[1,2]PascalHorton [3]CharlesObled [1]MichelJaboyedoff

[1]University of Lausanne, Institute of Earth Sciences, Lausanne, Switzerland [2]University of Bern, Oeschger Centre for Climate Change Research, Institute of Geography, Bern, Switzerland [3]Université de Grenoble-Alpes, LTHE, Grenoble, France

Finding better analogue situations at a sub-daily time step

P. Horton et al.

Pascal Horton (pascal.horton@alumnil.unil.ch)

1

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

02.04.2017

Abstract

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

The search for candidate situationssimilar to a given target day is usually undertaken by comparing the state of the atmosphere at fixed hours of the day for both the target day and the candidate analogues. This is a consequence of using standard daily precipitation time series, which are available over longer periods than subdaily data. However, it is unlikely for the best analogy to occur at the very same hour for the target and candidate situations. A better analogue situation may be found with a time shift of some hours since a better fit can occur at a different time of day. In order to assess the potential for finding better analogues at a different hour, a moving time window (MTW) has been introduced.

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

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

Both improvements of the analogy criterion and the performance scores on rainfall were found to be higher for MTWs with a smaller time step (3h). A 3h-MTW provides 8 times more candidate situations, even though not fully independent. As the MTW provides additional situations to the pool of possible analogues, it can be seen as an inflation of the meteorological archive. The technique being very simple and easily applicable, it should be considered for several applications in different contexts, may it be for operational forecasting or climate-related studies.

The Analogue method (AM) is based on the hypothesis that two relatively similar synoptic situations may produce similar local effects Lorenz1956, Lorenz1969. It is a statistical downscaling method [see][for other methods]Maraun2010 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 Duband1970, Zorita1999. Multiple variations of the method exist [a non-exhaustive listing can be found in][]BenDaoud2016. The AMs are used for operational precipitation forecasting, either in the context of weather forecasting, flood forecasting, or hydropower production [e.g.,][]Bontron2005, Hamill2006, Desaint2008a, GarciaHernandez2009b, Bliefernicht2010, Marty2010, Marty2012, Horton2012, Obled2014, Hamill2015, BenDaoud2016, as well as for precipitation downscaling from a climate perspective [e.g.][]Radanovics2013, Chardon2014, Dayon2015. AMs can provide multivariate predictions that are physically consistent Raynaud2016b. Their spatial transferability is analysed in Chardon2014 and Radanovics2013. Their temporal transferability has also been the topic of recent studies for past or future climates Dayon2015, Caillouet2016.

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, sampled at different hours (e.g. every 1é or 6 or 3 hour). The second is an archive of the target variable to be predicted, here precipitation cumulated over a given duration. It is most often made of standard ( 6hUTC to 30h UTC) daily precipitation totals, 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.

Ruosteenoja1988 and Vandendool1994 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. For that reason, limited 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. Given the cumulative aspect of the predictand, the corresponding meteorological situation is characterized by several predictors taken at different but fixed hours of the day, and the analogue situations are assessed by comparing these same predictors at the same fixed hours of the day. However, it can be expected that the best analogy of the synoptic situations does not occur systematically at the same time of the day and that better candidates can be found by shifting to different hours. With this assumption, a moving time window (MTW) was introduced to allow the search for candidates at different hours of the day. Previous tests 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 skillfor a specific predictand Finet2008.

One can question the interest of using moving daily totals when for example 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 or variability. In addition, sometimes a resolution finer than the daily time step is not needed and another disaggregation technique may be used afterwards. Finally, when using a reconstructed precipitation archive, the errors in intra-daily precipitation distributions have a smaller impact on 24h totals than 6h totals.

The proposed 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 an *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.

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.][]Horton2012a, BenDaoud2016, Caillouet2016. AMs can also be combined with other methods [see e.g.][]Chardon2014. 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 Horton2017. However, the MTW technique is not in competition with other approaches and can rather be combined with these.

