HESS-2019-40

Title: Uncovering shortcomings of a weather typing method statistical downscaling method

Authors: Els Van Uytven, Jan De Niel and Patrick Willems

We would like to thank the reviewer for the constructive feedback and comments. Answers to the comments have been made (italic font) and the manuscript has been changed accordingly (blue font). We would like to note that the manuscript has changed significantly to address the comments of all reviewers and to improve the readability.

Response to reviewer comments (Anonymous Referee #1)

Summary and Overall Quality: This research investigates the fidelity of a weather typing based statistical downscaling strategy used to generate hydrometeorological forcing with respect to several of the underlying assumption implicit to these methods. In particular, they evaluate assumptions relating to the robustness of predictor predictand relationships - their predictive power, stationarity, and sensitivity to greenhouse gas forcing - and how well those relationships are captured by coupled models. The focus of this research is a case study for downscaling of precipitation and temperature for a catchment within Belgium and makes use of an established weather typing based downscaling strategy that also includes use of Clausius-Clapeyron (CC) scaling adjustments. The authors find informative relationships between the chosen weather-type predictors and forcing variable. While the coupled models capture the general relationships, they exhibit significant biases in particular with respect to the frequency of the underlying weather types. The predictor-predictand relationships also exhibit non-stationarity. The authors find the use of CC-scaling adjustments result in the downscaling method being able to generate more extreme values and account for changes in variance. Overall, the manuscript is well organized, though the readability could be improved through more detailed formulation of the methods rather than the extensive narrative.

Specific Comments:

(1) There is very little direct formulation of the SDM within the manuscript; it is mostly left to either supplementary material or to an extensive list of references. This left the manuscript feeling less than "self-contained," and readability could be improved with more direct formulation of the methods. This should include moving the WT-formulation from supplementary material into the primary manuscript.

REPLY: The WT formulation in supplementary information has moved into the main manuscript. The readability has also been improved by removing some of the redundant information in the manuscript and the more direct formulation and explanation of the methods and the results.

(2) There are a number of different datasets that are being included. However, there is very little information/discussion on why these data were selected, and it is confusing how data are being used. Why were ERA-40 and NCEP/NCAR used when these are older-generation reanalyses?

REPLY: Precipitation time series are available for the station in Uccle since 1901. We have, however, only access to the time series between 1901 and 2000. The range of this precipitation time series has been compared with the range of different re-analysis datasets. The comparison points out that the older generation re-analysis datasets cover the largest part of the available observed precipitation time series. More specifically, the re-analysis datasets and the observations have data in common for the period 1957-2000.

The resolution of the data are disparate; how was weather typing applied to each dataset? Were they all resampled to the coarsest resolution data (5x5) to allow for consistent WT-metrics to be defined? If not, how might the fact that the finer resolution data are likely to capture more variability affect the frequency distributions of the different WT? Were all the CMIP models resampled to the same resolution?

REPLY: The WTs for the re-analysis datasets and for the climate model runs have been determined considering resolution of the WT classification system. This information was originally provided in supplementary information and has been moved into the main manuscript.

How is the in situ, station data, being used in the compositing? Are all of the precipitation composite information being drawn using only the station data? That is, are the reanalysis only being used for developing the WT-classification and the results are just different regroupings of the underlying precipitation; or are the reanalyses precipitation actually being *composited as well?*

REPLY: The WT classification system is applied to the re-analysis datasets. Next, the produced WTs are coupled to the observed precipitation amounts, providing the historical pool with WTs and their associated precipitation amounts. We have rewritten the methodology to better explain the coupling between the precipitation time series for the RMI station in Uccle and the associated WTs based on the output of the reanalysis datasets.

(3) It is not clear if the station precipitation data can be used together with the hydrologic model. Specifically, the hydrologic model appears to have been calibrated (i.e. tuned to) a different observational dataset with likely a different climatology compared to that of the climatology of a single station time series. This may limit the applicability of using downscaled forcing (to that of a single station) to a dataset with a different climatology than that used to calibrate the hydrologic model.

REPLY: In a study for the Flemish Environment agency, De Niel and Willems (2016) investigated the spatial and temporal variations in precipitation time series for 43 rain gauges in Flanders. Their results indicated significant differences between west (coastal area) and east (Antwerp, Flemish Brabant and Limburg). As Uccle is situated in central Belgium and is located approximately 100 km from the Grote Nete catchment, the application of the hydrometeorological time series for Uccle as input series for the hydrological model of the Grote Nete catchment involves small uncertainties. The calibration of the hydrological model is, however, performed using precipitation, potential evapotranspiration and discharge time series for stations in the Grote Nete catchment.

We remark that the hydrological climate change impact analysis is removed from the manuscript for sake of brevity.

(4) Results indicated super-CC scaling of precipitation changes. This indicates potentially significant components of non-thermodynamic generated forcing, either the frequency and/or intensity of weather types. The author's decomposition seems to only account for frequency changes of WT and/or precipitation changes, but is rolling-up covariant (deviation) terms into "other" forcing. A more detailed decomposition may be warranted to better understand the demonstrated super-CC scaling along with projected changes; specifically Figure 9 "other" should be more thoroughly decomposed.

REPLY: The decomposition of the precipitation changes into contributions arising from the dynamic and thermodynamic processes has been performed for the average daily precipitation amount, projected by the climate model output. Indeed, a more detailed decomposition of the precipitation changes could be performed, as for instance done by Kröner (2016) and Kroner et al. (2017).

We would like to point out that the decomposition is performed for direct climate model output. This means that the CC relation has been indirectly considered in the climate models and is not directly applied as done by the downscaling methodology. Moreover, in the case that the results for the downscaled time series would have been used, then it is questionable whether the CC relation influences the average daily precipitation amounts. More specifically, the CC relation influences the more extreme precipitation amounts, not the average precipitation amounts.

(5) Figure 10 is used to establish the lack of stationarity of the underlying relationships. However, the predictor-predictand relationship appears to only be evaluated with respect to temporal changes without any control for temperature changes. Given that the used SDM implements a temperature-dependent CC-scaling, it is possible that controlling for temperature changes (and CC-scaling) in addition to temporal changes may show that the utilized predictor-predictand relationship is actually stationary as long as temperature-dependency is also included. If accounting for temperature-dependent scaling related changes results in a stationary relationship, then this would provide a more robust justification for the use of CC-scaling as part of the SDM.

REPLY: The stationarity assumption has been evaluated using the re-analysis based WTs and the observed precipitation amounts. Hence, focus is solely put on the relation between WTs and precipitation and temperature is indeed not considered. As pointed out by the reviewer, the stationarity assumption could become more accurate when also considering temperature as a predictor. The latter could be verified by defining surrogate climate model runs and apply the SDM to the surrogate climate model runs. We however note that the empirical precipitation distributions for the W WTs for the periods 1981-1990 and 1931-1941 differ over the entire range of return periods. The application of the CC relation would thus not resolve the differences.

One of the other reviewers (Mohammad Sohrabi) wondered whether other strategies exist to test the stationarity assumption. After carefully re-reading some references, we modified the verification of the stationarity assumption. In summary, the stationarity assumption implies that the relation between the predictors and the predictand remains time-invariant. In other words, the predictors-predictand relation, which has been established using historical observations, should remain applicable under climate changes. Assuming the stationarity assumption is valid, the individual contributions by the dynamical and thermodynamic processes to the precipitation processes would not change.

In this context, the decomposition of the precipitation changes is also applied to surrogate climate model runs. The latter runs are defined by splitting the observed time series in different smaller time series. The decomposition of the surrogate climate model based precipitation amount changes is thereafter compared with the decomposition of the longterm global climate model based precipitation amount changes.

The decomposition of the surrogated based changes indicates that the contribution by the dynamical processes is important and is thus not negligible. The influence of the large scale atmospheric circulation on precipitation in winter season is comprehensively described in literature (Boé and Habets, 2014; Sousa et al., 2017; Tabari and Willems, 2018; Willems, 2013). The results furthermore indicate that the thermodymical processes gain importance at the end of the 20st century. The latter in agreement with results of Ntegeka and Willems (2008), identifying an intensification of the precipitation amounts due to the increasing temperatures.

In the modified manuscript, the approach to verify the stationarity assumption has been replaced and above discussion has been added.

(6) A potentially novel component of this work is related to the CC-scaling adjustments and implementation. However, it does not appear to be emphasized within the manuscript as much of the relevant material is placed in the supplementary manuscript. The scope of this work would be more novel with a stronger focus on these aspects and less on the general analysis of GCM biases in weather-type frequency and intensity; perhaps the former (CC) could be emphasized throughout the paper and the latter included in a more condensed fashion.

REPLY: In the modified version of the manuscript, more focus is put on the application of the CC relation and less on the biases.

Specific Comments:

(1) Line 16: “160% to 240%” : This is confusing. Is the increase 60% to 140% of current day’s values or is the increase truly 160% to 240% more than today’s values (i.e. increases of 100% is a doubling of today’s values). Please clearly state.

REPLY: The increase is estimated between 160% and 240%. Modified.

(2) Line 31: “downscaling and,” : There are several instances in the manuscript where the comma is placed after “and” in a compound sentence. In these cases, the “,” should be placed prior to the conjunction.

REPLY: Modified.

(3) Line 31: “by (Hewitson et al., . . .)” : There are multiple instances in the manuscript where the full references are encapsulated within parentheses but should instead only have the publication year within parentheses. For example, 2)31, 3)10, and 3)23. Please carefully proofread.

REPLY: Modified.

(4) Line 25: Redundant use of “independent”

REPLY: Modified.

(5) Line 27: Figure 3 is noted but it should be Figure 2. Note that all figure numbers in the narrative should be double-checked.

REPLY: Modified.

References

Boé, J. and Habets, F. (2014). Multi-decadal river flow variations in France. Hydrol. Earth Syst. Sci., 18, 691-708.

De Niel, J. and Willems, P. (2016). Setup of a procedure for validation and correction of meteorological data (original title: Opzetten van een validatie- en correctieprocedure voor meteorologische data), 59pp. Leuven: KU Leuven – Afdeling Hydraulica.

Kröner, N., Kotlarski, S., Fischer, E., Lüthi, D., Zubler, E. and Schär, C. (2017). Separating climate change signals into thermodynamic, lapse-rate and circulation effects: theory and application to the European summer climate. Climate Dynamics, 48 (9-10), 3425–3440.

Kröner, N. (2016). Identifying and quantifying large-scale drivers of European climate change. ETH, PhD manuscript, doi:10.3929/ethz-a-010793497.

Sousa, P. M., Trigo, R. M., Barriopedro, D., Soares, P. M. M., Ramos, A. M. and Liberato, M. L. R. (2017). Responses of European precipitation distributions and regimes to different blocking locations. Climate Dynamics, 48(3-4), 1141-1160.

Tabari, H. and Willems, P. (2018). Seasonally varying footprint of climate change on precipitation in the Middle East. Scientific Reports, 8 (4435), 1-10.

Willems, P. (2013). Multidecadal oscillatory behaviour of rainfall extremes in Europe. Climatic Change, 120 (4), 931-944.

HESS-2019-40

Title: Uncovering shortcomings of a weather typing method statistical downscaling method

Authors: Els Van Uytven, Jan De Niel and Patrick Willems

*We would like to thank the reviewer for the constructive feedback and comments. Answers to the comments have been made (see below) and the manuscript has been changed accordingly (*supplement to the comment, blue font*). We would like to note that the manuscript has changed significantly to address the comments of all reviewers and to improve the readability.*

Response to reviewer comments (Anonymous Referee #2)

The study aims to evaluate different common assumptions inherent statistical downscaling methods (SDMs). The overarching goal of this study is certainly an important one. However more needs to be done to increase the scientific significance of this study. My few main comments are below:

(1) Why focus on WT based SDM only? Why not choose at least one more type of statistical downscaling method such as Bias-correction and Spatial downscaling method, which is a widely used SDM. I think inclusion of at least one more SDM would allow the study to more appropriately address the overarching goal.

REPLY: The main objective of this study is to verify and evaluate the general and structural statistical downscaling assumptions in order to develop a statistical downscaling ensemble tailored to the case study and thus end-user needs. Most studies address the general and structural statistical downscaling assumptions independently. This results in studies addressing one or some of general statistical downscaling assumptions (Dixon et al., 2016; Fu et al., 2018; Haberlandt et al., 2015; Hertig et al., 2017; Mendoza et al., 2016; Merkenschlager et al., 2017; Salvi et al., 2016; Tabari et al., 2016) and other studies addressing the structural assumptions by statistical downscaling of surrogate climate model runs (Bürger et al., 2012; Gutmann et al., 2014; Hertig et al., 2018; Maraun et al., 2018; Roberts et al., 2019; Werner and Cannon, 2016; Widmann et al., 2019; Yang et al., 2019) or by statistical downscaling of the projected climate model output (Li et al., 2017; Sørup et al., 2018; Sunyer et al., 2015; Vaittinada Ayar et al., 2016; Wang et al., 2016; Wootten et al., 2017). However, to objectively identify shortcomings of statistical downscaling methods, the verification and evaluation of the general and structural assumptions should be carried out simultaneously. To the authors knowledge, there are yet no papers which simultaneously address the verification of both types of assumptions.

We agree that verifying the general and structural assumptions for another SDM can be of interest, but such investigation is not performed in the considered study for sake of brevity.

(2) The SDM method needs to be described better. Please consider including the section S1 into the main script.

REPLY: The supplementary information on the methodology of the weather typing method has been placed into the main script.

(3) Please also clarify in greater details the novelty of this study and implications beyond the study domain.

REPLY: We refer the reviewer to the modified introduction and conclusions.

(4) I also think the manuscript can benefit from a thorough copy editing.

REPLY: The quality of the manuscript has been improved.

References:

Bürger, G., Murdock, T. Q., Werner, A. T., Sobie, S. R., and Cannon, A. J.: Downscaling Extremes—An Intercomparison of Multiple Statistical Methods for Present Climate, Journal of Climate, 25, 4366–4388, https://doi.org/10.1175/JCLI-D-11-00408.1, 2012.

Dixon, K. W., Lanzante, J. R., Nath, M. J., Hayhoe, K., Stoner, A., Radhakrishnan, A., Balaji, V., and Gaitán, C. F.: Evaluating the stationarity assumption in statistically downscaled climate projections: is past performance an indicator of future results?, Climatic Change, 135, 395–408, https://doi.org/10.1007/s10584-016-1598-0, 2016.

Fu, G., Charles, S. P., Chiew, F. H., Ekström, M., and Potter, N. J.: Uncertainties of statistical downscaling from predictor selection: Equifinality and transferability, Atmospheric Research, 203, 130–140, https://doi.org/10.1016/j.atmosres.2017.12.008, 2018.

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., and Rasmussen, R. M.: An intercomparison of statistical downscaling methods used for water resource assessments in the United States, Water Resources Research, 50, 7167–7186,https://doi.org/10.1002/2014WR015559, 2014.

Haberlandt, U., Belli, A., and Bárdossy, A.: Statistical downscaling of precipitation using a stochastic rainfall model conditioned on circulation patterns - an evaluation of assumptions, International Journal of Climatology, 35, 417–432, https://doi.org/10.1002/joc.3989*, 2015.*

Hertig, E., Merkenschlager, C., and Jacobeit, J.: Change points in predictors–predictand relationships within the scope of statistical downscaling, International Journal of Climatology, 37, 1619–1633, https://doi.org/10.1002/joc.4801, 2017.

Hertig, E., Maraun, D., Bartholy, J., Pongracz, R., Vrac, M., Mares, I., Gutiérrez, J. M., Wibig, J., Casanueva, A., and Soares, P. M. M.: Comparison of statistical downscaling methods with respect to extreme events over Europe: Validation results from the perfect predictor experiment of the COST Action VALUE, International Journal of Climatology, pp. 1–22, https://doi.org/10.1002/joc.5469, 2018.

Li, J., Johnson, F., Evans, J., and Sharma, A.: A comparison of methods to estimate future sub-daily design rainfall, Advances in Water Resources, 110, 215–227, https://doi.org/10.1016/j.advwatres.2017.10.020, 2017.

Maraun, D., Widmann, M., and Gutiérrez, J. M.: Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment, International Journal of Climatology, pp. 1–12, https://doi.org/10.1002/joc.5877, 2018.

Mendoza, P. A., Mizukami, N., Ikeda, K., Clark, M. P., Gutmann, E. D., Arnold, J. R., Brekke, L. D., and Rajagopalan, B.: Effects of different regional climate model resolution and forcing scales on projected hydrologic changes, Journal of Hydrology, 541, 1003–1019, https://doi.org/10.1016/j.jhydrol.2016.08.010, 2016.

Merkenschlager, C., Hertig, E., and Jacobeit, J.: Non-stationarities in the relationships of heavy precipitation events in the Mediterranean area and the large-scale circulation in the second half of the 20th century, Global and Planetary Change, 151, 108 – 121, https://doi.org/10.1016/j.gloplacha.2016.10.009, 2017.

Roberts, D. R., Wood, W. H., and Marshall, S. J.: Assessments of downscaled climate data with a high-resolution weather station network, International Journal of Climatology, 39, 3091–3103, https://doi.org/10.1002/joc.6005, 2019.

Salvi, K., Ghosh, S., and Ganguly, A. R.: Credibility of statistical downscaling under nonstationary climate, Climate Dynamics, 46, 1991–2023, https://doi.org/10.1007/s00382-015-2688-9, 2016.

Sørup, H. J. D., Davidsen, S., Löwe, R., Thorndahl, S. L., Borup, M., and Arnbjerg-Nielsen, K.: Evaluating catchment response to artificial rainfall from four weather generators for present and future climate, Water Science and Technology, 77, 2578–2588, https://doi.org/10.2166/wst.2018.217, 2018.

