

## Reply to comments of referee #1

First of all, thank you for carefully reading the manuscript and for the constructive feedback.

Some referee comments are recalled in italics and followed by the authors' responses. Others are addressed here, but without being recalled. The technical corrections or rephrasing are not discussed here, but will be performed.

- Introduction: The structure will be changed and the description of the method moved into "Data and method". Additional references will be added, and different variations will be presented, instead of relying on another paper to cover the literature review. The introduction will thus be rewritten.
- *The issue discussed in the manuscript goes beyond this "fixed/moving" window issue. It more generally consider influence of the size of the data set (including the length of the archive) on the prediction skill. I am not aware of works that explored this issue but references to those, if any, and associated results have to be presented in the introduction.*  
The issue of the archive length will be better described, according to the relevant literature.
- *The final objective of the work is to increase the prediction skill of the method. Other means are possible for this. They should be also mentioned.*  
We will mention other means to increase the prediction skill in the introduction.
- MTW strategy vs longer archive: We agree that the MTW might not be as performant as a 4 times longer archive, for the reasons provided. We indeed observed a gain in performance similar to doubling the archive. Your point on the transient climate is also interesting. We will add a discussion on these aspects in the paper.
- Number of analogs: we totally agree with the referee that when the choice of the predictors or the parameters are improved, leading to a better prediction, the optimal number of analog situations thus decreases. However, when the length of the archive increases, the optimal number of analogs increases too, for a better performance, up to a certain threshold. This has been demonstrated in the PhD of G. Bontron (2004, p. 227), and we can also see it in Fig.7 of Hamill et al. (2006). This increase in the number of analog situations results objectively in better performance skills in our case. The MTW enriches the pool of available situations, even though they are not fully independent. We will improve the clarity of this analysis. It is true though that the choice of predictors can be improved and auxiliary predictors are missing. The chosen methods are not the most recent ones, but are benchmarks as they have been used by several studies as references, and the results are easier to interpret than more elaborated methods. A note on that will be added.
- *Explain both terms of the decomposition. Another well-known decomposition of the CRPSS is that of Herbasch. Please mention it also and clarify the advantage of that of Bontron. It is not clear at this stage.*  
We will remove the analysis of the sharpness and accuracy, as it brings unneeded complexity without being very instructive.

- *Why using a non conventional notation for the CRPS score. Please use the classical one (CRPSS) or justify*  
It is the journal convention for variables in equations (“Multi-letter variables should be avoided. Instead use single-letter variables with subscript (e.g.  $E_{\text{RMS}}$  instead of RMSE, or  $E_{\text{T}}$  instead of ET).”). It has been changed elsewhere.
- Resolution of the atmospheric reanalysis: it was shown by some studies that the resolution of the reanalysis dataset does not improve significantly the performance of the analogy of the atmospheric circulation. However, a full analysis of this aspect is out of the scope of this publication, and is a topic we are working on right now. Nevertheless, we can assume that it does not alter the fact that we can find better analog situations at different hours of the day.
- Seasonal stratification: yes, we consider a seasonal stratification. It is mentioned in l. 36-39. This should be more clear after the change in the paper structure.
- *You use the mean intensity of rainfall as a proxy of “dynamism of the atmospheric situation”. Another proxy could be the intensity of the variations within the geopotential fields (e.g. mean gradient value). Why did you not explore this? It would better fit to your “dynamism” concept.*  
Yes, that could be a possibility. However, even though the link between the dynamism and the precipitation amount is not direct, the interest in analyzing it this way is that it highlights improvements we are directly interested in: a better prediction of high precipitation amounts. This will be rephrased to focus more on the precipitation thresholds rather than the dynamism of the circulation pattern.
- Sharpness and accuracy: We will remove the analysis of the sharpness and accuracy, as mentioned previously.
- Figure 11: this figure is complex and not very instructive. It will be removed along with the sharpness and accuracy analysis.
- Tables: we agree that there are many tables. We will remove some unanalyzed data, such as the spatial windows in Tables 3, 4, 7, and 8 in order to group the remaining information. Tables 10 and 11 will also be removed along with the sharpness and accuracy analysis.
- The other unmentioned detailed issues will be fixed

## Reply to comments of referee #2

The authors would like to thank Referee 2 for his/her positive comments on the manuscript.

Some referee comments are recalled in italics and followed by the authors’ responses. Others are addressed here, but without being recalled. The technical corrections or rephrasing are not discussed here, but will be performed.

- Sampling of extremes: The topic of extreme values within the analog method should be addressed in details, we agree on that and plan to work on it. However, it is out of the scope

of this paper. The proposed MTW improvement does not change the limitation of the maximum observed values in the archive, but it is not the topic of the present paper. We will however add a note on that issue in the paper.

- Introduction and applications in hydrology: The introduction will be rewritten, as explained to referee #1, and references will be added. Some applications to hydrology will also be cited.
- *Given the focus of the paper on the sub-daily time step, I suggest to refer more on the needs in urban hydrology.*  
The sub-daily time step is introduced in the search of analog situations, so at the predictors level. The predictand remains a 24-h precipitation total, which limits the application to urban hydrology.
- How the sample sizes N1 and N2 are determined and optimized: Indeed, this has not been detailed here. We will provide more insight on the calibration procedure.
- The other unmentioned issues will be fixed

## Reply to comments of referee #3

We would like to thank Referee 3 for his detailed and relevant review.

Some referee comments are recalled in italics and followed by the authors' responses. Others are addressed here, but without being recalled. The technical corrections or rephrasing are not discussed here, but will be performed.

- Introduction: as replied to referee #1, the introduction will be rewritten, with an improved literature review and context description. The method description will be moved to the methods section.
- *The title wrongly suggests that these conclusions are valid only in a forecasting context, while they actually have much more general implications.*  
We agree that the application of the MTW is not limited to forecasting. The title will be changed to a more generic term.
- The structure will be improved, with a better separation of methods, results and discussion.
- *Notations: Please use conventional abbreviations for commonly used quantities: Teweles-Wobus Score → TWS or S1, continuous ranked probability (skill) score → CRP(S)S, root mean square error → RMSE*  
It is the journal convention for variables in equations ("Multi-letter variables should be avoided. Instead use single-letter variables with subscript (e.g.  $E_{\text{RMS}}$  instead of RMSE, or  $E_{\text{T}}$  instead of ET)."). It has been changed elsewhere.
- The number of tables will be decreased, as explained to referee #1: "We will remove some unanalyzed data, such as the spatial windows in Tables 3, 4, 7, and 8 in order to group the

remaining information. Tables 10 and 11 will also be removed along with the sharpness and accuracy analysis, which is not very informative.”

- *L15-17: Is it not rather because heavy precipitation events are rarer?*  
We cannot exclude this argument, and it might be a mix of both factors. We will add a note on that aspect.
- *L24-25: I don't understand*  
This reports to section 4, but might not be necessary in the abstract as it brings some confusion.
- *L63: What are the parameters to calibrate? Please list them.*  
More details will be provided on the calibration procedure.
- *L70: Please either provide a peer-reviewed reference for this decomposition or detail it here.*  
We will remove the analysis of the CRPS decomposition, as it brings complexity without being very informative.
- *L76: Please detail the computation of the climatological distribution.*  
A description will be added.
- *L93-94: I don't understand.*  
This will be removed as it relies on partial analysis.
- *L113-114: Please justify the use of such an outdated global reanalysis (I understand this is partly for having a long time coverage). And add also the potential of using more recent and products with higher quality to the discussion.*  
See answer to referee #1. These points will be addressed in the discussion.
- *L168-174: This analysis is done for different classes of precipitation values. Whether this relates to the intensity of circulation dynamics is another issue.*  
We will reformulate this section.
- *L192-193: Figure 8 is not necessary. Please remove of put it in a supplementary material.*  
The figure will be removed.
- *L219-220, "No relationship [ : : ] criteria": I don't understand.*  
This will be removed.
- *L246-250: Is it shown somewhere in the manuscript?*  
No it is not shown, as the figure globally is very similar to Fig. 10.
- *L350-354: Given recent studies on RCM biases, I have serious doubts that RCM precipitation is reliable enough for it to be use as observed series in this context.*  
We agree with the referee and will change this sentence.
- Figure 11: We will drop the analysis of the CRPS decomposition, as it brings complexity without being very informative.

- *Table 13: The choice for preselecting these 4 points should be somehow justified.*  
You are right. These are simply the points surrounding the catchment. It will be explained.
- The other unmentioned issues will be fixed

## Significant changes in the manuscript

The manuscript was heavily reprocessed, which makes the changes tracking (next pages) a bit useless... The major points are:

- The whole structure changed, following the referees' advices. It should be much clearer now.
- The introduction has been re-written and include more references to other works.
- The decomposition in sharpness/accuracy has been removed, along with its figure.
- 8 tables were removed and some information were integrated into a new figure.
- The discussion has been extended to cover several questions asked by the referees.
- Globally, there is almost not a sentence from the original document that hasn't changed.
- The document has also been through English editing.