Section 2 presents the context of the study as well as the data and methods, including the proposed MTW technique. The benefits of introducing an MTW were assessed first in regards to the improvement of the analogue dates selection (Sect. 3.1) and then in terms of precipitation prediction skill (Sect. 3.2). Finally, the results are discussed in Sect. 4, and the conclusions are found in Sect. 5.

## 1 Data and methods

### 1.1 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 Horton2012. The data were available at an hourly time step for 29 years (1982–2010) and were also collected at a standard daily time step (from 6:00 UTC to 6:00 UTC the next day) for 50 years (1961–2010). 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 two of the most recent global reanalysis datasets: the european ERA-20C Poli2016 with a 3h temporal resolution and a spatial resolution of 1°, and MERRA-2 from NASA with a 6h temporal resolution and a spatial resolution of 0.625°x 0.5°. The study was originally performed on the NCEP/NCAR reanalysis 1 dataset Kalnay1996, and the conclusions were similar. ERA-20C is build by assimilating only surface observations, and is thus available on a long period (1900–2010). On the contrary, MERRA-2 is build by assimilating observations at higher levels and from more sensors, including satellite data, but is then more limited in time (1980–present). It is of interest to assess the relevance of an MTW with these two datasets, as they represent different kind of products. The specific advantages of ERA-20C is that it allows testing an MTW with a 3h time step, and it covers a long period. On the other hand, MERRA-2 has a finer time and a higher spatial resolution and can be expected to be more accurate at higher levels of the atmosphere. The variables extracted from these datasets were geopotential heights at 500 hPa (Z500) and 1000 hPa (Z1000), as well as the total precipitable water (TPW) and the relative humidity at 850 hPa (RH850).

### 1.2 The considered analogue methods

The first considered AM is based on the analogy of the atmospheric circulation only [Table 1;][]Obled2002, Bontron2005. Searching for analogue situations to a target day starts by a seasonal stratification through a preselection step of the potential candidates for analogy. The restriction is a search for analogue days within a 4-month window centred on the target date for every year of the archive. The similarity of the atmospheric circulation of the target date with every day of the preselection is assessed by processing the Teweles1954 score (S1) that is a comparison of gradients on geopotential heights over a certain spatial window and at certain hours:

(1)

where is the forecast geopotential height difference between the *i*th pair of adjacent points from the grid of the target situation, is the corresponding observed geopotential height difference in the candidate situation, and *m* is the number of pairs of adjacent points in the grid. The differences are processed separately in both directions. With smaller S1 values, there is greater similarity in the pressure fields.

The predictor variables extracted from reanalysis datasets are considered at different hours of the day. Based on Bontron2005, geopotential heights are compared at 1000 hPa (Z1000) at 12:00 UTC and 500 hPa (Z500) at 24:00 UTC. The time of the day at which the predictors are selected is found by Bontron2004 to have a significant influence.

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

The second reference method [2Z-2MI, Table 2;][]Bontron2005 adds a subsequent level of analogy with moisture variables, compared by means of the root-mean-square error (RMSE):

(2)

where is the *i*th predictor value from the grid of the target situation, is the corresponding observed value in the candidate situation, and *n* is the number of points in the grid.

The additional predictor is a moisture index composed of the product of the total precipitable water (TPW) with the relative humidity at 850 hPa (RH850), taken at 12:00 UTC and 24:00 UTC Bontron2004. When adding a second level of analogy, dates are subsampled from the analogues on the atmospheric circulation, resulting in a smaller number of analogue situations. When a second level of analogy is added, a higher number of analogues is kept on the first level.

More complex AMs exist with additional predictors [see e.g.][]Horton2012a, BenDaoud2016, Caillouet2016. The MTW can also be applied to these. However, it is easier to interpret the impact of the MTW using more basic methods.