Sunyer, M. A., Hundecha, Y., Lawrence, D., Madsen, H., Willems, P., Martinkova, M., Vormoor, K., Bürger, G., Hanel, M., Kriaučiūnienė, J., Loukas, A., Osuch, M., and Yücel, I.: Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe, Hydrology and Earth System Sciences, 19, 1827–1847, https://doi.org/10.5194/hess-19-1827-2015, 2015.

Tabari, H., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Saeed, S., Brisson, E., Van Lipzig, N., and Willems, P.: Local impact analysis of climate change on precipitation extremes: are high-resolution climate models needed for realistic simulations?, Hydrology and Earth System Sciences, 20, 3843–3857, https://doi.org/10.5194/hess-20-3843-2016, 2016.

Vaittinada Ayar, P., Vrac, M., Bastin, S., Carreau, J., Déqué, M., and Gallardo, C.: Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: present climate evaluations, Climate Dynamics, 46, 1301–1329, https://doi.org/10.1007/s00382-015-2647-5, 2016.

Wang, L., Ranasinghe, R., Maskey, S., van Gelder, P. H. A. J. M., and Vrijling, J. K.: Comparison of empirical statistical methods for downscaling daily climate projections from CMIP5 GCMs: A case study of the Huai River Basin, China, International Journal of Climatology, 36, 145–164, https://doi.org/10.1002/joc.4334, 2016.

Werner, A. T. and Cannon, A. J.: Hydrologic extremes - an intercomparison of multiple gridded statistical downscaling methods, Hydrology and Earth System Sciences, 20, 1483–1508, https://doi.org/10.5194/hess-20-1483-2016, 2016.

Widmann, M., Bedia, J., Gutiérrez, J. M., Bosshard, T., Hertig, E., Maraun, D., Casado, M. J., Ramos, P., Cardoso, R. M., Soares, P. M. M.,Ribalaygua, J., Pagé, C., Fischer, A. M., Herrera, S., and Huth, R.: Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment, International Journal of Climatology, pp. 1–27, https://doi.org/10.1002/joc.6024, 2019.

Wootten, A., Terando, A., Reich, B. J., Boyles, R. P., and Semazzi, F.: Characterizing sources of uncertainty from global climate models and downscaling techniques, Journal of Applied Meteorology and Climatology, 56, 3245–3262, https://doi.org/10.1175/JAMC-D-17-0087.1, 2017.

Yang, Y., Tang, J., Xiong, Z., Wang, S., and Yuan, J.: An intercomparison of multiple statistical downscaling methods for daily precipitation and temperature over China: present climate evaluations, Climate Dynamics, https://doi.org/10.1007/s00382-019-04809-x, 2019.

HESS-2019-40

Title: Uncovering shortcomings of a weather typing method statistical downscaling method

Authors: Els Van Uytven, Jan De Niel and Patrick Willems

We would like to thank the reviewer for the constructive feedback and comments. Answers to the comments have been made (italic font) and the manuscript has been changed accordingly (blue font). We would like to note that the manuscript has changed significantly to address the comments of all reviewers and to improve the readability.

Response to reviewer comments (Mohammad Sohrabi)

This study aims at evaluating assumptions of a weather typing (WT) based statistical downscaling method (SDM) for precipitation and river peak flows in Belgium. The results of such studies provide an assessment of end user needs in choosing a right downscaling methods out of many available methods in terms of the cons and pros of each method and the intended use of the results. However, the current study for uncovering the shortcomings of a downscaling method has several serious shortcomings itself as listed below:

(1) Validity of the study

The results of this study showed that the synoptic changes (WT occurrence changes) contributed only 20% of the total change in daily precipitation and the change is mostly (80%) explained by other processes including the thermo-dynamical processes. It obviously indicates that a weather typing based statistical downscaling method shouldn't be used in the region for the downscaling of precipitation which is mainly originated from local moisture. So, what would be the point for evaluating a downscaling method which cannot be used in the region?

REPLY: The weather typing method has been published in 2011 in an international peer reviewed and highly ranked journal (Willems and Vrac, 2011). The paper elaborates on the downscaling methodology and the differences between different weather typing methods (see also comment 4d). The paper, however, does not address the accuracy of the downscaling assumptions. Those assumptions have neither been verified in follow-up papers. On the occasion of the European VALUE COST project, the (European) statistical downscaling community has put additional focus on the evaluation of downscaling methods. In this context, the verification and evaluation of the downscaling assumptions have been conducted. Until now, no follow-up paper has indicated that the studied weather typing method has several shortcomings. These have only recently been uncovered by the verifying and evaluating the assumptions. The abstract, discussions and conclusions have been modified. They now question the applicability of the weather typing method for case studies where the precipitation changes are driven by thermodynamic processes.

Indeed, Figure 9 in the original manuscript points out that the thermo-dynamical processes mainly explain the changes in the average daily precipitation amount. We note that to some extent the thermodynamic processes have been accounted for in the downscaling methodology, more specifically through the application of the Clausius-Clapeyron relation. This relation, however, only applies to the extreme precipitation amounts, not the average daily precipitation amount. As also *suggested by one of the anonymous referees, in the modified manuscript, more focus has been put to the added value of the Clausius-Clapeyron relation.*

(2) Novelty of the study

The statistical downscaling method (SDM) was taken from the literature without any modification: SD-B-7 method from Willems and Vrac (2011). The evaluation of statistical downscaling methods is not also new, although it is mentioned in the second line of the abstract that "Each statistical downscaling method (SDM) has strengths and limitations, but those are rarely evaluated". Nine more comprehensive studies were mentioned in P3L15 of this paper (some of them from EU COST Action VALUE project for the evaluation of statistical downscaling methods) along with several unmentioned. A clear objective was not defined for this study. While in the abstract, only extreme precipitation and river peak flows were stated, the majority of the paper is about daily precipitation and not extremes. As an example, it was claimed (P6L12-15) that compared to the previous study by Brisson et al. (2011) across different stations in Belgium, the current study focuses on the extreme precipitation amounts. However, the presented results are more related to winter precipitation accumulation and percentage of wet days per WT (Figures 2 and 3). Overall, there is not a consistent storyline in the paper. It appears that the paper is a combination of several small studies (leftover results actually) and then fitting a statistical downscaling to them.

REPLY:

- *The introduction has been modified. The modifications, amongst others, involve a different formulation of the objective.*

- *The main objective of this study is to verify and evaluate the general and structural statistical downscaling assumptions in order to develop a statistical downscaling ensemble tailored to the case study and thus end-user needs. Most studies address the general and structural statistical downscaling assumptions independently. This results in studies addressing one or some of general statistical downscaling assumptions (Dixon et al., 2016; Fu et al., 2018; Haberlandt et al., 2015; Hertig et al., 2017; Mendoza et al., 2016; Merkenschlager et al., 2017; Salvi et al., 2016; Tabari et al., 2016) and other studies addressing the structural assumptions by statistical downscaling of surrogate climate model runs (Bürger et al., 2012; Gutmann et al., 2014; Hertig et al., 2018; Maraun et al., 2018; Roberts et al., 2019; Werner and Cannon, 2016; Widmann et al., 2019; Yang et al., 2019) or by statistical downscaling of the projected climate model output (Li et al., 2017; Sørup et al., 2018; Sunyer et al., 2015; Vaittinada Ayar et al., 2016; Wang et al., 2016; Wootten et al., 2017). However, to objectively identify shortcomings of statistical downscaling methods, the verification and evaluation of the general and structural assumptions should be carried out simultaneously. To the authors knowledge, there are yet no papers which simultaneously address the verification of both types of assumptions.*

- *The results describing the relation between the weather types and precipitation accumulation or the number of wet days have been moved to supplementary information.*

(3) Statistical analysis and extreme event definition

Another major issue in the paper is that the results for extremes are based on a limited sample size. The extremes were separated based on 4 seasons and 11 weather types. How extremes are the selected extremes for each season and weather type? What threshold was used for defining extremes? Apparently, the return period of 0.1 year was chosen as the threshold. The question would be whether precipitation and streamflow that occur on average every month is really considered extreme in hydrology. Due to a small sample size after the separation per season and per weather type, even an extreme precipitation of a 10-year return period amounts 0.5 mm/hr (Figure 5).

REPLY:

- *The independent daily and sub-daily precipitation amounts are identified using a peak over threshold (POT) method. More specifically, the threshold is set at 0.1 mm/h and at least 12 hours are considered between successive events. Next the independent daily (or sub-daily precipitation amounts are classified in function of the occurrence (season) and associated weather type.*

- *The threshold in the POT method is not return period based, but precipitation amount based. As indicated in above explanation, the threshold is set at 0.1 mm/h.*

- *The seasonal time scale has been used instead of the monthly time scale as the monthly time scale would introduce more sampling uncertainties. In order to address the sampling uncertainty, the number of WTs could be further reduced. More specifically, 4 wind directions could be considered instead of 8.*

  We have restricted the analysis to the winter season as De Niel et al. (2019) have shown that peak flows in the studied catchment mainly occur during winter season. We remark that the hydrological impact analysis is removed from the manuscript, but focus remains on the winter season. By considering only winter season, we focus on solely stratiform precipitation events. By combining the results for all seasons, stratiform and convective events would have been considered together. This would have been inaccurate as these events have a strongly different nature and manner of modelling. Stratiform and convective events have moreover different precipitation drivers.

(4) Justification of the selected methods

(a) Why was the Lamb weather classification used in this study, while k-means clustering is regarded as one of the best-performing classification schemes over western Europe (e.g., Beck and Philipp, 2010; Casado et al., 2010; Garcia-Valero et al., 2012; Broderick and Fealy, 2015).

REPLY:

Philipp et al. (2016) provide an overview on large scale atmospheric circulation classification systems. In summary, the classification systems are divided into subjective methods, threshold-based methods, methods based on principal component analysis, leader algorithms, hierarchical clustering analysis, optimization algorithms, mixture models and methods based on random processes. K-means methods are optimization algorithms, whereas the Jenkinson-Collison modified Lamb weather types are threshold based methods. The Jenkinson-Collison weather types have the advantage to be easier understood, as indicated in the paper of Brisson et al. (2011) and

Otero et al. (2018). Brisson et al. (2011) also claim that the Jenkinson-Collison Lamb WTs are physically more correct. We would like to point out that these WTs remain presently frequently used (Ästrøm et al., 2016; Manola et al., 2019).

(b) De Niel et al. (2018) identified a minor uncertainty contribution by the hydrological models in the peak flow changes. Among the tested hydrological models in that study, why NAM was selected for the current study?

REPLY: As the uncertainty contribution arising from the hydrological models is small, the reliance of the peak flow discharges does not involve large uncertainties. This is however not the case for the simulation of low flow discharges in river catchments. This has also been observed by Vansteenkiste et al. (2014), who studied the influence of hydrological model structures for the same river catchment. Moreover, the NAM rainfall runoff model is applied in many parts of the world.

We would like to remark that the hydrological impact analysis has been removed from the manuscript.

(c) Why were these three reanalysis datasets selected for this study? Why didn't they use the E-OBS observations (1950-2018) which has similar data coverage to the ERA40 reanalysis dataset (1948-2002)?

REPLY: E-OBS data is a land-only re-analysis dataset. For the WT typing algorithm, mean sea level pressure is also required for locations in the Atlantic Ocean and North Sea.

(d) Several downscaling methods were developed and evaluated by Willems and Vrac (2011). Why was the SD-B-7 method selected for this study?

REPLY: Willems and Vrac (2011) present two types of statistical downscaling methods: precipitation change factor methods and weather typing methods. For each type, a set of methods is presented. For the weather typing methods, more specifically, differences in the methods involve the definitions for the analogue days. In total, 7 weather typing methods have been presented, of which only one method (SD-B-7) is able to produce precipitation amounts outside the range of observations. As an intensification of the precipitation extremes is expected, the extrapolation of the precipitation amounts outside the range of observations is a requirement. Consequently, Willems and Vrac (2011) advice the application of that method for climate change impact analysis.

(5) Weather types

(a) The climate model results for WTs are largely biased. Although the bias was reported in the text regarding the mean scenario of the climate models, the difference goes up to 20% for the anti-cyclonic (A) WT and 30% for west (W) WT: none of the climate models can even reach the frequency of A WT estimated by the reanalysis datasets. Considering these large biases, how reliable would be the downscaling results based on these WTs? The climate models with a coarse resolution are expected to reasonably simulate these large scale patterns, and so what might be the main reason for such a bias in climate model results? I would be interesting to investigate why the GCMs have the largest uncertainty for the W WT as the main large scale driver of winter precipitation over western Europe?

REPLY:

- Climate model ensembles are often designed based on the climate model performance for current climate. However, there is yet no proof that better performing models produce more realistic projections. As stated by Mendlik and Gobiet (2016):

"In the literature, models are often selected based only on their performance in the past, without regarding spread in the climate change signals, with the aim to use only the "best" models. However, correlations between past performance and future climate change signals are known to be very weak, which means that there is no clear indication that the best performing models in the past are most realistic with regard to climate change signal. In addition, the ranking of models with regard to performance in the past is highly dependent on the definition of the performance measure, which leads to a very subjective ranking."

- Indeed, as discussed in the introduction of Phitan et al. (2016), the circulation biases decrease at higher horizontal climate model resolution. In order to focus on the main objective of this paper, the influence of the horizontal resolution on the WTs is not studied in the manuscript. The influence of the resolution on the biases has, however, been added to the discussion on the origin of the biases.

- A brief discussion on the origin of the biases in the W and A WTs has been added to the manuscript. In summary, the North Atlantic storm track has in the climate model simulations a zonal orientation rather than SW-NE tilt (Phitan et al., 2016, Zappa et al., 2014). The zonal orientation results in a pronounced meridional pressure gradient, creating zonal westerly flows which in turn impede the occurrences of anticyclones (Stryhal and Huth, 2019). Biases in the blocking frequency might also be explained by the climate model resolution (Anstey et al., 2013; Scaife et al., 2011, Woollings et al., 2018).

(b) What is the driver for the undefined weather type or the atmospheric state characterized by a weak flow? A sensitivity test of unclassified days on grid sizes and resolutions by Demuzere et al. (2009) showed that the number of these days decreases with grid resolution. Was that the case for this study as well? And generally, how will the different resolutions of reanalysis data from 2◦ to 5◦ explain the discrepancy between reanalysis-based WTs?

REPLY:

- We note that the same WT classification system is applied to all seasons. The relative occurrence frequencies are, however, seasonally dependent. A difference between winter and summer would be the occurrence frequency of the undefined WTs. More specifically, the occurrence frequency of the undefined weather types in winter season is negligible, while for summer season the undefined weather types represent at least 10% of the summer days. This is in agreement with the results of Otero et al. (2018).

- No sensitivity analysis on the grid resolution has been conducted as this has already performed by Demuzere et al. (2009).

- The application of the 16 point grid with a 10° resolution in the zonal direction and a 5° resolution in meridional direction allows the comparison with previous studies. We note a similar motivation by Demuzere et al. (2008) and Otero et al. (2018).

(c) A separate set of WTs was produced for each season in the current study. What will be the influence of the seasonal cycle on the classification produced as the MSLP fields are clustered?

REPLY:

We note that the same WT classification system is applied to all seasons. The relative occurrence frequencies are, however, seasonally dependent. In order to keep the manuscript condense, only winter season is studied. We refer the reviewer also to our reply on comment 5(b).

(6) Statistical downscaling by analogues

The CMIP5 GCMs provide data at a daily time scale. Were daily precipitation data from the GCM scenario period corresponded to observed sub-daily precipitation? If so, how are the results influenced by the difference in the time scale? Were the climate change signals assumed to be time scale dependent?

REPLY:

- *In the case that the observed precipitation time series has a sub-daily time step, the sub-daily precipitation amounts are aggregated to daily precipitation amounts. Next, for each season and WT, the exceedance probabilities for the daily precipitation amounts of wet days are calculated . In a similar way, for each of the projected wet days, the exceedance probabilities of the projected daily precipitation amounts are calculated. Thereafter, an analogue wet day is defined as an observed wet day occurring in the same season, having the same weather type and best approximating the exceedance probability of the projected day precipitation amount. To produce the downscaled time series, the daily precipitation amount for the observed analogue day is resampled. If the observed precipitation time series had a sub-daily time step, the sub-daily precipitation amounts for the analogue day are re-sampled. In other words, analogues are defined by comparing the exceedance probabilities for the observed <u>daily</u> precipitation amounts with the exceedance probabilities for the projected <u>daily</u> precipitation amounts.*

- *Since no sub-daily temperature time series are available, the CC relation is investigated at daily time scale. This is also the case when sub-daily precipitation amounts are available. The scaling rates identified at daily time scale are thereafter applied to the sub-daily precipitation amounts, assuming the changes at daily time scale are applicable at sub-daily time scale.*

(7) Scaling by the Clausius-Clapeyron relation

(a) Clausius-Clapeyron relation assumes that extreme precipitation amounts are controlled by local moisture availability. How local is moisture availability? The developed extreme precipitation-temperature scaling relations for central Belgium were used for river peak flow simulations in a catchment in the northeast of Belgium. How representative would be the scaling relations developed in central Belgium for northeast Belgium considering the local moisture availability assumption?

REPLY:

In a study for the Flemish Environment agency, De Niel and Willems (2016) investigated the spatial and temporal variations in precipitation time series for 43 rain gauges in Flanders. Their results indicated significant differences between west (coastal area) and east (Antwerp, Flemish Brabant and Limburg). A study investigating the spatial and temporal variation in the temperature time series in Flanders has yet to be conducted. As Uccle is situated in central Belgium and is located approximately 100 km from the Grote Nete catchment, it is expected that application of the hydrometeorological time series of the RMI station in Uccle for case studies in the Grote Nete catchment involves only small uncertainties. Moreover, we would like to point out that the application of the short observed time series, as is the case for the stations in the Grote Nete catchment, also involves uncertainties.

We would like to remark that the hydrological impact analysis has been removed from the manuscript. Hence, this comment is not applicable to the modified manuscript.

