Reply to referee 1: Review of "The Analogue Method for Precipitation Prediction: Finding Better Analogue Situations at a Sub-Daily Time Step." (HESS-2016-246) by Horton et al. 2016.

This is my first time reviewing this manuscript, though it is my understanding that it was previously reviewed by others, with major revisions requested. I have not studied the comments of other reviewers, with the desire to provide guidance as independent as possible.

My recommendation, with regret, is rejection. The article is needs significant reorganization and rethought, above what is required for a major revision. Some of my most significant issues include: (a) it provided detail of a minor improvement to a rather antiquated post-processing methodology; (b) it uses data sets for which there are better alternatives; (c) it doesn't describe all of the procedures clearly; (d) it doesn't consider other alternatives as controls against which to evaluate the methodology. In sum, I think the authors need to re-evaluate their research from top to bottom.
All the analyses were performed again with a more recent dataset. It allowed us to change the structure of the study and then to restructure the paper to gain in clarity. The procedures were then better explained. The other mentioned issues are discussed below.

Here are some more details on my most significant concerns.

Poor organization. Journal articles are generally minor variations on a standard organization, with an introduction, data, methods, results, conclusion. Here are some of the issues with such sections:
We improved the structure and simplified the overall logic of the study.

Introduction. There is now a rather rich body of literature on the statistical post-processing to produce probabilistic precipitation forecasts. These include Gamma-distribution fitting methods, Bayesian Model Averaging, Extended Logistic Regression, and more. With the use of an older analog approach, one wonders why such an approach is considered, in the first instance, and how one should place in context the results to follow. Without a thorough review of other possible alternatives and some explanation of why your approach is being considered despite these others, the reader is left wondering why they should bother with continuing to read the rest of the manuscript.
The present use of the analogue method is not a statistical post-processing of the precipitation. It is a statistical adaptation technique, which is classified as a downscaling approach (which can be a language abuse), as it provides a statistical precipitation prediction (or ev. forecast) based on large scale predictors describing the synoptic scale conditions. The precipitation output from the NWP model (or GCM, RCM, … according to the context) is not even considered in the method, as it would be in statistical post-processing methods. The goal of using an analogue method is to predict local precipitation based on the outputs of a global NWP model, without the use of a limited-area model. It brings an alternative approach that can complement the forecast provided by limited-area models in operational forecasting e.g., or the RCMs in the context of climate change impact studies.
There was a European project (COST VALUE, http://www.value-cost.eu/) that aimed at comparing different downscaling methods. The results of the project are not published yet, but show that there is not an overall best downscaling technique, as the different techniques have some advantages and weaknesses in different characteristics. The analogue method was found to be strong in some characteristics and weaker in others. However, such comparison is out of the scope of the present study.

As we show in the introduction, there are several articles published in the very last years using the analogue method, and we are aware of its use in different ongoing studies. Thus, the goal of the paper is more to provide an improvement to the users of the method. We refer to the paper of Maraun et al (2010) for alternative downscaling methods.

Data. What information was being used for the forecast was not made clear, and this is crucial information to find out right away. See (2a) below for more.
We added this information. By the way, the study takes place in a perfect prog context, not a forecasting context.

Methods. Section 2.2 and 2.6 seem to be describing two different methods. There should be one, single, clear, unambiguous description of the methodology to be used.
Section 2.2 describes the analogue method itself, and section 2.6 (now 2.3) how the parameters of the method are calibrated. We tried to make it more clear and restructured section 2.
Also in this section, while verification using CRPSS is relatively standard, so are other verification diagnostics like reliability diagrams, which are not presented.
Reliability diagrams were added (Figure 12).

Choice of data sets.

Forecast data. What is used as the forecast, and why, is not clear. Is another analysis from the NCEP-NCAR reanalysis time series used as a surrogate for a numerical forecast? Is ECMWF, or other model forecast data used? I could not tell. If NCEP-NCAR reanalysis data is used, this then begs the question: why? This would not be a practical forecast methodology, where one needs information in advance of the event. If a numerical weather prediction forecast is used, then there are potential issues of forecast bias; the forecast model and the analysis may be different in character, leading to the issue of whether perfect-prog (your analog) type approaches are suitable or whether more model-output statistics approaches are needed.
This study takes place in the perfect prog context. There is no forecast here, only prediction on an independent period (the validation period) using the reanalysis data. Then, when applied in real-time forecasting, NWP forecasts are considered. The issue of the biases are known and for this reason: (1) the model used for the NWP forecast and the one used for the reanalysis product should be as similar as possible, and (2) the main predictors are geopotential heights, which are robust and not too much model-dependent (this is more an issue for moisture variables). Moreover, the geopotential heights are compared in terms of their shape and not absolute value. Some MOS use of AMs also exist, provided that a long enough archive is available. In that context, the MTW could also be used and provide improvements. When applied in a climate context, AMs rely on GCM or ev. RCMs outputs. In this study, we stick to the perfect prog context and show an improvement that can then be applied to the forecasting context or to climate impact studies.

Reanalysis data. There are more modern, more accurate reanalysis data sets available now such as ERA-20C, available at higher temporal resolution (3 hourly), which seems to be crucial for an article examining the usefulness of temporal shifts of the data. You dismiss this in your section 4.6 with an older reference, but I think given the focus in this article on temporal shifts, you need to reconsider higher temporal resolution reanalysis data.
You are right that it is important considering a 3 hourly dataset. We have done all analyzes again using ERA-20C and MERRA-2 reanalysis datasets. Doing so allowed us to consider another workflow and to simplify the analyses and then the paper.

Observation data. While the geographic details of the observation locations are definitely different from many other locations in Europe, there are still other locations in the mountains. Why not consider approaches that supplement the training data with other locations' observations, potentially allowing you to get more without needing a very lengthy reanalysis? You may have objections to this, but at the least it would be worth explaining your choices.
We are not sure to get your point. The AM exploit the intrinsic link between the synoptic-scale situation and the local precipitation. This relationship is very location-specific, as the meteorological influences will not be the same when considering a mountainous environment in another country. The location and characteristics of the driving elements, such as the low pressure centers and the fronts, will not be the same, neither the characteristics of the precipitation climatology. We can therefore not exploit remote data.

Methodology. Finding some other reference methodology, e.g., Bayesian Model Averaging, would certainly be desirable. Another logical control would be unadjusted ensemble forecasts from the ECMWF ensemble prediction system. This would allow you to have a point of comparison against which to judge the analog methodology.
As stated previously, the study does not take place in a forecast context, and thus cannot be compared to ensemble forecasts provided by NWP models.

But even considering the analog methodology in isolation, one wonders why you chose the particular approach to selecting analogs. In particular, I found myself wondering why, for example, you didn't use canonical correlation analysis approaches to determine what information in the reanalysis data set was most directly related to precipitation variability. Analog approaches, in my experience, are not very "efficient" with their training data; either a previous day's data is selected as an analog date, or not. This, then, drops on the floor all the other data, which may yet have some useful information. To use an analog approach, then, it seems especially incumbent to have demonstrated that you have chosen the most important predictors that will be used in selecting past dates.
Your version of the analogue method seems different to what we use here, so it is not easy to answer specifically to your points. However, here are some response elements:

- The choice of predictors is based on previous work (given as references) that made an intensive comparison of many variables. Moreover, the relevance of the geopotential heights and the moisture variables was shown by Horton (2012) to be the most relevant variables for the region of interest.
- A canonical correlation analysis cannot be performed here, as the geopotential heights are analyzed in terms of shape (Teweless-Wobus criteria) over a spatial window rather than absolute values.
- In the AM considered, the selection of analogue dates are not allowed within the same year as the target date. Thus, a selection of the previous date as analogue is not allowed.

In summary, I regret recommending rejection. I hope the authors will constructively use this feedback in the spirit intended.

Reply to referee 2: Review comments for “The analogue method for precipitation prediction: finding better analogues situations at a sub-daily time step” by Horton et al.

Recommendation: Major revision

The authors introduced a moving time window (MTW) for the analogue method so that better analogues at a different hour can be found for precipitation prediction in contrast to the use of analogues at fixed hours of the day in standard analogue method. They found that the MTW with the shorter archive on a sub-daily time step improved the analogy criterion values across the entire distribution of analogue dates and the skill of precipitation prediction in comparison with the standard analogue method with longer archive on a daily time step. In particular, the improvement in prediction skill is greater for days with heavy precipitation. The topic is important and has great implications for operational precipitation forecasting and impact studies associated with the hydrological community. The only constraint is that the implementation of such method requires the availability of sub-daily time series, which may not always exist.

I have several major comments. First, some necessary information regarding the the presented analyses should be provided. For example, what season are the results shown in Figures 4, 5,6,7,8, 10 based on?
We added a description of the calibration and validation periods. The results are for all seasons over the calibration or validation periods, covering several years.
The authors mentioned in Table 1 that the selection of analogue candidate is limited to the 4 months centered around the target date for every year. However, it is not clear what season the presented analyses focused on.
As explained above, the analyses are for all seasons over several years. It is true that the search for analogues is restricted to the 4 months centered around the target date, but the target dates cover all days over several years.
Also, it seems to me that the entire assessment is performed in the prognosis context. The authors mentioned “prediction” several times throughout the paper. No matter for a 47-year archive (1961-2008) or reduced 25-year archive (1982-2007), it is not clear if the authors used part of the archive for calibration and part of the independent period for validation. If it is real “prediction”, what period of data is the prediction performed on? All these details should be clearly described in the method section.
We now clarified the use of independent data in the section “2.3 Calibration of the analogue method”. It is indeed a prediction over a validation period, but still in the perfect prognosis framework, rather than a forecast.
Second, the paper, especially the results and discussion sections, is not well structured. These sections are divided into many small sub-sections. The content should be better organized and integrated to convey clear message. One example is, the discussion of Figure 4 and 5 appears in both section 3.1 and 3.3.
All the analyses were performed again with a more recent dataset. It allowed us to change the workflow of the study and then to restructure the paper to gain in clarity. Some results of secondary importance were moved to the discussion.
Third, the text needs to be improved in terms of the logic, transition, grammar and wording. Some sentences are really long, confusing, and quite hard to understand (see some specific comments below).

We tried to improve the language and better explain some points. The paper was corrected by Elsevier's English editing services.

Specific comments:

1. P1, line 6-7: confusing sentence, how about "the main reason for the use off daily precipitation time series is the length of their available archives, …
   This was changed.
2. P1, Line 7-9: "However, it is … at a different time of day". Long and confusing sentence. should rephrase it.
   This was changed.
3. P2, Line 22-23: "since they are based on observed situations with consistent spatial distribution" – consistent with what? Do you mean between target day and analogue dates? "as long as the analogue dates chosen for a region are the same" – same compared to what? When the target day changes, I think the analogue dates will change accordingly.
   This sentence was removed as it is a bit out of context.
4. P3, Line 2: "even for much higher orders of magnitude" – do you mean even longer archive?
   This was changed.
5. P3, Line 2-4: "Hopefully" – better to use "fortunately" based on context. Also, need reference for the statement "it appears that … 10° to 20°".
   This was changed and references added.
6. P3, Line 16-19: "Therefore, if the reduction of the archive … to an increase in performance". – very confusing sentence, please consider rephrasing it.
   This was removed.
7. P3, line 25: "similar conclusions" – what is the conclusions? - in creasing the prediction skill?
   This was removed.
8. P5, line 11-12: why MTW can not be applied to the 2nd level of analogy?
   Yes it can be applied to the 2nd level of analogy. Thus, this comment was removed.
9. P.6, line 9 for Figure 2: why not just keep candidate 24-h precipitation fixed from 6h to 30h, but allowed to choose the analogues on 6h, 12h, 18h, 24h, 30 h for both Z500 and Z1000? That allows you to choose the analogues on multiple time steps but within the 24-h window consistent with conventional method. What is the purpose to have the varying 24h precipitation totals if the main objective is to find the better analogues to predict the same target day precipitation?
   This section was rephrased for clarity.
10. P6, Line13: confusing sentence "no constraint … in order to restrict."
    This was removed.
11. P6, Line 27- 33, it is not clear how the method is implemented. The authors should provide a diagram to show the method. More details are preferred, such as do you just pick one best grid among four, what time lapse is allowed, how the temporal profile of best proxy is used to disaggregate? If you use the proxy variables from NCEP/NCAR reanalyses, why not directly use the precipitation from NCEP/NCAR reanalyses?
    The importance of this analysis was decreased and it moved to the discussion (Sect. 4.4), without providing all the details. It is indeed not the main message of the paper and brought some confusion. Some additional details on the time lapse and the non-consideration of precipitation were however added.

12. P7, line 13: Is the four points for geopotential height used to calculate the height gradient in both directions?
Yes. It has been clarified.
13. P8, Line 9: what does "globally significant" mean? Significant at what level?
This is not present anymore in the paper.
14. P8, section 3.2.1: It is not clear to me how the distribution of the analogy criterion for different analogue ranks is constructed. So for any target day, if 50 analogue dates are selected (50 ranks in total), each analogue date should have only one S1 value based on their similarity in geopotential fields.
Yes, each analogue date has a unique S1 value. These distributions for the different analogue ranks are obtained when applying the method on a long period. Then, for a considered analogue rank, we have multiple values of S1 corresponding to different target dates. We tried to clarify in the manuscript.
15. P9, line 6: "the number of candidate situations did not increase", but from table 3, N1 for 2Z-2MI is larger than N1 for 2Z.
Yes, but here it was with the original parameters, so without an increase in N1. It has been clarified.
16. P9, Line 10-11: could this because RMSE is not a good metric to assess the similarity for moisture fields?
No, it is because the use of an MTW does not increase the sample size in this case, the second level of analogy only subsamples in the dates provided by the first level. A comment was added.
17. P9, Line 13: why it is "prediction"? I think the entire assessment so far is in a prognosis context. Do you reselect the analogue dates for blue bar (MTW algorithm) in Fig. 4 and 5?
We use the term "prediction" instead of "forecast" because it is not operational forecast, but we stay in the perfect prog context. Predictions were established for a validation period with independent data. Fig. 4 and 5 do not exist anymore and were replaced by Fig. 9 and 10 that are different.
18. P9, Line 15: It will be good to test if the improvements of MTW and MTW-r over the static approach is significant?
The approach changed and the calculation are now done on a calibration period and an independent validation period in order to validate the potential gains.
19. P9, section 3.3.1: Fig.11 also indicates that the spread of difference of the CRPSS performance score is quite larger. It is not correct to say that the performance score was improved for days with high precipitation. The statement should be based on the average performance. Again, for Fig.11, it is not clear to me what each point represents. Do the points represent the analogue dates with precipitation amount in the specific categories? Then the total number of pints in figures are equal to the total number of analogues selected?
We improved the description of the analysis. It is indeed a difference on the scores between the conventional approach and the MTW. All points correspond to target dates over the whole period. It means e.g. that over the whole period, the prediction performance of most days with a target value above 30mm/d was improved. The mean is represented by the star and then shows an average improvement.
20. P9, section 3.3 and 3.3.1: when author say "prediction skill", does the author mean the use of calibrated parameters for independent data set?
Now we do use an independent dataset (the validation period, VP).

21. Same as #10, it would be good to show a map about the method 2 to help the reader understand what is concluded in section 3.4 and table 5.
As we wrote above, these results were removed to focus on the main results of the study.

List of relevant changes

- There is now a distinct calibration period and an independent validation period.
- All the analyses were performed again with more recent datasets: MERRA-2 and ERA-20C, which allowed a 3-h time step.
- The influence of the MTW time step (3, 6, 12 h) is analyzed throughout the paper.
- Reliability diagrams were added.
- The structure of the study changed, which resulted in a better organized paper gaining in clarity.
- The analyses of secondary importance were removed to focus more on the main results of the study.
- All the text has gone through significant editing, as one can see in the marked manuscript.