### 1.3 Calibration of the analogue method

AMs rely on parameters that need to be defined for every level of analogy. Here, the selection of the meteorological variables used as predictors, as well as their corresponding analogy criteria, pressure level, and hour of the day, were considered identical to the methods 2Z (Table 1) and 2Z-2MI (Table 2) described above. The parameters that were here calibrated are:

• The spatial windows, which are the domains on which predictors are compared. There is a spatial window specific to each level of analogy. It will thus differ between circulation and moisture variables.

• The optimal number of analogue situations for every level of analogy.

The semi-automatic sequential procedure elaborated by Bontron2004 was used to calibrate the AM [\s\do5(), see also]]Radanovics2013, BenDaoud2016. 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 Bontron2004, consists of the following steps:

1. For every level of analogy:

(a) Identification of the most skilled unitary cell (one point for moisture variables and four for geopotential heights in order to process the gradients) 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 until no improvement is reached.

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

2. 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, , is initiated at a chosen value. The procedure starts again from step 1 (calibration of the spatial window and the number of analogues) for the new level. The parameters calibrated on the previous analogy levels are fixed and do not change.

3. Finally, the numbers of analogues and 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.

The calibration is done in successive steps with a limited number of parameters in order to minimize the objective function (CRPS, Eq. 3). Previously calibrated parameters are generally not reassessed (except for the number of analogues). More advanced techniques, such as using genetic algorithms Horton2017, exist but are out of the scope of the present study.

The 29-yr hourly precipitation dataset was first aggregated into into a standard daily (6h UTC-30hUTC) dataset and then divided into a calibration period (CP) and a validation period (VP) in order to assess the robustness of the proposed improvements on independent data. The selection of the VP was evenly distributed over the entire series BenDaoud2010 to reduce potential biases related to trends linked to climate change or to the evolution in measurement techniques. Thus, one out of every five years was selected for validation, which represents a total of 6 yr for the VP and 23 yr for the CP.

The continuous ranked probability score [CRPS,][]Brown1974, Matheson1976, Hersbach2000 is often employed in order to assess the performance of AMs [see, e.g.,][]Bontron2004, Bontron2005, BenDaoud2008, Horton2012, Marty2012, Radanovics2013, Chardon2014, Junk2015, BenDaoud2016, Caillouet2016. It allows the evaluation of the predicted cumulative distribution functions *F*(*y*) of the precipitation values *y* from analogue situations compared to the observed value . A better prediction has a smaller score. It is defined as follows:

(3)

where is the Heaviside function that is null when and has the value of 1 otherwise.

### 1.4 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. The target situation (the day to predict) is still the same as in the conventional approach, that is a daily precipitation total between 06:00 and 30:00 UTC, characterized by predictors which are here Z1000 and Z500, taken at fixed hours 12:00 and 24:00 UTC respectively. The difference is that candidate situations are not only considered at the same time (12:00 and 24:00 UTC), but also at other hours by allowing a time shift. Therefore, instead of looking for analogues at a 24h time step, they are sought at every time step matching the predictor temporal resolution (Fig. 2). The ERA-20C dataset made here possible to test an MTW with a 12h, 6h, or 3h time step (named respectively 12h-MTW, 6h-MTW, 3h-MTW hereafter). Thus, the candidates are 2 (12h-MTW), 4 (6h-MTW), or 8 (3h-MTW) times as many (even though they are not fully independent) as in the conventional approach.

The target situations and their corresponding observed precipitation values do not change. We then aim at predicting the exact same precipitation time series both through the MTW and the conventional approach, so that the performance scores can be directly compared. In order to assess the benefit of searching for analogue situations at a sub-daily time step for quantitative precipitation prediction, an appropriate precipitation archive is required. On the basis of hourly time series (Sect. 2.1), 24h totals were processed at time steps matching the MTW by building moving 24h totals on the period 1982-2010, starting respectively at 0, 3, 9 h UTC…etc . instead of 6hUTC only for the standard ones.