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

Pascal Horton<sup>1,2</sup>, Charles Obled<sup>3</sup>, and Michel Jaboyedoff<sup>1</sup>

<sup>1</sup>University of Lausanne, Institute of Earth Sciences, Lausanne, Switzerland

<sup>2</sup>University of Bern, Oeschger Centre for Climate Change Research, Institute of Geography, Bern, Switzerland

<sup>3</sup>Université de Grenoble-Alpes, LTHE, Grenoble, France

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

**Abstract.** ~~The Analogue Method (AM) aims at forecasting~~ Analogue methods (AMs) predict local weather variables (predictands), such as ~~precipitations~~precipitation, by means of a statistical relationship with predictors at a synoptic scale. The analogy is generally assessed ~~in the first place on the geopotential field by mean of a comparison of the gradients on gradients of geopotential heights first~~, in order to sample ~~the~~ 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 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 ~~constraint being main reason is~~ the use of daily ~~time series;~~ precipitation time series due to the length of ~~available archives they provide~~ their available archives, and the unavailability of equivalent archives at a finer time step. However, it is unlikely ~~that for~~ the best analogy ~~happens to occur~~ at the very same hour, ~~but it may~~ while it may be found with a time shift of some hours as it can occur at a different time of ~~the~~ day. In order to assess the potential ~~of for~~ finding better analogues at a different hour, a moving time window (MTW) has been introduced ~~on a reduced archive of hourly precipitation totals.~~

The MTW resulted in a better analogy in terms of the atmospheric circulation, with improved values of the analogy ~~criteria on the whole~~ criterion on the entire distribution of analogue dates. The improvement was found to grow with the analogue ~~ranks~~ rank due to an accumulation of ~~more similar situations~~ better analogues in the selection. ~~Moreover, the improvement is even more important for days with heavy precipitation events, which are generally related to more dynamic atmospheric situations, where timing is more specific.~~

A seasonal effect has also been identified, with larger improvements in winter than in summer, supposedly due to ~~the stronger effect of the diurnal cycle in summer;~~ stronger diurnal cycles in summer that favour predictors at the same hour for ~~target and analogues~~ the target and analogue days.

The impact of the MTW on ~~the prediction performance~~ prediction skill has been assessed by means of a sub-daily precipitation series transformed into moving ~~24 h-totals at a 6-hourly time step.~~ This resulted in an improvement of the prediction skills, which were even larger 24h totals at 6h time steps. The prediction skill was found to have improved by the MTW, and even to a greater extent after recalibrating the AM parameters. Moreover, the improvement was greater for days with heavy

precipitation, which are generally related to more dynamic atmospheric situations where timing is more specific and which are fewer in the meteorological archive.

However, ~~attempts to reconstruct longer precipitation series of running 24 h-totals by means of simple methods failed. It emphasized~~ 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 ~~a the~~ near future, either by means of growing ~~observed archives, archives of measurements~~ or by the establishment of ~~precipitation reanalyses through regional modeling. Then, the use of a regional precipitation reanalysis data at sub-daily time step.~~

~~The use of the~~ MTW in the AM ~~should be considered for any application, especially when the prediction quality of extreme events is important~~ can already be considered now for several applications in different contexts, may it be for operational forecasting or climate-related studies.

## 1 Introduction

~~The Analogue Method (AM) is~~ Analogue methods (AMs) are based on the ~~principle hypothesis~~ that two relatively similar synoptic situations may produce ~~relatively~~ similar local effects (Lorenz, 1956, 1969). ~~They are statistical downscaling methods (Maraun et al., 2010) and consist of finding past situations that are similar to the target day of interest in terms of the atmospheric circulation or other synoptic predictors. The local weather variables of interest (predictand) that 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 methods exist and listings can be found in Horton et al. (2016) or ?. The versions that will be considered here are often used as references or benchmarks for various improvements.~~ method exist (a non-exhaustive listing can be found in ~~The 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 et al., 2008; García Hernando et al., 2010; García Hernando et al., 2011; García Hernando et al., 2012; García Hernando et al., 2013; García Hernando et al., 2014; García Hernando et al., 2015; García Hernando et al., 2016; García Hernando et al., 2017; García Hernando et al., 2018; García Hernando et al., 2019; García Hernando et al., 2020; García Hernando et al., 2021; García Hernando et al., 2022; García Hernando et al., 2023; García Hernando et al., 2024; García Hernando et al., 2025). Other applications focus on temperatures (Radinovic, 1975; Woodcock, 1980; Kruizinga and Murphy, 1983; Delle Monache et al., 2013; Chardon et al., 2014; Dayon et al., 2015; Dayon et al., 2016; Dayon et al., 2017; Dayon et al., 2018; Dayon et al., 2019; Dayon et al., 2020; Dayon et al., 2021; Dayon et al., 2022; Dayon et al., 2023; Dayon et al., 2024; Dayon et al., 2025). 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 (Obled and Good, 1980; Bolognesi, 1993), insolation (Bois et al., 1981), and the trajectories of tropical cyclones (Keenan and Woodcock, 1981; Sievers et al., 2000; Fraedrich et al., 2003).~~

~~The first method~~ 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 can provide multivariate predictions that are physically consistent (Raynaud et al., 2016). Their temporal transferability has also been the topic of 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. 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, 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 the quality of the analogy, the archive length, and the size of the spatial domain (or degrees of freedom): errors increase with a bigger 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 km<sup>2</sup>, there is no need to consider meteorological domains larger than 10° to 20°. For that reason, smaller spatial windows are always considered when searching for analogues, and the archive length is maximized.

Therefore, due 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 analogue situations are assessed by comparing predictors at fixed hours of the day. However, it can be expected that the 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 hour. With this assumption, a moving time window (MTW) was introduced to allow the search for candidates at different hours of the day. Previous tests showed the benefit, in terms of analogy criterion values, of searching for analogue synoptic situations at a finer time step, but without assessing the impact on the prediction skill (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). AMs can also be combined with other methods (see 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., 2016). However, the MTW technique is not in competition with other approaches and can be combined with these. Indeed, as most of them rely on the atmospheric circulation in the first level of analogy, the application of a MTW should lead to similar conclusions.

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 (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.

## 2 Data and methods

### 5 2.1 Study area and data

The study area is the upper Rhône catchment in Switzerland. Precipitation time series come 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 years (1982–2007) and were also obtained at a daily time step (from 6:00 UTC to 6:00 UTC the next day) for 47 years (1961–2008). Due 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 1 dataset (Kalnay et al., 1996) with a 6h temporal resolution, 17 pressure levels, and a spatial resolution of 2.5°. This dataset is now relatively old, but it is not expected to affect the conclusions of the present study (see discussion in Sect. 4.4).

### 15 2.2 The considered analogue method

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 analogue situations for (Table 1, Obled et al., 2002; Bontron and Obled, 2005). Searching for analogue situations to a target day starts by a seasonal stratification through a preselection step of the potential candidate candidates for analogy. Here, it has been limited to the 4 months centered around the target date for every year of the archive, in order to cope with seasonal effects. Then, 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 ( $S_{TW}$ ), which  $S_1$  that is a comparison of gradients on geopotential heights over a certain spatial window and at certain hours. The smaller the  $S_{TW}$ :

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

25 where  $\Delta \hat{z}_i$  is the forecast geopotential height difference between the  $i$ th pair of adjacent points from the grid of the target situation,  $\Delta z_i$  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  $S_1$  values, the more similar the pressure fields.

The predictand, here a 24 h precipitation total, results from meteorological situations that are continuous transitions between changing circulation patterns. Thus, the observed predictors fields, there is greater similarity in the pressure fields.

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

Then,  $N_1$  dates with the lowest values of  $S_{TW-S1}$  are considered as analogues to the target day,  $N_1$  being a parameter to calibrate. Then (see Sect. 2.1). Finally, the daily observed precipitation amount of the corresponding dates provide the empirical conditional distribution considered as the probabilistic forecast prediction for the target day. Following the nomenclature proposed by Horton et al. (2016), this This method will be named 2Z.

The second reference method (2Z-2MI, Table 2) adds a subsequent level of analogy on with moisture variables, compared by means of the root-mean-square error ( $E_{RMS}$ )(method 2Z-2MI, Table 2). This second RMSE:

$$S_{CRP} = \int_{-\infty}^{+\infty} [F(y) - H(y - y^0)]^2 dy, \quad (2)$$

where  $\hat{v}_i$  is the  $i$ th 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 made composed of the product of the total precipitable water (TPW) with the relative humidity at 850 hPa (RH850) (Bontron, 2004). When adding a second level of analogy,  $N_2$  dates are subsampled in from the  $N_1$  analogues on the atmospheric circulation, to end up with resulting in a smaller number of analogue situations. When a second level of analogy is added, a higher number of  $N_1$  analogues is kept on the first level.

To calibrate the method, the commonly used technique is a semi-automatic sequential procedure elaborated by Bontron (2004). AtmoSwing (Horton et al., 2016), which implements this technique, was used in order to calibrate the parameters in the present study. The score that is most often considered to assess an AM performance is the Continuous Ranked Probability Score ( $S_{CRP}$ , Brown, 1974; Matheson and Winkler, 1976; Hersbach, 2000) More complex AMs exist with additional predictors (see e.g. Horton, 2012; Ben Daoud et al., 2016; Caillouet et al., 2016). 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, it is easier to interpret the impact of the MTW using more basic methods.