The analogue method for precipitation prediction: finding better analogue situations at a sub-daily time step

Pascal Horton[1,2], Charles Obled[3], and Michel Jaboyedoff[1]

[1]Institute of Earth Sciences, University of Lausanne, Lausanne, Switzerland
[2]Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Bern, Switzerland
[3]Laboratoire d'étude des Transferts en Hydrologie et Environnement (LTHE), Université de Grenoble-Alpes, Grenoble, France

Correspondence to: Pascal Horton (pascal.horton@alumnil.unil.ch)

Abstract. Analogue methods (AMs) predict local weather variables (predictands) ~~,~~ such as precipitation ~~,~~ by means of a statistical relationship with predictors at a synoptic scale. The analogy is generally assessed on gradients of geopotential heights first ~~, in order~~ to sample days with a similar atmospheric circulation. Other predictors ~~,~~ such as moisture variables ~~,~~ can also be added in a successive level of analogy.

The search for candidate situations ~~for~~ similar to a given target day is usually undertaken by comparing the state of the atmosphere at fixed hours of the day for both the target day and the candidate analogues. ~~The main reason is the use of~~ This is a consequence of using standard daily precipitation time series~~due to the length of their available archives, and the unavailability of equivalent archives at a finer time step~~, which are available over longer periods than sub-daily data. However, it is unlikely for the best analogy to occur at the ~~very same hour , while it~~ exact same hour for the target and candidate situations. A better analogue situation may be found with a time shift of ~~some hours as it~~ several hours since a better fit can occur at ~~a different time of~~ different times of the day. In order to assess the potential for finding better analogues at a different hour, a moving time window (MTW) has been introduced.

The MTW resulted in a better analogy in terms of the atmospheric circulation ~~, with~~ and showed improved values of the analogy criterion on the entire distribution of the extracted analogue dates. The improvement was found to ~~grow~~ increase with the analogue rank ~~due~~ owing to an accumulation of better analogues in the selection. A seasonal effect has also been identified, with larger improvements shown in winter than in summer~~, supposedly due~~ . This may be attributed to stronger diurnal cycles in summer that favour predictors taken at the same hour for the target and analogue days.

The impact of the MTW on the precipitation prediction skill has been assessed by means of a sub-daily precipitation series transformed into moving ~~24h totals at 6h~~ 24 h totals at 12-h, 6-h, and 3-h time steps. The prediction skill was ~~found to have~~ improved by the MTW, ~~and even to a greater extent after recalibrating the AM parameters~~as was the reliability of the prediction. Moreover, the ~~improvement was~~ improvements were greater for days with heavy precipitation, which are generally related to more dynamic atmospheric situations ~~where~~ in which the timing is more specific and ~~which are fewer~~ fewer records are available in the meteorological archive.

~~However, in order to produce quantitative precipitation predictions the MTW requires sub-daily precipitation time series, which are generally available for a shorter period than daily precipitation archives. Therefore, two simple temporal disaggregation~~

~~methods were assessed in order to make longer archives usable with the MTW. The assessed approaches were not successful, emphasizing the need to use time series with an appropriate chronology. These should be available in the near future, either by means of growing archives of measurements or by the establishment of regional precipitation reanalysis data at sub-daily time step.~~

~~The use of the MTW in the AM can already be considered now~~ The improvements of the analogy criterion and the performance scores on precipitation were both found to be higher for MTWs with a smaller time step of 3 h. A 3-h MTW provides eight times more candidate situations even though they are not fully independent. Since the MTW provides additional situations to the pool of possible analogues, it can be considered as an inflation of the meteorological archive. Because this technique is very simple and easily applicable, it should be considered for several applications in different contexts, ~~may it be for~~ such as operational forecasting or climate-related studies.

1 Introduction

~~Analogue methods (AMs) are~~ The analogue method (AM) is a statistical downscaling method based on the hypothesis that two relatively similar synoptic situations may produce similar local effects (Lorenz, 1956, 1969). ~~They are~~ Other statistical downscaling methods ~~(Maraun et al., 2010) and consist~~ for climate studies can be found in Maraun et al. (2010). The AM consists of finding past situations that are similar to the target day of interest in terms of ~~the~~ atmospheric circulation or other synoptic predictors. ~~The~~ Referred to as the predictand, the local weather variables of interest ~~(predictand) that~~ which were observed at the analogue dates are used to construct a probabilistic prediction for the target day (Duband, 1970; Zorita and Storch, 1999). Multiple variations of ~~the method exist (a non-exhaustive listing can be found in Ben Daoud et al., 2016). The~~ this method have been reported (Ben Daoud et al., 2016). AMs are used for operational precipitation forecasting ~~, either~~ in the context of weather forecasting, flood forecasting, ~~or hydropower production (e.g., Guilbaud, 1997; Bontron and Obled, 2005; Hamill and Whitaker, 2006; Desaint~~ hydropower production (e.g., Bontron and Obled, 2005; Hamill and Whitaker, 2006; Desaint et al., 2008; García Hernández et al., 2009; B as well as for precipitation downscaling from a climate perspective (e.g. Radanovics et al., 2013; Chardon et al., 2014; Dayon et al., 2015). ~~Other applications focus on temperatures (Radinovic, 1975; Woodcock, 1980; Kruizinga and Murphy, 1983; Delle Monache e wind (Gordon, 1987; Delle Monache et al., 2013, 2011; Vanvyve et al., 2015; Alessandrini et al., 2015b; Junk et al., 2015b, a), solar power (Alessandrini et al., 2015a; Bessa et al., 2015), snow avalanches (??), insolation (?), and the trajectories of tropical cyclones (???).~~

~~The spatial transferability of the method is analysed in Chardon et al. (2014) and Radanovics et al. (2013). A great advantage of AMs is that they create realistic precipitation patterns, since they are based on observed situations with consistent spatial distribution, as long as the analogue dates chosen for a region are the same. Moreover, they~~ AMs can provide multivariate predictions ~~that~~ which are physically consistent (Raynaud et al., 2016). Their spatial transferability is analysed in Chardon et al. (2014) and Radanovics et al. (2013), and their temporal transferability has ~~also been the topic of~~ been discussed in recent studies for past or future climates (Dayon et al., 2015; Caillouet et al., 2016).

The method requires two different archives. The first is a meteorological archive describing the state of the atmosphere at a synoptic scale, such as reanalysis datasets, at different hours of the day. The second is an archive of the target variable to be forecast, here precipitation. This archive is made of precipitation cumulated over a certain time duration, most often daily but sometimes sub-daily, predicted, which is usually standard daily precipitation totals from 06:00 UTC to 06:00 UTC the following day either at a target station or integrated over a target catchment. Obviously, the period to be used is limited to the smallest period common to the two archives.

Ruosteenoja (1988) and Van Den Dool (1994) have analysed the influence of the length of the meteorological archive on the quality of the analogy. They highlighted a three-way relationship between among the quality of the analogy, the archive length, and the size of the spatial domain (or degrees of freedom) : and determined that errors increase with a bigger domain , larger domain but decrease with a longer archive. They demonstrated that it is not possible to find good analogues over the whole northern hemisphere with a 100 year archive (and even for much higher orders of magnitude). Hopefully, it appears that for the aim of predicting precipitation over point stations or catchments of some 100 or 1000 km², there is no need to consider meteorological domains larger than 10° to 20°. For that reason, smaller limited spatial windows are always considered when searching for analogues, and the archive length is maximizedmaximised.

Therefore, due owing to the availability of long precipitation archives at a daily time step that have no equivalent at a finer resolution, AMs are usually implemented on a daily basis. Consequently, the Given the cumulative aspect of the predictand, the corresponding meteorological situation is characterized by several predictors taken at different but fixed hours of the day, and the analogue situations are assessed by comparing predictors at these same predictors at the same fixed hours of the day. However, it can be expected that the best analogy of the synoptic situations does not occur systematically at the same time of the day and that better candidates can be found by shifting to a different hourdifferent hours. With this assumption, a moving time window (MTW) was introduced to allow the search for candidates at different hours of the day. Previous tests have showed the benefit, in terms of analogy criterion values, of searching for analogue synoptic situations at a finer time step, but without assessing although such research did not assess the impact on the prediction skill (Finet et al., 2008). of a specific predictand (Finet et al., 2008).

The MTW finds analogue situations at different hours of the day, which can also be seen as an inflation of the archive length. However, despite having x time more candidate situations, the quantity of new information is not expected to be as important as a x time longer archive due to significant correlation between successive situations within the same day. Nevertheless, if the MTW can improve the prediction skill of the AM, it means that it does extract new information from the archive. Therefore, if the reduction of the archive length needed by the MTW, due to the reduced availability of a sub-daily precipitation time series, is expected to decrease the AM performance, the archive inflation brought by the MTW is expected to contribute to an increase in performance.

Other possibilities exist for increasing the prediction skill of the AMs. A classical approach is to add new predictors or new successive levels of analogy (see e.g. Horton, 2012; Ben Daoud et al., 2016; Caillouet et al., 2016)(e.g. Horton, 2012; Ben Daoud et al., 20 AMs can also be combined with other methods (see e.g. Chardon et al., 2014)(e.g. Chardon et al., 2014). Another possibility is to use a global optimization technique, such as genetic algorithms, in order to better optimize the method and to add new parameters ~~(?)~~(Horton et al., 2017). However, the MTW technique is not in competition with other approaches and can be combined with ~~these. Indeed, as most of themrely on the atmospheric circulation in the first level of analogy, the application of a MTW should lead to similar conclusions~~them.

The outline of this paper is as follows. Section 2 presents the context of the study as well as the data and methods, including the proposed MTW technique. The ~~impact of the reduction of the archive and the improvements brought by the MTW are assessed in Sect. 3. The~~ benefits of introducing ~~a MTW were assessed first in regards to the analogy criterion improvement between synoptic situations~~ an MTW are assessed first regarding the improvement of the analogue date selection (Sect. ~~??~~3.1) and then in terms of precipitation prediction skill (Sect. 3.2). Finally, the results are discussed in Sect. 4, and the conclusions are found in Sect. 5. A list of the acronyms and their definitions is provided in appendix A.

2 Data and methods

2.1 Study area and data

The study area is the upper Rhône catchment in Switzerland. ~~Precipitation time series come~~ The precipitation time series were obtained from six automatic weather stations, ~~viz.,~~ Ulrichen, Zermatt, Visp, Montana, Sion, and Aigle (Fig. 1), that are subject to various meteorological influences (Horton et al., 2012). The data were available at an hourly time step for ~~25~~ 29 years (1982–~~2007~~2010) and were also obtained at a standard daily time step ~~(from 6~~of 06:00 ~~UTC to 6~~UTC to 06:00 ~~UTC the next day) for 47~~ UTC the following day for 50 years (1961–~~2008). Due~~ 2010). Owing to the low density of weather stations with high temporal resolution and long archives, no spatially aggregated rainfall was processed. ~~The results will hereafter be presented arbitrarily for the Ulrichen station but are equivalent for all stations.~~

Synoptic-scale variables ~~,~~ used as predictors ~~,~~ were extracted from ~~the NCEP/NCAR reanalysis~~ two of the most recent global reanalysis datasets: the European Center for Medium Range Weather Forecasting 20th century reanalysis (ERA-20C) (Poli et al., 2016), with 3-h temporal resolution and a spatial resolution of 1~~dataset (Kalnay et al., 1996) with a 6h temporal resolution , 17 pressure levels,~~°, and ~~a~~ Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), from the National Aeronautics and Space Administration (NASA), with a 6-h temporal resolution and a spatial resolution of ~~2.5°. This dataset is now relatively old,~~ 0.625°× 0.5°. MERRA-2 is an update for the first MERRA reanalysis (Rienecker et al., 2011). The study was originally performed on the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis 1 dataset (Kalnay et al., 1996); the conclusions were similar. ERA-20C is built by assimilating only surface observations and is thus available for a long period (1900–2010). On the contrary, MERRA-2 is built by assimilating observations at higher levels and from more sensors including satellite data, but it is ~~not expected to affect the conclusions of the present study (see discussion in Sect. ??)~~ more limited in time (1980–present). It is of interest to assess the relevance of an MTW with these two datasets because they represent different types of products. The specific advantages of ERA-20C are that it allows for testing an MTW with a 3-h time step, and it covers a long period. On the contrary, MERRA-2 has a higher spatial resolution and can be expected to be more accurate at higher levels of the atmosphere. The variables extracted from these datasets were geopotential heights at 500 hPa (Z500) and 1000 hPa (Z1000), the total precipitable water (TPW), and the relative humidity at 850 hPa (RH850).

2.2 The considered analogue methodmethods

The first considered AM is based on the analogy of the atmospheric circulation only (Table 1, Obled et al., 2002; Bontron and Obled, 2005). Searching (Table 1; Obled et al., 2002; Bontron and Obled, 2005). Before searching for analogue situations to for a target daystarts by a seasonal stratification through a preselection step , seasonal stratification is conducted through preselection of the potential candidates for analogy. The restriction is a search for search is restricted to analogue days within a 4-month four-month window centred on the target date for every year of the archive. The similarity of the atmospheric circulation of the target date with every day of the preselection is assessed by processing the Teweles and Wobus (1954) score (S1)that , which is a comparison of gradients on of geopotential heights over a certain particular spatial window and at certain hours:

$$S_1 = 100 \frac{\sum_{i}^{m} |\Delta\hat{z}_i - \Delta z_i|}{\sum_{i}^{m} max(|\Delta\hat{z}_i|, |\Delta z_i|)}, \tag{1}$$

where $\Delta\hat{z}_i$ is the forecast geopotential height difference between the ith pair of adjacent points from the grid of the target situation, Δz_i is the corresponding observed geopotential height difference in the candidate situation, and m is the number of pairs of adjacent points in the grid. The differences are processed separately in both directions. With smaller S1 values, there is greater similarity in the pressure fields.

The predictor variables extracted from reanalysis datasets are considered at different hours of the day. Based on Bontron and Obled (2005), geopotential heights are compared at 1000 -hPa (Z1000) at 12:00 -UTC and UTC and at 500 -hPa (Z500) at 24:00 -UTC. The time of the day at which the predictors are selected is found by Bontron (2004) to have has a significant influence (Bontron, 2004).

Then, N_1 dates with the lowest values of S1 are considered as analogues to the target day, where N_1 being a parameter to calibrate (see needs to be calibrated (Sect. 2.3). Finally, the daily observed precipitation amount of the corresponding dates provide the empirical conditional distribution considered as the probabilistic prediction for the target day. This method will be named is referred to as 2Z.

The second reference method (2Z-2MI, Table 2) (2Z-2MI, Table 2; Bontron and Obled, 2005) adds a subsequent level of analogy with moisture variables, which are compared by means of the root-mean-square root–mean–square error (RMSE) criteria:

$$E_{RMS} C_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{v}_i - v_i)^2}, \tag{2}$$

where $\hat{v}_i$ is the ith predictor value from the grid of the target situation, v_i is the corresponding observed value in the candidate situation, and n is the number of points in the grid.

The additional predictor is a moisture index composed of the product of the total precipitable water (TPW) with the relative humidity at 850 ~~-~~hPa (RH850) at 12:00 UTC and 24:00 UTC (Bontron, 2004). When adding a second level of analogy, N_2 dates are subsampled from the N_1 analogues on the atmospheric circulation, resulting in a smaller number of analogue situations. ~~When~~ Moreover, when a second level of analogy is added, a higher number of N_1 analogues is kept on the first level.

More complex AMs exist with additional predictors ~~(see e.g. Horton, 2012; Ben Daoud et al., 2016; Caillouet et al., 2016)~~(e.g. Horton, 2 The MTW can also be applied to these~~, as they generally rely on a similarity of the atmospheric circulation in the first level of analogy. However,~~ ; however, it is easier to interpret the impact of the MTW by using more basic methods.