## 2 Results

The AMs were applied to both ERA-20C and MERRA-2 datasets. Most results are presented for the ERA-20C dataset, as it allows testing a 3h-MTW, but would lead to similar conclusions with MERRA-2. The impact of an MTW approach was assessed on two different basis:

• Original parameters: The AMs were initially calibrated for each station based on the conventional fixed 24h approach. Then, the MTW was introduced into the AMs, but the parameters (spatial windows and number of analogues) were not reassessed. The analogy was then identical, with the only difference that the MTW had more candidate at disposal.

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

The MTW is expected to have an impact on both the selection of analogue dates and the prediction skill. Both are analyzed separately hereafter.

### 2.1 Influence of the MTW on the selection of analogue dates

First, the impact of the MTW on the selection of analogue dates was analyzed, independently of any precipitation data. This was done on the period 1982-2010, but could have been done on a longer period, as no sub-daily precipitation series is required at this stage. The results are presented for the Ulrichen station, but were very similar for all other stations. The original parameters were used here when assessing the impact of the MTW, no recalibration of the parameters took place.

#### 2.1.1 Analogy of the atmospheric circulation

When searching for analogues in the first level of analogy, such as on the geopotential heights in the 2Z method (Sect. 2.2), there were up to eight times as many candidates (even though not fully independent) with the MTW than before (Sect. 2.4). Figure 3 shows the changes in the distributions, for target dates over the CP, of the analogy criterion (S1) for the , , and analogue rank, due to introduction of an MTW with a 12h, 6h, or 3h time step. The shape of the distributions of the conventional approach and the MTWs were found to be similar, but the values of the analogy criterion were gradually reduced (shifted to the left) with smaller time steps MTWs and were, therefore, better. The circulation analogy was here regularly improved with the doubling of the MTW time step. For analogues with higher ranks (e.g. or ), the difference between the two distributions was larger than for the first rank, which means that the improvement increased with the rank of the analogues.

The improvements of the analogy with the rank of the analogues are summarized in Fig. 4, which shows (top) quantiles of the S1 criterion for the conventional method and the MTWs at different time steps, and (bottom) quantiles of the relative reduction (meaning improvement) due to the MTWs. This confirms that this improvement is constantly increasing from the first to the last analogue (Fig. 4 bottom). It can be explained by the accumulation of better analogues in the selection, with new better situations pushing previous analogues to higher ranks. One can also see the regularity of the improvements brought by decreasing time steps. The introduction of the MTW allows finding better analogue situations in the first level of analogy, resulting in a selection of days with here a better similarity in the atmospheric circulation. This improvement is superior for MTWs with a smaller time step, which obviously allows abetter matching.

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

The effect of the MTW on the S1 criterion has been decomposed per season and is presented in Fig. 6. Differences between the seasons are substantial, with greater improvements for winter than summer. An analysis of the selected hours for the geopotential height predictor was then performed for the MTW with different time steps (Fig. 7). It was found that the new choice of the temporal window in winter is more balanced between the different hours of the day (more regular repartition), and this for all MTW time steps. This means a change in selection of a greater portion of the analogue dates compared to the conventional approach. On the contrary, the days in summer had a preference for the initial temporal window (Z500 24 h & Z1000 12 h). It is likely due to more pronounced diurnal effects, which reduced the potential for improvement of the criteria. These effects are in phase with the daily cycle, and good analogues were essentially found for the same hours. When a 3h-MTW was used, the time step following the initial temporal window (so Z500 27 h & Z1000 15 h) was almost as often selected as the initial one for all seasons. It can be considered as relatively similar in terms of the solar influence. The other seasons were between these two extremes, which is consistent with their respective improvements in S1.