### 2.3 Performance score

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 evaluating the the evaluation of the predicted cumulative distribution functions  $F(y)$ , for example of the precipita-

tion values  $y$  from analogue situations, compared to the observed value  $y^0$ . ~~The better the forecast, the smaller the~~ A better prediction has a smaller score. It is defined as follows:

$$S_{\text{CRP}} = \int_{-\infty}^{+\infty} [F(y) - H(y - y^0)]^2 dy, \quad (3)$$

where  $H(y - y^0)$  is the Heaviside function that is null when  $y - y^0 < 0$ , and has the value of 1 otherwise. ~~Bonton (2004) proposed a decomposition of this score interpreted into a sharpness and an accuracy terms.~~

In order to compare the value of the score in regard to a reference, one often considers its skill score expression, ~~and use 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 ( $S_{\text{SCR P}}$ ) is thus defined as following~~ continuous ranked probability skill score (CRPSS) is defined as follows:

$$S_{\text{SCR P}} = \frac{S_{\text{CRP}} - S_{\text{CRP}}^r}{S_{\text{CRP}}^p - S_{\text{CRP}}^r} = 1 - \frac{S_{\text{CRP}}}{S_{\text{CRP}}^r} \quad (4)$$

where  $S_{\text{CRP}}^r$  is the  $S_{\text{CRP}}$  value for the reference (climatological distribution) and  $S_{\text{CRP}}^p$  ~~would be the one is~~ for a perfect ~~forecast prediction~~ (which implies  $S_{\text{CRP}}^p = 0$ ). A better ~~forecast prediction~~ is characterized by an increase in  ~~$S_{\text{SCR P}}$~~  CRPSS:  $S_{\text{SCR P}} = 1$  for a perfect prediction and  $S_{\text{SCR P}} < 0$  for a prediction with a lower skill than the reference.

~~The AM is usually implemented with a daily time step, due to the availability of long precipitation archives that have no equivalent at a finer resolution. Therefore, the analogue situations are assessed on the basis of a daily time step, by comparing predictors at fixed hours of the day, otherwise one would not know what precipitation values to assign to them. However, it can be expected that the 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 hour. On this assumption, a moving time window (MTW) was introduced to allow searching for candidates at different hours of the day~~

## 20 2.4 The moving time window (MTW) approach

The moving time window (MTW) technique aims at finding better analogue situations at different hours of the day rather than comparing the predictors at the same fixed hours. If the target situation is kept at 12:00 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).

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.

~~Previous tests showed the benefit, in terms of analogy criteria values, of In order to assess the benefit of searching for analogue synoptic situations at a finer time step, but without assessing the impact on the forecast skills (Finet et al., 2008). In order to assess this potential improvement of the prediction, precipitation data with a resolution of 10 minutes, on a respectable archive length, was used sub-daily time step for quantitative precipitation prediction, an appropriate precipitation series is~~  
5 ~~required. On the basis of high temporal resolution time series (Sect. 2.1). It was then aggregated in the form of 24 h-totals, but starting at different 6-hourly time step, 24h totals were processed at a 6h time step by means of a moving total (Sect. ??).~~

~~One can question the interest of using daily totals when a 6-hourly precipitation series can be used. The first reason is that the 6-hourly time series generated by the AM may not represent well the dynamic of the accurate precipitation (results not shown); 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 smoothed signal. Then, sometimes one just does not need finer resolution than the daily time step. Finally, when using a reconstructed precipitation archive, the errors in intra-day precipitation distributions have a lesser impact over 24 h daily totals smaller archive length and the potentially better analogy between target and MTW candidates needs to be assessed.~~

## 15 **2.5 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 MTW is explained in Sect. ?. As this method requires sub-daily time series, 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 available archive length is reduced compared to more common daily measurements, which has some consequences on the performance score (Sect. 3.1). The benefit of introducing a MTW was assessed first in regards to the analogy criteria improvement (Sect. 3.1), and then in terms of precipitation prediction skill original daily time series were allocated into a new moving average of 24h totals (Fig. 3).~~

~~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. 3.2). Finally, attempts were made to reconstruct longer 24 h-total precipitations series (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.~~

### 3 Data and method

#### 2.1 Calibration of the analogue method

#### 2.2 Study area and data

The semi-automatic sequential procedure elaborated by Bontron (2004) was used to calibrate the AM (see also Radanovics et al., 2013; Be

5 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:

The study area is the upper Rhône catchment in Switzerland (see also Horton, 2012). Due to the low density of weather stations with high temporal resolution and long archives, no spatially aggregated rainfall was processed. The time series (on  
10 the period 1982–2007) come from

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 automatic weather stations, namely Ulrichen, Zermatt, Visp, Montana,  
15 Sion and Aigle (Fig.1), which are subject to various meteorological influences (Horton et al., 2012). The results will thereafter be presented arbitrarily for the Ulrichen station, but they are equivalent for all stations. The precipitation time series were then aggregated by means of a moving 24 h-total (see next section); :00 or 12:00 UTC),

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

Predictors are extracted from the NCEP/NCAR reanalysis I (Kalnay et al., 1996) dataset with a 6-hourly temporal resolution,  
20 17 pressure levels, and a spatial resolution of 2.5.

#### 2.2 The moving time window (MTW) approach

In order to assess the benefit of searching analogue situations at a sub-daily time step, an appropriate precipitation series is required. On the basis of high resolution time series (Sect. 2.1), 24 h-totals were processed, but at a 6-hourly time step (temporal resolution of the reanalysis dataset), by means of a moving 24 h-total.

25 The target situations and their corresponding observed precipitation values (used for validation)

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) over a large domain. Every point (or cell) of the full domain is assessed on all predictors

of the 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 until no improvement is reached.

5 (c) The number of analogue situations  $N_1$  is then reconsidered and optimized 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 (calibration of the spatial window and the number of analogues) for the new level. The parameters calibrated on the previous analogue levels are fixed and do not change, because the forecast is always established for a fixed period of .

10

4. Finally, the target day (6–30 h), as before. The difference is that the candidates are 4 times more numerous (Fig. 2). Based on some preliminary tests (not shown), no constraint was added in order to restrict the selection of multiple analogues within the same candidate dates numbers of analogues  $N_1$  and  $N_2$  for the different levels of analogy are reassessed. This is done iteratively by varying the number of analogues of each level in a systematic way.

15

The calibration is done in successive steps with a limited number of parameters. Previously calibrated parameters are generally not reassessed (except for the number of analogues). More advanced techniques, such as using genetic algorithms (Horton et al., 2016), exist but are out of the scope of the present study.

### 3 Results

#### 3.1 Consequences of the archive reduction

20 ~~As sub-daily precipitation time series are usually available on a shorter period than traditional daily time steps, the first assessment consists in assessing the loss of performance resulting from a~~ The performance loss resulting from the reduction of a 47 years-year archive (1961to–2008, corresponding to the daily time series) to 25 years (1982to–2007). This change is, corresponding to the hourly time series in Sect. 2.4) was assessed with the original method ~~without MTW.~~

25 Both 2Z (Table 1) and 2Z-2MI (Table 2) methods were considered. The AM parameters were calibrated ~~using AtmoSwing (Horton et al., 2016) on the original archive (Tables ?? and ??) and will be~~ see resulting analogue numbers in Table 3) and were used thereafter.

The impact of the ~~change in the archive length is summarized in Table ?? for both~~ archive reduction can be seen in Fig. 4 for the 2Z and method and in Fig. 5 for the 2Z-2MI methods method. As expected, a loss of performance ~~can be was~~ observed for each station, except for that of Aigle, which ~~seems seemed~~ relatively indifferent to this change. This loss ~~is was~~ globally significant, with up to ~~-1.89 points–1.89 points (% of CRPSS)~~ for Visp and the 2Z method.

30

## 4 Results and discussion

### 3.1 Influence of the MTW on the analogy criteria criterion

#### 3.1.1 Changes in the atmospheric circulation analogy

When searching for analogues in the first level of analogy, such as on the geopotential heights  $\zeta$  as in the 2Z method (Sect. 2.2), there are now 4 times more candidates than before, which obviously allows to find better matches.

Figure ?? presents as many candidates (even though not fully independent) with the MTW than before (Sect. 2.4). Figure 6 shows the changes in the distributions of the  $S_{TW}$  criterion analogy criterion (S1) for the 1<sup>st</sup>, 5<sup>th</sup>, 20<sup>th</sup>, and 40<sup>th</sup> analogue for analogue rank at the Ulrichen station on over the whole calibration period  $\zeta$  (1961–2008 – the full period could be used here, as precipitation was not considered at this stage), due to introduction of the MTW. The precipitation target remains as before, that is centered on 18 h UTC (6 h UTC to 6 h UTC the next day). The shapes shape of the distributions of the conventional approach and the MTW are were found to be similar, but the values of the analogy criteria are now criterion were reduced (shifted to the left) and therefore better. An increase in and were, therefore, better. For analogues with higher ranks (e.g. 20<sup>th</sup> or 40<sup>th</sup>), the difference between a fixed window and a moving window is identifiable the two distributions was larger than for the first ranks, which means that the last analogues are further improved. The latter effect is due to the accumulation of improvements brought by the new analogue situations in the selection improvement increased with the rank of the analogues.

The improvements of the  $S_{TW}$  criteria analogy with the rank of the analogues are summarized in Fig. 7, which shows (top) quantiles of the  $S_{TW}$  criteria according to the analogue rank S1 criterion for the conventional method and the MTW, and (bottom) quantiles of the relative reduction (meaning improvement) due to the MTW. This confirms that all quantile seem quantiles were similarly reduced ( $S_{TW}$  S1 distributions keep their shape), and that this improvement is was constantly increasing from the first to the last analogue (Fig. 7 bottom).

The median of the  $S_{TW}$  values reduction (Fig. 7 bottom) starts approximately at 5 % for the first analogue and reaches more than 10 % for the last one. This increasing trend with the analogue rank can be explained by the accumulation of better analogues in the distribution. The minimum improvement starts from 0 and reaches 2–3%, meaning that the criteria have been improved on most analogues for every day of the calibration period. All other stations have a similar improvement of the  $S_{TW}$  criteria, both in terms of distribution shape and amplitude.