2.3 ~~Performance score~~Calibration of the analogue method

~~In order to assess the performance of AMs , the continuous ranked probability score (CRPS, Brown, 1974; Matheson and Winkler, 1976; He often employed (see, e.g., Bontron, 2004; Bontron and Obled, 2005; Ben Daoud et al., 2008; Horton et al., 2012; Marty et al., 2012; Radar It allows the evaluation of the predicted cumulative distribution functions $F(y)$ of the precipitation values y from analogue situations compared to the observed value y^0. A better prediction has a smaller score. It is defined as follows:~~

$$S_{\text{CRP}} = \int_{-\infty}^{+\infty} \left[F(y) - \text{H}(y - y^0)\right]^2 \mathrm{d}y,$$

~~where $\text{H}(y - y^0)$ is the Heaviside function that is null when $y - y^0 < 0$ and has the value of 1 otherwise.~~

~~In order to compare the value of the score in regard to a reference, one often considers its skill score expression and uses the climatological distribution (i.e., the distribution of all precipitation values from the corresponding archive) as the reference. The continuous ranked probability skill score (CRPSS) is defined as follows:~~

$$S_{\text{SCRP}} = \frac{S_{\text{CRP}} - S^r_{\text{CRP}}}{S^p_{\text{CRP}} - S^r_{\text{CRP}}} = 1 - \frac{S_{\text{CRP}}}{S^r_{\text{CRP}}}$$

~~where S^r_{CRP} is the S_{CRP} value for the reference (climatological distribution) and S^p_{CRP} is for a perfect prediction (which implies $S^p_{\text{CRP}} = 0$). A better prediction is characterized by an increase in CRPSS: $S_{\text{SCRP}} = 1$ for a perfect prediction and $S_{\text{SCRP}} < 0$ for a prediction with a lower skill than the reference.~~

2.4 ~~The moving time window (MTW) approach~~

~~The moving time window (MTW) technique aims at finding better analogue situations at different hours~~ AMs rely on parameters that need to be defined for every level of analogy. Here, the selection of the meteorological variables used as predictors, as well as their corresponding pressure level and hour of the day~~rather than comparing the predictors at the same fixed hours. If the~~

~~target situation is kept at 12:00~~ , were considered to be identical to those used in methods 2Z (Table 1) and ~~24:00 UTC for Z1000 and Z500 respectively, candidate situations are not only considered at 12:00 and 24:00 UTC, but at other hours by allowing a time shift. Therefore, instead of looking for analogues at a 24h time step, they are sought at the time step matching the predictor temporal resolution which is a 6h time step in this study (Fig. 2) .~~ 2Z-2MI (Table 2) described above. The parameters calibrated in this study are listed below.

~~The target situations and their corresponding observed precipitation values do not change because the prediction is still established on a daily basis for a fixed period of the target day (6:00 UTC to 6:00 UTC the next day) as before. The difference is that the candidates are 4 times as many (even though they are not fully independent) as in the conventional approach. No constraint was added in order to restrict the selection of multiple analogues within the same candidate dates.~~

- Spatial windows, which are the domains in which the predictors are compared. A spatial window is specific to each level of analogy; thus, its extent differs among circulation and moisture variables.
- The optimal number of analogue situations for every level of analogy.

~~In order to assess the benefit of searching for analogue situations at a sub-daily time step for quantitative precipitation prediction, an appropriate precipitation series is required. On the basis of high temporal resolution time series (Sect. 2.1), 24h totals were processed at a 6h time step by means of a moving 24h total. Since sub-daily precipitation time series are available on a shorter period than daily ones, the loss of performance resulting from a reduction of a 47 year archive (1961–2008) to 25 years (1982–2007) can be expected. However, the competition between the loss of performance due to a smaller archive length and the potentially better analogy between target and MTW candidates needs to be assessed.~~

2.4 ~~Reconstruction of a longer precipitation archive~~

~~It would be profitable to be able to apply the MTW to a longer archive (here, 1961–2008), rather than being limited to the reduced period where the sub-daily precipitation data are available (cf. previous section). Therefore, in order to reconstruct a longer archive of moving 24h totals, two simple disaggregation approaches of the daily precipitation time series were assessed.~~

~~The first technique was a proportional distribution, where the observed daily precipitations were considered constant during the measuring period (6:00 UTC to 6:00 UTC the next day). Proportional parts of the original daily time series were allocated into a new moving average of 24h totals (Fig. **??**).~~

~~The second approach aimed at getting closer to the chronology of the actual precipitation by relying on some informative variables during the reconstruction procedure. Data from the NCEP/NCAR reanalysis 1 (Sect. 2.1) were used for this purpose, despite their rough resolution (2.5°). Precipitable water and relative humidity (at 1000 hPa, 925 hPa or 850 hPa) were assessed at the four points surrounding the catchment, and a time lapse between both series was allowed to take into account the significant distance separating the weather stations and the reanalysis grid point. The best proxy variable was identified by means of a correlation analyses (on non-zero values) with the 6-hourly precipitation time series over the period 1982–2007.~~

~~Once the best proxy had been selected, its temporal profile was used in order to disaggregate the daily precipitation time series.~~

2.4 ~~Calibration of the analogue method~~

The semi-automatic sequential procedure ~~elaborated~~ discussed by Bontron (2004) was used to calibrate the AM ~~(see also Radanovics et al.,~~ The analogy levels (e.g. the atmospheric circulation or moisture index) are calibrated sequentially. The ~~parameters calibrated by this approach, for every level of analogy, are the spatial windows on which the predictors are compared and the number of analogues. The~~ procedure, as defined by Bontron (2004), consists of the following steps:

1. ~~Manual selection of the following parameters:~~

   (a) ~~meteorological variable (e.g. geopotential height, temperature, relative humidity, etc.),~~

   (b) ~~pressure level (e.g. 500 hPa, ...),~~

   (c) ~~temporal window (hour of the day – e.g. 6:00 or 12:00 UTC),~~

   (d) ~~initial analogue numbers (e.g. $N_1 = 50$).~~

2. For every level of analogy ~~:~~

   (a) Identification of the most skilled unitary cell ~~(one point for moisture variables and four for geopotential heights when using the S1 criterion)~~ of the predictor data over a large domain. Every point ~~(or cell)~~ or cell of the full domain is ~~assessed on all~~ jointly assessed on the predictors of the current level of analogy~~, jointly (consisting generally of the same variable but on different pressure levels and at different hours)~~.

   (b) From this most skilled point, the spatial window is expanded by successive iterations in the direction of greater performance gain. The spatial window ~~grows by repeating the previous steps~~ increases until no improvement is reached.

   (c) The number of analogue situations N_1, which was initially set to an arbitrary value, is then reconsidered and ~~optimized~~ optimised for the current level of analogy.

3. A new level of analogy can then be added ~~,~~ based on other variables ~~(~~such as the moisture index ~~)~~ at chosen pressure levels and hours of the day. The number of analogues for the next level of analogy, N_2, is initiated at a chosen value. The procedure starts again from step ~~2~~ 1 (calibration of the spatial window and the number of analogues) for the new level. The parameters calibrated on the previous ~~analogue~~ analogy levels are fixed and do not change.

4. Finally, the numbers of analogues N_1 and N_2 for the different levels of analogy are reassessed. This is ~~done~~ performed iteratively by varying the number of analogues of each level in a systematic ~~way~~manner.

The calibration is ~~done~~ performed in successive steps ~~with a limited number of parameters. Previously~~ and aims at minimizing the objective function (CRPS, Eq. 3). Except for the number of analogues, previously calibrated parameters are generally not re-assessed~~(except for the number of analogues)~~. More advanced techniques ~~,~~ such as using genetic algorithms ~~(?),~~ (Horton et al., 2017) exist but are ~~out of~~ beyond the scope of the present study.

3 ~~Results~~

The 29-year hourly precipitation dataset was divided into a calibration period (CP) and a validation period (VP) in order to assess the robustness of the proposed improvements on independent data. The selection of the VP was evenly distributed over the entire series (Ben Daoud, 2010) to reduce potential biases related to trends linked to climate change or to evolution in the measurement techniques. Thus, 1 out of every 5 years was selected for validation, which represents a total of 6 years for the VP and 23 years for the CP.

2.1 ~~Consequences of the archive reduction~~

~~The performance loss resulting from the reduction of a 47 year archive (1961–2008, corresponding to the daily time series) to 25 years (1982–2007, corresponding to the hourly time series in Sect. 2.1)was assessed with the original method without MTW. Both 2Z (Table 1)and 2Z-2MI (Table 2)methods were considered.~~ The ~~AM parameters were calibrated on the original archive~~ continuous ranked probability score (CRPS, Brown, 1974; Matheson and Winkler, 1976; Hersbach, 2000) is often employed to assess the performance of AMs (e.g. Bontron, 2004; Bontron and Obled, 2005; Ben Daoud et al., 2008; Horton et al., 2012; Marty et al., This allows for evaluation of the predicted cumulative distribution functions $F(y)$ of the precipitation values y from analogue situations compared with the observed value y^0; a better prediction has a smaller score. This score is defined as

$$S_{\text{CRPS}} = \int_{-\infty}^{+\infty} \left[F(y) - \mathrm{H}(y - y^0)\right]^2 \mathrm{d}y, \quad (3)$$

where $\mathrm{H}(y - y^0)$ is the Heaviside function, which is null when $y - y^0 < 0$ and has the value of 1 otherwise.

2.1 The moving time window approach

The MTW technique is used to find better analogue situations at different hours of the day rather than comparing the predictors at the same fixed hours. The target situation (the day to predict) is still the same as that in the conventional approach: a daily precipitation total between 06:00 and 30:00 UTC characterised by predictors, which in this case are Z1000 and Z500 at 12:00 and 24:00 UTC, respectively. The difference is that candidate situations are considered not only at the same time (~~see resulting analogue numbers in Table ??)and were used thereafter.~~ 12:00 and 24:00 UTC), but also at other hours by allowing a time shift. Therefore, instead of looking for analogues at a 24-h time step, they are sought at the time step matching the predictor temporal resolution (Fig. 2). The ERA-20C dataset was used here to test an MTW with 12-h, 6-h, or 3-h time steps, which are referred to as 12-h MTW, 6-h MTW, and 3-hr MTW, respectively. Thus, the candidates are two (12-h MTW), four (6-h MTW), or eight (3-h MTW) times as many as in the conventional approach even though they are not fully independent.

The target situations and their corresponding observed precipitation values do not change. The predictions are asessed on the exact same precipitation time series for both the MTW and the conventional approach so that the performance scores can be directly compared. In order to assess the benefit of searching for analogue situations at a sub-daily time step for quantitative precipitation prediction, an appropriate precipitation archive is required. On the basis of the hourly time series (Sect. 2.1), 24-h totals were processed at time steps matching the MTW by building moving 24-h totals on the period 1982–2010, for example starting at 0, 3, ..., 21 h UTC instead of 6 h UTC as for the standard archive.

The impact of the archive reduction can be seen in Fig. **??** for the 2Z method and in Fig. **??** for the 2Z-2MI method. As expected, a loss of performance was observed for each station

3 Results

The considered AMs were applied to both ERA-20C and MERRA-2 datasets. Most of the results are presented for the ERA-20C dataset because it allows for testing a 3-h MTW, although similar conclusions can be found by using MERRA-2. The impact of an MTW approach was assessed on the two bases given below.

- Original parameters: The AMs were initially calibrated for each station based on the conventional fixed 24-h approach. Then, the MTW was introduced into the AMs, but the parameters (spatial windows and number of analogues) were not reassessed. The analogy was then identical, except for that of Aigle, which seemed relatively indifferent to this change . This loss was globally significant, with up to −1.89 points (% of CRPSS) for Visp and the 2Z method. the fact that the MTW had more candidates at disposal.

- Recalibrated: The parameters of the AMs were reassessed after the introduction of the MTW at different time steps. Indeed, one can assume that the introduction of the MTW might change the optimum value of some parameters (different spatial windows and number of analogues). The calibration (Sect. 2.3) was then reprocessed to adapt the AM parameters to the new information available through the MTW. The main difference was the increase in the optimal number of analogues to retain, which is discussed in Sect. 4.2.

The MTW is expected to affect both the selection of analogue dates and the prediction skill. Both are analysed separately hereafter. The results are presented for the Ulrichen station, but were very similar for all other stations.

3.1 Influence of the MTW on the analogy criterionselection of analogue dates

3.1.1 Changes in the atmospheric circulation analogy

First, the impact of the MTW on the selection of analogue dates was analysed independently of precipitation data. This was conducted for the period 1982–2010, although a longer period could have been used because no sub-daily precipitation series is required at this stage. The original parameters were used when assessing the impact of the MTW; no recalibration of the parameters was conducted.

3.1.1 Analogy of the atmospheric circulation

When searching for analogues in the first level of analogy, such as on the geopotential heights in the 2Z method (Sect. 2.2), there are 4 up to eight times as many candidates (even though not fully independent) were provided with the MTW than before even though they are not fully independent (Sect. 2.1). Figure 3 shows the changes in the distributions, for target dates over the CP, of the analogy criterion (S1) for the 1^{st}, 5^{th}, 20^{th}, and 40^{th} analogue rank at the Ulrichen station over the whole calibration period (1961–2008 – the full period could be used here, as precipitation was not considered at this stage), due to introduction 15^{th}, and 30^{th} analogue rank due to the introduction of an MTW with a 12-h, 6-h, or 3-h time step. The shapes of the MTW . The shape of the distributions of the conventional approach and the MTW MTWs were found to be similar, but . However, the values of the analogy criterion were gradually reduced (shifted to the left) and were, therefore , better. with smaller MTW time steps and were therefore better. The circulation analogy was regularly improved with the doubling of the MTW time step. For analogues with higher ranks (e.g. 20^{th} or $40^{th}$$15^{th}$ or 30^{th}), the difference between the two distributions was larger than that for the first ranksrank, which means that the improvement increased with the rank of the analoguesanalogue.

The improvements of the analogy with the rank of the analogues are summarized summarised in Fig. 4, which shows (top) the quantiles of the S1 criterion for the conventional method and the MTW, and (bottom) quantiles MTWs at different time steps (top) and those of the relative reduction (meaning improvement) due to the MTWindicating improvement owing to the MTWs (bottom). This confirms that all quantiles were similarly reduced (S1 distributions keep their shape), and that this improvement was constantly increasing the improvement constantly increased from the first to the last analogue (Fig. 4 bottom). This can be explained by the accumulation of better analogues in the selection, whereby the new, better situations push the previous analogues to higher ranks. The regularity of the improvements brought by decreasing time steps is also evident. The introduction of the MTW enables better analogue situations to be found in the first level of analogy, resulting in the selection of days with better similarity in the atmospheric circulation. This improvement was superior for MTWs with a smaller time step, which obviously allows better matching.

3.1.2 Changes per precipitation class

The impact of the MTW on the analogy criterion has been analysed per precipitation classes (class, for the target day), for 6-h and 3-h MTWs. The results are summarised in Fig. 5 by the median reduction of S1 for the days with precipitation (organized into classes) between two thresholds. With a reduction in the number of cases per classbeing reduced, the curves are were not as smooth as in those in the previous analyses. It is nevertheless clear that the improvements were larger for days with higher precipitation . Once again, the results for all the other stations were similar. for both the 6-h and 3-h MTW.

