Summary of paper:

The paper presents a framework for detecting drought and categorizing it into seven types based on climatic drivers. It then summarizes the severity, duration, and frequency of each drought type at each grid box. This framework is applied to mHM hydrologic model outputs based on observed climate forcing data and GCM forcing data over the European domain, and is used to evaluate historical GCM-based drought types and their projected future changes.

Overall comments:

Overall, the drought-type analysis presented in the paper is interesting and offers deeper insight into drought mechanisms and how they are likely to evolve in the future. I have a few comments below, including line-by-line comments, but I think they are all minor. I think the paper would be ready for publication after one revision.

Major comments

The mHM model use temperature-based ET and snow accumulation and melt modeling approaches. I would say this is acceptable for the historical evaluations, given that calibration was performed using historical climate data. However, because these critical processes are oversimplified and may be predominantly sensitive to temperature rather than to other meteorological forcing variables, such as humidity, longwave radiation, shortwave radiation, and wind speed, this type of process parameterization could be questionable for future periods, when climate conditions may be outside the historical range. In other words, there is no guarantee that the calibrated parameters will remain optimal, and ET and snow dynamics may be sensitive to changes in other climate variables in the future. In the discussion, there are brief comments on this PET limitations, but I would suggest expanding this discussion to explain why the current PET formulation, and possibly the snow model, have potential limitations for analyses involving climate change.

Even though the GCM data are bias corrected, the reference meteorological data used for GCM bias correction are not EOBS data (please clarify this). Different observed or reference datasets may have systematic differences among them. Therefore, even though mHM works well with EOBS, simulations forced by the reference meteorological dataset used for GCM bias correction, as well as by bias-corrected GCM data, may produce systematic differences relative to mHM simulations forced by EOBS. The question is: how different are temperature and precipitation from the bias-corrected GCMs from EOBS during the historical period? If there are systematic differences, how do these forcing differences affect hydrologic fluxes, states, and drought detection? Do these forcing differences contribute to the underestimation of drought events presented in Section 3.1?

I wonder whether the CMIP5 GCM forcing is the mean of ensemble members or is taken from a particular ensemble member for each GCM. Note that this is not the same use of the term "ensemble" as in the paper: each CMIP GCM run may have several different historical-future

traces based on GCM initial conditions, in addition to different emission scenarios. I do not think this was mentioned in Section 2.1. I doubt that the GCM forcing is an ensemble-member mean, but I wanted to check. If it is an ensemble-member mean, extremes would be reduced, which could contribute to fewer drought detections (or less severity). Also, since drought detection was percentile based, how is the percentile at each location computed? Is it computed separately for each forcing dataset at each location?

Specific line-by-line comments (minor)

L45: Please explain what "warm" and "cold" snow mean. I guess "season" is missing? I would suggest explaining these terms here, rather than only referencing past papers.

L58-59: Regarding the rain-to-snow mechanism: From the explanation in parentheses, I understand that this refers to a deficit of snowfall, leading to low snow accumulation and affecting spring-summer water availability. Is this interpretation correct? Or does this have something to do with the snow to rain transition due to weather patterns in specific years and/or climate change? The term "rain-to-snow" does not seem to match the explanation in parentheses. Reading further down, there are also snow-related processes. From this part of the paragraph, it is unclear how the rain-to-snow and snow-related processes are different.

Ok, Section 2.4 explain the drought types in detail. But I still suggest revising short description on drought types in the introduction. It might be helpful to include the paper outline in the end of introduction, so the readers expect the details of drought types in later section.

L65: GCM -> General Circulation Model or Global Climate Model.

Figure 2: It is a little strange to see a few white boxes that do not have a split. My understanding is that a white box represents a decision-node-like step, so it should have a split. Basically, is the box labeled "Transition to dry period" is description of "a wet to a dry season"? The same question applies to the warm snow season. For a decision tree like this, I would suggest removing these white boxes and characterizing the drought type, or terminal node, somewhere in the text.

Section 3.1: I suggest adding the total number of drought events and the bias for the GCM-based results.

L211: It appears to me that wet-to-dry drought has the lowest biases overall, not rain-to-snow drought.

L224: Please specify what is meant by temperature-dependent drought. Does this include all types except rainfall-deficit drought?

L252: Here, the temperature-dependent droughts are specified. I would suggest describing this in Section 2.4.

L317-318: This sentence, stating that timing is stable, somewhat contradicts the statement that duration is generally longer. To me, there are some shifts in the starting and ending dates, with earlier starts and later endings; for example, WCE-rainfall. This overall statement does not seem to be accurate.

L379-380: Could this framework be used with other hydrologic or land-surface model outputs, in addition to its ability to examine other climate forcings? If so, explicitly mentioning this would promote the usability of this framework by allowing users to examine uncertainty arising from hydrologic model choices.

L389-390: This sentence on the Zenodo data repository should be moved to the Data availability section.