Hydrol. Earth Syst. Sci. Discuss., 12, C1062–C1065, 2015
www.hydrol-earth-syst-sci-discuss.net/12/C1062/2015/

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Downscaling future precipitation extremes to urban hydrology scales using a spatio-temporal Neyman–Scott weather generator" *by* H. J. D. Sørup et al.

Anonymous Referee #1

Received and published: 16 April 2015

From what I understand this study has two main parts: In the first part, the Neyman-Scott-Rectangular-Pulses weather generator (WG) is introduced, and existing approaches extended to a high-resolution, urban hydrology context, with a model grid adapted to an observational database of 60 gauges in hourly time resolution. In the second part the WG parameters are perturbed through change factors coming from 6 different RCMs, and from that future scenarios are developed.

The topic of the study is thus quite ambitious. While the overall presentation generally appears to meet the required standards, crucial details are still often missing, and a final judgement whether everything adds up is difficult.

The study crucially depends on hourly RCM data. RCM bias, but also direct validation on present data, are evaded through the use of change factors. Some problems or questions remain, however: Two of the six RCMs are from the ENSEMBLES project. The smallest available time resolution of ENSEMBLES is 6 hours, see e.g. van der Linden and Mitchell (2009), or directly http://ensemblesrt3.dmi.dk:8080/opendap/hyrax/data/A1B/DMI_ECHAM5. The authors should clarify how Gregersen et al. (2013) come up with a HIRHAM-ECHAM of 1 hourly resolution (dry hours are moreover defined via a fairly large threshold of 0.22mm). For the other 4 RCMs Mayer et al. (2015) is referenced, but that source uses only historic simulations (based on ERA interim). - To sum up: the crucial hourly time resolution of RCMs, and their reliability, should be discussed or should at least be mentioned.

The focus of the paper is on extremes and how they may change in the future. This usually requires thorough validation of the methodology on present and past. I mentioned RCM reliability, but also the WG validation itself is poor, with too many judgements of the kind 'reasonably well', 'realistic levels', etc., and too few statements based on rigorous statistics. Two figures (5 and 6) serve to demonstrate WG performance. For Fig. 5 (basic temporal statistics) the authors explicitly state that WG does not reproduce extremes, and Fig. 6 (spatial decorrelation) failed to convince me that the WG was really site specific.

To conclude, the study is well written and important, but exactly because of the latter it requires a much better and more rigorous validation.

Some more specific comments:

I wonder whether the science might have been better served by leaving the two parts separated, since each of them - going to higher spatio-tmporal detail and producing climate change scenarios - have their own, intrinsic issues.

Related to that, it seems easy to apply change factors *directly* to the high-resolution

observations. Could this not be used as a first guess and used in comparison to the WG results?

The term 'change factor', if I am not mistaken, is used ambiguously. First, by referring to the ratios of certain RCM statistics from which climate change information is derived; second, as a WG statistic characterizing the change in extreme events. This gave me a real headache to follow the text.

The literature seems to prefer the name (ST)NSRP for WG. The authors should consider to use that name (NSRP).

There are strong trends in the observational database from the growing number of available stations. Any possible climate trend cannot clearly be separated from this trend, so the fitted sensitivity to climate remains unclear.

From the text (numbers refer to lines):

2566, 28: ...scenarios *driving* the RCM....

2567, 26ff: please define the core components of WG: rain cell origins, storm origin, cell center, etc.

2568, 7: should come before parameter list

2568, 12: Can intensity and duration be treated as independent? IDF curves prove otherwise.

2568, 25: The CGD dataset might just be too short.

2570, 21: Perhaps a word on WG computing times is in place here.

2571, 3: Put this before the definition of Y_i,j,k. - What exactly are the 'key statistics'? If e.g. the dry-day probability is one, one should not use a multiplicative change factor.

2571, 4: Again, it is confusing if things are discussed *before* they are defined.

2571, 5: I have not found anything appropriate in Burton et al (2010). Calculating

change factors is a crucial point, and so far it is completely unclear how to derive a change factor e.g. for dry hours from RCMs whose temporal resolution is (probably) daily.

2574, 1: Why 'unconditional'?

2575, 6: This paragraph shows that the fitting is quite unstable, which is bad news. The reader needs to know how this affects the simulation results.

2576, 4: Fig. 5 a-e are undefined. What is "reasonable"? - For 1h variance and skewness the results are pretty bad. Based on this figure, if representing extremes is desired I would judge this to be a failure. For the probabilities the WG seems to be stuck at specific levels (horizontal lines).

2576, 11: Likewise: What means "seem to reflect the overall behaviour"? - The reader will not be able to judge, based on these figures alone, whether the decay seen in WG compares well to SVK or not, and whether it might potentially compare equally well to decorrelations from *any* dataset.

2577, 1: Fig. 7 shows 100y events (not 50y). How does the accuracy in this figure relate to the discrepancies seen in Fig. 5?

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 2561, 2015.