This paper compares three different weather generators (WG) based on flow analogues and stochastic weather generators, including a new technique. They analyse the properties of the time series produced by these WG, with a focus on their ability to simulate extremes of precipitation and temperature. I find this paper very interesting and well written.

We thank the reviewer for this very positive feedback.

I just have a few minor questions for the authors.

We appreciate these comments and we will modify the manuscript accordingly.

C1: I am not sure I understand how you link the station data with the reanalyses data. If you use them to calculate daily MAP and MAT of your analogues, how do you calculate daily MAP and MAT for the 1900-1930 period for which you have ERA-20C data but no station data ? Do you only use station data as your observation to which you compare the WG's simulations ?

For the period 1900-1930, it is true that only atmospheric reanalyses are available, as station data are available from 1930 to 2014. As a consequence, possible analogue candidates are only seeked during the period 1930 to 2014, for which reanalysis and station data are available. For the period 1900-1930, analogues are thus sampled among possible dates in the period 1930-2014. The statistical link between the large-scale atmospheric configuration and the small-scale weather is thus inferred from 84 years and used in a "temporal extrapolation mode" to simulate the weather of a period where we only have the large-scale information. This will be clarified in Section 3.3.1.

As a consequence, also, the results of our simulations are compared to observations that do not correspond to the same period. The simulations cover 110 years and the observations 84 years (for all figures from 4 to 8). This is clarified in the figures.

C2: Could you specify the distance you use to calculate analogues ?

Analogues are selected according to the Teweles–Wobus score (TWS) proposed by Teweles and Wobus (1954). This score has been found to lead to higher performances than a more classical Euclidian or Malahanobis distance (Kendall et al. 1983; Guilbault et Obled, 1998; Wetterhall et al., 2005). It quantifies the similarity between two geopotential fields comparing their spatial gradients. It allows selecting dates that have the most similar spatial patterns in terms of atmospheric circulation at a given (or several) geopotential level(s). As a consequence, it does not compare the absolute values of the geopotential fields between 2 days. This paragraph will be included in the revised version of the manuscript.

Kendall, M., Stuart, A., Ord, J.K., 1983. The Advanced Theory of Statistics. Design and Analysis, and Time-series, vol. 3. Oxford Univ Press, New York. 780 p.

Teweles J, Wobus H. 1954. Verification of prognosis charts. Bulletin of the American Meteorological Society 35: 2599–2617.

Guilbaud S, Obled C. 1998. Prévision quantitative des précipitations journalières par une technique de recherche de journées antérieures analogues: optimisation du critère d'analogie (Daily quantitative precipitation forecast by an analogue technique: optimisation of the analogy criterion). Comptes Rendus de l'Académie des Sciences – Series IIA, Earth and Planetary Science Letters 327: 181–188. doi:10.1016/S1251-8050(98)80006-2.

Wetterhall F, Halldin S, Xu CY. 2005. Statistical precipitation downscaling in central Sweden with the analogue method. Journal of Hydrology 306: 174–190. doi:10.1016/j.jhydrol.2004.09.008.

C3: Why do you use HGT1000 rather than SLP ?

Raynaud et al. 2017 analysed the predictive skills of numerous atmospheric predictors and more particularly of geopotentials. They selected HGT1000 rather than SLP to compare the sensibility of the results to different geopotential levels (500, 700 or 1000hPa). However, when using the Teweles–Wobus distance SLP and HGT1000 would give very similar results as the positions of highs, of lows and the gradients are quite comparable between 1000hPa and ground level.

Raynaud, D., Hingray, B., Zin, I., Anquetin, S., Debionne, S., & Vautard, R. (2017). Atmospheric analogues for physically consistent scenarios of surface weather in Europe and Maghreb. International Journal of Climatology, 37(4), 2160-2176.

C4: The first time you introduce the term "scenario", I would define what you mean by scenario right away. As you surely know this word has several different meanings in climate science so it can be a bit confusing if you do not define it clearly.

We thank the reviewer for this suggestion. In this paper, a scenario indicates a possible realization of the climate for the considered period. Indeed, it differs from climate projections which are also including prescribed projections of socioeconomic global changes. A comment will be added in the revised manuscript.

C5: In the discussion, when you discuss the problem to produce temperature extremes because of climate change, you can expend this to extreme precipitations. There is an observed and projected trend on precipitation extremes related to anthropogenic climate change related to the Clausius-Clapeyron relationship.