#### 3.1.2 Influence of the weather situation

##### 3.1.2 Changes per precipitation class

It can be assumed that the atmospheric conditions with a low dynamism, such as the frequent anticyclonic situations, will not be radically improved by the introduction of the MTW. Conversely, dynamic situations, such as weather disturbances, have a well marked temporal evolution. Indeed, the position of the driving elements such as the low-pressure center and the fronts change

significantly during a day. We can therefore expect to improve more significantly these situations with a higher dynamism when introducing a MTW, as better matches to the target situation may be found.

The dynamism of a given atmospheric condition cannot be easily quantified. A basic assumption is considered here, which stipulates that the more a day is rainy, the more dynamic the situation is

5 The impact of the MTW on the analogy criterion has been analysed per precipitation classes (for the target day). The results of this analysis are summarized are summarised in Fig. 8 by the median reduction of  $S_{TW}$  ~~for S1~~ for the days with precipitation (organized into classes) between two thresholds. ~~The~~ With the number of cases per class being reduced, the curves are not as smooth as in previous ~~analyzes~~ analyses. It is nevertheless clear that the ~~improvement tends to increase on~~ improvements were larger for days with higher precipitation. ~~This is true for all our stations and confirms our intuition~~ Once again, the results for  
10 all the other stations were similar.

### 3.1.3 ~~Seasonal effect~~ Changes per season

Atmospheric dynamics ~~varies~~ vary greatly from one season to another, ~~which reflects~~ reflecting on the performance of the AM that is generally lower between June and August (Bliefernicht, 2010). ~~It therefore makes sense to verify the~~ The effect of the MTW ~~separately per season.~~ on the S1 criterion per season is presented in Fig. 9. Differences between the seasons were  
15 substantial, with greater improvements for winter than summer.

~~A seasonal effect can be observed on the reduction of the  $S_{TW}$  criteria due to the MTW (Fig. 9). The improvements are greater 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 based on the daily cycle and good analogues are essentially found for the same hours or time windows.~~

20 ~~An analysis of the selected hours for the geopotential predictor seems to confirm this assumption (Fig. 12). It was found that the new choice of the temporal window in winter, when using the MTW approach, is well balanced between the 4 options. This means a change of 75% of the analogues selection compared to the conventional approach, which improves the circulation analogy.~~

~~On the contrary, the summer months have a preference for the initial temporal window (Z500 24 h & Z1000 12 h), due to more pronounced diurnal effects which reduces the potential for improvement of the criteria. Other seasons are 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 generally even with a slightly larger amplitude than for Ulrichen (see example of the Visp station in Fig. ??).~~

### 3.1.4 Changes in the moisture analogy

30 When ~~considering~~ adding the second level of analogy of the 2Z-2MI method (Table 2), the number of ~~candidates~~ candidate situations did not increase, as they ~~are subsampled in the previously~~ were conditioned by the  $N_1$  previously selected analogues, but their dates ~~have had~~ changed. In contrast to ~~earlier~~ the AM on the atmospheric circulation only, both a reduction or an increase of the  ~~$E_{RMS}$~~  RMSE analogy criterion values ~~are possible~~ were possible with the MTW compared to the static approach.



Indeed, Fig. 10 shows an almost insignificant improvement of the  $E_{RMS}$ -RMSE values. Unlike for the first level of analogy, the relative improvement of the  $E_{RMS}$ -values are changes of the RMSE values were distributed relatively symmetrically around zero, with improvements and losses of the same amplitude. Once again, the results for the other stations are similar.

This result of a globally null improvement of the  $E_{RMS}$ -values does not mean that the 2Z-2MI method cannot be improved by the MTW. It means that after the selection of the analogues situations in terms of the synoptic circulation, the new candidate dates do not allow to find better analogues in terms of moisture. However, the selected dates have changed in the first level of analogy, and thus also in the final selection, which can potentially improve the performance scores on the predictand.

### 3.2 Impact on performance scores

#### 3.2 Impact of the MTW on prediction skill

A systematic improvement of the  $S_{TW}$  values was previously observed. However, finding better analogue situations does not obligatory imply better skills to predict the precipitation. The new performance scores (CRPSS) of the precipitation prediction are provided in Fig. 4 and 5 for the 2Z and 2Z-2MI methods, respectively, for the reduced archive (1982–2007). Therefore, the impact of the MTW introduction, and thus the selection of other analogue dates, has to be assessed on the performance scores. In order to perform this assessment, the 24 h-totals (moving average) at a 6-hourly time step were used (see Sect. 2.1). The target dates remain unchanged, since the original time slot (6 h UTC–6 h UTC the next day) is kept, and thus the performance scores can be directly compared with the former ones.

The new performance scores are provided in Table ??, along with the differences regarding the conventional method with the same archive length. The differences ranges. 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 introduction of the MTW brings an improvement of the performance that is not very large, but that is nevertheless significant. Moreover, it requires no additional predictor. No relationship was found between the improvement of the score and the reduction of the  $S_{TW}$  criteria, neither with the season.

##### 3.2.1 Improvement by per precipitation classes

The  $S_{TW}$  criteria S1 criterion was previously found to be improved to a greater extent for the most dynamic situations related with show greater improvement for higher precipitation values (Sect. 3.1.2). The changes in terms of performance scores will now be assessed regarding precipitation thresholds 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 behavior.

The increasing positive trend of skill improvements regarding the precipitation threshold is significant and shows that the prediction of higher precipitation totals is further improved. Thus, both the analogy criteria and the performance scores are improved to a greater extent for heavier precipitation events. On the contrary, the 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 accumulations are precipitation values, the performance scores were not improved.

### 3.2.2 Effect Recalibration of recalibrating the 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 ~~may-might~~ change the optimum value of some parameters. The calibration ~~has-then-been~~ (see Sect. 2.1) ~~should then be~~ reprocessed.

5 After recalibrating, ~~some~~ changes in the ~~optimal parameters can~~ AM parameters could be observed for both ~~the~~ 2Z (Table ??) and 2Z-2MI methods (Table ??). Among these, the ~~east-west dimension zonal extent~~ of the spatial windows of the circulation analogy ~~tends to decrease~~ decreased slightly (not shown). More significantly, the optimal number of analogues ~~increases~~  $N_1$  and  $N_2$  increased after introducing the MTW ~~, of a significant order of magnitude by a considerable magnitude~~ (Table 3): 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  
10 the first analogy level of the 2Z-2MI method ~~even~~ reached three times its previous value for the Visp station. ~~It seems as if the method perceive the increase in the archive length and diversity by allowing to extract more analogues. As we saw in Fig. 7, the~~ ( $N_1 = 135$  instead of 45; Table 3).

The values of the CRPSS scores for both methods (Fig. 4 and 5) have significantly increased after recalibration. When analysing the change in performance per precipitation class, the results (not shown) were very similar to the observations in  
15 Sect. 3.2.1, 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  
20 in performance score.

When using the first reconstructed time series based on the proportional distribution (see Sect. 2.5), the AM performance was degraded and was even below the conventional method without MTW (Table 4 to compare to Fig. 4 and 5).

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. 2.5), the best  
25 proxy was the precipitable water at 45° N and 7.5° E, with a time offset of -6 h (Table 5). Table 6 presents the CRPSS scores obtained by the disaggregated series using the proxy of precipitable water which are also lower than the conventional method.

## 4 Discussion

### 4.1 Improvement of the selection of analogue situations

For the first level of analogy, the improvement of the  $S_{TW}$  criteria grows along with the rank of the analogue, which shows an  
30 accumulation of S1 criterion (Fig. 7 bottom) started approximately at 5 % for the first analogue and reached more than 10 % for the last (40<sup>th</sup>) (Sect. 3.1.1). This increase in improvement with the analogue rank can be explained by the accumulation of

better analogues in 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 introduction of the MTW allows finding better analogue situations in the distributions.

5 ~~It seems thus profitable to widen~~ 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. 3.1.4), 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

10 ~~MTW. It means that after the selection of analogues in order to keep also some whose rank has increased, as they appear to be also relevant to predict precipitation values~~ the analogue situations for 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).

15 The introduction of the MTW improved the selection of synoptic analogues. ~~The number of good analogues is thus globally increased.~~

~~This increase in number of analogues has a slight effect on~~ 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 for application on long meteorological archives. For example, when processing forecasts for a target day with synoptic similarity with situations from

20 the past related to extreme weather, even though 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 know that the situation at hand has such analogue in the past. Another possible application is a quality assessment of the performance of the different precipitation classes. ~~The same analysis as in Sect. 3.2.1 has been performed again on the basis of the newly calibrated parameters. The results are generally very similar, but a slight performance increase of small~~

25 ~~precipitation values can be observed at the expense of higher amounts. The change in the analogues numbers is likely to be responsible for this difference in the performance distributions~~ 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

30 easier to use with the MTW.

~~The values of the  $S_{SCRIP}$  scores for both methods (Table ??) have significantly increased after recalibration.~~

## 4.2 Improvement of precipitation prediction

The MTW was found to improve the skill of the precipitation prediction for both the 2Z and 2Z-2MI methods (Fig. 4 and 5, Sect. 3.2). Moreover, it required no additional predictor. With the introduction of the MTW, the performance loss related

to the reduction of the archive length ~~is compensated~~ was negated (considering a 25 year archive of sub-daily data; the difference could be higher with a shorter archive). It was in this 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.