(b) Dry-bulb temperature was used here for developing scaling relations, whereas several recent studies (e.g., Wasko et al., 2018) recommended to use dew point temperature than dry-bulb temperature for Clausius-Clapeyron relation, as it is a better measure of precipitation changes because of increases in the moisture holding capacity of the atmosphere (Lenderink et al., 2011).

REPLY:

Indeed, several studies have pointed out that dew point temperature is a better predictor than the average daily temperature (Van de Vyver et al., 2019; Wasko et al., 2018). However, compared to average daily temperature, time series for the dew point temperature are not readily available for hydrological impact modellers.

The consideration of average daily temperatures rather than dew point temperatures has been added as a potential shortcoming of this study.

(8) Evaluation of greenhouse gas scenario assumption

To evaluate the greenhouse gas scenario assumption of the SDM, changes in the WT occurrences and average daily temperature as a function of the four RCPs were analyzed. However, this assumption might be tested for precipitation which is statistically downscaled in this study. Besides, the increase of the change in air temperature with greenhouse gas emissions is trivial. What is the relation of changes in warm extremes (with return periods ranging between 0.1 and 10 years) (Figures 7 and 8) to the downscaling of precipitation performed in this study? Though irrelevant, for warm extremes the changes in maximum temperature should be analyzed instead of mean daily temperature. P7L9-10: “To check whether the predictor simulation results are adequate and accurate, a comparison is made between the climate model simulated and observed daily average temperature statistics”. Is predictor warm extremes? It seems that the full range of temperature was used for the scaling (Figure 6).

REPLY:

As suggested, the sensitivity of the predictand to the greenhouse gas scenarios and the increase in greenhouse gas scenarios is verified for the predictand. Schoof (2013) points out that some predictor variables do not respond to the greenhouse gas scenarios, while others do. This means that a predictand response is achieved by a smart choice of predictors. We note that the definition of the greenhouse gas scenario assumption has therefore been changed. The response of the WTs to the different greenhouse gas scenarios has been moved to supplementary information, but is used as background information for the discussion of the results.

(9) Evaluation of the stationarity assumption

For the evaluation of the stationarity assumption, the extreme precipitation per weather type was compared for different sub-periods of 10 years length between 1901 and 1991. Use of a 10-year sub-period for this purpose is questionable as it is far smaller than a natural climate cycle, and hence the results are greatly influenced by natural climate variability. Isn't trend analysis a more robust approach for testing the stationarity? How large is the uncertainty in the presented results? In my view, a random selection of a dry weather type or and performing the same analysis or doing the would reveal the reliability of this results? Would performing the decadal analysis of extreme precipitation without considering the WTs lead to similar results? This is because based on the results of this study, weather types shouldn't be related to precipitation formation in the region (see comment 1).

REPLY:

- *We agree that 10 years of data is rather short with respect to the natural cycle. In that context, the results are indeed influenced by the climate variability.* We note that the methodology for the verification of the stationarity assumption has been altered in the modified manuscript. In the modified manuscript, the surrogate climate model runs are 20years long.

(10) Evaluation of the method specific assumptions

The evaluation of the SDM method specific assumptions was only performed for winter as the peak flows in the selected catchment mainly occur in the winter (P5L12-14). As mentioned in P4L30, "Application of this scaling rate to precipitation intensities is valid assuming that extreme precipitation amounts are controlled by local moisture availability and are not influenced by large scale atmospheric circulation patterns." However, the influence of the large scale circulations on winter precipitation in western Europe is well documented in the literature.

- *The influence of large scale circulation patterns on the historical precipitation amounts is indeed well documented in literature (Tabari and Willems, 2018; Willems, 2013). The discussion has been extended and references have been added to the discussion of the stationarity assumption.*

(11) Interpretation and discussion of the results

(a) The results were interpreted and discussed in a way that the authors expected the results be. For example, in P12L20-25, the authors attribute the difference between their results and those of Otero et al. (2018) for changes in the anti-cyclonic WT to considered climate model ensemble, the location of the 16 points grid for the WT classification system and, the reference and scenario periods. I am wondering why these differences between the two studies (this study and Otero et al., 2018) are only important for the changes in the anti-cyclonic WT and not for the changes in cyclonic, west and southwest WTs! Another example is in P15L28-31, where the authors mention only the results for RCP4.5 and RCP8.5 and not all RCPs to show that changes in the WT occurrences and precipitation are magnified under increasing greenhouse gas scenarios. Looking at the results, changes in the occurrence of W WTs are far smaller for RCP8.5 compared to RCP6.0 (P12L5). Also, there is not a clear pattern for the changes in cyclonic (C) WTs where changes are equal to 5% for RCP 2.6, - 3% for RCP 4.5, -6% for RCP 6.0 and -5% for RCP 8.5. It was speculated in P12L8-10 that these discontinuities in the uni-directionality of the changes may be explained by the smaller ensemble size for RCP 6.0 compared to the other RCPs, and/or by the different RCP sub-ensemble compositions. This issue can be easily checked by selecting the same GCMs for different RCPs. Why isn't this an issue for temperature changes (P12L28-30).

REPLY:

- Under global warming, climate models project a poleward shift of the Northern Hemisphere jet-streams and storm tracks, resulting increase occurrence of zonal flows and less blocking occurrences (Barnes and Screen, 2015; Santos et al., 2016; Stryhal and Huth, 2019; Woollings et al., 2018). These projections correspond with an increased occurrence of W and SW WTs and a decreased occurrence of A WTs. This physical background information has been added to the discussion of the response of the predictors and the predictand to the greenhouse gas scenarios.

- Schoof (2013) points out that not all variables respond to the greenhouse gas scenarios. While the absolute temperature values increase, the absolute values of the mean sea level pressure do not change. For mean sea level pressure, changes occur in the spatial patterns.

- The monotonicity of the changes for increasing greenhouse gas scenarios is not guaranteed when the ensemble compositions and sizes differ. Moreover, in some cases, for the same ensemble composition, the monotonicity of the changes for increasing greenhouse gas scenarios might be masked by the climate model uncertainties and the stochastic uncertainty, this is the uncertainty related to the variability of the climate system (Van Uytven and Willems, 2018).

(b) In several places in the text, internal variability was argued as the reason for unexplained behaviors in the results: for example, internal variability as the reason for the large bias of climate models for WT simulations and also internal variability among the climate models as the reason for discontinuities in the uni-directionality of the changes with greenhouse gas scenarios. These statements might be supported by the results or the literature.

- References on the influence of the influence of the internal variability and thus the choice of the reference period have been added to the manuscript.

- We refer the reviewer also to our reply on comment 11a

(c) The results for the changes in the frequency of WTs showed that the frequency of the wet WTs will increase under climate change and the frequency of the dry WTs will decrease. It is worth discussing what physical explanations are behind the increasing frequency of westerlies and the decreasing frequencies of easterlies under climate change. How might global warming decrease the frequencies of cyclonic and anticyclonic WTs? The response to these important questions would be beneficial for improving weather typing based SDMs.

REPLY:

Under global warming, climate models project a poleward shift of the Northern Hemisphere jet-streams and storm tracks, resulting increase occurrence of zonal flows and less blocking occurrences (Barnes and Screen, 2015; Santos et al., 2016; Stryhal and Huth, 2019; Woollings et al., 2018). Woolings et al. (2018) also list references reporting an eastward shift of the blocking activity. These physical explanations have been added to the discussion of the response of the predictand to the greenhouse gas scenarios.

(d) "Stationary is dead" is now a fact for hydrologists. What would be the contribution of this part of analysis to the existing knowledge? Rather than discussing the natural climate anomalies for this single location in western Europe in section 4.5, it would be more useful to discuss the drivers for such anomalies, as several global studies regarding these historical natural anomalies of the climate system have already been published. P13L21-22: "the 10 minutes precipitation amounts with a 1 year return period measures 6mm/h for the negative anomaly, whereas it is 14mm/h for the positive anomaly". What might be the reason for the positive and negative anomalies in W WT and the related extreme precipitation? P13L20 & P13L25: "The difference between the positive and negative precipitation anomaly is especially visible for the W WTs and this for all aggregation levels between 10 minutes to 1 day. For the C WTs, no differences appear between the amounts for the positive and negative anomaly". Why is W WT based extreme precipitation timescale-dependent, but not the C WT-based extreme precipitation?

REPLY: A discussion on the drivers of the precipitation anomalies has been added to the results of the stationarity assumption. We note that the verification of the stationarity assumption has been conducted changed in the modified manuscript.

(12) Reanalysis data uncertainty

The uncertainty related to the choice of reanalysis data for WTs was considered. Given that the reliability of reanalysis products sharply decreases back in time (Ferguson and Villarini, 2013; Krueger et al., 2013) due to assimilating sparse observations and starting from a more uncertain initial state (Delaygue et al., 2019), a larger uncertainty of reanalysis data is expected for earlier years of the study period. Was the uncertainty calculated for the entire analysis period in this study? Does the uncertainty decrease as time progresses from far past to near past?

REPLY:

The re-analysis datasets provide data for different ranges of time. However, the comparison between the different re-analysis datasets is consistently performed using the same range of data. The manuscript has been modified to better indicate the prior.

The time evolution of the assimilation uncertainties in the re-analysis datasets has not been studied. We agree that this can be of interest, but such investigation is not performed in the considered study for sake of brevity. We also think it would divert the scope from the current main focus.

Specific comments

P5L19: The evapotranspiration data are not daily measured data for 100 years, are they? Is it reference evapotranspiration or potential evapotranspiration? If the former is the case, with grass or alfalfa as the reference crop?

REPLY: We remark that the hydrological climate change impact analysis is removed from the manuscript for sake of brevity.

P7L15: 33 unique control runs were used in this study. I think the number of independent model runs should be lower than that!? How was the dependency between climate models investigated?

REPLY: We agree that the number of independent model runs is smaller than the actual number of included model runs in the ensemble. The dependency between the climate model runs has not been investigated. As the word "unique" is confusing, this word has been removed from the text.

P7L16-17: The authors found that the choice of the reference period (control period) influences the evaluation of the perfect prognosis assumption. It would be interesting to present the results of the sensitivity analysis to the control period in the Supplementary Information.

REPLY: While we do agree that this study can be of interest, it is not performed in the considered manuscript for sake of brevity as well as since it would divert the scope from the current main focus. References on this topic have been added to the main manuscript.

Figure 12: Does grey area show the 5th-95th percentile range of climate change uncertainty?

REPLY: The 5^{th}-95^{th} percentile range of Figure 12 in the original manuscript represents the uncertainties associated with the surrogate climate model runs. The surrogate climate model runs are subsets of the observed time series. Hence, the 5^{th}-95^{th} percentile range represents the stochastic uncertainty, i.e. uncertainty related to the internal variability of the climate system.

Table 2: Were the same GCMs used for all climate variables studied?

REPLY: The same GCMs were used for all climate variables studied.

Used references not present in the reference list of the paper

REPLY: The reference list has been updated and completed.

References:

Anstey, J. A., Davini, P., Gray, L. J., Woollings, T. J., Butchart, N., Cagnazzo, C., Christiansen, B., Hardiman, S. C., Osprey, S. M., and Yang, S.: Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution, Journal of Geophysical Research Atmospheres, 118, 3956–3971, https://doi.org/10.1002/jgrd.50231, 2013.

Åström, H. L., Sunyer, M., Madsen, H., Rosbjerg, D., and Arnbjerg-Nielsen, K.: Explanatory analysis of the relationship between atmospheric circulation and occurrence of flood-generating events in a coastal city, Hydrological Processes, 30, 2773–2788, https://doi.org/10.1002/hyp.10767, 2016.

Barnes, E. A. and Screen, J. A.: The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?, Wiley Interdisciplinary Reviews: Climate Change, 6, 277–286, https://doi.org/10.1002/wcc.337, 2015.

Bürger, G., Murdock, T. Q., Werner, A. T., Sobie, S. R., and Cannon, A. J.: Downscaling Extremes—An Intercomparison of Multiple Statistical Methods for Present Climate, Journal of Climate, 25, 4366–4388, https://doi.org/10.1175/JCLI-D-11-00408.1, 2012.

Brisson, E., Demuzere, M., Kwakernaak, B., and Van Lipzig, N. P.: Relations between atmospheric circulation and precipitation in Belgium, Meteorology and Atmospheric Physics, 111, 27–39, https://doi.org/10.1007/s00703-010-0103-y, 2011.

Demuzere, M., Werner, M., van Lipzig, N. P. M., and Roeckner, E.: An analysis of present and future ECHAM5 pressure fields using a classification of circulation patterns, International Journal of Climatology, 29, 1796–1810, https://doi.org/10.1002/joc.1821, 2009.

De Niel, J., Van Uytven, E., and Willems, P.: On the correlation between precipitation and potential evapotranspiration climate change signals for hydrological impact analyses. Hydrological Sciences Journal, 64 (4), 420-433. https://doi.org/10.1080/02626667.2019.1587615. 2019.

De Niel, J., Van Uytven, E., and Willems, P.: Uncertainty analysis of climate change impact on river flow extremes based on a large multi model ensemble, under review, 2018.

De Niel, J. and Willems, P.: Opzetten van een validatie- en correctieprocedure voor meteorologische data. Leuven: KU Leuven - Afdeling Hydraulica. 2016.

Dixon, K. W., Lanzante, J. R., Nath, M. J., Hayhoe, K., Stoner, A., Radhakrishnan, A., Balaji, V., and Gaitán, C. F.: Evaluating the stationarity assumption in statistically downscaled climate projections: is past performance an indicator of future results?, Climatic Change, 135, 395–408, https://doi.org/10.1007/s10584-016-1598-0, 2016.

Fu, G., Charles, S. P., Chiew, F. H., Ekström, M., and Potter, N. J.: Uncertainties of statistical downscaling from predictor selection: Equifinality and transferability, Atmospheric Research, 203, 130–140, https://doi.org/10.1016/j.atmosres.2017.12.008, 2018.

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., and Rasmussen, R. M.: An intercomparison of statistical downscaling methods used for water resource assessments in the United States, Water Resources Research, 50, 7167–7186,https://doi.org/10.1002/2014WR015559, 2014.

Haberlandt, U., Belli, A., and Bárdossy, A.: Statistical downscaling of precipitation using a stochastic rainfall model conditioned on circulation patterns - an evaluation of assumptions, International Journal of Climatology, 35, 417–432, https://doi.org/10.1002/joc.3989*, 2015.*

Hertig, E., Merkenschlager, C., and Jacobeit, J.: Change points in predictors–predictand relationships within the scope of statistical downscaling, International Journal of Climatology, 37, 1619–1633, https://doi.org/10.1002/joc.4801, 2017.

Hertig, E., Maraun, D., Bartholy, J., Pongracz, R., Vrac, M., Mares, I., Gutiérrez, J. M., Wibig, J., Casanueva, A., and Soares, P. M. M.: Comparison of statistical downscaling methods with respect to extreme events over Europe: Validation results from the perfect predictor experiment of the COST Action VALUE, International Journal of Climatology, pp. 1–22, https://doi.org/10.1002/joc.5469, 2018.

Li, J., Johnson, F., Evans, J., and Sharma, A.: A comparison of methods to estimate future sub-daily design rainfall, Advances in Water Resources, 110, 215–227, https://doi.org/10.1016/j.advwatres.2017.10.020, 2017.

Manola, I., Steeneveld, G.-J., Uijlenhoet, R. and Holtslag, A. A. M.: Analysis of urban rainfall from hourly to seasonal scales using high-resolution radar observations in the Netherlands, International Journal of Climatology, 1-19, https://doi.org/10.1002/joc.6241*, 2019.*

Manola, I., van den Hurk, B., De Moel, H., and Aerts, J. C. J. H.: Future extreme precipitation intensities based on a historic event, Hydrology and Earth System Sciences, 22, 3777–3788, https://doi.org/10.5194/hess-22-3777-2018, 2018.

Maraun, D., Widmann, M., and Gutiérrez, J. M.: Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment, International Journal of Climatology, pp. 1–12, https://doi.org/10.1002/joc.5877, 2018.

Mendlik, T., and Gobiet, A.: Selecting climate simulations for impact studies based on multivariate patterns of climate change, Climatic Change, 135, 3-4, 381–393, https://doi.org/10.1007/s10584-015-1582-0, 2016.

Mendoza, P. A., Mizukami, N., Ikeda, K., Clark, M. P., Gutmann, E. D., Arnold, J. R., Brekke, L. D., and Rajagopalan, B.: Effects of different regional climate model resolution and forcing scales on projected hydrologic changes, Journal of Hydrology, 541, 1003–1019, https://doi.org/10.1016/j.jhydrol.2016.08.010, 2016.

Merkenschlager, C., Hertig, E., and Jacobeit, J.: Non-stationarities in the relationships of heavy precipitation events in the Mediterranean area and the large-scale circulation in the second half of the 20th century, Global and Planetary Change, 151, 108 – 121, https://doi.org/10.1016/j.gloplacha.2016.10.009, 2017.

Ntegeka, V. and Willems, P.: Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium, Water Resources Research, 44, 1–15, https://doi.org/10.1029/2007WR006471, 2008.

Otero, N., Sillmann, J., and Butler, T.: Assessment of an extended version of the Jenkinson–Collison classification on CMIP5 models over Europe, Climate Dynamics, 50, 1559–1579, https://doi.org/10.1007/s00382-017-3705-y, 2018

Philipp, A., Beck, C., Huth, R., and Jacobeit, J.: Development and comparison of circulation type classifications using the COST 733 dataset and software, International Journal of Climatology, 36, 2673–2691, https://doi.org/10.1002/joc.3920, 2016.

Pithan, F., Shepherd, T. G., Zappa, G., and Sandu, I.: Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag, Geophysical Research Letters, 43, 7231–7240, https://doi.org/10.1002/2016GL069551, 2016.

Roberts, D. R., Wood, W. H., and Marshall, S. J.: Assessments of downscaled climate data with a high-resolution weather station network, International Journal of Climatology, 39, 3091–3103, https://doi.org/10.1002/joc.6005, 2019.