3.1.2 Changes per season

~~Atmospheric dynamics vary greatly from one season to another, reflecting on the performance of the AM that is generally lower between June and August (Bliefernicht, 2010).~~ The effect of the MTW on the S1 criterion ~~per season~~ was decomposed per season and is presented in Fig. 6. ~~Differences between~~ The differences among the seasons were substantial, with greater improvements shown for winter than summer. Analysis of the selected hours for the geopotential height predictor was then performed for the MTW with different time steps (Fig. 7). The new choice of the temporal window in winter was more balanced between the different hours of the day, showing more regular repartition, for all MTW time steps. This indicates a change in selection of a greater portion of the analogue dates for winter. On the contrary, the days in summer showed a preference for the initial temporal window (Z500, 24 h and Z1000, 12 h). This is likely attributed to the more pronounced diurnal effects, which reduced the potential for improvement of the criteria. These effects are in phase with the daily cycle, and good analogues were essentially found for the same hours. When a 3-h MTW was used, the time step following the initial temporal window (Z500, 27 h and Z1000, 15 h) was selected almost as often as the initial values for all seasons. This temporal window can be considered as relatively similar in terms of the solar influence. The other seasons were between these two extremes, which is consistent with their respective improvements in S1.

3.1.2 ~~Changes in the moisture~~ Moisture analogy

When adding the second level of analogy of the 2Z-2MI method (Table 2), the number of candidate situations did not increase ~~, as~~ when using the original parameters because they were conditioned by the N_1 previously selected analogues~~, but their dates had~~ ; however, their dates changed. In contrast to the AM on the atmospheric circulation only, both a reduction ~~or an increase of~~ and an increase in the RMSE analogy criterion values were possible with the MTW compared ~~to the static approach. Indeed,~~ with the conventional fixed approach. The results of the second level of analogy, presented in Fig. 8~~shows an almost insignificant improvement of the RMSE values~~ , showed no improvement in RMSE values regardless of the MTW time step. Unlike the first level of analogy, the relative changes ~~of~~ in the RMSE values were distributed relatively symmetrically around zero~~, with improvements and losses of the same amplitude~~.

These criterion values were not improved by the MTW because the number of candidates (N_1) was not higher. The use of an MTW did not increase the sample size in this case because the second level of analogy consists in subsampling in the dates provided by the first level. However, this result of a globally null improvement in the RMSE values does not mean that the 2Z-2MI method cannot be improved by the MTW. After selection of the analogue dates for the synoptic circulation, the new candidate dates did not provide better analogues in terms of moisture. However, the selected dates changed in the first level of analogy and in the final selection; thus, the distributions of precipitation values were different, which can impact the performance of the precipitation prediction.

3.2 Impact of the MTW on the precipitation prediction~~skill~~

The ~~new performance scores (CRPSS~~changes in the performance score (CRPS) of the precipitation prediction are provided in Fig. ~~?? and ??~~ 9 and 10 for the 2Z and 2Z-2MI methods, respectively~~, for the reduced archive (1982–2007). The gains relative to the static approach on the same archive ranged from 0.57 to 2.14 points (% of CRPSS) for the 2Z method and from 1.53~~

to 2.20 points for the 2Z-2MI method. . The AMs were recalibrated on the CP and were assessed on the independent VP. The score was processed for each station and both reanalysis datasets. For ERA-20C, the 12-h MTW, 6-h MTW, and 3-h MTW could be assessed, whereas only the two first could be processed with the MERRA-2 dataset.

3.2.1 Improvement per precipitation classes

The S1 criterion was previously found to show greater improvement for higher precipitation values (Sect. ??). The changes in terms of prediction skill were also assessed for the precipitation classes. Fig. 11 synthesizes these differences for the Ulrichen station, with the other stations having the same behaviour. The performance score was improved for days with higher precipitation after the introduction of the MTW. In contrast, for non-rainy days and small precipitation values, MTW did improve the precipitation prediction because the CRPS globally decreased for all stations, both for the CP and the VP. The VP showed greater variability, which is likely related to its shorter period of six years. The prediction skill for the performance scores were not improvedCP was almost always improved further by reducing the time step of the MTW, but not of the same magnitude. For ERA-20C, the magnitude of the improvement of 2Z for the CP was relatively reduced when using a 3-h MTW instead of a 6-h MTW, and the variability for the VP increased.

3.2.1 Recalibration of AM parameters

The previous assessment of the performance improvement was established with the original parameters optimized with a fixed time window. However, one can assume that the introduction of the MTW might change the optimum value of some parameters. The calibration (see Sect. 2.3) should then be reprocessed.

After recalibrating, changes in the AM parameters could be observed for both the However, as shown in Fig. 4, the selection of analogue dates was improved relatively equally for every reduction of the MTW time step. A possible reason is that the ERA-20C dataset was built by assimilating only surface observations, whereas the AMs rely on the geopotential height at 500 hPa. Thus, the timing of the atmospheric circulation at higher pressure levels might not be completely in phase with the actual perturbation systems which caused the observed precipitations. The MTW with the MERRA-2 dataset and the 2Z and method also showed a certain slope break after the 12-h MTW that could not be explained and was less important for the 2Z-2MI method. However, the 12-h MTW resulted in important variability for the 2Z-2MI methods. Among these, the zonal extent of the spatial windows of the circulation analogy decreased slightly (not shown). More significantly, the optimal number of analogues N_1 and N_2 increased after introducing the MTW by a considerable magnitude (Table ??): 25 % to 83 % for the 2Z method and 20 % to 67 % for the final selection of the 2Z-2MI method. The number of analogues N_1 of the first analogy level of the 2Z-2MI method reached three times its previous value for the Visp station (N_1 = 135 instead of 45; Table ??).

The values of the CRPSS scores for both methods (Fig. ?? and ??) have significantly increasedafter recalibration. When analysing the change in performance per precipitation class, the results (not shown) were very similar to the observations in Sect. ??, with a slight performance increase for small precipitation values that can be observed at the expense of higher amounts, due to the higher number of analogues.

3.3 ~~Using a reconstructed archive~~

~~The reconstructed time series cover the full period. However, they were first assessed using the MTW on the reduced period in order be comparable with the real sub-daily precipitation time series, and thus to evaluate their relevance and the possible loss in performance score.~~

~~When using the first reconstructed time series based on the proportional distribution (see Sect. 4.4), the AM performance was degraded and was even below the conventional methodwithout MTW (Table **??** to compare to Fig. **??** and **??**).~~

~~For the second approach, the synoptic proxy for the temporal disaggregation of the daily time series had to be identified first. The results are illustrated this time for the Zermatt station. Among the considered moisture variables (see Sect. 4.4), the best proxy was the precipitable water at 45° N and 7.5° E, with a time offset of –6 h (Table **??**). Table **??** presents the CRPSS scores obtained by the disaggregated series using the proxy of precipitable water which are also lower than the conventional method~~the MERRA-2 dataset.

4 ~~Discussion~~

3.1 ~~Improvement of the selection of analogue situations~~

~~For the first level of analogy, the~~ The introduction of an MTW was previously found to show greater improvement of the S1 criterion ~~(Fig. 4 bottom) started approximately at 5 % for the first analogue and reached more than 10 % for the last (40^{th}) (Sect. **??**). This increase in improvement with the analogue rank can be explained by the accumulation of better analogues in~~ for days with higher precipitation values (Sect. 3.1.1). The impact of the MTW was then assessed in terms of changes in prediction performance per precipitation class for the target day. Figure 11 synthesises these differences for the ~~selection, with new better situations pushing previous analogues to higher ranks. The curve representing the minimum improvement is mostly superior to zero, meaning that the criteria have been improved on most analogue ranks for every day of the calibration period. All other stations had a similar improvement in the S1 criterion, both in terms of the distribution shape and amplitude. The~~ Ulrichen station by using the original parameters with the MTW. After the introduction of the MTW~~allows finding better analogue situations in the first level of analogy, resulting in a selection of days with better similarity in atmospheric circulation.~~

~~When adding a second level of analogy with moisture variables, the criteria values (RMSE) were not improved by the MTW(Sect. **??**), as the number of candidates was not higher (still the N_1 days selected in the first level of analogy). However, this result of a globally null improvement of the RMSE values does not mean that the 2Z-2MI method cannot be improved by the MTW. It means that after the selection of the analogue situations for~~ , the performance score was generally further improved with reduced CRPS for days with higher precipitation than for non-rainy days and small precipitation values. When recalibrating the parameters using the MTW, the ~~synoptic circulation, the new candidate dates did not provide better analogues in terms of moisture. However, the selected dates have changed in the first level of analogy and also in the final selection, and thus the distributions of precipitation values were different, which did improve the prediction skill (Sect. 3.2).~~

The introduction of the MTW improved the selection of synoptic analogues. Independently of its impact on the prediction skill for precipitation, or the availability of a predictand time series with sub-daily time step, this improvement has a potential in itself for application on long meteorological archives. For example, when processing forecasts for a target day showing synoptic similarity with situations from the past related to extreme weather, even if for them no precipitation archive is available. Indeed, some of those days with strong precipitation events may be documented, either qualitatively in the daily press or more quantitatively in flood reports. It is thus worth to known that the situation at hand has had such analogue in the past. Another possible application is a quality assessment of the selection of analogues on a shorter period, where precipitation data with a high temporal resolution is available. Indeed, if the selection of analogues with the MTW on the long period, for a specific target day, differ from the selection on the short period, this may point out a sub-optimal forecast. Finally, other predictands might not need sub-daily total values, but point observations (e.g. hail, or extreme wind gusts), which make them easier to use with the MTW .

3.1 Improvement of precipitation prediction

The MTW was found to improve the skill of the precipitation prediction for both spread of the changes in CRPS tended to decrease, although a slight loss in performance was noted for the most rainy days (not shown).

Reliability diagrams were also created for the VP for both 2Z and 2Z-2MI methods (Fig. ?? and ??, Sect. 3.2). Moreover, it required no additional predictor. With the introduction of the MTWmethods and for the exceedance of all days percentiles 80 % (P80, 4 mm), 90 % (P90, the performance loss related to the reduction of the archive length was negated (considering 9.5 mm), and 95 % (P95, 17.5 mm) at the Ulrichen station (Fig. 12). These diagrams plot the observed frequency against the predicted probability of a 25 year archive of sub-daily data; the difference could be higher with a shorter archive). It was in this binary event, which in this case is threshold exceedance. For perfectly calibrated predictions, the curve should be along the diagonal. The VP contains only six years (Sect. 2); therefore, higher variability is present for higher thresholds, and the curves are smoother and closer to the diagonals for the CP, which is longer (not shown). Figure 12 shows that the MTW improved the reliability for both 2Z and 2Z-2MI methods and for all thresholds because the curves moved towards the diagonal with decreasing MTW time steps. For P80, the 3-h MTW predictions fit very well the observed frequency. When considering higher thresholds, the 2Z-2MI method was generally better than 2Z. The case of the same magnitude as if the length of the archive doubled. Note, that despite the number of candidate situations being 4 times as many, the gains seem to be lower than for a quadrupled archive length. The likely reason is that an actual longer archive would contain more atmospheric situations that might have been observed less frequently during a shorter period. Moreover, high correlations between sub-daily circulation patterns are expected. The MTW, therefore, did not create 4 times as many independent data, but the increase in information was nevertheless substantial. In this regards, a reanalysis dataset with a higher temporal resolution might not improve the performance significantly more than a 6-hourly dataset. In contrast, reducing the MTW to a 12-hour time step might reduce redundancy in the archive, but might also reduce the chance of finding better analogue situations, as circulation patterns can evolve substantially in 12 hours.

~~Moreover, in a transient climate, an eventual nonstationarity of the link between predictors and precipitation might discard the relevance of analogues from the distant past and increase the relevance of using a more recent and shorter archive rather than a long one. In such cases, the archive inflation brought by the MTW is also relevant~~P95 threshold showed a conditional dry bias associated with larger forecast probabilities for the considered AMs; however, this was substantially reduced by the 3-h MTW. This conditional dry bias means that the AM did not predict the event with a probability equal to 1 for 2Z. Obtaining a predicted probability of exceedance of 1 would mean that all analogue dates for a certain target date would be above the threshold. The number of analogue dates being 30 in this case, it is less likely to have 30 values above the threshold when this one increases. This issue has been addressed by Marty (2010) who proposed a bias correction method for the AM. However, as shown in Figure 12, the introduction of an MTW, especially with a 3-h time step, improves this mismatch to some degree.

4 Discussion

4.1 Better prediction of heavy precipitation

Both the analogy criteria (Sect. ~~??~~3.1.1) and the performance scores (Sect. ~~??~~3.2) were improved to a greater extent for days with heavier precipitation. This is likely ~~due~~ attributed to the fact that higher precipitation events are a consequence of atmospheric conditions with greater dynamics ~~,~~ such as weather disturbances, which have ~~a~~ well-marked temporal evolution. Indeed, the position of the driving elements, such as the ~~low pressure~~ low-pressure centres and the fronts, change significantly during the course of a day. These situations are less numerous than anticyclonic situations, which makes it less likely to find very good analogues at the same time of the day. We can, therefore, expect to more significantly improve these situations with greater dynamics ~~more significantly when introducing a MTW , as~~ when introducing an MTW because better matches to the target situation may be found. In contrast, days with low dynamics in ~~the~~ atmospheric circulation, such as anticyclonic situations, will not be radically improved by the introduction of the MTW.

The MTW improved the prediction for days with heavier precipitation~~, and~~ . Therefore, it should improve the prediction of extremes ~~due to~~ owing to selection of better analogue situations ~~, but also due to~~ and to addition of possible new extreme precipitation values resulting from ~~24h~~ moving 24 h totals with a certain time shift. However, even though the distribution of analogue precipitation values should move towards the targeted extreme ~~, providing a~~ to provide better prediction, the MTW itself does not ~~allow to predict~~ enable prediction of extreme events that were not yet observed and are ~~, therefore ,~~ therefore not present in the archive. ~~The~~ However, the extremes in AMs can be modelled by ~~, for example,~~ extrapolation of a truncated exponential or gamma distribution fitted to the analogue values (Obled et al., 2002). Another possible approach is ~~by combining~~ to combine AMs with other methods (e.g. Chardon et al., 2014). From this perspective, the MTW might improve the prediction of extremes ~~as~~ because it improves the distribution of precipitation values for days with higher precipitation, on which post-treatment techniques rely. However, this ~~goes~~ topic is beyond the scope of the present study.

4.2 ~~Seasonal effect~~The relationship between the MTW time step and the number of analogues

Section ?? and Fig. 6 revealed a difference in the improvement of When recalibrating the AMs with MTWs of different time steps, the S1 criterion according to the season, with greater improvements for winter than summer. One hypothesis is that the diurnal effects of the summer months have an influence on the atmospheric circulation at least in the lower layers. This effect is in phase with the daily cycle, and good analogues are essentially found for the same hours.

An analysis of the selected hours for the geopotential height predictor seems to confirm this assumption optimal number of analogues changed for both 2Z and 2Z-2MI methods. N_1 and N_2, from the first and the second level of analogy, tended to significantly increase with a reduced MTW time step (Fig. 7). It was found that the new choice of the temporal window in winter, when using the MTW approach, is well balanced between the four options. This means a change of 75 % of the analogues selected compared to the conventional approach. On the contrary, the days during the summer months had a preference for the initial temporal window (Z500 24 h & Z1000 12 h), likely due to more pronounced diurnal effects, which reduced the potential for improvement of the criteria. The other seasons were between these two extremes, which is consistent with their respective improvements. This seasonal effect was observed for each station in a very similar way and even with slightly larger amplitude than for Ulrichen.

4.3 On the increase in the number of analogues

After the recalibration of the AMs with the MTW, the optimal analogue numbers were significantly higher than in the original methods (see Sect. ?? and Table ??) 13). When using ERA-20C, all optimal numbers of analogues for the different analogy levels tended to double when using a 6-h MTW instead of the classic approach and to even triple when using a 3-h MTW. These higher numbers of analogues were objectively chosen by using the calibration procedure (Sect. 2.3) in order to increase the prediction skill of the methods.

