All your comments are found below, with our answers written in **bold**.

1. Regarding the precipitation occurrence modeling, the authors employ a Bernoulli distribution while acknowledging that temporal dependence structures cannot be adequately incorporated into the precipitation occurrence process due to the infeasibility of transforming binary precipitation occurrence into Gaussian random variables. This represents a notable limitation, as the temporal structure in precipitation occurrence may not be sufficiently preserved under this approach. Given this acknowledged technical constraint, I would encourage the authors to discuss why traditional Markov chain-based methods, which are well-established for modeling temporal dependencies in precipitation occurrence, were not adopted as an alternative.

   Thank you for pointing this out. Based on your suggestion, we developed an additional downscaling model, where precipitation occurrences are modelled using a 1. order Markov chain, with time- and location-dependent transition probabilities that are modelled using the same covariates as in the original occurrence model. The new model strongly outperforms the old occurrence model, and will be added into the revised manuscript. In the new occurrence model, we use two distinct global models for modelling precipitation occurrences. One of these describe the probability that day $t+1$ is wet, given that day t is dry. The other describes the probability that day $t+1$ is wet, given that day t is wet. All local occurrence models are also exchanged with a pair of local models, describing the same two probabilities. More details will be provided in the revised manuscript. Figure 1 displays the difference in skill for the full precipitation downscaling model when we replace the old occurrence model with the new Markov chain model. The upper row is one of the Figures in the submitted manuscript, although slightly changed, because we now use more years of data, after a comment from reviewer #1. The lower row will be one of the Figures in the revised manuscript.

2. While the authors acknowledge that their proposed downscaling method treats precipitation and temperature independently and thus cannot capture multivariate relationships between these variables, it remains unclear whether the framework is capable of preserving spatial or inter-site dependence structures within the downscaled precipitation and temperature fields themselves. The ability to maintain realistic spatial coherence is critical for many hydrological applications, and I recommend that the authors provide additional results and discussion addressing the performance of their method in reproducing spatial dependence patterns across the downscaled domain.

   Thank you for this comment. We have now performed additional experiments to evaluate the spatial coherence when downscaling temperature and precipitation to multiple nearby locations.
   Our downscaling method is built to simulate an ensemble of temperature/precipitation time series at any one specific location. It is possible to repeat this procedure for any number of locations, e.g., for a regular grid of any resolution, covering a catchment. However, to turn these sets of ensembles into a proper spatiotemporal ensemble, one would need to find a way of matching ensemble members from each location in a way that reproduces the internal spatiotemporal variability of local temperature and precipitation. The method for reproducing the spatiotemporal variability would essentially be a spatiotemporal version of the Schaake Shuffle (Clark et al., 2004), which would work for ensembles of time series, at locations without any available observations. To the best of our knowledge, no