Salvi, K., Ghosh, S., and Ganguly, A. R.: Credibility of statistical downscaling under nonstationary climate, Climate Dynamics, 46, 1991–2023, https://doi.org/10.1007/s00382-015-2688-9, 2016.

Santos, J. A., Belo-Pereira, M., Fraga, H., and Pinto, J. G.: Understanding climate change projections for precipitation over western Europe with a weather typing approach, Journal of Geophysical Research: Atmospheres, 121, 1170–1189, https://doi.org/10.1002/2015JD024399*, 2016.*

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M., and Williams, K.: Improved Atlantic winter blocking in a climate model, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL049573, 2011.

Schoof, J. T.: Statistical Downscaling in Climatology, Geography Compass, 7, 249–265, https://doi.org/10.1111/gec3.12036, 2013.

Sørup, H. J. D., Davidsen, S., Löwe, R., Thorndahl, S. L., Borup, M., and Arnbjerg-Nielsen, K.: Evaluating catchment response to artificial rainfall from four weather generators for present and future climate, Water Science and Technology, 77, 2578–2588, https://doi.org/10.2166/wst.2018.217, 2018.

Stryhal, J. and Huth, R.: Trends in winter circulation over the British Isles and central Europe in twenty-first century projections by 25 CMIP5 GCMs, Climate Dynamics, 52, 1063–1075, https://doi.org/10.1007/s00382-018-4178-3, 2019.

Stryhal, J. and Huth, R.: Classifications of winter Euro-Atlantic circulation patterns: An intercomparison of five atmospheric reanalyses, Journal of Climate, 30, 7847–7861, https://doi.org/10.1175/JCLI-D-17-0059.1, 2017.

Sunyer, M. A., Hundecha, Y., Lawrence, D., Madsen, H., Willems, P., Martinkova, M., Vormoor, K., Bürger, G., Hanel, M., Kriaučiūnienė, J., Loukas, A., Osuch, M., and Yücel, I.: Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe, Hydrology and Earth System Sciences, 19, 1827–1847, https://doi.org/10.5194/hess-19-1827-2015, 2015.

Tabari, H. and Willems, P.: Lagged influence of Atlantic and Pacific climate patterns on European extreme precipitation, Scientific Reports, https://doi.org/10.1038/s41598-018-24069-9, 2018.

Tabari, H., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Saeed, S., Brisson, E., Van Lipzig, N., and Willems, P.: Local impact analysis of climate change on precipitation extremes: are high-resolution climate models needed for realistic simulations?, Hydrology and Earth System Sciences, 20, 3843–3857, https://doi.org/10.5194/hess-20-3843-2016, 2016.

Vaittinada Ayar, P., Vrac, M., Bastin, S., Carreau, J., Déqué, M., and Gallardo, C.: Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: present climate evaluations, Climate Dynamics, 46, 1301–1329, https://doi.org/10.1007/s00382-015-2647-5, 2016.

Van de Vyver, H., Van Schaeybroeck, B., De Troch, R., Hamdi, R., and Termonia, P.: Modeling the scaling of short-duration precipitation extremes with temperature, Earth and Space Science, https://doi.org/10.1029/2019EA000665, 2019.

Van Uytven, E. and Willems, P.: Greenhouse gas scenario sensitivity and uncertainties in precipitation projections for central Belgium, Journal of Hydrology, 558, 9–19, https://doi.org/10.1016/j.jhydrol.2018.01.018, 2018.

Vansteenkiste, T., Tavakoli, M., Ntegeka, V., De Smedt, F., Batelaan, O., Pereira, F., and Willems, P.: Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections, J. Hydrol., 519, 743–755, https://doi.org/10.1016/j.jhydrol.2014.07.062, 2014.

Wang, L., Ranasinghe, R., Maskey, S., van Gelder, P. H. A. J. M., and Vrijling, J. K.: Comparison of empirical statistical methods for downscaling daily climate projections from CMIP5 GCMs: A case study of the Huai River Basin, China, International Journal of Climatology, 36, 145–164, https://doi.org/10.1002/joc.4334, 2016.

Wasko, C., Lu, W. T., and Mehrotra, R.: Relationship of extreme precipitation, dry-bulb temperature, and dew point temperature across Australia, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aad135, 2018.

Werner, A. T. and Cannon, A. J.: Hydrologic extremes - an intercomparison of multiple gridded statistical downscaling methods, Hydrology and Earth System Sciences, 20, 1483–1508, https://doi.org/10.5194/hess-20-1483-2016, 2016.

Widmann, M., Bedia, J., Gutiérrez, J. M., Bosshard, T., Hertig, E., Maraun, D., Casado, M. J., Ramos, P., Cardoso, R. M., Soares, P. M. M.,Ribalaygua, J., Pagé, C., Fischer, A. M., Herrera, S., and Huth, R.: Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment, International Journal of Climatology, pp. 1–27, https://doi.org/10.1002/joc.6024, 2019.

Willems, P.: Multidecadal oscillatory behaviour of rainfall extremes in Europe, Climatic Change, 120, 931–944, https://doi.org/10.1007/s10584-013-0837-x, 2013.

Willems, P. and Vrac, M.: Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change, Journal of Hydrology, 402, 193–205, https://doi.org/10.1016/j.jhydrol.2011.02.030, 2011.

Woollings, T., Barriopedro, D., Methven, J., Son, S.-W., Martius, O., Harvey, B., Sillmann, J., Lupo, A. R., and Seneviratne, S.: Blocking and its Response to Climate Change, Current Climate Change Reports, 4, 287–300, https://doi.org/10.1007/s40641-018-0108-z, 2018.

Wootten, A., Terando, A., Reich, B. J., Boyles, R. P., and Semazzi, F.: Characterizing sources of uncertainty from global climate models and downscaling techniques, Journal of Applied Meteorology and Climatology, 56, 3245–3262, https://doi.org/10.1175/JAMC-D-17-0087.1, 2017.

Yang, Y., Tang, J., Xiong, Z., Wang, S., and Yuan, J.: An intercomparison of multiple statistical downscaling methods for daily precipitation and temperature over China: present climate evaluations, Climate Dynamics, https://doi.org/10.1007/s00382-019-04809-x, 2019.

Zappa, G., Masato, G., Shaffrey, L., Woollings, T., and Hodges, K.: Linking Northern Hemisphere blocking and storm track biases in the CMIP5 climate models, Geophysical Research Letters, 41, 135–139, https://doi.org/10.1002/2013GL058480, 2014.

Uncovering shortcomings of weather typing method

Els Van Uytven[1], Jan De Niel[1], and Patrick Willems[1,2]

[1]KU Leuven, Department of Civil Engineering - Hydraulics Section, Kasteelpark 40 box 2448, 3001 Leuven, Belgium
[2]Vrije Universiteit Brussel, Department of Hydrology and Hydraulic Engineering, Building T Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium

Correspondence: Els Van Uytven (els.vanuytven@kuleuven.be)

Abstract. In recent years many methods for statistical downscaling of the precipitation climate model outputs have been developed. Statistical downscaling is performed under general and method specific (structural) assumptions, but those are rarely evaluated simultaneously. This paper illustrates the verification and evaluation of the downscaling assumptions for a weather typing method. Using observations and the output of a global climate model ensemble, the skill of the method is evaluated for precipitation downscaling in central Belgium during winter season (December to February). Shortcomings of the studied method have been uncovered and are identified as biases and a time-variant predictors-predictand relation. More specifically, the predictors-predictand relation is found informative for the historical observations, but becomes inaccurate for the projected climate model output. The latter inaccuracy is explained by the increased importance of the thermodynamical processes in the precipitation changes. The results therefore question the applicability of the weather typing method for the case study location. Besides shortcomings, the results also demonstrate the added value of the Clausius-Clapeyron relation for precipitation amount scaling. The verification and evaluation of the downscaling assumptions are a tool to design a statistical downscaling ensemble tailored to end-user needs.

1 Introduction

For a 1.5°C temperature rise, the worldwide direct flood damage is estimated to increase by 160%-240% (Dottori et al., 2018). To minimize that potential impact, our society opts for two complementary strategies: climate mitigation and climate adaptation (Stocker et al., 2013). Consequently, vulnerability, impact and adaptation studies find ground in our society (Alfieri et al., 2016; Åström et al., 2016; Brekke et al., 2009; Termonia et al., 2018; Vansteenkiste et al., 2014; Vermuyten et al., 2018; Willems, 2013). These studies require projected hydro-meteorological time series, using the output of global climate models as primary information. However, the direct application of this output for impact modelling is hindered by climate model biases (Kotlarski et al., 2014; Tabari et al., 2016) and by the mismatch in temporal and spatial resolutions between the climate model output and the time series required for impact modelling (Cristiano et al., 2018; Salvadore et al., 2015). Therefore, statistical downscaling or dynamical downscaling is applied. The statistical downscaling approach bridges the resolution gap through statistical relations between the predictors and predictand, whereas in the dynamical downscaling approach regional and limited area climate models (RCMs and LAMs, respectively) are developed. Despite the refined resolution of RCMs and LAMs, their climate model output remains biased and requires bias correction (Ehret et al., 2012; Maraun, 2016; Teutschbein and Seibert, 2012). Both downscaling approaches have strengths and shortcomings, arising from their underlying assumptions (Casanueva et al., 2016; Flaounas et al., 2013; Le Roux et al., 2018; Maraun et al., 2010; Vaittinada Ayar et al., 2016).

The statistical downscaling approach builds on four **general assumptions** (Benestad et al., 2008; Maraun et al., 2010; Maraun and Widmann, 2018; Schoof, 2013):

- **the relation between the predictors and predictand is relevant** (referred to as **the informative assumption**). This is of importance in the development of new statistical downscaling methods (SDMs), which requires the selection of predictors (Fu et al., 2018; Sachindra et al., 2018; Wilby and Wigley, 2000; Yang et al., 2017). The selected predictors should relate to the physical processes explaining the predictand changes. Precipitation, more specially, responds to large scale atmospheric circulation and thermodynamical laws (Emori and Brown, 2005; Kröner et al., 2017; Santos et al., 2016) and, hence, sea level pressure, geopotential height, relative humidity and/or (dew point) temperature are common predictors (Maraun and Widmann, 2018).
- **the predictors are adequately and accurately simulated by the climate model runs** (referred to as **the perfect prognosis assumption**. The evaluation of this assumption is foremost performed under the name *bias analysis*. The bias in the predictors depends, among others, on the model resolution, the parametrisation schemes, the internal variability and the choice of the reference period (Anstey et al., 2013; Arakawa, 2004; Davini et al., 2017; Deser et al., 2012; Fadhel et al., 2017; Hartung et al., 2017; Prein et al., 2015; Rybka and Tost, 2014; Tabari et al., 2016; Vanden Broucke et al., 2018; Watterson et al., 2014).
- **the relation between the predictors and the predictand remains time-invariant** (referred to as the **the stationarity assumption**). This means that the relation between the predictors and the predictand, which has been established using historical observations, remains applicable under climatic changes. Of all assumptions, this assumption is the most difficult one to validate as no future observations are yet available (Dixon et al., 2016; Lanzante et al., 2018; Maraun and Widmann, 2018; Salvi et al., 2016; Wang et al., 2018).
- **the predictand is sensitive to the greenhouse gas scenarios**. Schoof (2013) has pointed out that one predictor variable could strongly respond to the greenhouse gas scenarios, while another variable would not. This observation is, for instance, applicable to changes in temperature and mean sea level pressure, respectively. Moreover, due to the internal variability of the climate system and the climate model related uncertainties, the response of the predictor to the greenhouse gas scenarios is often masked (Van Uytven and Willems, 2018). Hence, the response of the predictand to the greenhouse gas scenarios is governed by a smart choice of predictors.

Alongside the general statistical downscaling assumptions, each SDM has **method specific** or **structural assumptions**. They are encapsulated in the downscaling methodology, create the method strengths and limitations, and are responsible for the statistical downscaling uncertainty contribution. An overview of commonly applied SDMs for precipitation downscaling and their strengths and limitations is provided by Hewitson et al. (2014), Maraun et al. (2010), Maraun and Widmann (2018) and Sunyer et al. (2015).

The main objective of this paper is to simultaneously verify and evaluate the general and structural statistical downscaling assumptions. Most studies address the general and structural statistical downscaling assumptions independently. Hence, there are studies addressing one or some of general statistical downscaling assumptions (Dixon et al., 2016; Fu et al., 2018; Haberlandt et al., 2015; Hertig et al., 2017; Mendoza et al., 2016; Merkenschlager et al., 2017; Salvi et al., 2016; Tabari et al., 2016) and other studies addressing the structural assumptions by statistical downscaling of surrogate climate model runs (Bürger et al., 2012; Gutmann et al., 2014; Hertig et al., 2018; Maraun et al., 2018; Roberts et al., 2019; Werner and Cannon, 2016; Widmann et al., 2019; Yang et al., 2019) or by statistical downscaling of the projected climate model output (Li et al., 2017; Sørup et al., 2018; Sunyer et al., 2015; Vaittinada Ayar et al., 2016; Wang et al., 2016; Wootten et al., 2017). To objectively identify shortcomings of statistical downscaling methods, the verification and evaluation of the general and structural assumptions should however be performed simultaneously. To the authors knowledge, there are yet no papers which simultaneously address the verification of both types of assumptions.

In this paper, the verification and evaluation of the general and structural assumptions is illustrated for a weather typing (WT) SDM for the purpose of climate change impact modelling on precipitation in Belgium during winter (December to February). The studied WT SDM is referred to as *SD-B-7* by Willems and Vrac (2011). Downscaling is performed in three steps. In the first step, weather types are identified based on the mean sea level pressure patterns. In the second step, the relation between the predictors (weather types) and predictand (point precipitation) is established using analogues. In the last step, the precipitation amounts are scaled following the Clausius-Clapeyron relation. Overall strengths would emerge from the physical background of the SDM (Shepherd et al., 2018).

This paper is organised as follows. Section 2 introduces the studied SDM and the hydro-meteorological data. Section 3 outlines the verification of the downscaling assumptions and corresponding results and discussions are included in Section 4. Section 5 summarizes the main findings and makes suggestions for future research.

2 Statistical downscaling methods, case study and data

2.1 The weather typing method

The considered WT method is the method referred to as *SD-B-7* by Willems and Vrac (2011). This method has been selected over the other WT methods as it anticipates the intensification of extreme precipitation amounts. The method downscales the daily gridded climate model output to point time series with a time step equal to the observed time series using a three step approach. In the first step, the Jenkinson-Collison automated Lamb WT classification system is applied and the WTs are identified. In the second step, downscaled precipitation time series are produced using WT analogues. In the last step, the precipitation amounts are scaled using the Clausius-Clapeyron (CC) relation.

Step 1: Jenkinson-Collison automated Lamb WT classification scheme

As shown in Figure 1, a 16-point grid is centered around the study area. Assuming p_i is the mean sea level pressure (MSLP) in point i of the 16 point grid and ψ the latitude of the study area, then the southerly flow (SF), westerly flow (WF), total flow (F), southerly shear vorticity (ZS), westerly shear vorticity (ZW) and total shear vorticity (Z) are calculated as follows (Jenkinson and Collison, 1977; Jones et al., 1993; Philipp et al., 2016):

$$
\begin{aligned}
&SF = \frac{1}{\cos(\psi)}(0.25(p_5+2p_9+p_{13})) - \frac{1}{\cos(\psi)}(0.25(p_4+2p_8+p_{12})),\\
&WF = 0.5(p_{12}+p_{13}) - 0.5(p_4+p_5),\\
&F = (SF^2+WF^2)^{1/2},\\
&ZS = \frac{\sin(\psi)}{\sin(\psi-5)}(0.25(p_6+2p_{10}+p_{14})) - \frac{\sin(\psi)}{\sin(\psi-5)}(0.25(p_5+2p_9+p_{13})) - \frac{\sin(\psi)}{\sin(\psi-5)}(0.25(p_4+2p_8+p_{12})) +\\
&\quad \frac{\sin(\psi)}{\sin(\psi-5)}(0.25(p_3+2p_7+p_{11})),\\
&ZW = \frac{\sin(\psi)}{\sin(\psi+5)}(0.50(p_{15}+p_{16}) - 0.50(p_8+p_9)) \; - 0.50\times\cos^2(\psi)\times(0.50(p_8+p_9) - 0.50(p_1+p_2)),\\
&Z = ZS + ZW
\end{aligned}
\tag{1}
$$

The flow direction is based on an 8 direction compass (N, NE, E, SE, S, SW, W, NW) and is calculated as follows:

$$\frac{1}{\tan(WF/SF)} \tag{2}$$

If the outcome of Equation (2) is positive, 180° is added.

Based on a comparison of the flow indices and the flow direction, 27 different WTs are identified. The comparison of the flow indices considers following criteria:

- $|Z| < F$: **pure directional WTs** (W, NW, N, NE, E, SE, S, SW)
- $|Z| > 2F$ and Z >0: **pure cyclonic WT** (C)
- $|Z| > 2F$ and Z <0: **pure anti-cyclonic WT** (A)
- $F < |Z| < 2F$ and Z >0: **hybrid cyclonic WTs** (HCW, $HCNW$, HCN, $HCNE$, HCE, $HCSE$, HCS, $HCSW$)

– $F < |Z| < 2F$ and $Z < 0$: **hybrid anti-cyclonic WTs** (HAW, $HANW$, HAN, $HANE$, HAE, $HASE$, HAS, $HASW$)

– $F < 6$ and $Z < 6$: **undefined WT** (U)

These 27 WTs are regrouped to 11 WTs by equally dividing the hybrid WTs over the corresponding non-directional WTs (cyclonic or anti-cyclonic) and directional WTs. Although this might lead to information loss (Schiemann and Frei, 2010), it is in line with previous case studies for Belgium (Brisson et al., 2011; De Niel et al., 2017; Demuzere et al., 2009; Willems and Vrac, 2011).