As shown in Fig. 4, the improvement of the S1 criterion grew increases along with the rank of the analogue, which shows an accumulation of better analogue situations in the distributions. It seems profitable to widen Widening the selection of analogues in order to also keep some whose rank has increased , as appears to be beneficial for keeping some with increased rank because they appear to be relevant to the prediction of the in predicting the precipitation values. The Thus, the number of good analogues was globally increased.

A higher number of analogues generally means, implies a poorer analogy with an archive of fixed length, a poorer analogy. Indeed, when the choice of the predictors or the parameters are improvedimprove, leading to a better prediction, the optimal number of analogue situations decreasestends to decrease. However, when the length of the archive increases, the optimal number of analogues increases too for a also tends to increase for better performance up to a certain threshold (Bontron, 2004; Hamill et al., 2006). The observed increase in the number of analogue situations with the MTW resulted resulting in better performance skills for the given methods , as it can be seen can therefore be considered as an inflation of the archive. However, if new relevant predictors were added to the method, the number of analogues would then decrease.

4.3 ~~Why not just make 6-hourly predictions?~~

4.3 Inflation of the archive

Because sub-daily precipitation time series are often available over a shorter period than daily, a reduction of the overall archive length at disposal is necessary. This usually has a negative consequence on the skill of precipitation prediction. The role of the archive length was assessed on the ERA-20C dataset and is presented in Fig. 14 for both 2Z and 2Z-2MI methods using the conventional approach of fixed 24-h windows. As expected, the performance of the method globally increased with the length of the archive. The best performances were found with a 44-year archive, which corresponds to the period 1961–2010 minus the 6 years of independent VP. This addition of 21 years of archive to the original 23 years resulted in an improvement of both 2Z and 2Z-2MI methods that was of the same order of magnitude as the introduction of a 6-h MTW (Fig. 9 and 10). The 3-h MTW resulted in similar performance for the 2Z method and slightly superior improvements for the 2Z-2MI method.

As previously stated, the MTW can be considered as an inflation of the archive because it provides additional situations to the pool of possible analogues. With the introduction of the MTW, the performance loss related to an eventual reduction of the archive length to meet the length of the sub-daily precipitation archive was indeed compensated. A 6-h MTW brought gains of the same magnitude similar to an almost double archive length. However, it should be noted that despite the number of candidate situations, which in this case is four times the original number, the gains appear to be lower than those for a quadrupled archive length. The likely reason is that an actual longer archive would contain more atmospheric situations that might have been observed less frequently during a shorter period. Moreover, rather strong serial correlations between successive sub-daily circulation patterns are expected.

Many reanalysis datasets begin in the 80s owing to availability of satellite data and are thus more accurate than reanalysis datasets based on conventional data only; however, they provide shorter archives. When using such an archive, the MTW approach shows a certain potential for enriching the pool of possible analogues. Moreover, in a transient climate, the eventual nonstationarity of the link between predictors and precipitation might discard the relevance of analogues from the distant past and can increase the relevance of using a more recent and shorter archive rather than a long one. In such cases, the archive inflation brought by the MTW is also relevant.

One can question the interest of using moving daily totals when~~6-hourly~~ , for example, 6-h precipitation series can be predicted instead. However, the ~~6-hourly~~ 6-h time series generated by the AM might not ~~represent accurately~~ accurately represent the intra-daily precipitation distribution ~~(results not shown)~~or variability. In addition, sometimes a resolution finer than the daily time step is not needed, and another disaggregation technique may be used afterwards. Finally, when using a reconstructed precipitation archive, the errors in intra-daily precipitation distributions have a smaller impact ~~than 24h daily~~ on 24-h totals than on 6-h totals.

4.4 ~~On the use~~ Reconstruction of ~~an old reanalysis dataset~~a long precipitation archive suitable for the MTW

~~The considered reanalysis dataset, which is the NCEP/NCAR reanalysis 1 (Kalnay et al., 1996), is relatively old and has a coarse resolution (2.5°). Newer reanalysis datasets could have been used in this study and might have resulted in higher~~

~~performance scores. However, Ben Daoud et al. (2009) showed that the sensitivity of~~ As discussed in Sect. 4.3, the ~~method to~~ skills of AMs improved with the length of the archive. It would then be even more profitable to apply the MTW to the longest archive possible rather than being limited to the period in which sub-daily precipitation data are available. Therefore, the idea is to reconstruct longer archives of moving 24-h totals from existing standard daily precipitation series. For this purpose, disaggregation techniques can be used. In this study, two simple disaggregation approaches of the daily precipitation time series were assessed.

The first technique is a proportional distribution in which the observed daily precipitations were considered to be constant during the measurement period of 06:00 h UTC to 06:00 h UTC the following day. The proportional parts of the ~~reanalysis dataset is rather small. This is particularly true for simple AMs that mainly rely on the atmospheric circulation, which is already well defined at a coarser resolution. That might not be the case in more elaborate methods relying on thermodynamic data. A comparative analysis of several reanalysis datasets within the AM is being conducted and will be the topic of a dedicated study~~original daily time series were allocated into a new moving average of 24 h totals (e.g. a 00:00 – 24:00 h UTC total would be made of ¼ of the standard precipitation of day one and ¾ of day two).

~~Moreover, one can assume that the gain observed here by introducing a MTW should also be found for a better reanalysis dataset. Indeed, a better dataset does not negatively influence the fact that one can find better analogue situations at other hours of the day rather than at fixed hours.~~ The second approach was used to obtain data that are closer to the chronology of the actual precipitation by relying on informative proxy variables during the reconstruction procedure. Data from NCEP/NCAR reanalysis 1 were used for this purpose. Precipitable water and relative humidity at 1000 hPa, 925 hPa, or 850 hPa were assessed at the four points surrounding the catchment. Precipitation time series from the reanalysis were not considered due to the presence of zeros, which may not always match the prediction; this can lead to an undefined temporal repartition. Time lapses from -12 h to +12 h between both series were introduced to consider the significant distance separating the weather stations and the reanalysis grid point. The best proxy variable, precipitable water, was identified through correlation analyses on non-zero values with the 6-h precipitation time series. When the best proxy was selected, its temporal profile was used to disaggregate the daily precipitation time series.

4.5 ~~Relevance of the reconstructed precipitation archives~~

~~The two simple methods assessed to reconstruct precipitation archives (Sect. 4.4)~~ These two simple methods did not result in valuable outputs (~~Sect. ??~~results not shown). Indeed, the performance improvement brought by the MTW was lost due to the poor quality of the reconstructed precipitation archives. A slight improvement ~~could be~~ was obtained for the second method ~~relying~~ which relied on a proxy variable ~~compared to~~ rather than the proportional distribution method~~, but it was still~~. However, this improvement was relatively small, and most of the benefit of the MTW was lost. ~~A more recent reanalysis archive~~ Another reanalysis dataset with more accurate moisture variables ~~might~~ could produce better proxies.

These attempts to transpose the MTW on the ~~total archive highlighted~~ (usually) longest meteorological archive highlight the importance of the actual rainfall chronology. The MTW is beneficial ~~, provided that~~ if the precipitation series are close to the observed one. Without values. In the absence of a precipitation series with an accurate sub-daily chronology, the introduction of a an MTW does not improve the precipitation prediction.

5 Conclusions

The AMs are most often based on a daily time step due owing to the availability of long precipitation archives. However, it is unlikely that two analogue synoptic situations , that which evolve relatively quickly , would correspond would match optimally at the same hours of the day. It is probable that better matches can may be found at another timea different hour, which can change the selection of the analogue dates.

As Finet et al. (2008) had previously shown shown in previous research (Finet et al., 2008), the introduction of a MTW allows finding an MTW enables better analogue situations in terms of the atmospheric circulation . It has been demonstrated in this studythat the improvement of the atmospheric circulation to be obtained. Using recent reanalysis datasets, MTWs with 12-h, 6-h, and 3-h time steps were assessed in the present study. Improvement in the S1 criterion values was growing increased with the rank of the analogue . This was due owing to the accumulation of better analogues within the predicted distributions.

The improvement of the circulation analogy was found to be more important for days with heavier precipitation, which are generally related to more dynamic atmospheric situations and are less frequent in the archive. These situations have more specific circulation patterns that are evolving more rapidly. Therefore, a MTW was found to be of particular interest in this kind of situation, benefiting the prediction of heavier precipitation eventsMoreover, the improvement in the S1 criterion increased further for smaller MTW time steps.

A seasonal effect has been highlighted , as was highlighted such that the MTW was more profitable for winter than summer. The reason is likely for summer. A likely reason is that the diurnal cycle has a bigger greater effect in summer than in winter, which results in better analogues at the same time of the day. The preference for the same hours in summer has been demonstrated; whereas, 75 % , whereas a large part of the analogue situations were selected at a different time in winter.

The impact of the MTW on the prediction skill was never assessed before due to the shortcoming of long precipitation series at a sub-daily time step. Here, it was assessed for a 25 year time series with a high temporal resolution. After the introduction MTW improved the precipitation prediction, with the CRPS showing a global decrease for all stations for both the CP and VP. The prediction skill for the CP was almost always improved further by reducing the time step of the MTW, but the magnitude differed. The results for the VP showed the same global trends but had more variability. Moreover, the performance scores increased of the same magnitude as if the length of the archive doubled.

The parameters were then calibrated again, using the reliability of the considered AMs improved for the prediction of different threshold exceedances, especially with a 3-h MTW. Some parameters changed, with the main difference being the number of analogues, which systematically and significantly increased compared to the original set. A wider selection of analogue situations, containing those whose rank decreased, seemed beneficial for the prediction performance. The number of good analogues was globally increased in the same way as if the archivelength increased. This change seems to benefit

Improvements of the analogy criterion and the performance scores were both found to be higher for MTWs with a smaller time step of 3 h. A 3-h MTW provided eight times more candidate situations, although they were not fully independent. Because the MTW provides additional situations to the pool of possible analogues, it can be considered as an inflation of the archive, which can be interesting in several applications.

The improvement in the circulation analogy and the change in performance of the precipitation prediction was found to be more important for days with heavier precipitation. These days are generally related to more dynamic atmospheric situations and are less frequent in the archive. These situations have more specific circulation patterns that are evolving more rapidly. Therefore, an MTW was found to be of particular interest in this type of situation and is beneficial for the prediction of ~~days with small precipitation totals~~heavier precipitation events.

The importance of the quality of the precipitation archive was also demonstrated ~~, as~~ because simplistic reconstructions of ~~6-hourly~~ sub-daily time series led to a loss of all the improvement brought by the MTW. The precipitation prediction ~~is~~ improved only when the precipitation chronology ~~is~~ was close to the accurate ~~one~~values. Attempts to reconstruct longer time series based on simplistic proportional distributions or by using meteorological variables from the NCEP/NCAR reanalysis 1 dataset as ~~proxy did not succeed~~a proxy were not successful. Other reanalysis datasets with more accurate moisture variables could ~~eventually perform better~~ show better performance.

The use of the MTW relies partly on the availability of long precipitation series at a sub-daily time step and with high accuracy. ~~First, these~~ The precipitation data archives of high temporal resolution ~~precipitation data are increasing~~ have increased over time. ~~Another possible source~~ Other possible sources of such archives is the establishment of precipitation reanalysis at a regional scale ~~or~~ in addition to the use of reanalysis-driven regional climate models or limited area models over a long period. Even though outputs from these models might be biased or not accurate enough, information regarding the timing of the precipitation events could be useful in disaggregating the station time series.

~~Finally, since long meteorological archives (reanalysis datasets) are more and more available, the improvements proposed by the MTW especially for days with heavy precipitation may be interesting even without long continuous precipitation~~ The introduction of the MTW improved the selection of synoptic analogues. Regardless of its impact on the prediction skill for precipitation or the availability of a predictand time series with a sub-daily time step, this improvement has the potential for application to long meteorological archives. For example, ~~recent target day may have~~ when processing forecasts for a target day showing synoptic similarity with known situations from the ~~early twentieth century, for which no continuous daily~~ past related to extreme weather, it can be used even if no precipitation archive is available. ~~However, some of those days with strong precipitation events may be documented, either qualitatively in the daily press or more quantitatively in flood reports. Nevertheless,~~ Thus, it is worth ~~to known that~~ knowing whether the situation at hand ~~has had such~~ had such an analogue in the ~~far past.~~ past. Another possible application is quality assessment of the selection of analogues on a shorter period, where sub-daily precipitation data are available. Indeed, if the selection of analogues with the MTW on a long period for a specific target day differs from the selection for the shorter period, a sub-optimal forecast could be identified. Finally, some other predictands might not need sub-daily total values but point observations such as hail or extreme wind gusts, which make them easier to use with the MTW.

The use of the MTW in the AM can already be considered now The technique is very simple and easily applicable. Therefore, it should be considered for several applications in different contexts, may it be for such as operational forecasting or climate-related studies.

Appendix A: Acronyms

2Z Name of the analogue method of the atmospheric circulation

2Z-2MI Name of the analogue method composed of a first level on the atmospheric circulation and a second level on a moisture index

AM Analogue method

CP Calibration period

CRPS Continuous ranked probability score

CRPSSERA-20C Continuous ranked probability skill score European Center for Medium Range Weather Forecasting 20th century reanalysis

MERRA-2 Modern-Era Retrospective Analysis for Research and Applications, Version 2

MTW Moving time window

NCAR National Center for Atmospheric Research

NCEP National Center for Environmental Prediction

RH850 Relative humidity at 850 hPa

RMSE Root-mean-square Root–mean–square error

S1 Teweles and Wobus (1954) score

TPW Total precipitable water

VP Validation period

Z1000 Geopotential height at 1000 hPa

Z500 Geopotential height at 500 hPa

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We The authors thank the Swiss Federal Office for Environment (FOEV), the Roads and Water Courses Service, Energy and Water Power Service of the Wallis Canton, and the Water, Land, and Sanitation Service of the Vaud Canton which financed the MINERVE (Modélisation des Intempéries de Nature Extrême des Rivières Valaisannes et de leurs Effets) project that which started this research. Thanks are also extended to Dominique Bérod for his support.

The fruitful collaboration with the Laboratoire d'Etude des Transferts en Hydrologie et Environnement of the Grenoble Institute of Technology (G-INP) was made possible thanks to by the Herbette Foundation.

NCEP reanalysis data provided by the were obtained from the Web site of the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. Precipitation time series were provided by MeteoSwiss.

The authors would also like to acknowledge the work of anonymous reviewers that who contributed to improving this manuscript.

References

Alessandrini, S., Delle Monache, L., Sperati, S., and Cervone, G.: An analog ensemble for short-term probabilistic solar power forecast, Appl. Energ., 157, 95–110, doi:10.1016/j.apenergy.2015.08.011, 2015a.

Alessandrini, S., Delle Monache, L., Sperati, S., and Nissen, J. N.: A novel application of an analog ensemble for short-term wind power forecasting, Renew. Energ., 76, 768–781, doi:10.1016/j.renene.2014.11.061, 2015b.

Ben Daoud, A.: Améliorations et développements d'une méthode de prévision probabiliste des pluies par analogie., Ph.D. thesis, Université de Grenoble, 2010.

Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., and Bontron, G.: La prévision des précipitations par recherche d'analogues: état de l'art et perspectives, in: Colloque SHF-191e CST - «Prévisions hydrométéorologiques», Lyon, 18-19 novembre 2008, 6, 2008.

Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., and Bontron, G.: La prévision des précipitations par recherche d'analogues : état de l'art et perspectives, La Houille Blanche, pp. 60–65, doi:10.1051/lhb/2009079, 2009.

