. Figures 2 and 3. I don't see the point of including ERA5 information as points, as if each grid-point was a station. Why not including an additional panel showing the map from ERA5 data. Additionally, I think the figures would benefit from including some contours at different elevations, such as 1500 m and 2000 m. I suggest those values since they are used in lines 220-226.
6. Some parts of the text mention that it is difficult to say with the available stations if some precipitation amounts are realistic or not (e.g., line 212). To investigate more on that, why some information from gridded precipitation datasets based on satellite data have not been included in the comparison? I think the manuscript would benefit from a complementary source of information about precipitation.
7. Figures 4 and 6: If I am not mistaken, these figures show the mean of the annual accumulated precipitation (mean of 30 values since 30 years are available). However, I think that the analysis would benefit from including some kind of information about the variability during the entire period. This information is included in Figure 7, but I think it should be included too in these figures.

### **Minor comments:**

Line 7: the WRF 4.2.2 model -> the WRF model (version 4.2.2)

Lines 70, 142, 413: use the n-dash for the periods stated in the manuscript. Example: 1991–2020

Lines 85-87: I would reference only once the Table 1 in the sentence.

Lines 130-131: I think there is a typo in the labels of the domains. WRF1 should be D1 and the 3 km resolution domain D2.

Line 221: appreciable (orange boxes; Fig. 4) precipitation -> appreciable precipitation (orange boxes; Fig. 4)

Figure 8. The labels of the bottom right panel are too small.