Step 2: Statistical downscaling by analogues

Downscaled time series are produced by finding analogues for the projected climate model output. In a first step, the bias in the number wet days is removed using a climate model dependent and seasonal dependent wet day threshold. In the next step, the downscaled precipitation time series are constructed by WT analogues.

The first criteria defining an analogue wet day are the season and WT. Consider day *d* of the projected climate model output, corresponding with season *s*, WT *wt* and a daily precipitation amount *p*. Then, the search for an analogue day is conducted among the observed wet days in season *s* for which the WT equals *wt*. Besides the season and the WT, the exceedance probability of the daily precipitation amount *p* is considered. More precisely, the exceedance probability is calculated using the total daily precipitation amounts of wet days occurring in season *s* and corresponding with the WT *wt*. As such, the analogue precipitation amount for day *d* equals the daily precipitation amount of the observed time series with the closest exceedance probability.

In the case that the observed precipitation time series has a sub-daily time step, the sub-daily precipitation amounts are aggregated to daily precipitation amounts. Next, for each season, for each WT, the exceedance probabilities for the observed daily precipitation amounts of wet days are calculated based on the total daily precipitation amount. After determining the analogue day, the sub-daily precipitation amounts of the analogue day are re-sampled to produce the downscaled time series.

Step 3: Precipitation scaling by the Clausius-Clapeyron relation

Besides large scale circulation patterns, precipitation also responds to thermodynamical processes. The latter processes are accounted for by precipitation scaling following the CC relation. The CC relation describes the water holding capacity in air masses, which more specifically increases by 7% per degree warming. Application of this scaling rate to precipitation intensities is valid assuming that extreme precipitation amounts are controlled by the local moisture availability and are not influenced by the large scale atmospheric circulation patterns. In reality, however, physical processes interact and also higher scaling rates are found (Barbero et al., 2018; Blenkinsop et al., 2018; Manola et al., 2018; Lenderink et al., 2017; Zhang et al., 2017). The CC relation is determined on annual time scale. The temperature rise, to be applied for the CC scaling, is computed using a seasonal quantile based approach.

Although several studies have pointed out that dew point temperature is a better predictor than the average daily temperature (Van de Vyver et al., 2019; Wasko et al., 2018), average daily temperatures have been considered. Remark that data for average daily temperatures are more readily available for hydrological impact modellers.

2.2 Meteorological data

For the main station of the Royal Meteorological Institute in Uccle, precipitation and average temperature time series are available for the period 1901-2000 with a 10-minutes and daily time step, respectively. The historical WTs are identified using the daily gridded MSLP output for the EMULATE, ERA40 and NCEP/NCAR re-analysis datasets(Table 1). Hence, this study accounts for the recent findings of Horton and Brönnimann (2018) and Stryhal and Huth (2017). Both studies indicate that re-analysis datasets introduce uncertainties in the classification of WTs and the statistical downscaling step. By using daily WTs, rapidly occurring changes in the large scale atmospheric circulation might be neglected (Åström et al., 2016). However, the winter season is of interest and for this season no rapidly evolving circulation changes, i.e. within one day, are expected.

The climate model ensemble, presented in Table 2, includes 93 CMIP5 climate model runs of which 33 control runs. For the climate change impact analysis, all four representative concentration pathways (RCPs) are considered, where the RCP 2.6, 4.5, 6.0 and 8.5 sub-ensembles include 20, 28, 15 and 30 climate model runs, respectively. For each climate model run, daily MSLP, precipitation and average temperature output are extracted for 1961-1990 (control period) and 2071-2100 (scenario period). The precipitation and temperature data are extracted for the grid cell covering Uccle, whereas MSLP, required for the WT identification, is extracted for a larger area covering Uccle using the 16-point grid of the WT classification system (Figure 1).

3 Verification of the statistical downscaling assumptions

The verifications of following assumptions are performed for winter season, including the months December, January and February.

3.1 Informative assumption

The informative assumption defines the existence of an informative and physically based relationship between the predictors and predictand. The predictors of the WT method are the average daily temperatures and WTs.

In order to examine the informative assumption for the WTs, the WT occurrences and the precipitation statistics related to the individual WTs are determined for the period 1961-1990. The studied precipitation statistics involve the precipitation accumulation, the number of wet days and the empirical distribution of independent extreme precipitation amounts. The independent 10-minutes, hourly and daily precipitation amounts are determined using a peak-over-threshold method, setting the threshold at 0.1 mm/h and defining at least 12 hours between successive events (Willems, 2000).

In order to examine the informative assumption for the average daily temperatures, the existence of the CC relation is verified. The independent precipitation amounts are determined using the 10-minutes precipitation amount time series and the daily average temperature time series between 1901 and 2000. First, the 10-minutes precipitation events in the time series are identified using a peak-over-threshold method (threshold $=$ 0.1 mm/h and time between successive events $>$12 hours). Next, the precipitation events and corresponding temperatures are classified in moving temperature bins and per bin sorted from low to high (Manola et al., 2018). Finally, the magnification of the 90^{th}, 95^{th} and 99^{th} percentile precipitation amount for increasing temperature bins is investigated.

3.2 Perfect prognosis assumption

The verification of the perfect prognosis assumption is especially of importance for the WT method. More specifically, the application of the WT analogues follows the principle of perfect prognosis methods. This means that a statistical relation is first defined between observed predictors and observed predictand. Thereafter, the statistical relation is applied to the projected climate model output. Consequently, the calibrated statistical relation is not tailored to biases in the climate model output. The scaling of the precipitation amounts by the CC relation, on the contrary, follows the principles of model output statistical methods. Those methods implicitly assume that the climate model biases are time-invariant and that through the application of changes the biases in the projected climate model output are cancelled by the biases in the historical climate model output.

The verification of the perfect prognosis assumption involves a comparison between the observed and climate model simulated WT occurrences. The verification is conducted over the period 1961-1990 using the historical output of 33 global climate model runs (Table 2).

3.3 Stationarity assumption

To verify the stationarity assumption, the contributions of the dynamical and thermodynamical processes governing precipitation changes are studied over time. To this end, the observed period (1901-2000) is split into different sub-periods. The sub-periods are 20 years long and range between 1901 and 1920, 1921 and 1940, 1941 and 1960, 1961 and 1981 and 1981 and 2000. Each sub-period is thereafter considered as the scenario period for surrogate climate model runs. For instance, when 1901-1920 is selected as scenario period, then the periods 1921-1940, 1941-1960, 1961-1981 and 1981-2000 act as control periods. When 1981-2000 is selected as the scenario period, then the periods 1901-1920, 1921-1940, 1941-1960 and 1961-1980 act as control periods. The combination of the different sub-periods yields an ensemble of 20 surrogate climate models. For each of the surrogate climate model runs, the change in the average daily precipitation amounts of wet days is decomposed. More specifically, the precipitation amount changes ΔP are governed by the changes in the WT occurrence changes, i.e. the contribution by the dynamical processes $\Delta P_{dynamical}$, and thermodynamical and local/mesoscale feedback changes, i.e. ΔP_{other} (Souverijns et al., 2016). The contributions are calculated as follows:

$$
\begin{aligned}
&\Delta P = \Delta P_{dynamica} + \Delta P_{other}\\
&\Delta P_{dynamica} = \sum_{j=1}^{11} \left(N_{j,scen} - N_{j,contr}\right) P_{j,contr}\\
&\Delta P_{other} = \sum_{j=1}^{11} \left(P_{j,scen} - P_{j,contr}\right) N_{j,scen}
\end{aligned}
\tag{3}
$$

with

- $N_{j,contr}$ the absolute occurrence frequency of wet days with WT j in the climate model output for the control period,
- $N_{j,scen}$ the absolute occurrence frequency of wet days with WT j in the climate model output for the scenario period,
- $P_{j,contr}$ the average daily precipitation amount of the wet days with WT j in the climate model output for the control period and
- $P_{j,scen}$ the average daily precipitation amount of the wet days with WT j in the climate model output for the scenario period.

The decomposition is also performed using the historical and projected output of the 93 membered global climate model ensemble (Table 2).

3.4 Response to greenhouse gas scenarios

In order to verify the response of the predictand to the greenhouse gas scenarios, the WT method is applied to the output of 93 global climate model runs (Table 2). Next, the daily precipitation amounts for the downscaled time series are compared against the observed precipitation amounts and the intensification of the extreme precipitation amounts for increasing greenhouse gas scenarios is investigated. The intensification is visually inspected and focus is put on the magnification of the changes for increasing greenhouse gas scenarios. Furthermore, a comparison is made between the coarse global climate model changes and the changes for the downscaled time series. For sake of brevity, the changes in the 30-year return period and the average winter precipitation accumulation are investigated.

3.5 Structural downscaling assumptions

To investigate the added value of the CC relation, the original SDM (with CC scaling) and the SDM without CC scaling are applied to the projected output of 93 global climate models (Table 2). The control period and range of observations are defined as 1961-1990 and the scenario period as 2071-2100. A comparison is made between the projected changes for the SDM with CC scaling and the SDM without CC scaling. The added value of the CC relation will be discussed in combination with the response of the predictand to the greenhouse gas scenarios.

4 Results and discussions

4.1 The informative assumption

Figure 2 presents the relative WT occurrence frequencies during winter season. The results show that the A WTs occur most frequently and represent approximately 30% of the winter days. Also the W, SW and C WTs are identified as frequently occurring. The occurrence frequency of each of these WTs is approximately 12%. Despite for some details, there are no differences between the different re-analysis datasets. The WT occurrence patterns are generally in agreement with the recent findings of Otero et al. (2018). They identified the A WTs as the overall dominant winter WT in Europe. The A WTs, more specifically, represent on average 25% of the winter days. The average occurrence frequency of the C WTs in Europe is estimated at approximately 15%, the W WTs at approximately 8% and the SW at approximately 5%. Remark that the WT occurrences presented by Otero et al. (2018) have been averaged out over the European domain.

Apart from the A WTs, the W, SW and C WTs are associated with a high precipitation accumulation and together explain up to 71% of the total winter precipitation accumulation and between 73% and 93% of the WT occurrences are wet days (Supplementary Information, Figures A1 and A2). Additionally, these WTs are associated with higher precipitation amounts, as for instance shown for the NCEP/NCAR re-analysis dataset in Figure 3. More specifically, the 1-year daily precipitation amount for the W WTs measures 0.51 mm/h and is twice as large as the corresponding amount for the A WTs, which measures 0.19 mm/h. Also the NW WTs are characterized by higher precipitation amounts and a higher wet day frequency. However, compared to the W, SW and C WTs, their occurrence is rather low and contributes less to the precipitation accumulation.

The relation between precipitation and temperature is presented in Figure 4. This figure demonstrates the intensification of the independent precipitation amounts with increasing temperature. For instance, the 90^{th} percentile precipitation amount increases by 7% per 1°C and this increase follows the CC relation. For temperatures higher than 10°C, the scaling rate increases up to 14% per 1°C. Similar scaling rates are obtained for the higher precipitation percentiles. For percentiles smaller than 90%, the scaling rate of 7% per 1°C is not identified. The identified CC relation is compared with other studies for Belgium (De Troch, 2016; Van de Vyver et al., 2019) and for neighbouring regions (Lenderink and van Meijgaard, 2008). Although the CC relation in those other studies has been established using dew point temperature rather than average daily temperature, similar scaling rates are obtained. Considering the findings of recent studies, the application of dew point temperature is expected to better estimate the increases in the atmospheric moisture capacity and, thus, the precipitation changes (Van de Vyver et al., 2019; Wasko et al., 2018).

4.2 The perfect prognosis assumption

Figure 2 compares the WT occurrences for the historical climate model outputs with those for the re-analysis datasets. The comparison reveals large biases, in particular for the W and A WTs. More precisely, the climate models overestimate the occurrence of W WTs by approximately 11%, whereas the A WTs are underestimated by 14%. Moreover, different to the re-analysis datasets, the W WTs are the most prominently occurring WTs in the climate model outputs. These findings are in agreement with the recent study by Stryhal and Huth (2018). Using different atmospheric classification patterns, they found an overall overestimation of the westerly circulation, which is estimated to be approximately 7% for the British isles and increases towards central Europe up to 21%. Otero et al. (2018) and Stryhal and Huth (2018) also indicate that climate models have a poor performance in reproducing the occurrence of A WTs.

The overestimation of the W WTs is explained by the orientation of the North Atlantic storm track in the climate models. It has, more specifically, a zonal orientation instead of a SW-NE tilt (Pithan et al., 2016; Zappa et al., 2014). The zonal orientation results in a pronounced meridional pressure gradient, creating zonal westerly flows, which in turn impede the occurrences of anticyclones (Stryhal and Huth, 2019). Biases in the blocking frequency are also arising from the climate model resolution (Anstey et al., 2013; Scaife et al., 2011; Woollings et al., 2018).

Although it would be possible to remove the bias in WT occurrences, for instance through re-sampling (Mehrotra and Sharma, 2019), the studied WT method does not do that. Note that such bias correction would require a technique that simultaneously accounts for the bias in the WT occurrences, the WT persistence and the relation between the WTs and other hydro-meteorological variables.

4.3 The stationarity assumption

The contribution of the dynamical processes to the precipitation amount changes for the surrogate climate model runs is presented in Figure 6. Based on the median results, the dynamical processes are responsible for 35 to 55% of the changes. These high contributions are explained by the findings of Ntegeka and Willems (2008) and Willems (2013). More specifically, oscillations in the 100-years precipitation amount time series for Uccle have been identified. Some periods are characterised by higher precipitation amounts and are referred to as positive anomalies. The periods characterised by smaller precipitation amounts are referred to as negative anomalies. Willems (2013) observed that the precipitation anomalies coincide with anomalies in the number of W WTs and with anomalies in the pressure difference between the Azores and Scandinavia. The coincidence of large scale atmospheric circulation patterns and precipitation amounts has also been studied for other locations in Europe. In this context, Tabari and Willems (2018) identified the North Atlantic Oscillation and the ENSO signal as dominant drivers for the extreme winter precipitation amounts. Hence, the findings of Tabari and Willems (2018) and Willems (2013) imply that large scale atmospheric circulation influences winter precipitation in Europe. For the end of the 20^{th} century, the dynamical processes explain only 20% of the precipitation amount changes. This lower contribution is compensated by a higher contribution by the thermodynamical processes. More specifically, Ntegeka and Willems (2008) point out that the higher precipitation amounts are governed by an intensification of the positive anomaly. The intensification arises from the increasing temperatures, which in turn are attributed to the climatic changes.

Figure 7 shows the contributions by the dynamical and the thermodynamcial processes to the long term projected changes. The changes in the WTs account for 18% of the total change and, hence, they are primarily driven by the thermodynamical processes. This is in agreement with the findings of Kröner (2016), who investigated the drivers for precipitation changes in Europe. As the thermodynamical processes are only to some extent included in the downscaling methodology, the applicability of the SDM is questioned. Note that the application of the CC relation is limited to the extreme precipitation amounts, while the thermodynamical processes also influence the average precipitation amounts.

4.4 Response to the greenhouse gas scenarios and the added value of the CC relation

Climate models project a poleward shift of the Northern Hemisphere jet-streams and storm tracks, resulting increased occurrence of zonal flows and fewer blocking occurrences (Barnes and Screen, 2015; Santos et al., 2016; Stryhal and Huth, 2019; Woollings et al., 2018). As a consequence, an increasing occurrence of W and SW WTs and a decreasing occurrence of A WTs are projected (Supplementary Information, Figure A3). More specifically, under the total uncertainty range, i.e. all RCPs combined, the occurrence of W WTs is projected to increase by 7%. For the RCP sub-ensembles, the increase in W WTs is magnified from 6% for RCP 4.5 to 11% for RCP 8.5. The A WTs, on the contrary, decrease by 10% for RCP 4.5 and 12% for RCP 8.5. Using the same climate model ensemble, the median change in the average temperature is estimated at 1.6°C for

RCP 2.6, 2.1°C for RCP 4.5, 2.6°C for RCP 6.0 and 3.7°C for RCP 8.5 (Termonia et al., 2018).

The changes in the 30-year daily winter precipitation amounts are shown in Figure 8. While the studied SDM without CC scaling does not project any changes in the 30-year daily precipitation amount, the coarse GCM projections do. This indicates that the reliance of the projections on analogues involves significant shortcomings. Those shortcomings can however be overcome by the application of the CC relation. As a result, an intensification of the extreme precipitation amounts is obtained for the studied SDM. The intensification is, moreover, in agreement with the theoretical estimations. The projected changes for RCP 8.5 are, for instance, estimated at 25.8% and equal the theoretical change values (3.7°C $\times$ 7%). Remark that the CC relation is applied at daily time scale and that at daily time scale the super CC scaling rate (14% per 1°C) is not observed. To some extent, the estimated median change values for the studied SDM differ from the coarse climate model projections. More specifically, the differences in the median change values range between 3% and 5%. Besides differences in the change values, there are differences in the monotonicity of the change intensification for increasing GHSs. For the statistically downscaled changes, the intensification of the change values is monotonic due the monotonic increase in the temperature predictor. For the coarse global climate model changes, on the contrary, the monotonicity is masked by random uncertainties, climate model uncertainties and the stochastic uncertainty arising from the internal variability of the climate system (Van Uytven and Willems, 2018).

The changes in average winter precipitation accumulation are shown in Figure 9. The comparison between the coarse climate model changes and the statistically downscaled changes indicates that the statistical downscaling step increases the changes. More specifically, under the total uncertainty range, the statistically downscaled changes are approximately 15% larger. The latter is explained by the absence of bias correction schemes, the overestimation of the W WTs (Figure 5), which have been identified as one of the wetter WTs (Section 4.1), and the projected increase of these WTs (Figure A3). Figure 9 furthermore shows that the statistically downscaled changes for the winter precipitation accumulation are not monotonic for increasing greenhouse gas scenarios. The latter is observed as the CC relation only applies to the extreme precipitation amounts and, hence, the monotonicity of the temperature changes is not transferred to the changes in the average winter precipitation accumulation.