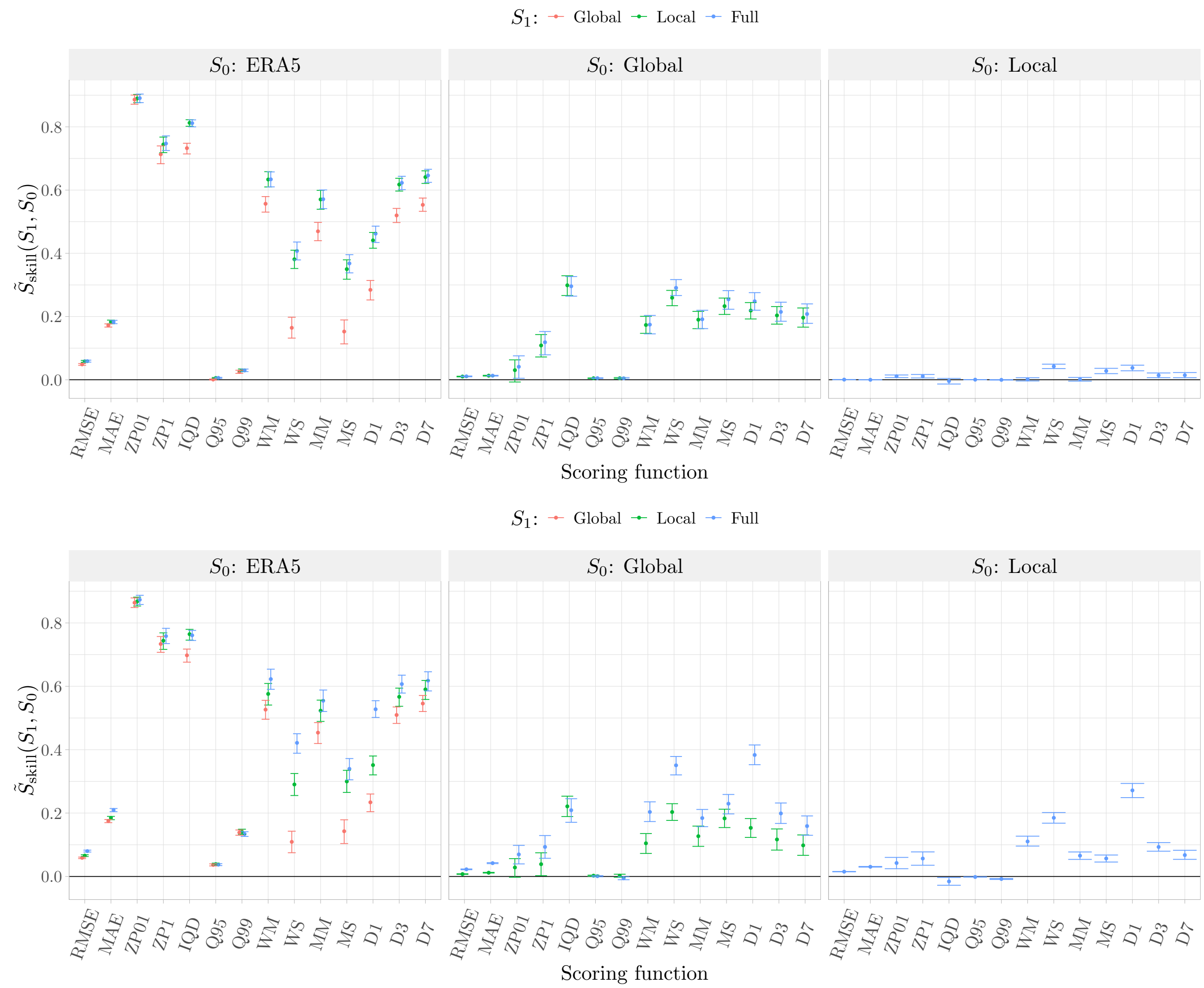

Figure 1: *Upper row: Precipitation skill scores with the old occurrence model. Lower row: Precipitation skill scores where the occurrence model in the "Full" model has been replaced by the new Markov chain model. The "Global" and "Local" models are the same as in the upper row, i.e., using the old occurrence model.*

such method exists in the literature. Instead, we can examine the performance of the simplest possible matching method, where the different ensemble members are randomly ordered and then matched such that all of the first ensemble members together form the first spatiotemporal ensemble member, all of the second ensemble members form the second spatiotemporal ensemble member, and so on. This method is not expected to perform well if the spatial variability is dominated by small-scale variability that cannot be described using ERA5 together with high-resolution elevation data. However, if most of the spatial variability can be explained using elevation data and large-scale weather patterns, the spatiotemporal ensemble should perform well.

In our new experiment, we create a large selection of spatial domains with 100 km radii, all containing at least 8 weather stations. Within each domain, we create spatiotemporal precipitation and temperature ensembles, using the locations of all available weather stations inside the given domains. Then, we aggregate the

simulated weather in space, to create time series of the average, median, minimum and maximum temperature or precipitation inside the domain. We also create time series of the standard deviation of temperature or precipitation inside the domain. These time series are then compared with observed time series, created using the available weather stations. This allows us to evaluate the spatial properties of the spatiotemporal ensemble, and to compare with the spatiotemporal properties when not downscaling, i.e., only using ERA5. Figure 2 displays overall skill scores for these spatially aggregated statistics, using ERA5 as the base model and our downscaling models as the competitors. These will be added to the revised manuscript, together with an explanation of how they were created, and a discussion spatial consistency for downscaling.