Ben Daoud, A., Sauquet, E., Bontron, G., Obled, C., and Lang, M.: Daily quantitative precipitation forecasts based on the analogue method: improvements and application to a French large river basin, Atmos. Res., 169, 147–159, doi:10.1016/j.atmosres.2015.09.015, 2016.

Bessa, R., Trindade, A., Silva, C. S., and Miranda, V.: Probabilistic solar power forecasting in smart grids using distributed information, Int. J. Elec. Power, 72, 16–23, doi:10.1016/j.ijepes.2015.02.006, 2015.

Bliefernicht, J.: Probability forecasts of daily areal precipitation for small river basins, Ph.D. thesis, Universität Stuttgart, 2010.

Bontron, G.: Prévision quantitative des précipitations: Adaptation probabiliste par recherche d'analogues. Utilisation des Réanalyses NCEP/NCAR et application aux précipitations du Sud-Est de la France., Ph.D. thesis, Institut National Polytechnique de Grenoble, 2004.

Bontron, G. and Obled, C.: L'adaptation probabiliste des prévisions météorologiques pour la prévision hydrologique, La Houille Blanche, 1, 23–28, doi:10.1051/lhb:200501002, 2005.

Brown, T.: Admissible Scoring Systems for Continuous Distributions., Tech. rep., http://eric.ed.gov/?id=ED135799, 1974.

Caillouet, L., Vidal, J.-P., Sauquet, E., and Graff, B.: Probabilistic precipitation and temperature downscaling of the Twentieth Century Reanalysis over France, Clim. Past, 12, 635–662, doi:10.5194/cp-12-635-2016, 2016.

Chardon, J., Hingray, B., Favre, A.-C., Autin, P., Gailhard, J., Zin, I., and Obled, C.: Spatial Similarity and Transferability of Analog Dates for Precipitation Downscaling over France, J. Climate, 27, 5056–5074, doi:10.1175/JCLI-D-13-00464.1, 2014.

Dayon, G., Boé, J., and Martin, E.: Transferability in the future climate of a statistical downscaling method for precipitation in France, J. Geophys. Res.-Atmos., 120, 1023–1043, doi:10.1002/2014JD022236, 2015.

Delle Monache, L., Nipen, T., Liu, Y., Roux, G., and Stull, R.: Kalman Filter and Analog Schemes to Postprocess Numerical Weather Predictions, Mon. Weather Rev., 139, 3554–3570, doi:10.1175/2011MWR3653.1, 2011.

Delle Monache, L., Eckel, F. A., Rife, D. L., Nagarajan, B., and Searight, K.: Probabilistic Weather Prediction with an Analog Ensemble., Mon. Weather Rev., 141, 3498–3516, doi:10.1175/MWR-D-12-00281.1, 2013.

Desaint, B., Nogues, P., Perret, C., and Garçon, R.: La prévision hydrométéorologique opérationnelle: l'expérience d'Electricité de France, in: Colloque SHF-191e CST - «Prévisions hydrométéorologiques», Lyon, 18-19 novembre 2008, p. 8 p., doi:10.1051/lhb/2009054, 2008.

Duband, D.: Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations., Ph.D. thesis, Thèse de 3ème cycle de la faculté des sciences de Paris., 1970.

Finet, T., Marty, R., Zin, I., and Obled, C.: Developing and transferring the ANALOG approach for PQPF's from French OHMCV catchments to quick responding catchments in Venetia and Romania., Tech. rep., LTHE & CNRS, Grenoble, France, 2008.

García Hernández, J., Horton, P., Tobin, C., and Boillat, J.: MINERVE 2010: Prévision hydrométéorologique et gestion des crues sur le Rhône alpin., Wasser Energie Luft – Eau Energie Air, 4, 297–302, 2009.

Gordon, N. D.: Statistical very short-range forecasting via analogues., European Space Agency Special Publication, pp. 307–312, 1987.

Guilbaud, S.: Prévision quantitative des précipitations journalières par une méthode statistico-dynamique des recherche d'analogues - Application à des bassins du pourtour méditerranéen, Ph.D. thesis, Institut National Polytechnique de Grenoble, 1997.

Hamill, T. and Whitaker, J.: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application, Mon. Weather Rev., 134, 3209–3229, doi:10.1175/mwr3237.1, 2006.

Hamill, T. M., Whitaker, J. S., and Mullen, S. L.: Reforecasts: An Important Dataset for Improving Weather Predictions, B. Am. Meteorol. Soc., 87, 33–46, doi:10.1175/BAMS-87-1-33, 2006.

Hamill, T. M., Scheuerer, M., and Bates, G. T.: Analog Probabilistic Precipitation Forecasts Using GEFS Reforecasts and Climatology-Calibrated Precipitation Analyses, Mon. Weather Rev., 143, 3300–3309, doi:10.1175/MWR-D-15-0004.1, 2015.

Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather Forecast., 15, 559–570, doi:10.1175/1520-0434(2000)015<0559:dotcrp>2.0.co;2, 2000.

Horton, P.: Améliorations et optimisation globale de la méthode des analogues pour la prévision statistique des précipitations. Développement d'un outil de prévision et application opérationnelle au bassin du Rhône à l'amont du Léman, Thèse de doctorat, Université de Lausanne, https://tel.archives-ouvertes.fr/tel-01441762, 2012.

Horton, P., Jaboyedoff, M., Metzger, R., Obled, C., and Marty, R.: Spatial relationship between the atmospheric circulation and the precipitation measured in the western Swiss Alps by means of the analogue method, Nat. Hazard Earth Sys., 12, 777–784, doi:10.5194/nhess-12-777-2012, 2012.

Horton, P., Jaboyedoff, M., and Obled, C.: Global Optimization of an Analog Method by Means of Genetic Algorithms, Mon. Weather Rev., 145, 1275–1294, doi:10.1175/MWR-D-16-0093.1, 2017.

Junk, C., Delle Monache, L., and Alessandrini, S.: Analog-based Ensemble Model Output Statistics, Mon. Weather Rev., 143, 2909–2917, doi:10.1175/MWR-D-15-0095.1, 2015a.

Junk, C., Delle Monache, L., Alessandrini, S., Cervone, G., and von Bremen, L.: Predictor-weighting strategies for probabilistic wind power forecasting with an analog ensemble, Meteorol. Z., 24, 361–379, doi:10.1127/metz/2015/0659, 2015b.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., and Woollen, J.: The NCEP/NCAR 40-year reanalysis project, B. Am. Meteorol. Soc., 77, 437–471, doi:10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2, 1996.

Kruizinga, S. and Murphy, A.: Use of an analogue procedure to formulate objective probabilistic temperature forecasts in the Netherlands, Mon. Weather Rev., 111, 2244–2254, doi:10.1175/1520-0493(1983)111<2244:uoaapt>2.0.co;2, 1983.

Lorenz, E.: Empirical orthogonal functions and statistical weather prediction, Tech. rep., Massachusetts Institute of Technology, Department of Meteorology, Massachusetts Institute of Technology, Dept. of Meteorology, 1956.

Lorenz, E.: Atmospheric predictability as revealed by naturally occurring analogues, J. Atmos. Sci., 26, 636–646, doi:10.1175/1520-0469(1969)26<636:aparbn>2.0.co;2, 1969.

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, Rev. Geophys., 48, 1–34, doi:10.1029/2009RG000314, 2010.

Marty, R.: Prévision hydrologique d'ensemble adaptée aux bassins à crue rapide. Elaboration de prévisions probabilistes de précipitations à 12 et 24 h. Désagrégation horaire conditionnelle pour la modélisation hydrologique. Application à des bassins de la région Cév, Ph.D. thesis, Université de Grenoble, 2010.

Marty, R., Zin, I., Obled, C., Bontron, G., and Djerboua, A.: Toward real-time daily PQPF by an analog sorting approach: Application to flash-flood catchments, J. Appl. Meteorol. Clim., 51, 505–520, doi:10.1175/JAMC-D-11-011.1, 2012.

Matheson, J. and Winkler, R.: Scoring rules for continuous probability distributions, Manage. Sci., 22, 1087–1096, doi:10.1287/mnsc.22.10.1087, 1976.

Obled, C.: Daniel Duband - cinquante ans de contributions scientifiques à l'hydrologie (1962-2011), La Houille Blanche, pp. 55–68, doi:10.1051/lhb/2014017, 2014.

Obled, C., Bontron, G., and Garçon, R.: Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, Atmos. Res., 63, 303–324, doi:10.1016/S0169-8095(02)00038-8, 2002.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J. N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An atmospheric reanalysis of the twentieth century, J. Climate, 29, 4083–4097, doi:10.1175/JCLI-D-15-0556.1, 2016.

Radanovics, S., Vidal, J.-P., Sauquet, E., Ben Daoud, A., and Bontron, G.: Optimising predictor domains for spatially coherent precipitation downscaling, Hydrol. Earth Syst. Sc., 17, 4189–4208, doi:10.5194/hess-17-4189-2013, 2013.

Radinovic, D.: An analogue method for weather forecasting using the 500/1000 mb relative topography, Mon. Weather Rev., 103, 639–649, doi:10.1175/1520-0493(1975)103<0639:aamfwf>2.0.co;2, 1975.

Raynaud, D., Hingray, B., Zin, I., Anquetin, S., Debionne, S., and Vautard, R.: Atmospheric analogues for physically consistent scenarios of surface weather in Europe and Maghreb, Int. J. Climatol., doi:10.1002/joc.4844, 2016.

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G. K., Bloom, S., Chen, J., Collins, D., Conaty, A., Da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, J. Clim., 24, 3624–3648, doi:10.1175/JCLI-D-11-00015.1, 2011.

Ruosteenoja, K.: Factors affecting the occurrence and lifetime of 500 mb height analogues: a study based on a large amount of data, Mon. Weather Rev., 116, 368–376, 1988.

Teweles, S. and Wobus, H. B.: Verification of prognostic charts, B. Am. Meteorol. Soc., 35, 455–463, 1954.

Van Den Dool, H. M.: Searching for analogues, how long must we wait?, Tellus, 46A, 314–324, 1994.

Vanvyve, E., Delle Monache, L., Monaghan, A. J., and Pinto, J. O.: Wind resource estimates with an analog ensemble approach, Renew. Energ., 74, 761–773, doi:10.1016/j.renene.2014.08.060, 2015.

Woodcock, F.: On the use of analogues to improve regression forecasts, Mon. Weather Rev., 108, 292–297, doi:10.1175/1520-0493(1980)108<0292:otuoat>2.0.co;2, 1980.

Zorita, E. and Storch, H. V.: The analog method as a simple statistical downscaling technique: comparison with more complicated methods, J. Climate, 12, 2474–2489, doi:10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2, 1999.

~~Illustration of the generation of 24h total moving averages by means of a proportional distribution. The colours refer to the corresponding day of the daily time series.~~

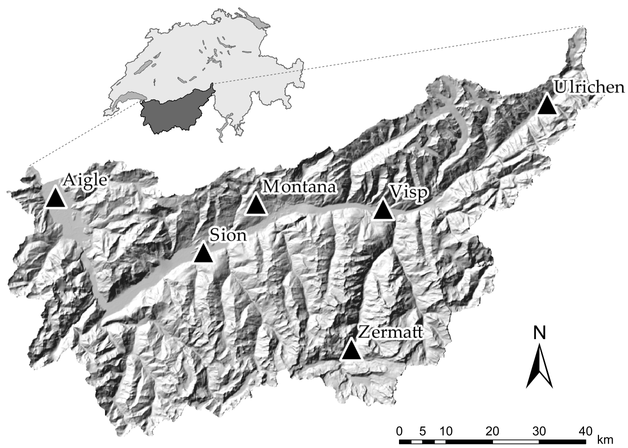

Figure 1. Position of the six weather stations of interest (~~Ulrichen~~Aigle, ~~Zermatt, Visp, ~~Montana, Sion, Ulrichen, Visp, and ~~Aigle~~Zermatt) in the upper Rhône catchment in Switzerland.

~~Performance score (CRPSS) of the AM on the atmospheric circulation at the different stations for (dashed) the full archive, i.e. 1961–2008; (yellow) the reduced archive, i.e. 1982–2007; (blue) the introduction of the MTW on the reduced archive; and (green) the recalibrated parameters of the AM with the MTW.~~

~~Same as Figure **??**, but for the analogue method with a second level with the moisture variables.~~

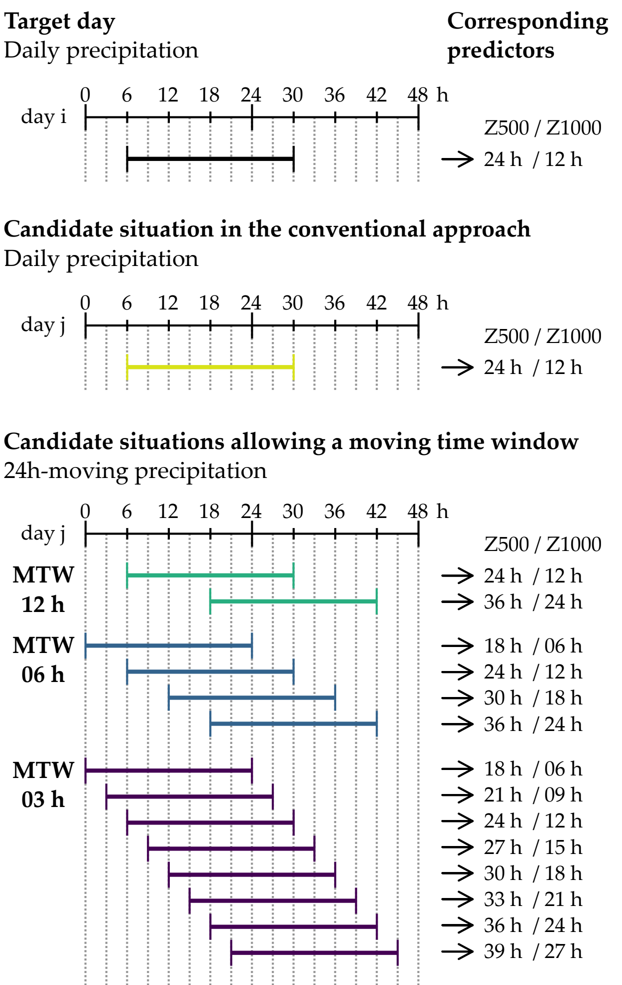

Figure 2. Illustration of the principle of a moving time window (MTW). The target situation is the same for the conventional approach and the MTW, while although the candidate situations with the MTW are 4 two, four, or eight times as many with the MTW. The larger horizontal bars represent the 24h precipitation totals; their associated predictors are listed on the right-hand side.