5 Conclusions

The studied SDM does not meet all assumptions. It is shown that the SDM has limitations and its skill could be improved. The WT method fails, among other assumptions, the perfect prognosis assumption. As the method is applied in a perfect prognosis context, improvements should involve the bias correction of the WT occurrences. Since the simulation of large scale atmospheric circulation patterns remains biased in RCMs (Addor et al., 2016; Jury et al., 2018), the application of a statistical bias correction method is suggested. A potential method would be the recently developed re-sampling approach of Mehrotra and Sharma (2019). Further extensions to the latter approach are, however, required to also address the biases in the WT persistence and the biases in the relation between WTs and other hydro-meteorological variables. Although the implementation of bias correction methods, it remains questionable whether the WT method is applicable for precipitation downscaling of the climate model output. The observations confirm an informative relation between the predictors and the predictand, but this relation is time-variant. The WT occurrences explain between 35% and 55% of the historical precipitation amount changes, but their contribution decreases to less than 20% for the end of the 21^{st} century. This means that the precipitation changes for the case study location are controlled by thermodynamical processes rather than dynamical processes (i.e. changes in WT occurrences). As the extreme precipitation amounts are scaled following the CC relation, the thermodynamical processes are to some extent accounted for in the downscaling methodology, but yet insufficiently. The CC relation produces extreme precipitation amounts outside the range of historical observations and thus anticipates the intensification of extreme events. The latter indicates the potential of the CC relation for improving non-parametric precipitation models. The standalone application of the CC relation as a SDM has recently been demonstrated by Manola et al. (2018) and Van de Vyver et al. (2019), but also those SDMs involve shortcomings (Zhang et al., 2017).

Uncovering the shortcomings of SDMs does not mean that their use is discouraged. One should not forget that other SDMs may also fail assumptions and, thus, also have shortcomings. By considering an ensemble of SDMs, the uncertainties introduced by those shortcomings can be taken into account. When SDM ensembles would be considered, ensemble members could be weighted based on their skill. The latter would be similar to the existing climate model weighing techniques (Sanderson et al., 2017). A first step towards a weighted SDM ensemble is still to be made by the statistical downscaling community.

Data availability. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 2) for producing and making available their model output. The climate model data is available through the website of the Earth System Grid Federation, https://esgf.llnl.gov/. The ERA40 re-analysis data set is available through the ECMWF Public Datasets Web Interface http://apps.ecmwf.int/datasets/ and the NCEP/NCAR data set through https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.re-analayis.html. The observed precipitation time series for the Uccle station were made available by Royal Meteorological Institute.

Author contributions. EVU conceptualised and developed the approaches. Under the supervision of PW, EVU and JDN performed the formal analysis and investigation. EVU prepared the visualisation and wrote the initial draft, which was critically reviewed and revised by PW and JDN.

Competing interests. The authors declare that there are no competing interests.

Acknowledgements. Els Van Uytven is funded by a doctoral grant from the Research Foundation – Flanders for her PhD entitled “Evaluation of statistical downscaling methods for climate change impact analysis on hydrological extremes” (F.W.O., grant number 11ZY418N). This financial support is gratefully acknowledged.

References

Addor, N., Rohrer, M., Furrer, R., and Seibert, J.: Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment, Journal of Geophysical Research: Atmospheres, 121, 2075–2089, https://doi.org/10.1002/2015JD024040, 2016.

Alfieri, L., Feyen, L., and Di Baldassarre, G.: Increasing flood risk under climate change: a pan-European assessment of the benefits of four adaptation strategies, Climatic Change, pp. 1–15, https://doi.org/10.1007/s10584-016-1641-1, 2016.

Ansell, T. J., Jones, P. D., Allan, R. J., Lister, D., Parker, D. E., Brunet, M., Moberg, a., Jacobeit, J., Brohan, P., Rayner, N. a., Aguilar, E., Alexandersson, H., Barriendos, M., Brandsma, T., Cox, N. J., Della-Marta, P. M., Drebs, a., Founda, D., Gerstengarbe, F., Hickey, K., Jónsson, T., Luterbacher, J., Nordli, Oesterle, H., Petrakis, M., Philipp, a., Rodwell, M. J., Saladie, O., Sigro, J., Slonosky, V., Srnec, L., Swail, V., García-Suárez, a. M., Tuomenvirta, H., Wang, X., Wanner, H., Werner, P., Wheeler, D., and Xoplaki, E.: Daily mean sea level pressure reconstructions for the European-North Atlantic region for the period 1850-2003, Journal of Climate, 19, 2717–2742, https://doi.org/10.1175/JCLI3775.1, 2006.

Anstey, J. A., Davini, P., Gray, L. J., Woollings, T. J., Butchart, N., Cagnazzo, C., Christiansen, B., Hardiman, S. C., Osprey, S. M., and Yang, S.: Multi-model analysis of Northern Hemisphere winter blocking: Model biases and the role of resolution, Journal of Geophysical Research Atmospheres, 118, 3956–3971, https://doi.org/10.1002/jgrd.50231, 2013.

Arakawa, A.: The Cumulus Parameterization Problem: Past, Present, and Future, Journal of Climate, 17, 2493–2525, https://doi.org/10.1175/1520-0442(2004)017<2493:RATCPP>2.0.CO;2, 2004.

Åström, H. L., Sunyer, M., Madsen, H., Rosbjerg, D., and Arnbjerg-Nielsen, K.: Explanatory analysis of the relationship between atmospheric circulation and occurrence of flood-generating events in a coastal city, Hydrological Processes, 30, 2773–2788, https://doi.org/10.1002/hyp.10767, 2016.

Barbero, R., Westra, S., Lenderink, G., and Fowler, H. J.: Temperature-extreme precipitation scaling: a two-way causality?, International Journal of Climatology, 38, e1274–e1279, https://doi.org/10.1002/joc.5370, 2018.

Barnes, E. A. and Screen, J. A.: The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it?, Wiley Interdisciplinary Reviews: Climate Change, 6, 277–286, https://doi.org/10.1002/wcc.337, 2015.

Benestad, R. E., Hanssen-Bauer, I., and Chen, D.: Empirical-Statistical Downscaling, WORLD SCIENTIFIC, https://doi.org/10.1142/6908, 2008.

Blenkinsop, S., Fowler, H. J., Barbero, R., Chan, S. C., Guerreiro, S. B., Kendon, E., Lenderink, G., Lewis, E., Li, X.-F., Westra, S., Alexander, L., Allan, R. P., Berg, P., Dunn, R. J. H., Ekström, M., Evans, J. P., Holland, G., Jones, R., Kjellström, E., Klein-Tank, A., Lettenmaier, D., Mishra, V., Prein, A. F., Sheffield, J., and Tye, M. R.: The INTENSE project: using observations and models to understand the past, present and future of sub-daily rainfall extremes, Advances in Science and Research, 15, 117–126, https://doi.org/10.5194/asr-15-117-2018, 2018.

Brekke, L. D., Maurer, E. P., Anderson, J. D., Dettinger, M. D., Townsley, E. S., Harrison, A., and Pruitt, T.: Assessing reservoir operations risk under climate change, Water Resources Research, 45, 1–16, https://doi.org/10.1029/2008WR006941, 2009.

Brisson, E., Demuzere, M., Kwakernaak, B., and Van Lipzig, N. P.: Relations between atmospheric circulation and precipitation in Belgium, Meteorology and Atmospheric Physics, 111, 27–39, https://doi.org/10.1007/s00703-010-0103-y, 2011.

Bürger, G., Murdock, T. Q., Werner, A. T., Sobie, S. R., and Cannon, A. J.: Downscaling Extremes—An Intercomparison of Multiple Statistical Methods for Present Climate, Journal of Climate, 25, 4366–4388, https://doi.org/10.1175/JCLI-D-11-00408.1, 2012.

Casanueva, A., Herrera, S., Fernández, J., and Gutiérrez, J. M.: Towards a fair comparison of statistical and dynamical downscaling in the framework of the EURO-CORDEX initiative, Climatic Change, 137, 411–426, https://doi.org/10.1007/s10584-016-1683-4, 2016.

Cristiano, E., ten Veldhuis, M.-C., Gaitan, S., Ochoa Rodriguez, S., and van de Giesen, N.: Critical scales to explain urban hydrological response, Hydrology and Earth System Sciences Discussions, pp. 1–40, https://doi.org/10.5194/hess-2017-715, 2018.

Davini, P., Corti, S., D'Andrea, F., Rivière, G., and von Hardenberg, J.: Improved Winter European Atmospheric Blocking Frequencies in High-Resolution Global Climate Simulations, Journal of Advances in Modeling Earth Systems, 9, 2615–2634, https://doi.org/10.1002/2017MS001082, 2017.

De Niel, J., Demarée, G., and Willems, P.: Weather Typing-Based Flood Frequency Analysis Verified for Exceptional Historical Events of Past 500 Years Along the Meuse River, Water Resources Research, 53, 8459–8474, https://doi.org/10.1002/2017WR020803, 2017.

De Troch, R.: The application of the ALARO-0 model for regional climate modeling in Belgium : extreme precipitation and unfavorable conditions for the dispersion of air pollutants under present and future climate conditions, https://biblio.ugent.be/publication/7247081, 2016.

Demuzere, M., Werner, M., van Lipzig, N. P. M., and Roeckner, E.: An analysis of present and future ECHAM5 pressure fields using a classification of circulation patterns, International Journal of Climatology, 29, 1796–1810, https://doi.org/10.1002/joc.1821, 2009.

Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: The role of internal variability, Climate Dynamics, 38, 527–546, https://doi.org/10.1007/s00382-010-0977-x, 2012.

Dixon, K. W., Lanzante, J. R., Nath, M. J., Hayhoe, K., Stoner, A., Radhakrishnan, A., Balaji, V., and Gaitán, C. F.: Evaluating the stationarity assumption in statistically downscaled climate projections: is past performance an indicator of future results?, Climatic Change, 135, 395–408, https://doi.org/10.1007/s10584-016-1598-0, 2016.

Dottori, F., Szewczyk, W., Ciscar, J.-C., Zhao, F., Alfieri, L., Hirabayashi, Y., Bianchi, A., Mongelli, I., Frieler, K., Betts, R., and Feyen, L.: Increased human and economic losses from river flooding with anthropogenic warming, Nature Climate Change, 8, https://doi.org/10.1038/s41558-018-0257-z, 2018.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "should we apply bias correction to global and regional climate model data?", Hydrology and Earth System Sciences, 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Emori, S. and Brown, S. J.: Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate, Geophysical Research Letters, 32, 1–5, https://doi.org/10.1029/2005GL023272, 2005.

Fadhel, S., Rico-Ramirez, M. A., and Han, D.: Uncertainty of Intensity–Duration–Frequency (IDF) curves due to varied climate baseline periods, Journal of Hydrology, 547, 600–612, https://doi.org/10.1016/j.jhydrol.2017.02.013, 2017.

Flaounas, E., Drobinski, P., Vrac, M., Bastin, S., Lebeaupin-Brossier, C., Stéfanon, M., Borga, M., and Calvet, J. C.: Precipitation and temperature space-time variability and extremes in the Mediterranean region: Evaluation of dynamical and statistical downscaling methods, Climate Dynamics, 40, 2687–2705, https://doi.org/10.1007/s00382-012-1558-y, 2013.

Fu, G., Charles, S. P., Chiew, F. H., Ekström, M., and Potter, N. J.: Uncertainties of statistical downscaling from predictor selection: Equifinality and transferability, Atmospheric Research, 203, 130–140, https://doi.org/10.1016/j.atmosres.2017.12.008, 2018.

Gutmann, E., Pruitt, T., Clark, M. P., Brekke, L., Arnold, J. R., Raff, D. A., and Rasmussen, R. M.: An intercomparison of statistical downscaling methods used for water resource assessments in the United States, Water Resources Research, 50, 7167–7186, https://doi.org/10.1002/2014WR015559, 2014.

Haberlandt, U., Belli, A., and Bárdossy, A.: Statistical downscaling of precipitation using a stochastic rainfall model conditioned on circulation patterns - an evaluation of assumptions, International Journal of Climatology, 35, 417–432, https://doi.org/10.1002/joc.3989, 2015.

Hartung, K., Svensson, G., and Kjellström, E.: Resolution, physics and atmosphere–ocean interaction – How do they influence climate model representation of Euro-Atlantic atmospheric blocking?, Tellus A: Dynamic Meteorology and Oceanography, 69, 1406 252, https://doi.org/10.1080/16000870.2017.1406252, 2017.

Hertig, E., Merkenschlager, C., and Jacobeit, J.: Change points in predictors–predictand relationships within the scope of statistical downscaling, International Journal of Climatology, 37, 1619–1633, https://doi.org/10.1002/joc.4801, 2017.

Hertig, E., Maraun, D., Bartholy, J., Pongracz, R., Vrac, M., Mares, I., Gutiérrez, J. M., Wibig, J., Casanueva, A., and Soares, P. M. M.: Comparison of statistical downscaling methods with respect to extreme events over Europe: Validation results from the perfect predictor experiment of the COST Action VALUE, International Journal of Climatology, pp. 1–22, https://doi.org/10.1002/joc.5469, 2018.

Hewitson, B. C., Daron, J., Crane, R. G., Zermoglio, M. F., and Jack, C.: Interrogating empirical-statistical downscaling, Climatic Change, 122, 539–554, https://doi.org/10.1007/s10584-013-1021-z, 2014.

Horton, P. and Brönnimann, S.: Impact of global atmospheric reanalyses on statistical precipitation downscaling, Climate Dynamics, https://doi.org/10.1007/s00382-018-4442-6, 2018.

Jenkinson, A. and Collison, F.: An initial climatology of gales over the North Sea, Synoptic Climatology Branch Memorandum, 62, 1977.

Jones, P. D., Hulme, M., and Briffa, K. R.: A comparison of Lamb circulation types with an objective classification scheme, International Journal of Climatology, 13, 655–663, https://doi.org/10.1002/joc.3370130606, 1993.

Jury, M. W., Herrera, S., Gutiérrez, J. M., and Barriopedro, D.: Blocking representation in the ERA-Interim driven EURO-CORDEX RCMs, Climate Dynamics, https://doi.org/10.1007/s00382-018-4335-8, 2018.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bulletin of the American Meteorological Society, 77, 437–472, https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., Van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble, Geoscientific Model Development, 7, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014, 2014.

Kröner, N.: Identifying and quantifying large-scale drivers of European climate change, https://doi.org/10.3929/ethz-a-010793497, 2016.

Kröner, N., Kotlarski, S., Fischer, E., Lüthi, D., Zubler, E., and Schär, C.: Separating climate change signals into thermodynamic, lapse-rate and circulation effects: theory and application to the European summer climate, Climate Dynamics, 48, 3425–3440, https://doi.org/10.1007/s00382-016-3276-3, 2017.

Lanzante, J. R., Dixon, K. W., Nath, M. J., Whitlock, C. E., and Adams-Smith, D.: Some Pitfalls in Statistical Downscaling of Future Climate, Bulletin of the American Meteorological Society, 99, 791–803, https://doi.org/10.1175/BAMS-D-17-0046.1, 2018.

Le Roux, R., Katurji, M., Zawar-Reza, P., Quénol, H., and Sturman, A.: Comparison of statistical and dynamical downscaling results from the WRF model, Environmental Modelling and Software, 100, 67–73, https://doi.org/10.1016/j.envsoft.2017.11.002, 2018.

Lenderink, G. and van Meijgaard, E.: Increase in hourly precipitation extremes beyond expectations from temperature changes, Nature Geoscience volume, 1, 511–514, https://doi.org/doi.org/10.1038/ngeo262, 2008.

Lenderink, G., Barbero, R., Loriaux, J. M., and Fowler, H. J.: Super-Clausius-Clapeyron scaling of extreme hourly convective precipitation and its relation to large-scale atmospheric conditions, Journal of Climate, 30, 6037–6052, https://doi.org/10.1175/JCLI-D-16-0808.1, 2017.

Li, J., Johnson, F., Evans, J., and Sharma, A.: A comparison of methods to estimate future sub-daily design rainfall, Advances in Water Resources, 110, 215–227, https://doi.org/10.1016/j.advwatres.2017.10.020, 2017.

Manola, I., van den Hurk, B., De Moel, H., and Aerts, J. C. J. H.: Future extreme precipitation intensities based on a historic event, Hydrology and Earth System Sciences, 22, 3777–3788, https://doi.org/10.5194/hess-22-3777-2018, 2018.

Maraun, D.: Bias Correcting Climate Change Simulations - a Critical Review, Current Climate Change Reports, 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

Maraun, D. and Widmann, M.: Statistical Downscaling and Bias Correction for Climate Research, Cambridge University Press, https://doi.org/10.1017/9781107588783, 2018.

Maraun, D., Wetterhall, F., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developements to bridge the gap between dynamical models and the end user, Reviews of Geophysics, 48, 1–38, https://doi.org/10.1029/2009RG000314.1.INTRODUCTION, 2010.

Maraun, D., Widmann, M., and Gutiérrez, J. M.: Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment, International Journal of Climatology, pp. 1–12, https://doi.org/10.1002/joc.5877, 2018.

Mehrotra, R. and Sharma, A.: A Resampling Approach for Correcting Systematic Spatiotemporal Biases for Multiple Variables in a Changing Climate, Water Resources Research, 55, 754–770, https://doi.org/10.1029/2018WR023270, 2019.

Mendoza, P. A., Mizukami, N., Ikeda, K., Clark, M. P., Gutmann, E. D., Arnold, J. R., Brekke, L. D., and Rajagopalan, B.: Effects of different regional climate model resolution and forcing scales on projected hydrologic changes, Journal of Hydrology, 541, 1003–1019, https://doi.org/10.1016/j.jhydrol.2016.08.010, 2016.