#### 4.2.1 Changes in sharpness and accuracy

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.

The  $S_{CRP}$  score can be decomposed into two components, namely sharpness and accuracy (see Bontron, 2004). The impact of the MTW has been analyzed on these components and the results can be found in Table ?? for the original parameters and in Table ?? for the re-calibrated methods. The changes are expressed relative to the total  $S_{CRP}$ . Both components do not have the same value ranges: accuracy is in our case almost twice the sharpness values. Since the  $S_{CRP}$  is considered here and not the  $S_{SCRIP}$ , improved prediction capacity results in a lower score. A decrease of the score is thus desirable.

It appears that the sharpness is a bit inferior with the MTW in favor of accuracy (Table ??). When using the recalibrated parameters, the components are a bit more balanced, but still the accuracy prevails on the sharpness (Table ??). Fig. ?? illustrates the changes (relative to the total  $S_{CRP}$ ) in sharpness and accuracy for different precipitation thresholds at the Ulrichen station. ~~Be it for~~ Both the analogy criteria (Sect. 3.1.2) and the ~~2Z or 2Z-2MI method~~, performance scores (Sect. 3.2.1) were improved to a greater extent for days with heavier precipitation. This is likely due 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 centres and the fronts, change significantly during 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 improve these situations with greater dynamics more significantly when introducing a MTW, as better matches to the ~~original or the recalibrated parameters~~, 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 should improve the prediction of extremes due to better analogue situations, but also due to possible new extreme values resulting from 24h totals with a certain time shift. However, even though the distribution of analogue precipitation values should move towards the targeted extreme, providing

a better prediction, the MTW itself does not allow to predict extreme events that were not yet observed and are, therefore, not present in the archive. The extremes in AMs can be modelled by, for example, extrapolation of a truncated exponential or gamma distribution fitted to the changes are similar: the improvements concern the accuracy over the sharpness, and this to a greater extent for significant precipitation events. The MTW does not allow to improve the sharpness, but the improvement in accuracy is significant, and this especially for the days with heavy precipitation. In terms of the predicted precipitation distribution, it means that the median of the prediction is closer to the observed values, while the distribution is not more condensed than previously. analogue values (Obled et al., 2002). Another possible approach is by combining AMs with other methods (e.g. Chardon et al., 2014). From this perspective, the MTW might improve the prediction of extremes as it improves the distribution of precipitation values for days with higher precipitation, on which post-treatment techniques rely. However, this goes beyond the scope of the present study.

## 5 Application attempts on the full archive

The improvement provided by the MTW is interesting, mainly for heavy precipitation events, and thus it would be profitable to be able to apply it to the complete archive. Unfortunately, there is no long precipitation time series available with a sub-daily time step. In order to reconstruct a longer archive of moving 24 h-totals, different disaggregation approaches of the daily time series were assessed and are presented in the following sections.

### 4.1 Seasonal effect

### 4.2 Proportional distribution

Section 3.1.3 and Fig. 9 revealed a difference in the improvement of 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.

A proportional distribution is certainly the simplest approach that can be performed. It consists in allocating proportional parts of the original daily time series into a new moving average of 24 h-totals. An analysis of the selected hours for the geopotential height predictor seems to confirm this assumption (Fig. 3). When using this reconstructed time series on the shorter period, 12). It was found that the new choice of the temporal window in winter, when using the method performance was degraded (Table 4 to compare to Table ??) and was even below the conventional method without MTW (Table ??). The benefit of a better selection of the analogue situations is lost due to a precipitation archive of poor quality. 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.

## 4.2 Use of a proxy variable

### 4.2 On the increase in the number of analogues

5 ~~As demonstrated in the previous section,~~ After the recalibration of the AMs with the MTW, the optimal analogue numbers were significantly higher than in the original methods (see Sect. 3.2.2 and Table 3). As shown in Fig. 7, the performance improvement brought by the MTW is lost when using a simplistic precipitation time series. In order to get closer to the chronology of the actual precipitation, the reconstruction procedure must rely on some informative variable. Data from the improvement of the S1 criterion grew along with the rank of the analogue, which shows an accumulation of better analogue situations in the distributions. It seems profitable to widen the selection of analogues in order to also keep some whose rank has increased, as they appear to be relevant to the prediction of the precipitation values. The number of good analogues was globally increased.

15 A higher number of analogues generally means, with an archive of fixed length, a poorer analogy. Indeed, when the choice of the predictors or the parameters are improved, leading to a better prediction, the optimal number of analogue situations decreases. However, when the length of the archive increases, the optimal number of analogues increases too for a 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 in better performance skills for the given methods, as it can be seen 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?

20 One can question the interest of using moving daily totals when 6-hourly precipitation series can be predicted instead. However, the 6-hourly time series generated by the AM might not represent accurately the intra-daily precipitation distribution (results not shown). 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 totals.

### 4.4 On the use of an old reanalysis dataset

25 The considered reanalysis dataset, which is the NCEP/NCAR reanalysis 1 (Seet1 (Kalnay et al., 1996), is relatively old and has a coarse resolution ( $2.5^\circ$ ). 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 the method to 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.

30

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. 2.1) was used in this perspective, despite their rough resolution.

5 The first step consisted in determining which variables from the precipitable water or the relative humidity (at 1000 hPa, 925 hPa or 850 hPa) produces the precipitation chronology that is the most correlated (on non-zero values) with the accurate 6-hourly precipitation time series on the period 1982–2007. The location of the most predictive point from the reanalysis grid had to be assessed (closest points: 5–7.5° E and 45–47.5° N), as well an eventual time lapse between both series, due to the significant distance separating the weather stations and the reanalysis grid point. The most relevant point will probably not be  
10 the same for all stations, since one can expect to find the optimum in the direction of the main moisture inflows for the station. This assessment must be carried out for each station.

#### 4.5 Relevance of the reconstructed precipitation archives

The results of the correlation analyses (on non-zero values) between the new series reconstituted using the proxy and the actual series are illustrated in Table 5 for Zermatt (on the short period, 1982–2007). The best proxy is the precipitable water at  
15 45° N and 7.5° E, with a time offset of -6 h. The determination coefficient of 0.721 is higher than that of the series from the proportional distribution (0.698), which confirms that a little information was added to the series of precipitation. The two simple methods assessed to reconstruct precipitation archives (Sect. 2.5) did not result in valuable outputs (Sect. 3.3). Indeed, the performance improvement brought by the MTW was lost due to the poor quality of the precipitation archives.

Table 6 presents the  $S_{SCRIP}$  scores obtained by the reconstituted series using the proxy of precipitable water at the optimal  
20 point (Table 5). A slight improvement can be obtained could be obtained for the second method relying on a proxy variable compared to the proportional distribution method, but it is was still relatively small, and most of the benefit of the MTW is again lost. was lost. A more recent reanalysis archive with more accurate moisture variables might produce better proxies.

These attempts to transpose the MTW on the total archive highlight highlighted the importance of the actual rainfall chronology. The MTW is profitable beneficial, provided that the precipitation series are close to the observed one. Without a pre-  
25 cipitation series with an accurate sub-daily chronology, the introduction of a MTW remains fruitless does not improve the precipitation prediction.

## 5 Conclusions

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

As Finet et al. (2008) had previously shown, the introduction of a MTW allows finding better analogue situations in terms of the atmospheric circulation. It has been here demonstrated demonstrated in this study that the improvement of the  $S_{TW}$  criteria

~~values is~~ S1 criterion values was growing with the rank of the analogue. This ~~is was~~ due to the accumulation of better analogues within the predicted distributions.

The improvement of the circulation analogy was found to be more important for ~~heavy precipitation events~~ days with heavier precipitation, which are generally related to more dynamic atmospheric situations. ~~This dynamism implies and are less frequent in the archive. These situations have~~ more specific circulation patterns that are evolving more rapidly. ~~Thus~~ Therefore, a MTW ~~is was~~ found to be of particular interest in this kind of situation, ~~and benefits benefiting~~ the prediction of ~~extreme heavier~~ precipitation events.

A seasonal effect has been highlighted, as the MTW ~~is was~~ more profitable for winter ~~months~~ than summer. The reason is likely that the diurnal cycle has a bigger 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, whereas,~~ 75 % of the analogue situations are were selected at a different time in winter.

The impact of the MTW on the prediction ~~performance has never been assessed before,~~ skill was never assessed before due to the shortcoming of long precipitation series at a sub-daily time step. ~~The impact of the MTW approach could be assessed on~~ Here, it was assessed for a 25 years- year time series with a high temporal resolution. After the introduction of the MTW, the performance scores increased of the same ~~order of~~ magnitude as if the length of the archive doubled, ~~in this case study.~~

The parameters were then calibrated again, using the 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 ~~profitable beneficial~~ for the prediction performance. The number of good analogues ~~is thus globally increased was globally increased in the same way as if the archive length increased.~~ This change seems to benefit ~~mainly~~ the prediction of days with small precipitation totals.

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

~~Perspectives of using the MTW rely~~ The use of the MTW relies partly on the availability of long precipitation series at a sub-daily time step and with high accuracy. ~~The first source of such data is the simple fact that with time~~ First, these archives of high temporal resolution precipitation data are ~~getting longer. The second source increasing over time. Another possible source of such archives~~ is the establishment of precipitation reanalysis at a ~~country scale, regional scale~~ or the use of reanalysis-driven ~~RCMs (Regional Climate Models) or LAMs (Limited Area Models) on regional climate models or limited area models over~~ a long period. ~~When such precipitation data will be~~ 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 use of a improvements proposed by the MTW especially for days with heavy precipitation may be interesting even without long continuous precipitation



archives. For example, recent target day may have synoptic similarity with situations from the early twentieth century, for which no continuous daily 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, it is worth known that the situation at hand has such analogue in the far past.