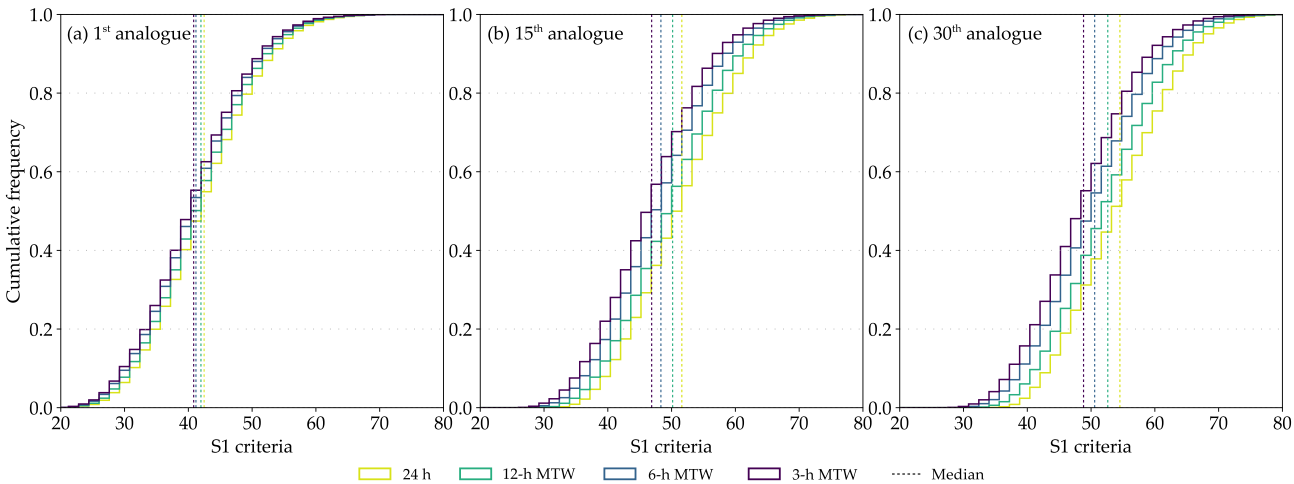

Figure 3. Changes in ~~the~~ S1 criterion distributions ~~due~~ owing to the ~~MTW~~ introduction of an MTW with a 12-h, 6-h, or 3-h time step. Distributions are provided for (a) the 1^{st}, (b) ~~5^{th}~~ 15^{th}, and (c) ~~20^{th}, and (d) 40^{th}~~ 30^{th} analogue ranks for all days of the CP at the Ulrichen station~~over the whole calibration period (1961–2008)~~.

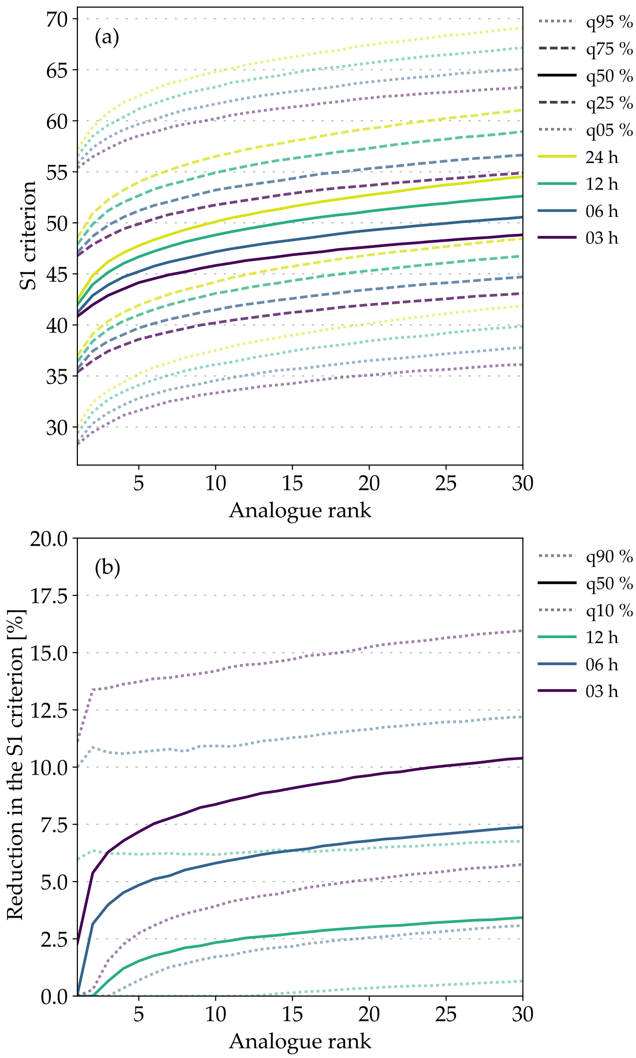

Figure 4. Synthesis of the changes in the S1 criterion ~~due~~ owing to the MTW with 12-h, 6-h, and 3-h time steps for the Ulrichen station depending on the rank of the analogue. (a) Quantiles of the S1 distributions corresponding to the MTW with ~~and without~~ different time steps; 24-h is the conventional approach without MTW. (b) Quantiles of the relative improvements of the S1 criterion when using the MTW compared with the conventional approach.

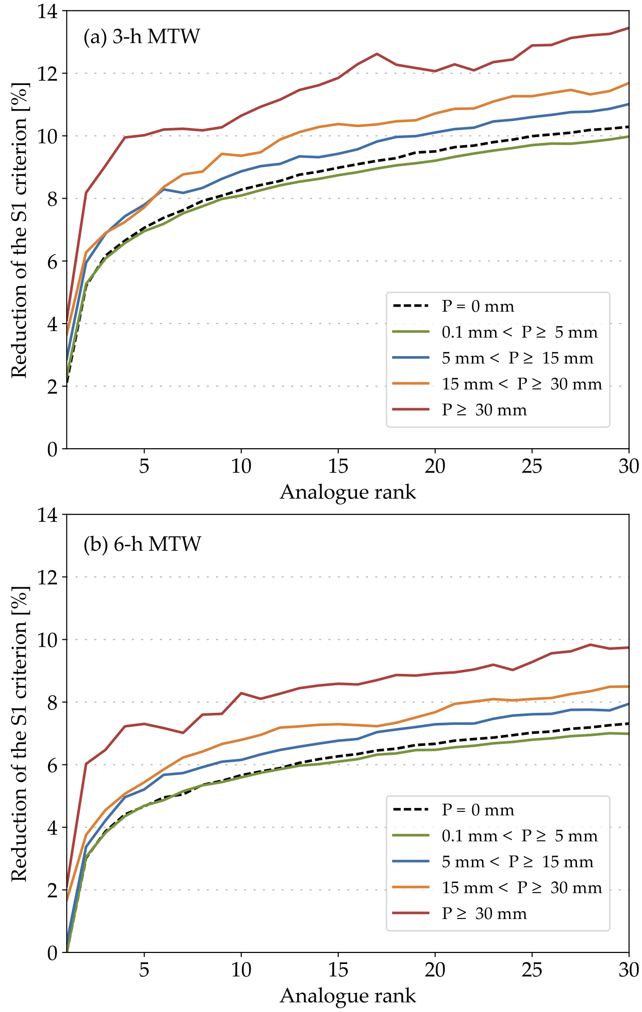

Figure 5. Distribution of the median improvements of the S1 criterion ~~due~~ owing to the (a) 3-h MTW and (b) 6-h MTW, depending the on daily precipitation thresholds at the Ulrichen station.

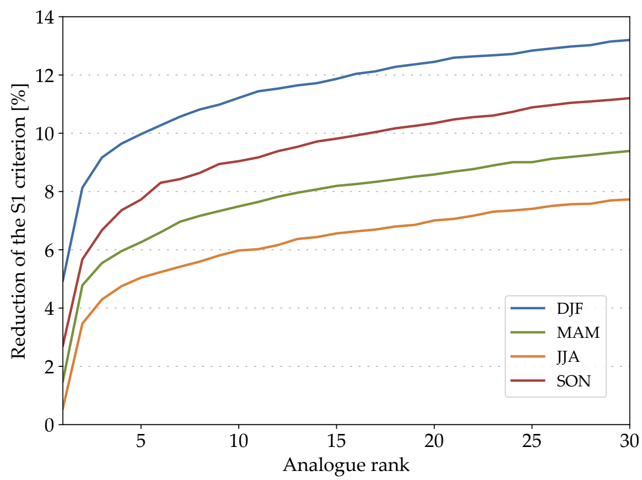

Figure 6. Seasonal effect on the median reduction of the S1 criterion for the Ulrichen station ~~due~~ owing to the MTW. DJF: winter~~,~~; MAM: spring~~,~~; JJA: summer~~, and~~; SON: fall.

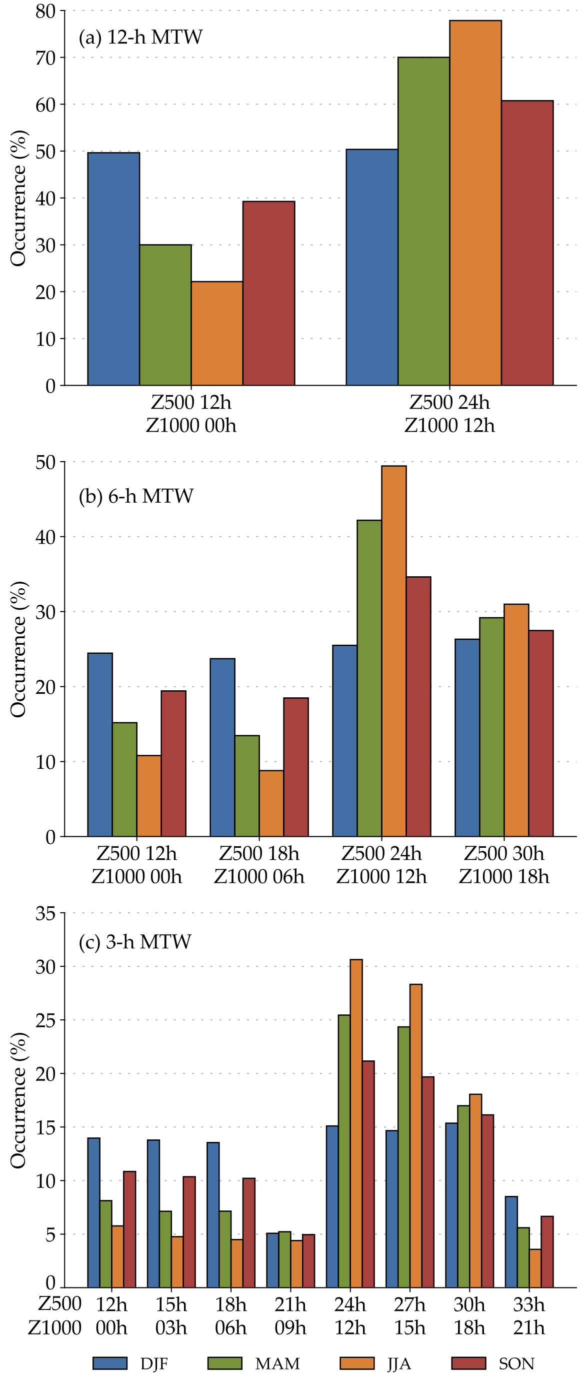

Figure 7. Distribution of the predictor hours on the selected analogue dates for the Ulrichen station when using a (a) 12-h MTW, (b) 6-h MTW, and (c) 3-h MTW, depending on the season.

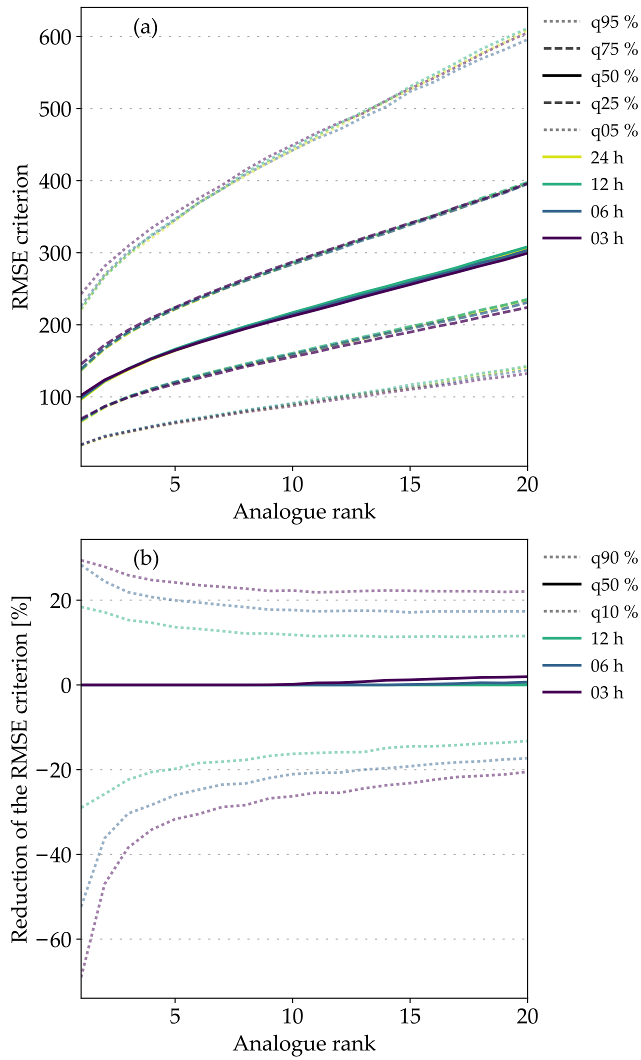

Figure 8. Synthesis of the changes in the RMSE criterion ~~due~~ owing to the MTW for the Ulrichen station, depending on the rank of the analogue. (a) Quantiles of the RMSE distributions with and without the MTW. (b) Quantiles of the relative improvements of the RMSE criterion when using the MTW.

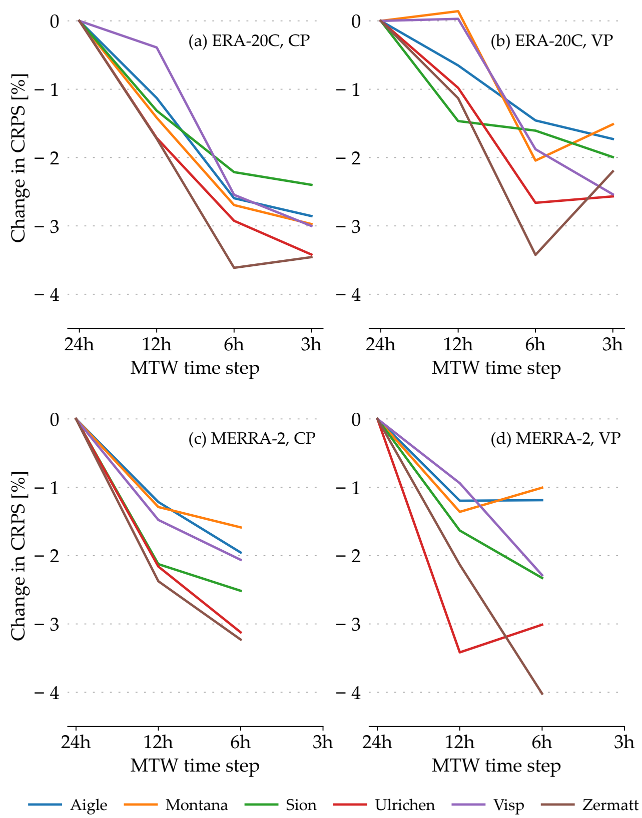

Figure 9. Differences of the CRPSS Changes in performance score due (CRPS) of the 2Z method at the different stations for MTW of varying time steps relative to the introduction conventional approach of 24 h. Results are provided for ERA-20C for the MTW as (afunction of daily precipitation thresholds at) CP and (b) VP as well as for MERRA-2 for the Ulrichen station(c) CP and (d) VP. The stars represent averagesLower CRPS values indicate better performance.

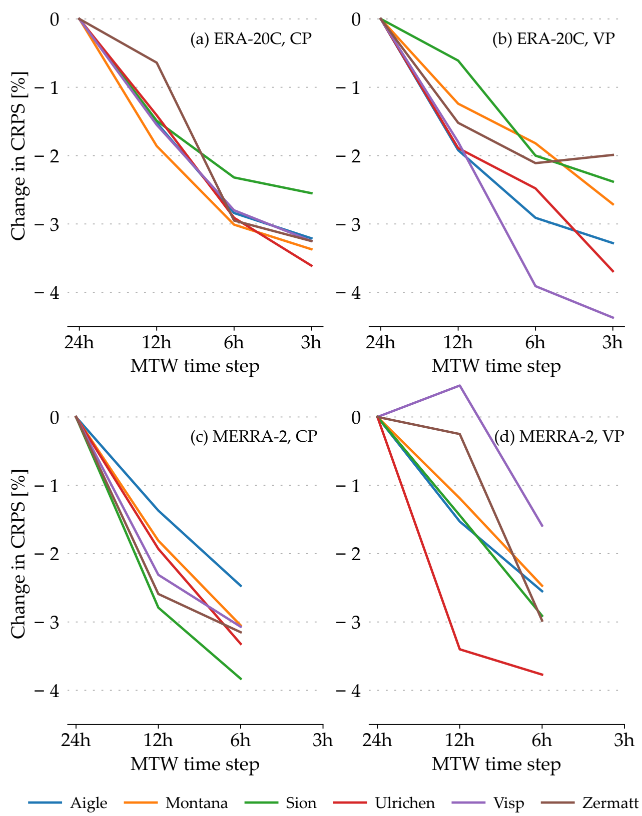

Figure 10. Same as Figure 9 but for the 2Z-2MI method.