Merkenschlager, C., Hertig, E., and Jacobeit, J.: Non-stationarities in the relationships of heavy precipitation events in the Mediterranean area and the large-scale circulation in the second half of the 20th century, Global and Planetary Change, 151, 108 – 121, https://doi.org/10.1016/j.gloplacha.2016.10.009, 2017.

Ntegeka, V. and Willems, P.: Trends and multidecadal oscillations in rainfall extremes, based on a more than 100-year time series of 10 min rainfall intensities at Uccle, Belgium, Water Resources Research, 44, 1–15, https://doi.org/10.1029/2007WR006471, 2008.

Otero, N., Sillmann, J., and Butler, T.: Assessment of an extended version of the Jenkinson–Collison classification on CMIP5 models over Europe, Climate Dynamics, 50, 1559–1579, https://doi.org/10.1007/s00382-017-3705-y, 2018.

Philipp, A., Beck, C., Huth, R., and Jacobeit, J.: Development and comparison of circulation type classifications using the COST 733 dataset and software, International Journal of Climatology, 36, 2673–2691, https://doi.org/10.1002/joc.3920, 2016.

Pithan, F., Shepherd, T. G., Zappa, G., and Sandu, I.: Climate model biases in jet streams, blocking and storm tracks resulting from missing orographic drag, Geophysical Research Letters, 43, 7231–7240, https://doi.org/10.1002/2016GL069551, 2016.

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., Keller, M., Tölle, M., Gutjahr, O., Feser, F., Brisson, E., Kollet, S., Schmidli, J., Van Lipzig, N. P., and Leung, R.: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges, Reviews of Geophysics, 53, 323–361, https://doi.org/10.1002/2014RG000475, 2015.

Roberts, D. R., Wood, W. H., and Marshall, S. J.: Assessments of downscaled climate data with a high-resolution weather station network, International Journal of Climatology, 39, 3091–3103, https://doi.org/10.1002/joc.6005, 2019.

Rybka, H. and Tost, H.: Uncertainties in future climate predictions due to convection parameterisations, Atmospheric Chemistry and Physics, 14, 5561–5576, https://doi.org/10.5194/acp-14-5561-2014, 2014.

Sachindra, D. A., Ahmed, K., Shahid, S., and Perera, B. J. C.: Cautionary note on the use of genetic programming in statistical downscaling, International Journal of Climatology, pp. 1–17, https://doi.org/10.1002/joc.5508, 2018.

Salvadore, E., Bronders, J., and Batelaan, O.: Hydrological modelling of urbanized catchments: A review and future directions, Journal of Hydrology, 529, 62–81, https://doi.org/10.1016/j.jhydrol.2015.06.028, 2015.

Salvi, K., Ghosh, S., and Ganguly, A. R.: Credibility of statistical downscaling under nonstationary climate, Climate Dynamics, 46, 1991–2023, https://doi.org/10.1007/s00382-015-2688-9, 2016.

Sanderson, B. M., Wehner, M., and Knutti, R.: Skill and independence weighting for multi-model assessments, Geoscientific Model Development, 10, 2379–2395, https://doi.org/10.5194/gmd-10-2379-2017, 2017.

Santos, J. A., Belo-Pereira, M., Fraga, H., and Pinto, J. G.: Understanding climate change projections for precipitation over western Europe with a weather typing approach, Journal of Geophysical Research: Atmospheres, 121, 1170–1189, https://doi.org/10.1002/2015JD024399, 2016.

Scaife, A. A., Copsey, D., Gordon, C., Harris, C., Hinton, T., Keeley, S., O'Neill, A., Roberts, M., and Williams, K.: Improved Atlantic winter blocking in a climate model, Geophysical Research Letters, 38, https://doi.org/10.1029/2011GL049573, 2011.

Schiemann, R. and Frei, C.: How to quantify the resolution of surface climate by circulation types: An example for Alpine precipitation, Physics and Chemistry of the Earth, 35, 403–410, https://doi.org/10.1016/j.pce.2009.09.005, 2010.

Schoof, J. T.: Statistical Downscaling in Climatology, Geography Compass, 7, 249–265, https://doi.org/10.1111/gec3.12036, 2013.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., Senior, C. A., Sobel, A. H., Stainforth, D. A., Tett, S. F. B., Trenberth, K. E., van den Hurk, B. J. J. M., Watkins, N. W., Wilby, R. L., and Zenghelis, D. A.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, Climatic Change, 151, 555–571, https://doi.org/10.1007/s10584-018-2317-9, 2018.

Sørup, H. J. D., Davidsen, S., Löwe, R., Thorndahl, S. L., Borup, M., and Arnbjerg-Nielsen, K.: Evaluating catchment response to artificial rainfall from four weather generators for present and future climate, Water Science and Technology, 77, 2578–2588, https://doi.org/10.2166/wst.2018.217, 2018.

Souverijns, N., Thiery, W., Demuzere, M., and Lipzig, N. P. M. V.: Drivers of future changes in East African precipitation Drivers of future changes in East African precipitation, Environmental Research Letters, 11, 1–9, https://doi.org/10.1088/1748-9326/11/11/114011, 2016.

Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P.: Summary for Policymakers, book section SPM, p. 1–30, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, https://doi.org/10.1017/CBO9781107415324.004, www.climatechange2013.org, 2013.

Stryhal, J. and Huth, R.: Classifications of winter Euro-Atlantic circulation patterns: An intercomparison of five atmospheric reanalyses, Journal of Climate, 30, 7847–7861, https://doi.org/10.1175/JCLI-D-17-0059.1, 2017.

Stryhal, J. and Huth, R.: Classifications of winter atmospheric circulation patterns: validation of CMIP5 GCMs over Europe and the North Atlantic, Climate Dynamics, pp. 1–24, https://doi.org/10.1007/s00382-018-4344-7, 2018.

Stryhal, J. and Huth, R.: Trends in winter circulation over the British Isles and central Europe in twenty-first century projections by 25 CMIP5 GCMs, Climate Dynamics, 52, 1063–1075, https://doi.org/10.1007/s00382-018-4178-3, 2019.

Sunyer, M. A., Hundecha, Y., Lawrence, D., Madsen, H., Willems, P., Martinkova, M., Vormoor, K., Bürger, G., Hanel, M., Kriaučiūnienė—, J., Loukas, A., Osuch, M., and Yücel, I.: Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe, Hydrology and Earth System Sciences, 19, 1827–1847, https://doi.org/10.5194/hess-19-1827-2015, 2015.

Tabari, H. and Willems, P.: Lagged influence of Atlantic and Pacific climate patterns on European extreme precipitation, Scientific Reports, https://doi.org/10.1038/s41598-018-24069-9, 2018.

Tabari, H., De Troch, R., Giot, O., Hamdi, R., Termonia, P., Saeed, S., Brisson, E., Van Lipzig, N., and Willems, P.: Local impact analysis of climate change on precipitation extremes: are high-resolution climate models needed for realistic simulations?, Hydrology and Earth System Sciences, 20, 3843–3857, https://doi.org/10.5194/hess-20-3843-2016, 2016.

Termonia, P., Van Schaeybroeck, B., De Cruz, L., De Troch, R., Caluwaerts, S., Giot, O., Hamdi, R., Vannitsem, S., Duchêne, F., Willems, P., Tabari, H., Van Uytven, E., Hosseinzadehtalaei, P., Van Lipzig, N., Wouters, H., Vanden Broucke, S., van Ypersele, J. P., Marbaix, P., Villanueva-Birriel, C., Fettweis, X., Wyard, C., Scholzen, C., Doutreloup, S., De Ridder, K., Gobin, A., Lauwaet, D., Stavrakou, T., Bauwens, M., Müller, J. F., Luyten, P., Ponsar, S., Van den Eynde, D., and Pottiaux, E.: The CORDEX.be initiative as a foundation for climate services in Belgium, Climate Services, https://doi.org/10.1016/j.cliser.2018.05.001, 2018.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, Journal of Hydrology, 456-457, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052, 2012.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A., Jenne, R., McNally, A. P., Mahfouf, J. F., Morcrette, J. J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961–3012, https://doi.org/10.1256/qj.04.176, 2005.

Vaittinada Ayar, P., Vrac, M., Bastin, S., Carreau, J., Déqué, M., and Gallardo, C.: Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: present climate evaluations, Climate Dynamics, 46, 1301–1329, https://doi.org/10.1007/s00382-015-2647-5, 2016.

Van de Vyver, H., Van Schaeybroeck, B., De Troch, R., Hamdi, R., and Termonia, P.: Modeling the scaling of short-duration precipitation extremes with temperature, Earth and Space Science, https://doi.org/10.1029/2019EA000665, 2019.

Van Uytven, E. and Willems, P.: Greenhouse gas scenario sensitivity and uncertainties in precipitation projections for central Belgium, Journal of Hydrology, 558, 9–19, https://doi.org/10.1016/j.jhydrol.2018.01.018, 2018.

Vanden Broucke, S., Wouters, H., Demuzere, M., and van Lipzig, N. P. M.: The influence of convection-permitting regional climate modeling on future projections of extreme precipitation: dependency on topography and timescale, Climate Dynamics, https://doi.org/10.1007/s00382-018-4454-2, 2018.

Vansteenkiste, T., Tavakoli, M., Ntegeka, V., Smedt, F. D., Batelaan, O., Pereira, F., and Willems, P.: Intercomparison of hydrological model structures and calibration approaches in climate scenario impact projections, Journal of Hydrology, 519, 743 – 755, https://doi.org/https://doi.org/10.1016/j.jhydrol.2014.07.062, 2014.

Vermuyten, E., Meert, P., Wolfs, V., and Willems, P.: Combining Model Predictive Control with a Reduced Genetic Algorithm for Real-Time Flood Control, Journal of Water Resources Planning and Management, 144, 04017 083, https://doi.org/10.1061/(ASCE)WR.1943-5452.0000859, 2018.

Wang, L., Ranasinghe, R., Maskey, S., van Gelder, P. H. A. J. M., and Vrijling, J. K.: Comparison of empirical statistical methods for downscaling daily climate projections from CMIP5 GCMs: A case study of the Huai River Basin, China, International Journal of Climatology, 36, 145–164, https://doi.org/10.1002/joc.4334, 2016.

Wang, Y., Sivandran, G., and Bielicki, J. M.: The stationarity of two statistical downscaling methods for precipitation under different choices of cross-validation periods, International Journal of Climatology, 38, e330–e348, https://doi.org/10.1002/joc.5375, 2018.

Wasko, C., Lu, W. T., and Mehrotra, R.: Relationship of extreme precipitation, dry-bulb temperature, and dew point temperature across Australia, Environmental Research Letters, 13, https://doi.org/10.1088/1748-9326/aad135, 2018.

Watterson, I. G., Bathols, J., and Heady, C.: What influences the skill of climate models over the continents?, Bulletin of the American Meteorological Society, 95, 689–700, https://doi.org/10.1175/BAMS-D-12-00136.1, 2014.

Werner, A. T. and Cannon, A. J.: Hydrologic extremes - an intercomparison of multiple gridded statistical downscaling methods, Hydrology and Earth System Sciences, 20, 1483–1508, https://doi.org/10.5194/hess-20-1483-2016, 2016.

Widmann, M., Bedia, J., Gutiérrez, J. M., Bosshard, T., Hertig, E., Maraun, D., Casado, M. J., Ramos, P., Cardoso, R. M., Soares, P. M. M., Ribalaygua, J., Pagé, C., Fischer, A. M., Herrera, S., and Huth, R.: Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment, International Journal of Climatology, pp. 1–27, https://doi.org/10.1002/joc.6024, 2019.

Wilby, R. L. and Wigley, T. M. L.: Precipitation predictors for downscaling: Oberserved and General Circulation Model relationships, Int. J. Climatol., 20, 641–661, https://doi.org/Doi 10.1002/(Sici)1097-0088(200005)20:6<641::Aid-Joc501>3.0.Co;2-1, 2000.

Willems, P.: Compound intensity / duration / frequency-relationships of extreme precipitation for two seasons and two storm types, Journal of Hydrology, 233, 189–205, 2000.

Willems, P.: Multidecadal oscillatory behaviour of rainfall extremes in Europe, Climatic Change, 120, 931–944, https://doi.org/10.1007/s10584-013-0837-x, 2013.

Willems, P. and Vrac, M.: Statistical precipitation downscaling for small-scale hydrological impact investigations of climate change, Journal of Hydrology, 402, 193–205, https://doi.org/10.1016/j.jhydrol.2011.02.030, 2011.

Woollings, T., Barriopedro, D., Methven, J., Son, S.-W., Martius, O., Harvey, B., Sillmann, J., Lupo, A. R., and Seneviratne, S.: Blocking and its Response to Climate Change, Current Climate Change Reports, 4, 287–300, https://doi.org/10.1007/s40641-018-0108-z, 2018.

Wootten, A., Terando, A., Reich, B. J., Boyles, R. P., and Semazzi, F.: Characterizing sources of uncertainty from global climate models and downscaling techniques, Journal of Applied Meteorology and Climatology, 56, 3245–3262, https://doi.org/10.1175/JAMC-D-17-0087.1, 2017.

Yang, C., Wang, N., and Wang, S.: A comparison of three predictor selection methods for statistical downscaling, International Journal of Climatology, 37, 1238–1249, https://doi.org/10.1002/joc.4772, 2017.

Yang, Y., Tang, J., Xiong, Z., Wang, S., and Yuan, J.: An intercomparison of multiple statistical downscaling methods for daily precipitation and temperature over China: present climate evaluations, Climate Dynamics, https://doi.org/10.1007/s00382-019-04809-x, 2019.

Zappa, G., Masato, G., Shaffrey, L., Woollings, T., and Hodges, K.: Linking Northern Hemisphere blocking and storm track biases in the CMIP5 climate models, Geophysical Research Letters, 41, 135–139, https://doi.org/10.1002/2013GL058480, 2014.

Zhang, X., Zwiers, F. W., Li, G., Wan, H., and Cannon, A. J.: Complexity in estimating past and future extreme short-duration rainfall, Nature Geoscience, 10, 255–259, https://doi.org/10.1038/ngeo2911, 2017.

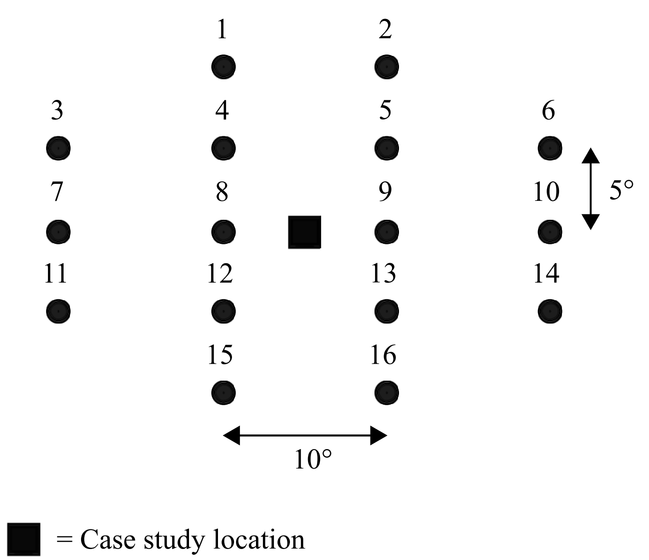

Figure 1. Spacing and numbering in the 16-point grid for the Jenkinson-Collision automated Lamb weather typing classification scheme.

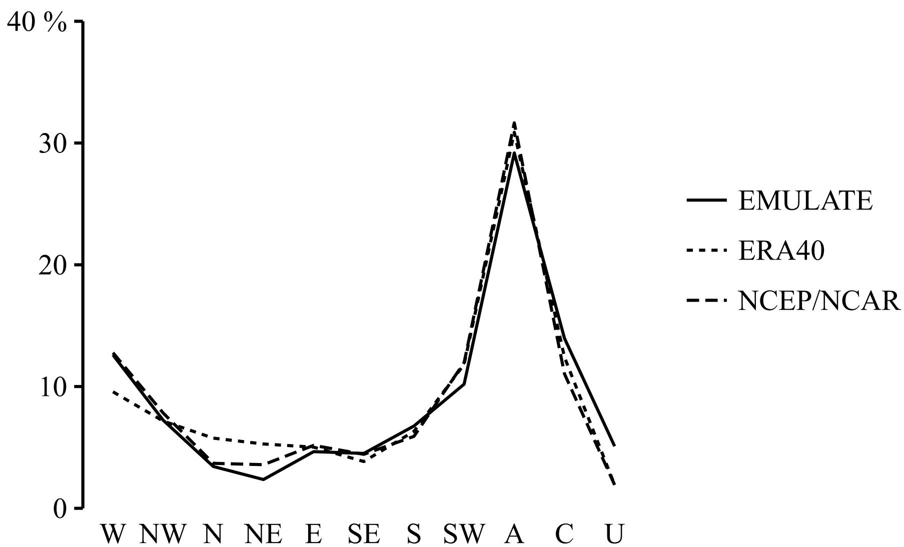

Figure 2. Relative WT occurrence in the winter season for different re-analysis datasets. The results are obtained for the reference period 1961-1990.