#### 2.1.2 Moisture analogy

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

These criterion values were not improved by the MTW because the number of candidates was not higher (still the days selected in the first level of analogy). The use of an MTW did not increase the sample size in this case, as the second level of analogy only subsamples in the dates provided by the first level. However, this result of a globally null improvement of the RMSE values does not mean that the 2Z-2MI method cannot be improved by the MTW. It means that after the selection of the analogue dates 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 can have an impact on the performance of the precipitation prediction.

### 2.2 Impact of the MTW on the precipitation prediction

The changes in the performance score (CRPS) of the precipitation prediction are provided in Fig. 9 and 10 for the 2Z and 2Z-2MI methods, respectively. The AMs were here recalibrated on the CP, and assessed on the independent VP. The score was processed for each station and both reanalysis datasets. For ERA-20C, the 12h-MTW, 6h-MTW, and 3h-MTW could be assessed, while only the two first could be processed with the MERRA-2 dataset.

The MTW did improve the precipitation prediction, as the CRPS globally decreased for all stations, both on the CP and the VP. The VP showed a greater variability, likely related to the shorter period it represents (6 yr). The prediction skill on the CP was almost always improved further by reducing the time step of the MTW, but not of the same magnitude. For ERA-20C, the magnitude of the improvement of 2Z on the CP was relatively reduced when using a 3h-MTW instead of a 6h-MTW, and the variablility on the VP increased. However, as it is shown in Fig. 4, the selection of analogue dates was improved relatively equally for every reduction of the MTW time step. A possible reason could be that the ERA-20C dataset was build by only assimilating surface observations, while the AMs rely on the geopotential at 500 hPa. Thus, the timing of the atmospheric circulation at higher pressure levels might not be completely in phase with the real perturbation systems that caused the observed precipitations. The MTW with the MERRA-2 dataset and the 2Z method also showed a certain slope break after the 12h-MTW, that could not be explained, and that is less important for the 2Z-2MI method. However, the 12h-MTW resulted in important variability for the 2Z-2MI method and the MERRA-2 dataset.

The introduction of an MTW was previously found to show greater improvement of the S1 criterion for days with higher precipitation values (Sect. 3.1.1). The impact of the MTW was then assessed in terms of change of prediction performance per precipitation classes (for the target day). Figure 11 synthesizes these differences for the Ulrichen station, using the original parameters with the MTW. It shows that, after the introduction of the MTW, the performance score was in average further improved (reduced CRPS) for days with higher precipitation than for non-rainy days and small precipitation values. When recalibrating the parameters using the MTW, the spread of the changes in CRPS tended to decrease, but with a slight loss in performance for the most rainy days (not shown).

Reliability diagrams were also performed on the VP for both the 2Z and 2Z-2MI methods and for the exceedence of percentiles 80 % (P80 – 4 mm), 90 % (P90 – 9.5 mm), and 95 % (P95 – 17.5 mm) at the Ulrichen station (Fig. 12). Reliability diagrams plot the observed frequency against the predicted probability of a binary event – a threshold exceedence in this case. For perfectly calibrated predictions, the curve should be as close as possible to the diagonal. The VP only contains six years (Sect. 2), and thus higher variability is present for higher thresholds, the curves being smoother and closer to the diagonals for the CP, which is longer (not shown). Figure 12 shows that the MTW improved the reliability for both 2Z and 2Z-2MI methods and for all thresholds, as the curves moved toward the diagonal with a decreasing MTW time step. For P80, the 3h-MTW predictions fited very well the observed frequency. When considering higher thresholds, the 2Z-2MI method is generally better than 2Z. The case of a P95 threshold shows a conditional dry bias associated with larger forecast probabilities for the considered AMs, which was however substantially reduced by the 3h-MTW. This conditional dry bias means for example here that the AM did not predict the event with a probability equal to 1 for 2Z. Obtaining a predicted probability of exceedence of 1 would mean that all analogue dates for a certain target date would be above the threshold. As the number of analogue dates is 30 in this case, it becomes less likely to have 30 values above the threshold when it increases. This issue is the topic of an ongoing work. However, as one can see in Figure 12 (bottom), the introduction of an MTW, especially with a 3h timse step, improves partly this missmatch.