The use of the MTW in the AM ~~should be considered for any application, especially when the prediction quality of extreme events is important~~ can already be considered now for several applications in different contexts, may it be for operational forecasting or climate-related studies.

~~Thanks to Dominique Béroud for his support and to Renaud Marty for his fruitful collaboration over the years.~~

## 10 Appendix A: Acronyms

~~Thanks to-~~

ZZ Name of the analogue method of the atmospheric circulation

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

15 AM Analogue method

CRPS Continuous ranked probability score

CRPSS Continuous ranked probability skill score

MTW Moving time window

NCAR National Center for Atmospheric Research

20 NCEP National Center for Environmental Prediction

RH850 Relative humidity at 850 hPa

RMSE Root-mean-square error

S1 Teweles and Wobus (1954) score

TPW Total precipitable water

25 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 thank the Swiss Federal Office for Environment (FOEV), the Roads and Water ~~ourses~~ Courses Service, Energy and Water Power Service of the Wallis Canton, and the Water, Land, and Sanitation Service of the Vaud Canton ~~who~~ which financed the MINERVE (Modélisation des Intempéries de Nature Extrême des Rivières Valaisannes et de leurs Effets) project ~~which~~ that started this research. Thanks to Dominique Béro d for his support.

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

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

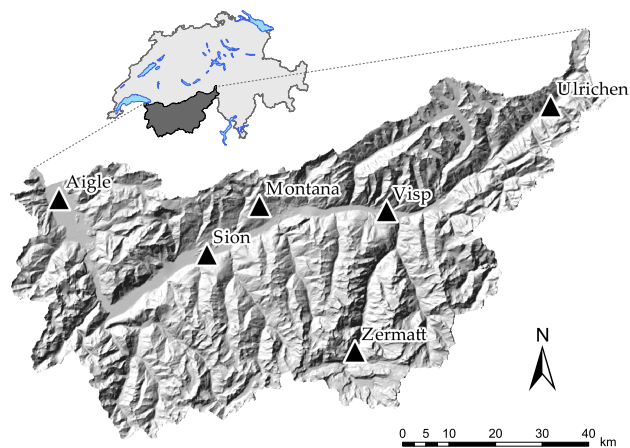
The authors would also like to acknowledge the work of anonymous reviewers that contributed to improving this manuscript.

## 10 References

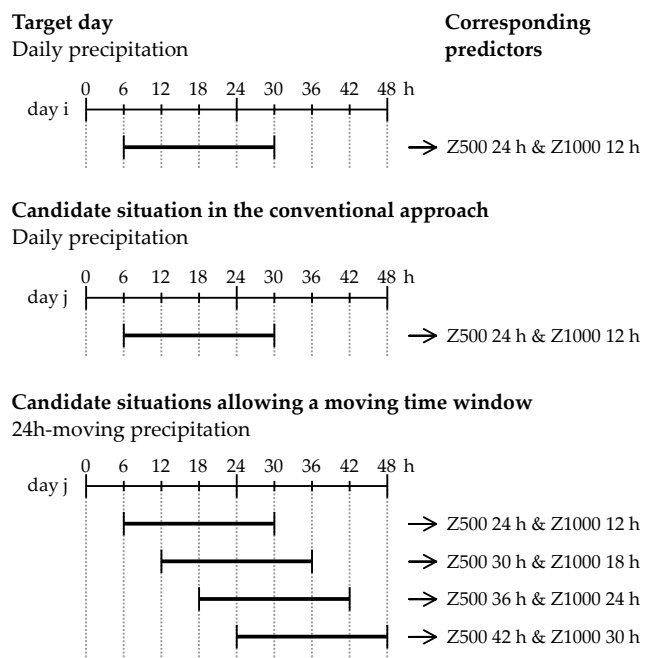
- Alessandrini, S., Delle Monache, L., Sperati, S., and Cervone, G.: An analog ensemble for short-term probabilistic solar power forecast, *Appl. Energ.*, 157, 95–110, doi:10.1016/j.apenergy.2015.08.011, 2015a.
- Alessandrini, S., Delle Monache, L., Sperati, S., and Nissen, J. N.: A novel application of an analog ensemble for short-term wind power forecasting, *Renew. Energ.*, 76, 768–781, doi:10.1016/j.renene.2014.11.061, 2015b.
- 5 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.
- 10 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.
- Bois, P., Obled, C., and Thalamy, J.: Etude des liaisons entre champs de pression et températures sur l’Europe avec la durée d’insolation: Application à la possibilité de la prévision d’insolation, in: C.R. du colloque Météorologie de l’énergie solaire, PIRDES-CNRS, Toulouse, pp. 145–177, 1981.
- 15 Bolognesi, R.: Premiers développements d’un modèle hybride pour le diagnostic spatial des risques d’avalanches, *La Houille Blanche*, 8, 551–553, doi:10.1051/lhb/1993045, 1993.
- 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.
- 20 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.
- 25 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.
- 30 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.
- 35

- 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.
- Fraedrich, K., Raible, C. C., and Sielmann, F.: Analog Ensemble Forecasts of Tropical Cyclone Tracks in the Australian Region, *Weather Forecast.*, 18, 3–11, doi:10.1175/1520-0434(2003)018<0003:AEFOTC>2.0.CO;2, 2003.
- 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, 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.*, Accepted, doi:10.1175/MWR-D-16-0093.1, 2016.
- 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.
- Keenan, T. D. and Woodcock, F.: Objective Tropical Cyclone Movement Forecasts Using Synoptic and Track Analogue Information, Tech. rep., Bureau of Meteorology, 1981.
- 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.

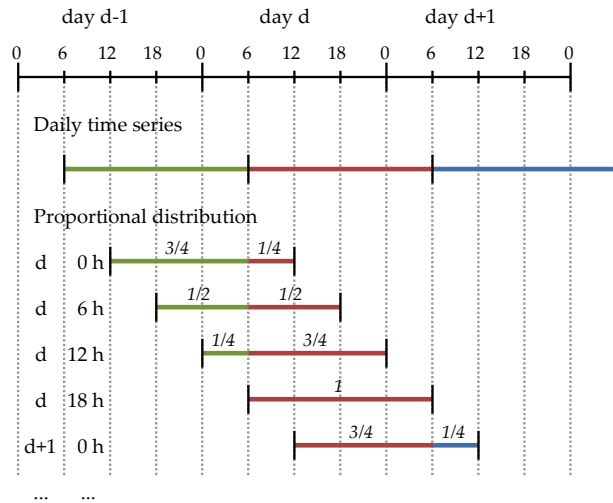
- Lorenz, E.: Empirical orthogonal functions and statistical weather prediction, Tech. rep., Massachusetts Institute of Technology, Department of Meteorology, Massachusetts Institute of Technology, Dept. of Meteorology, 1956.
- Lorenz, E.: Atmospheric predictability as revealed by naturally occurring 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 developments to bridge the gap between dynamical models and the end user, *Rev. Geophys.*, 48, 1–34, doi:10.1029/2009RG000314, 2010.
- Marty, R.: Préviation 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 QPPF 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. and Good, W.: Recent developments of avalanche forecasting by discriminant analysis techniques: a methodological review and some applications to the Parsenn area (Davos, Switzerland), *J. Glaciol.*, 25, 315–346, doi:10.3198/1980JoG25-92-315-346, 1980.
- 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.
- 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.
- 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.
- Sievers, O., Fraedrich, K., and Raible, C. C.: Self-Adapting Analog Ensemble Predictions of Tropical Cyclone Tracks, *Weather Forecast.*, 15, 623–629, doi:10.1175/1520-0434(2000)015<0623:SAAEPO>2.0.CO;2, 2000.
- 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.



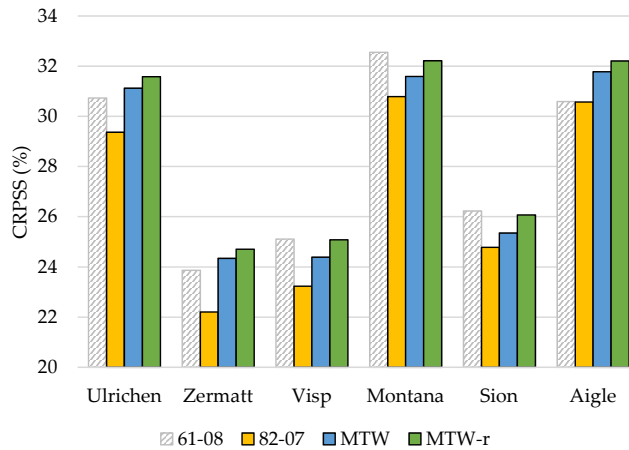
**Figure 1.** Position of the six weather stations of interest (Ulrichen, Zermatt, Visp, Montana, Sion, and Aigle) in the study area upper Rhône catchment in Switzerland.



**Figure 2.** Illustration of the principle of a moving time window. The target situation is the same for the conventional approach and the MTW, while the candidate situations are 4 times as many with the MTW. The larger horizontal bars represent the 24-h-24h precipitation totals.