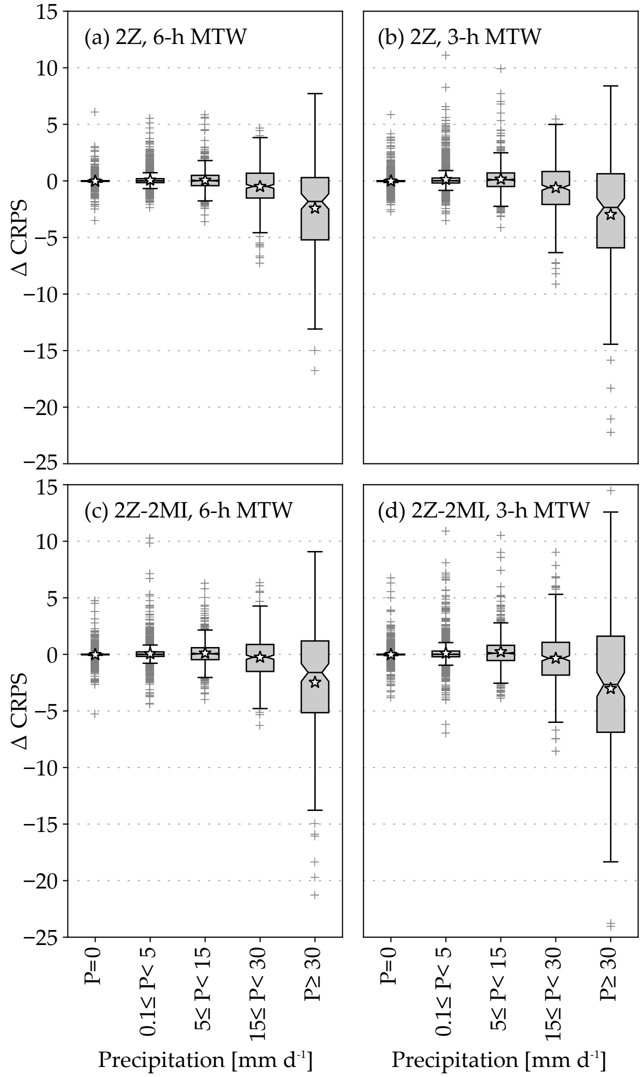

Figure 11. Differences in the CRPS values at the Ulrichen station owing to the introduction of the MTW as a function of different daily precipitation thresholds for the target date. The results are provided for the 2Z method (top) and the 2Z-2MI method (bottom) with a 6-h MTW (left) and a 3-h MTW (right). Stars represent averages. Lower CRPS values indicate better improvements.

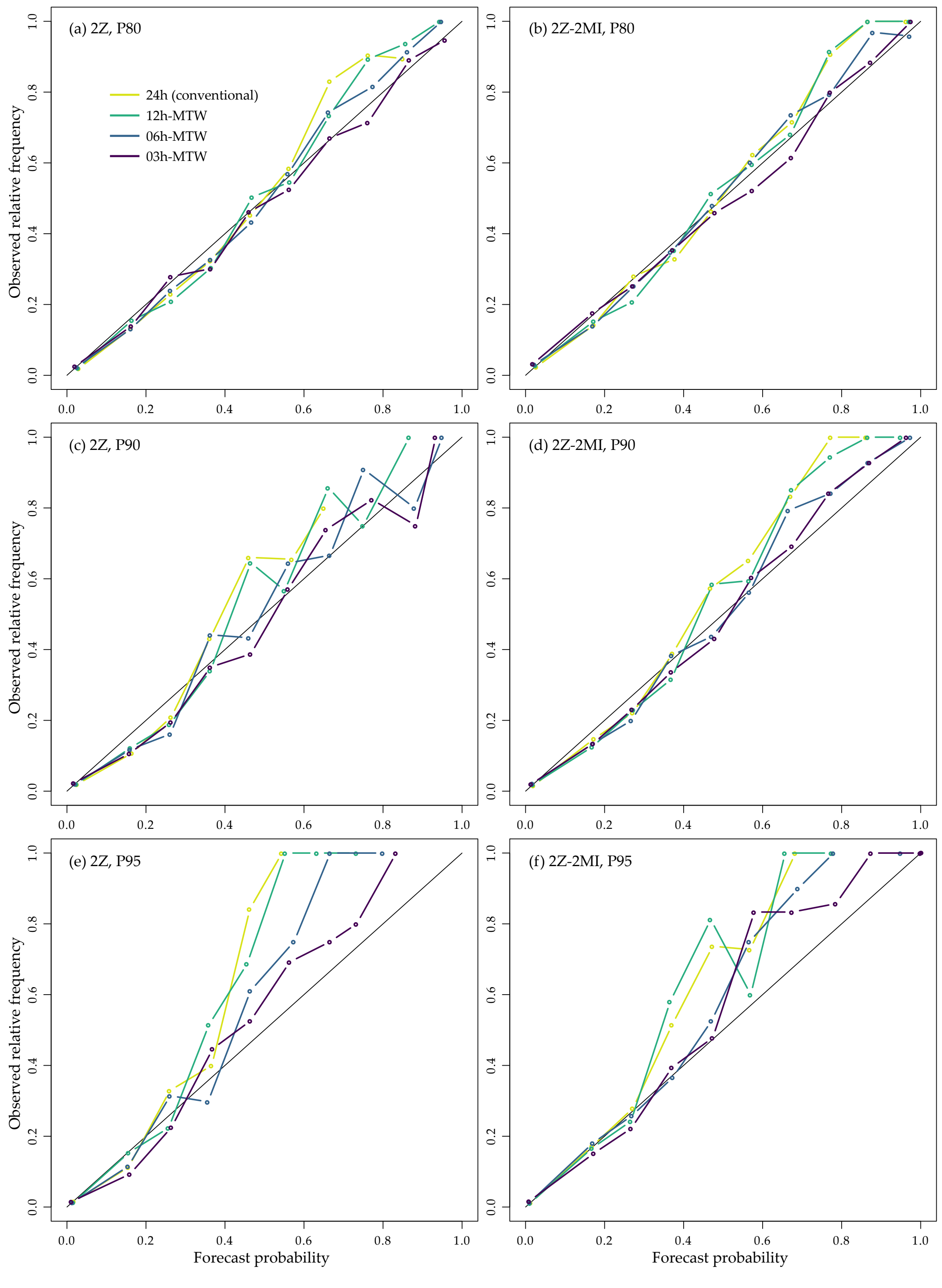

Figure 12. Reliability diagrams for the 2Z (left) and 2Z-2MI (right) methods and the prediction of the exceedance of percentiles 80 % (top), 90 % (middle), and 95 % (bottom) at the Ulrichen station for the VP. The conventional approach of 24 h is provided as well as the 12-h, 6-h, and 3-h MTWs.

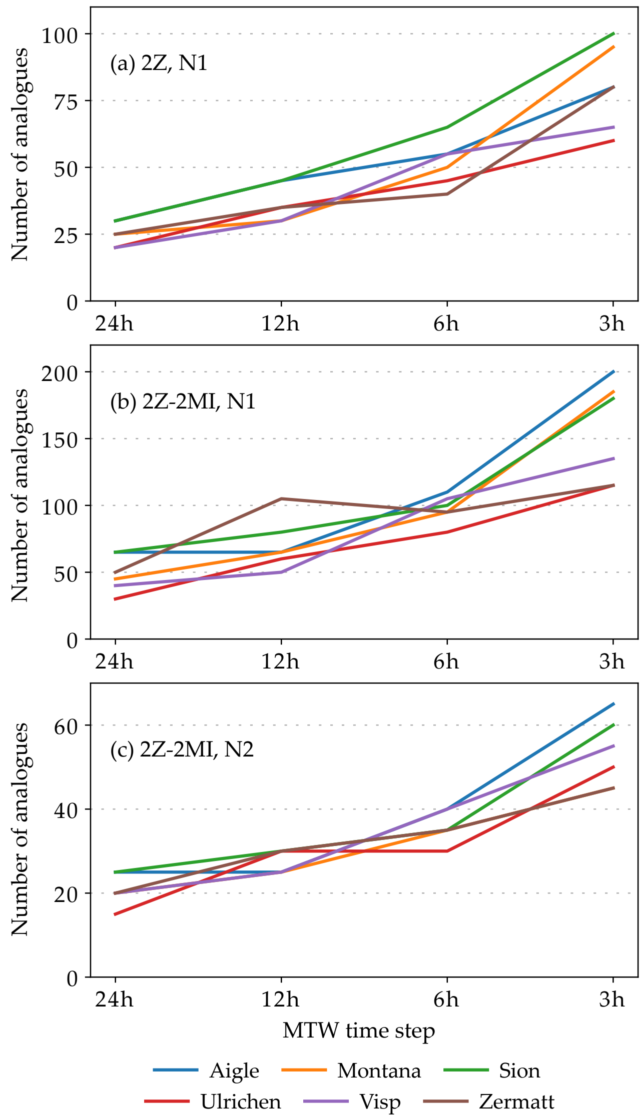

Figure 13. ~~Distribution~~ Optimal number of analogues of the ~~predictor hours on~~ first and second levels of analogy, N_1 and N_2, respectively, for the ~~selected analogue dates~~(a) 2Z and (b), ~~when~~ (c) 2Z-2MI methods after recalibration using the MTW~~, depending on the season for the Ulrichen station~~.

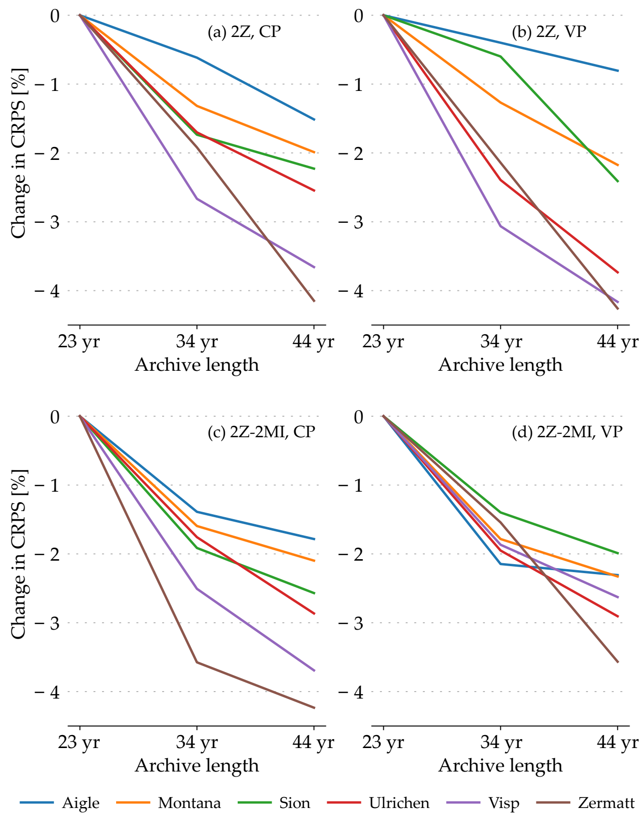

Figure 14. Change in performance score (CRPS) of the 2Z (top) and 2Z-2MI (bottom) methods at the different stations for the CP (left) and VP (right) for an increasing archive length with the conventional approach of 24-h window. Lower CRPS scores indicate better results.

Table 1. Parameters for the reference method on the atmospheric circulation (2Z). The first column ~~is~~ shows the level of analogy (0 for preselection)~~,~~. The following columns shown the ~~second column is the~~ meteorological variable ~~,~~ and ~~then~~ its hour of observation~~(temporal window). The~~ , the criterion used for the current level of analogy~~is then provided~~, ~~as well as~~ and the number of analogues.

Level	Variable	Hour	Criterion	Nb
0	±60 days around the target date			
1	Z1000	12 h	S1	N_1
	Z500	24 h		

Table 2. Parameters of the reference method with moisture variables (2Z-2MI). ~~Same~~ The conventions are the same as those in Table 1

Level	Variable	Hour	Criterion	Nb
0	±60 days around the target date			
1	Z1000	12 h	S1	N_1
	Z500	24 h		
2	TPW * RH850	12 h	RMSE	N_2
	TPW * RH850	24 h		

~~Optimal number of analogues (of the first and second level of analogy, respectively, N_1 and N_2) of the method based on the atmospheric circulation only (method 2Z) and the method with a second level of analogy with moisture variables (2Z-2MI) on the full archive (Standard), and after recalibration using the MTW (MTW-r). 2Z 2Z N_1 N_1 N_2 N_1 N_1 N_2Ulrichen 40 60 25 50 110 35Zermatt 35 55 25 55 80 30Visp 30 45 25 55 135 35Montana 40 55 30 55 110 40Sion 40 90 30 55 140 50Aigle 50 100 35 75 135 45~~

~~Values of the CRPSS (%) score for the original and the recalibrated parameters (with the sequential method, as described in Sect. 2.3) using the MTW approach on the disaggregated precipitation time series (short period) by means of the proportional distribution. original recalib. original recalib. Ulrichen 29.13 29.61 33.15 33.45 Zermatt 22.17 22.80 26.72 27.43 Visp 22.32 22.89 27.01 28.04 Montana 29.41 30.24 33.83 34.55 Sion 22.98 23.41 28.57 29.15 Aigle 29.07 29.46 34.66 35.09~~

~~Value of the coefficient of determination between the reconstructed 6-hourly precipitation time series using the listed variables and the accurate time series for the period 1982–2007. The grid points surrounding the region are: 1) 5° E, 47.5° N; 2) 5° E, 45° N; 3) 7.5° E, 47.5° N; and 4) 7.5° E, 45° N. The highest coefficient of determination is indicated in bold. –12 h –6 h 0 h +6 h +12 h 1 0.668 0.669 0.684 0.683 0.670 2 0.669 0.669 0.683 0.681 0.669 3 0.662 0.673 0.691 0.682 0.673 4 0.666 0.671 0.688 0.681 0.668 1 0.672 0.673 0.684 0.684 0.675 2 0.674 0.674 0.683 0.682 0.672 3 0.662 0.673 0.691 0.682 0.673 4 0.666 0.671 0.689 0.681 0.668 1 0.675 0.675 0.679 0.678 0.671 2 0.681 0.690 0.691 0.677 0.664 3 0.665 0.680 0.693 0.683 0.675 4 0.675 0.694 0.706 0.681 0.659 1 0.688 0.687 0.667 0.655 0.652 2 0.697 0.699 0.669 0.644 0.644 3 0.686 0.708 0.689 0.655 0.648 4 0.696 **0.721** 0.696 0.643 0.636~~

~~Values of the CRPSS (%) score for Zermatt for the original and the recalibrated parameters (with the sequential method, as described in Sect. 1) using the MTW approach on the disaggregated precipitation time series (short and long periods) by means of proxy variables from the reanalysis dataset. original recalib. original recalib. 1982–2007 22.57 23.14 27.11 27.71 1961–2008 23.81 24.38 28.42 28.86~~