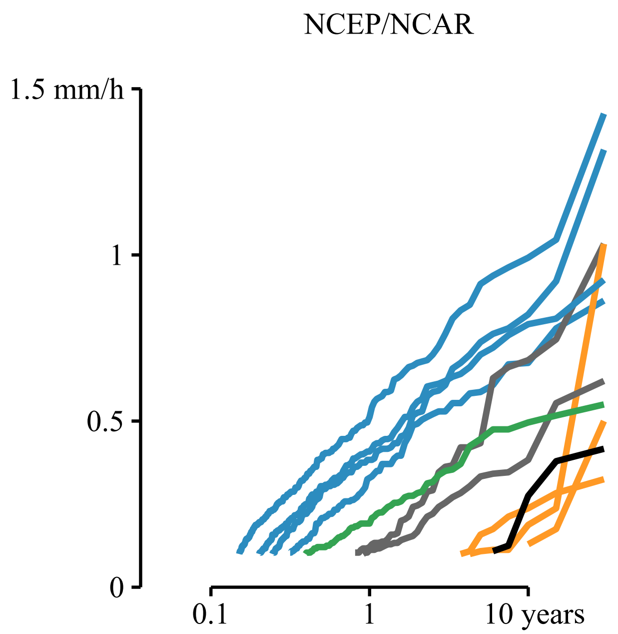

Figure 3. Daily winter precipitation amounts per weather type for the NCEP/NCAR re-analysis dataset. The blue lines indicate the results for the W, NW, SW and C WTs, the green line for the A WT, the grey lines for the N and S WTs, the orange lines for the NE, E, SE WTs and the black line for the U WTs. The results are obtained for the reference period 1961-1990.

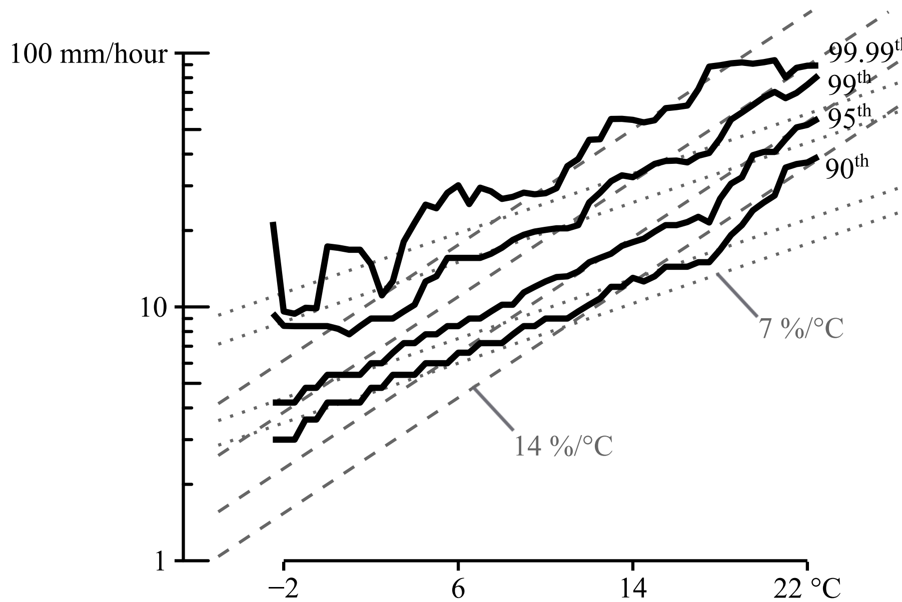

Figure 4. The relation between daily average temperature and independent 10 minutes precipitation amounts. The relation is defined on annual time scale and this by using the entire Uccle time series (1901-2000). The CC relation (+7%/°C) is indicated by the grey dotted lines, whereas the 2 × CC relation (+14%/°C) by the grey dashed lines. The black lines show the magnification of the 90^{th}, 95^{th}, 99^{th} and 99.9^{th} percentile precipitation amount for increasing temperatures.

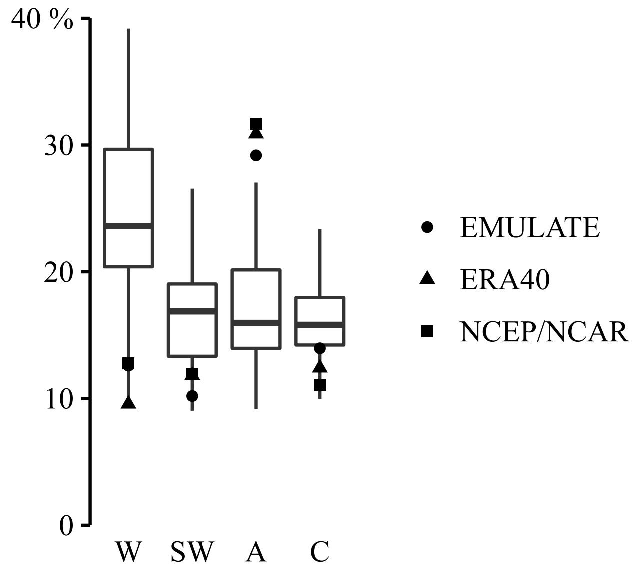

Figure 5. Relative WT occurrence in the winter season for different re-analysis datasets (dots) and climate model runs (boxplots). The results are obtained for the reference period 1961-1990.

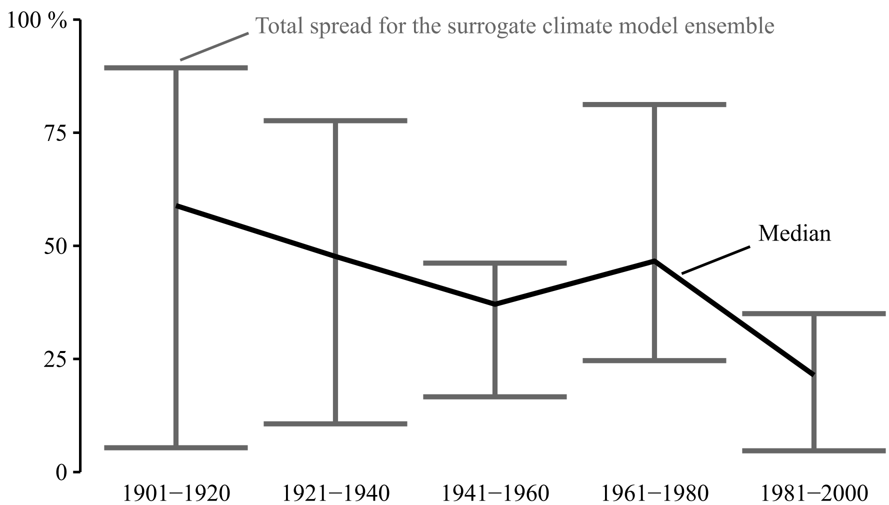

Figure 6. Contribution of the dynamical processes to the change in the average precipitation amount of wet days in function of the choice of the scenario period.

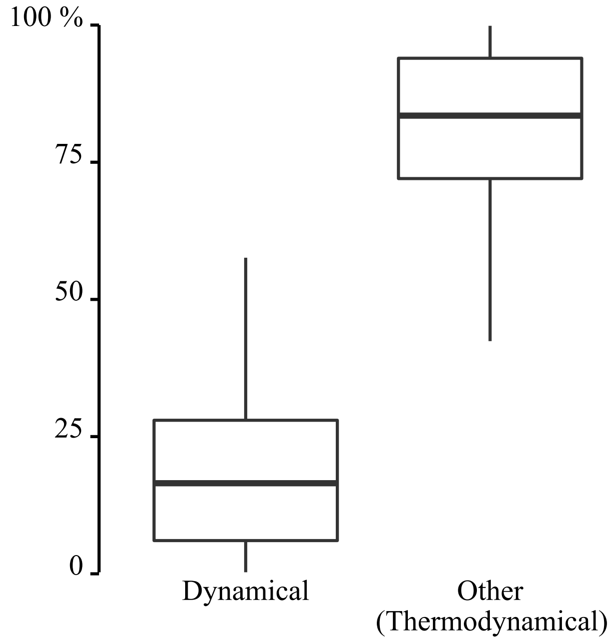

Figure 7. Contribution of the dynamical processes, i.e. changes in the WT occurrence frequencies, and other effects, for instance due to thermodynamical processes, to the change in the average daily precipitation amount of wet winter days. The results are based on the climate model output for the scenario (2071-2100) and control period (1961-1990).

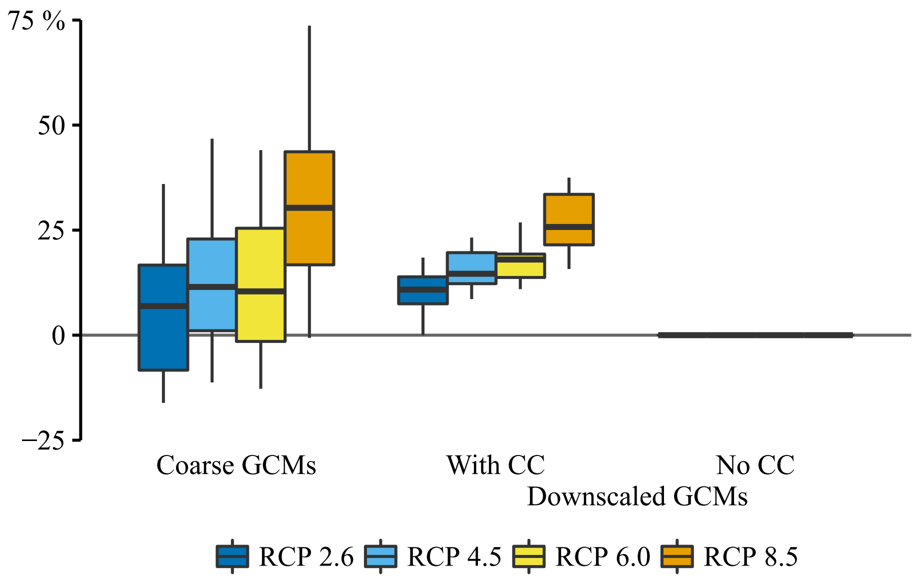

Figure 8. Changes in the 30-year daily precipitation winter amount in function of the different RCPs. The changes are obtained using the climate model ouput for 2071-2100 with respect to the output for 1961-1990.

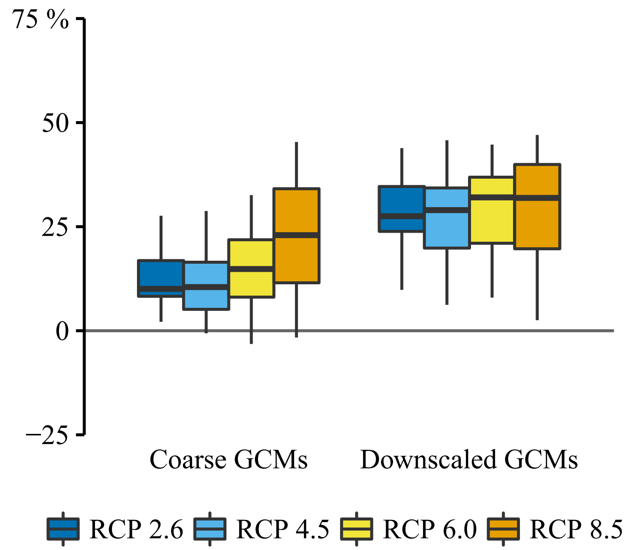

Figure 9. Changes in the average winter precipitation accumulation in function of the different RCPs. The changes are obtained using the climate model ouput for 2071-2100 with respect to the output for 1961-1990.

Table 1. Overview of the re-analysis dataset ensemble employed in this study.

Re-analysis dataset	Resolution Lon [$^\circ$] $\times$ Lat [$^\circ$]	Time range	Reference
EMULATE	5.0×5.0	1881-2000	Ansell et al. (2006)
ERA40	2.0×2.0	1948-2019	Uppala et al. (2005)
NCEP/NCAR	2.5×2.5	1957-2002	Kalnay et al. (1996)

Table 2. Overview of the climate model ensemble employed in this study.

Model	Resolution Lon [°] × Lat [°]	RCP 8.5	RCP 6.0	RCP 4.5	RCP 2.6
ACCESS1-0	1.9 × 1.3	1			
bcc-csm1-1-m	1.1 × 1.1	1	1	1	1
BNU-ESM	2.8 × 2.8	1		1	1
CanESM2	2.8 × 2.8	5		5	5
CMCC-CESM	3.8 × 3.7	1			
CMCC-CMS	1.9 × 1.8	1		1	
CNRM-CM5	1.4 × 1.4	1		1	1
CSIRO-Mk3-6-0	1.9 × 1.9	10	10	10	10
GFDL-CM3	2.5 × 2.0	1	1	1	1
GFDL-ESM2G	2.5 × 1.5	1	1	1	1
GFDL-ESM2M	2.5 × 1.5	1	1	1	
GISS-E2-R	2.5 × 2.0			2	
HadGEM2-AO	1.9 × 1.3	1	1	1	1
IPSL-CM5A-LR	3.8 × 1.9	4	1	4	1
IPSL-CM5A-MR	2.5 × 1.3	1	1	1	1
IPSL-CM5B-LR	3.8 × 1.9	1		1	
MIROC-ESM-CHEM	2.8 × 2.8	1	1	1	1
MIROC-ESM	2.8 × 2.8	1	1	1	1
MIROC5	1.4 × 1.4	3	3	2	3
MPI-ESM-LR	1.9 × 1.8	1		1	1
MPI-ESM-MR	1.9 × 1.8	1		1	1
MRI-CGCM3	1.1 × 1.1	1	1	1	1
NorESM1-M	2.5 × 1.9	1	1	1	1
Total number of runs:		30	15	28	20

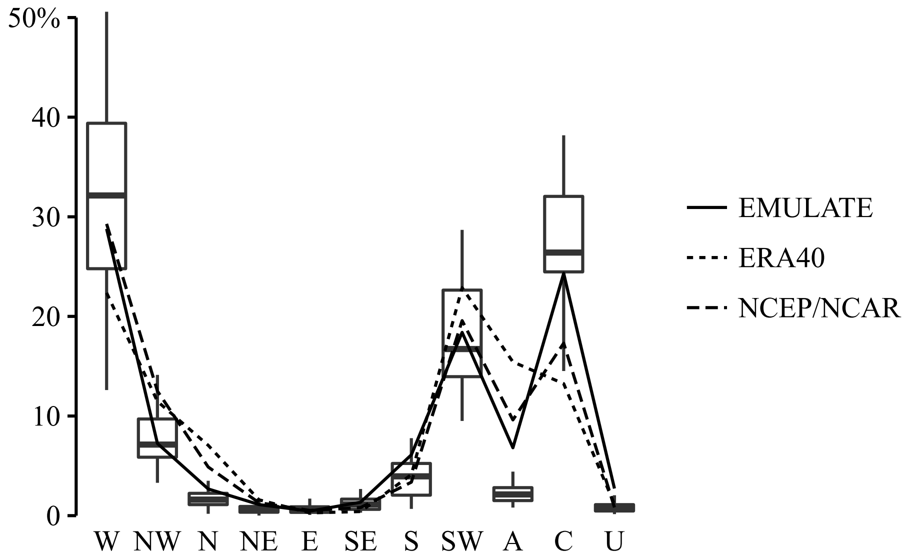

Figure A1. Relative winter precipitation accumulation per WT for different re-analysis datasets (lines) and climate model runs (boxplots). The results are obtained for the reference period 1961-1990.

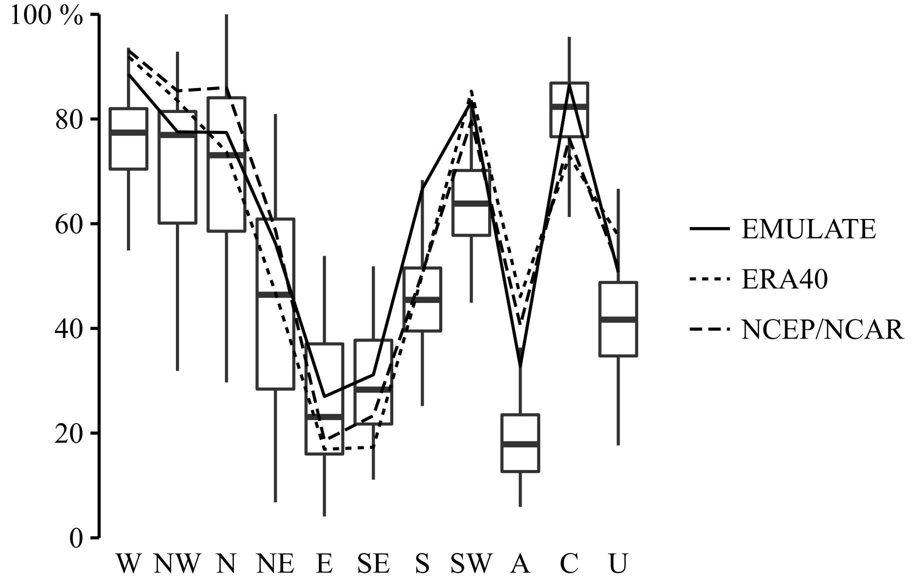

Figure A2. Percentage wet days per WT in the winter season for different re-analysis datasets (lines) and climate model runs (boxplots). The results are obtained for the reference period 1961-1990.

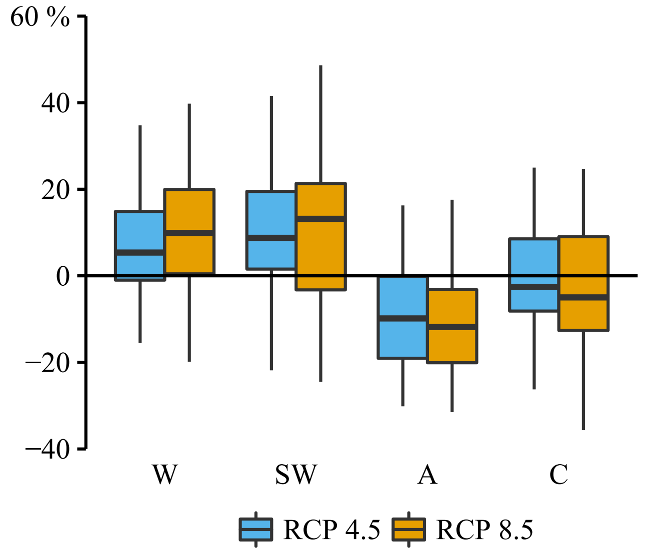

Figure A3. Changes in the winter WT occurrences for RCP 4.5 and RCP 8.5. The changes are obtained using the climate model ouput for 2071-2100 with respect to the output for 1961-1990.