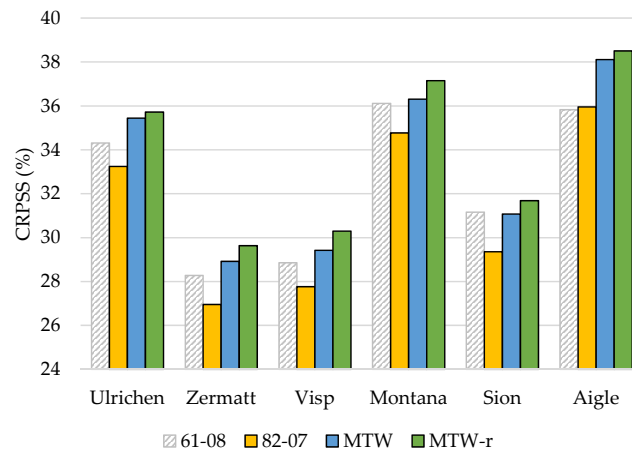


**Figure 3.** Illustration of the generation of 24h total moving averages by means of a proportional distribution. The colours refer to the corresponding day of the daily time series.



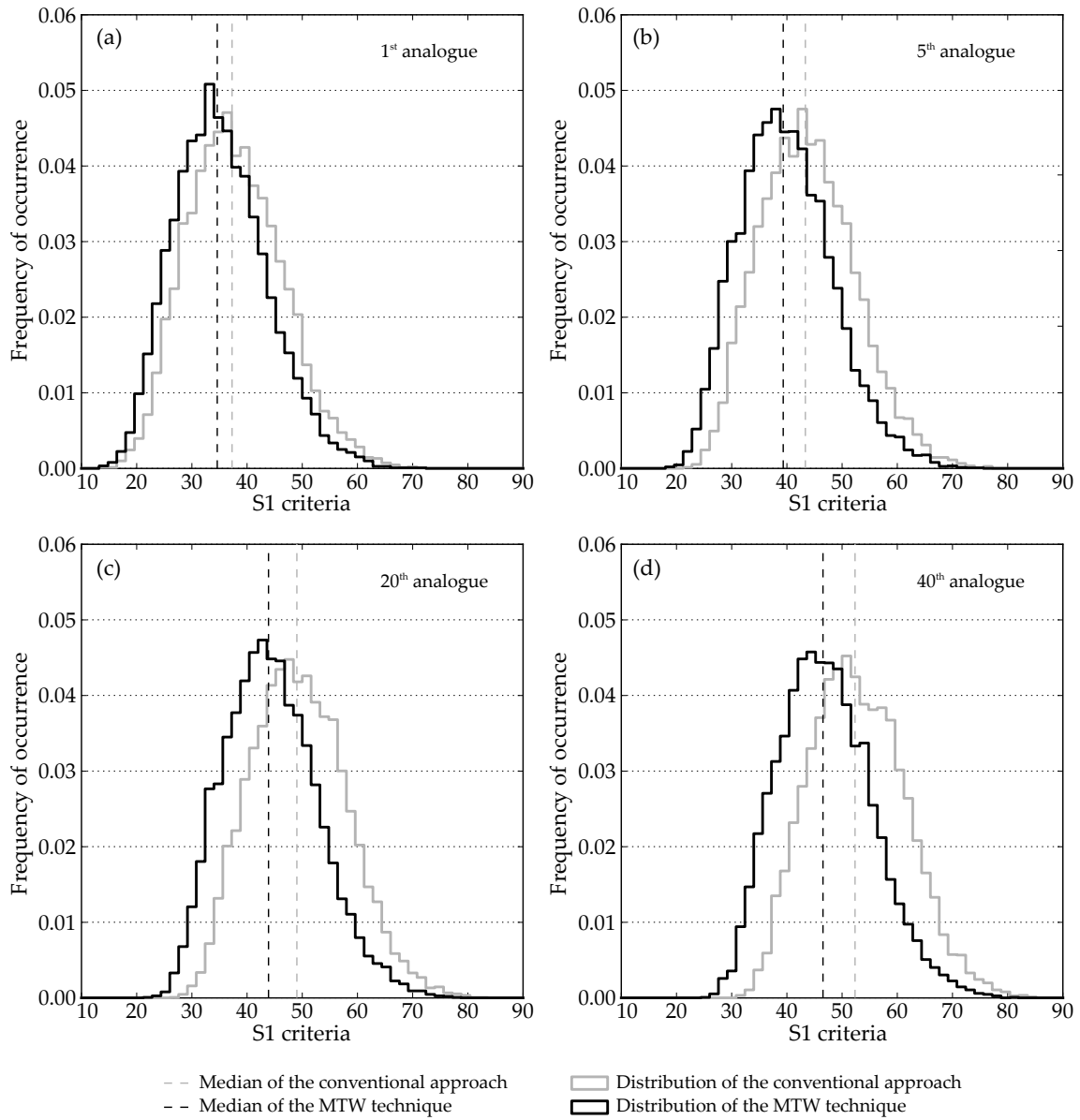
**Figure 4.** Performance score (CRPSS) of the AM on the atmospheric circulation at the different stations for (dashed) the full archive, i.e. 1961–2008; (yellow) the reduced archive, i.e. 1982–2007; (blue) the introduction of the MTW on the reduced archive; and (green) the recalibrated parameters of the AM with the MTW.

Influence (%) of the MTW on the  $S_{CRP}$  components (accuracy and sharpness) relatively to the total  $S_{CRP}$ , for the 2Z and 2Z-2MI methods, according to different precipitation thresholds. The results are presented for (top) the original parameters, and (bottom) the recalibrated parameters. Improved prediction results in a lower score.

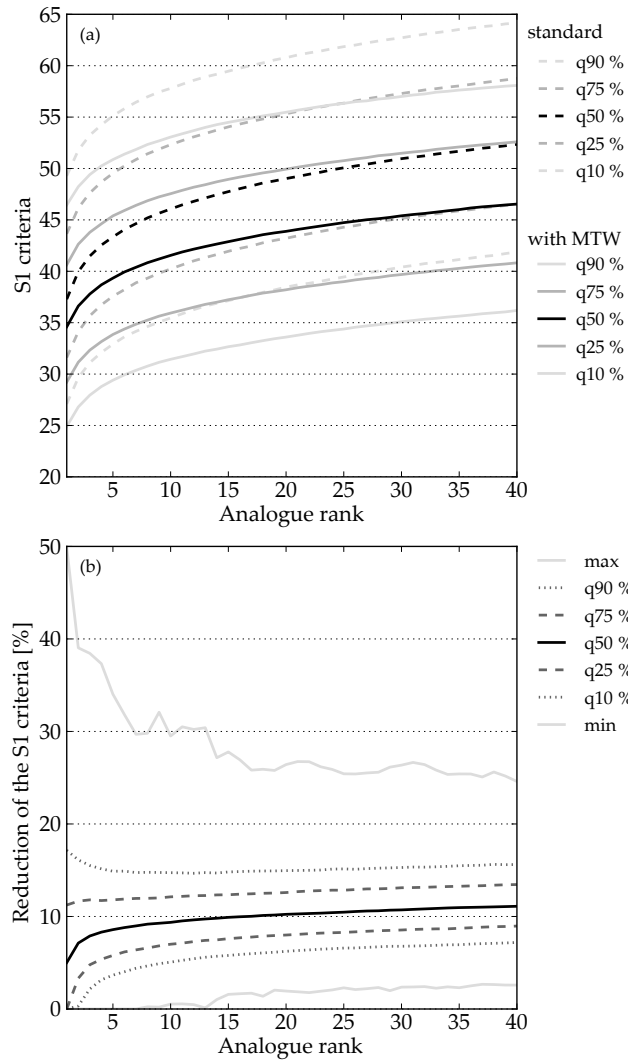


**Figure 5.** Same as Figure 4, but for the analogue method with a second level with the moisture variables.

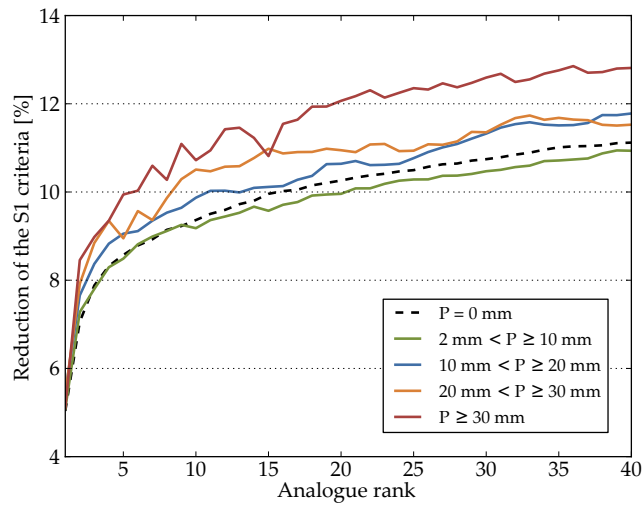




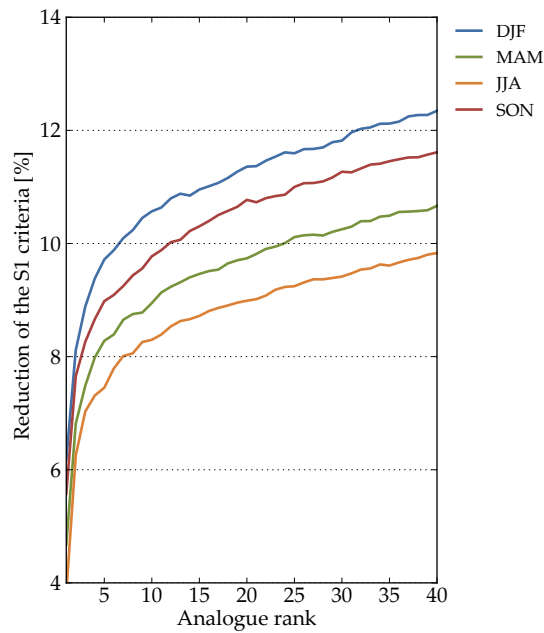
**Figure 6.** Changes in the  $S_{TW}$ -criteria-S1 criterion distributions due to the MTW of (a) the 1<sup>st</sup>, (b) 5<sup>th</sup>, (c) 20<sup>th</sup>, and (d) 40<sup>th</sup> analogue ranks for the Ulrichen station, due to over the MTW whole calibration period (1961–2008).