References

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., & Wilby, R. (2004). The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields. *Journal of Hydrometeorology*, *5*(1), 243–262. https://doi.org/10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2

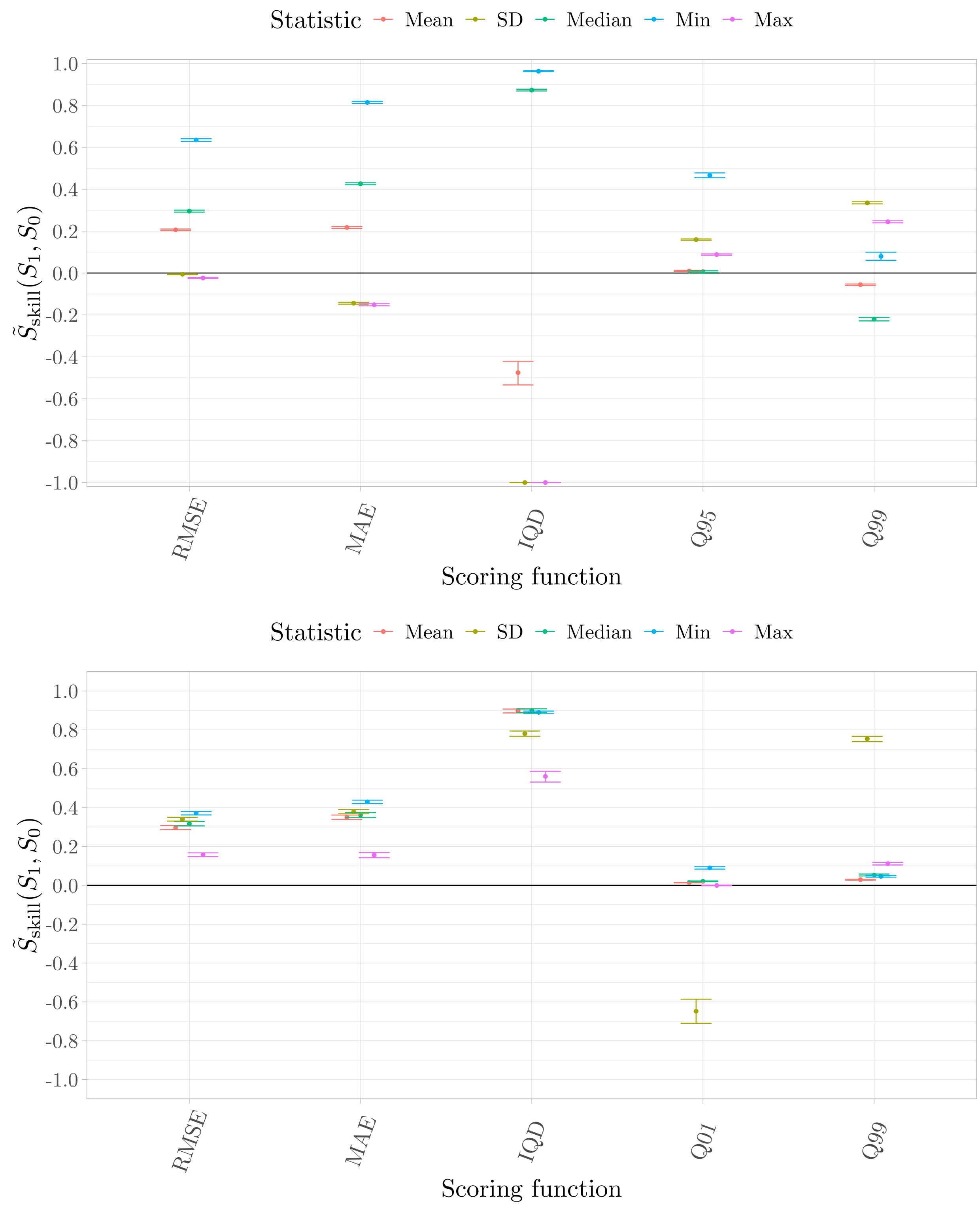

Figure 2: *Upper row: Skill scores for the five spatially aggregated precipitation statistics, using ERA5 as the base model, and the new full precipitation model as the competitor. Lower row: Skill scores for the five spatially aggregated temperature statistics, using ERA5 as the base model, and the full temperature model as the competitor.*