We thank the reviewer for this comment. Indeed, many evidences have shown that temperature and precipitation extremes are expected to increase in many regions of the world as a result of the global warming.

Concerning the increase of temperature extremes, as discussed at l. 420-428 of the current manuscript, the limitations of SCAMP+ concerning the generation of hot summers and cold winters are very likely related to the temperature increase experienced over the 20th century, which appears clearly when looking at the hottest summers and the coldest winters. Additional experiments will be performed in order to verify this assumption. In details, we will detrend observed temperatures using a regional linear long-term trend, as done in Evin et al. (2018). We will then redo all the analyses on these detrended temperature observations. We expect this pre-processing to solve (at least partly) this particular issue.

Concerning the increase of precipitation extremes, it is difficult to illustrate this aspect with our case study. Figure 1 below shows the evolution of annual maxima of temperature (absolute differences compared to the period 1930-1960 for the 26 stations) and precipitation (relative differences compared to the period 1930-1960 for the 105 stations). While the trend for temperature extremes is clearly observed, the trend for precipitation extremes seems absent. However, as the literature indicates siginificant trends for precipitation extremes in other regions of the world, as well as expected trends in the future, we agree that this aspect deserves an extended comment in the discussion.

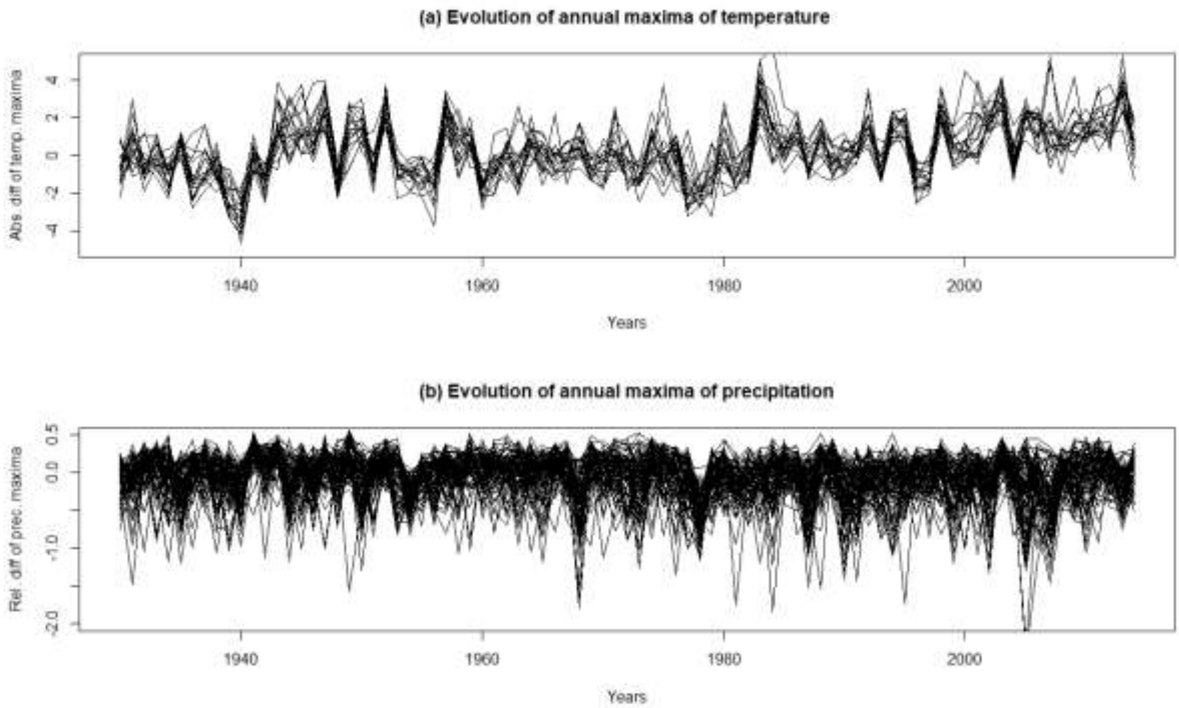

Fig. 1: Evolution of annual maxima of temperature (absolute differences compared to the period 1930-1960 for the 26 stations) and precipitation (relative differences compared to the period 1930-1960 for the 105 stations).

Evin, Guillaume, Anne-Catherine Favre, and Benoit Hingray. 2018. "Stochastic Generators of Multi-Site Daily Temperature: Comparison of Performances in Various Applications." *Theoretical and Applied Climatology*, February, 1–14. https://doi.org/10.1007/s00704-018-2404-x.