**Figure 7.** Synthesis of the changes in the  $S_{TW}$  criteria, S1 criterion due to the MTW, for the Ulrich station, depending on the ranks rank of the analogue. (a) Quantiles of the  $S_{TW}$ -S1 distributions with and without the MTW. (b) Quantiles of the relative improvements of the S1 criteria-criterion when using the MTW.



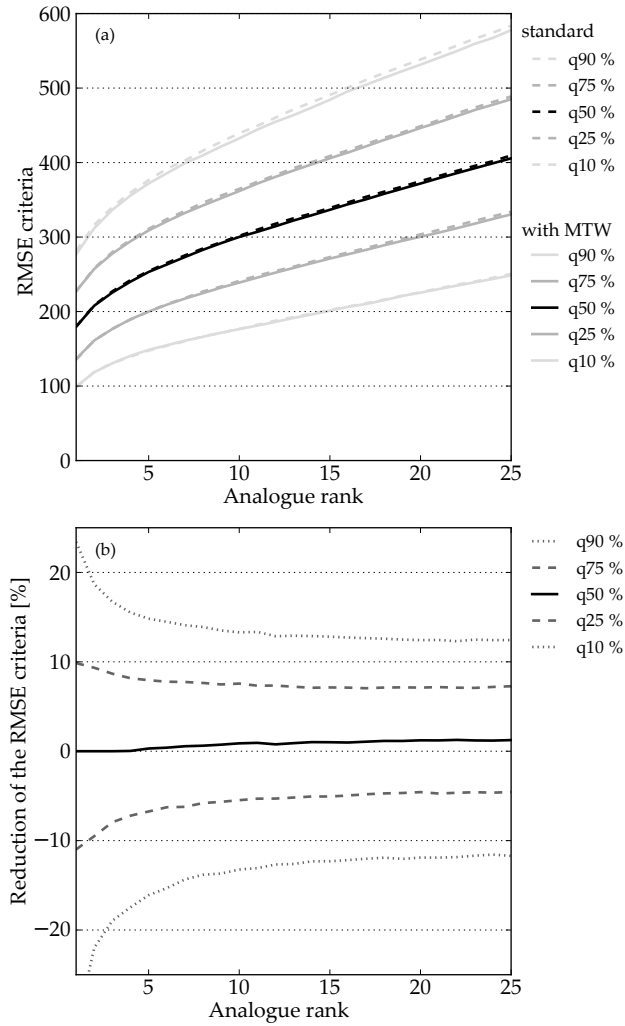
**Figure 8.** Distribution of the median improvements of the  $S_{TW}$ -criteria,  $S_1$  criterion due to the MTW, depending on daily precipitation thresholds at the Ulrichen station.



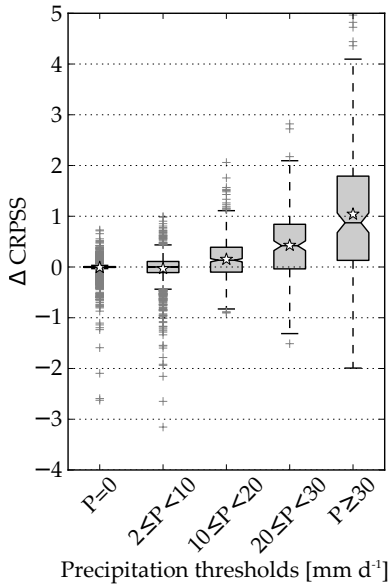
**Figure 9.** Seasonal effect on the median reduction of the  $S_{TW}$ -criteria,  $S_1$  criterion for the Ulrichen station due to the MTW. DJF: winter, MAM: spring, JJA: summer, and SON: fall.

Distribution of the predictors hours in the selected analogue dates depending on the season, for the Ulriehen station.

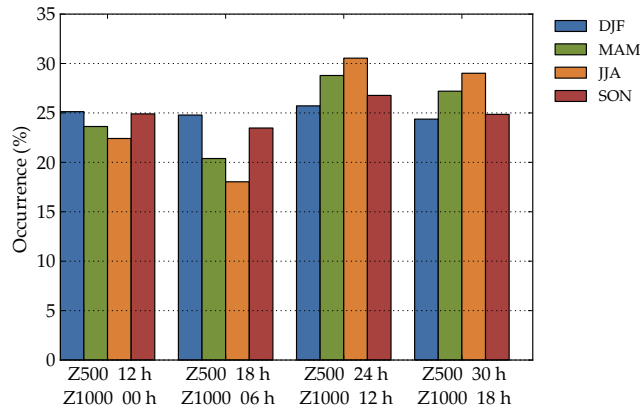
Distribution of the predictors hours in the selected analogue dates depending on the season, for the Visp station.



**Figure 10.** Synthesis of the changes in the  $E_{RMS}$  criteria, RMSE criterion due to the MTW  $\tau$ , for the Ulriehen station  $\tau$ , depending on the ranks rank of the analogue. (a) Quantiles of the  $E_{RMS}$ -RMSE distributions with and without the MTW. (b) Quantiles of the relative improvements of the  $E_{RMS}$  criteria-RMSE criterion when using the MTW.



**Figure 11.** Differences of the  $S_{\text{SCRIP-CRPSS}}$  performance score, due to the introduction of the MTW, as a function of daily precipitation thresholds at the Ulrichen station. The stars represent averages.



**Figure 12.** Illustration-Distribution of the generation of 24-h-totals moving averages by means of a proportional distribution. The colours refer to predictor hours on the corresponding day of selected analogue dates, when using the daily time series MTW, depending on the season for the Ulrichen station.

**Table 1.** Parameters ~~of~~ for the reference method on the atmospheric circulation (2Z). The first column is the level of analogy (0 for pre-selection), ~~then comes~~ the second column is the meteorological variable, and then its hour of observation (temporal window). The ~~criteria~~ criterion used for the current level of analogy is then provided, as well as the number of analogues.

Level	Variable	Hour	<del>Criteria</del> <u>Criterion</u>	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 conventions as Table 1

Level	Variable	Hour	<del>Criteria</del> <u>Criterion</u>	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		

**Table 3.** ~~Calibrated parameters (spatial windows and~~ Optimal number of analogues ~~)for~~ (of the first and second level of analogy, respectively,  $N_1$  and  $N_2$ ) of the method based on the geopotential at 500 hPa and 1000 hPa atmospheric circulation only (method 2Z) and skill score (%) ~~of~~ of 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).

Station	Standard			MTW-r		
	2Z		2Z-2MI	2Z		2Z-2MI
	$N_1$	$N_1$	$N_2$	$N_1$	$N_1$	$N_2$
Ulrichen	40	60	25	50	110	35
Zermatt	35	55	25	55	80	30
Visp	30	45	25	55	135	35
Montana	40	55	30	55	110	40
Sion	40	90	30	55	140	50
Aigle	50	100	35	75	135	45

**Table 4.** Values of the  $S_{\text{SCRIP-CRPS}} (\%)$  skill-score for the ~~newly-calibrated-original and the recalibrated~~ parameters ~~(with the sequential method, as described in Sect. 2.1)~~ using the MTW approach ~~. The differences are expressed regarding on the conventional method with disaggregated precipitation time series (short period) by means of the same archive-lengthproportional distribution.~~

Station	2Z		2Z-2MI	
	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

**Table 5.** Value of the coefficient of determination between the reconstructed 6-hourly precipitation time series using the listed variables ~~and~~ the accurate time series ~~on~~ for the period ~~1982-2007-1982-2007~~. The grid points ~~are surrounding the following region are:~~ 1)  $5^{\circ}\text{E}$ ,  $47.5^{\circ}\text{N}$ ; 2)  $5^{\circ}\text{E}$ ,  $45^{\circ}\text{N}$ ; 3)  $7.5^{\circ}\text{E}$ ,  $47.5^{\circ}\text{N}$ ; and 4)  $7.5^{\circ}\text{E}$ ,  $45^{\circ}\text{N}$ . The highest coefficient of determination is indicated in bold.

Variable	Point	Time lapse				
		<del>-12-12 h</del>	<del>-6-6 h</del>	0 h	+6 h	+12 h
RH1000	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
RH925	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
RH850	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
TCW	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	<b>0.721</b>	0.696	0.643	0.636

**Table 6.** Values of the  $S_{\text{SCRIP-CRPSS}}$  (%) skill 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.

Period	2Z		2Z-2MI	
	original	recalib.	original	recalib.
<del>1982-2007</del> <u>1982-2007</u>	22.57	23.14	27.11	27.71
<del>1961-2008</del> <u>1961-2008</u>	23.81	24.38	28.42	28.86