## 4 Discussion

### 4.1 A better prediction of heavy precipitation

Both the analogy criteria (Sect. 3.1.1) and the performance scores (Sect. 3.2) were improved to a greater extent for days with heavier precipitation. This is likely due 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 capture these situations with greater dynamics more significantly when introducing an MTW, as better matches to the target situation may be found. In contrast, days with low dynamics in the atmospheric circulation, such as anticyclonic situations, will not be radically improved by the introduction of the MTW.

The MTW improved the prediction for days with heavier precipitation, and should improve the prediction of extremes due to better analogue situations extraction, but also due to possible new extreme precipitation values resulting from moving 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. However, the extremes in AMs can be modelled by, for example, extrapolation of a truncated exponential or gamma distribution fitted to the analogue values Obled2002. Another possible approach is by combining AMs with other methods [e.g.][]Chardon2014. 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.

### 4.2 The relationship between the MTW time step and the number of analogues

When recalibrating the AMs with MTWs of different time steps, the optimal number of analogues changed for both 2Z and 2Z-2MI methods. and , from the first and the second level of analogy, tend to significantly increase with the reduction of the time step of the MTW (Fig. 13).This may interpreted as the equivalent of an expansion of the meteorological archive that an MTW approach provides. When using ERA-20C, all optimal numbers of analogues (for the different analogy levels) tend in average to double when using a 6h-MTW instead of the classic approach, and to even triple when using a 3h-MTW. These higher numbers of analogues were objectively chosen by the calibration procedure (Sect. 2.3) in order to increase the prediction skill of the methods.

As shown in Fig. 4, the improvement of the S1 criterion grows 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.

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 tends to decrease. However, when the length of the archive increases, the optimal number of analogues also tends to increase for a better performance up to a certain threshold [][]Bontron2004, Hamill2006a. The observed increase in the number of analogue situations with the MTW resulting in better performance skills for the given methods can therefore 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 again.

### 4.3 Inflation of the archive

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

However, since the MTW provides additional situations to the pool of possible analogues, it can be seen as an inflation of the archive. With the introduction of the MTW, the performance loss related to an eventual reduction of the archive length (to meet the length of the subdaily precipitation archive) is effectively compensated. A 6h-MTW brought gains of the same magnitude as if the length of the archive almost doubled. Note, that despite the number of candidate situations being in this case 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, rather strong serial correlations between sub-daily successive circulation patterns are expected.

Many reanalysis datasets start in the 80s due to availability of satellite data, and are thus more accurate than reanalysis datasets based on conventional data only, but they provide shorter archives. When using such an archive, the MTW approach shows a certain potential to enrich the pool of possible analogues. 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.

### 4.4 Reconstruction of a long precipitation archive suitable for the MTW

As we saw in Sect. 4.3, the skills of AMs improved with the length of the archive. It would then be even more profitable to apply the MTW to the longest archive possible, rather than being limited to the period where sub-daily precipitation data are available. Therefore, the idea is to reconstruct longer archives of moving 24h totals, from existing standard (fixed hour) precipitation series. For this purpose, disaggregation techniques might be used. Here, two simple disaggregation approaches of the daily precipitation time series were assessed.

The first technique was a proportional distribution, where the observed daily precipitations were considered constant during the measuring period (6:00 UTC to 6:00 UTC the next day). Proportional parts of the original daily time series were allocated into a new moving average of 24h totals.(e.g. a 0-24hUTC total would be made of ¼ of day one and ¾ of day 2 standard precipitation)

The second approach aimed at getting closer to the chronology of the actual precipitation by relying on some informative proxy variables during the reconstruction procedure. Data from the NCEP/NCAR reanalysis 1 were used for this purpose. Precipitable water and relative humidity (at 1000 hPa, 925 hPa or 850 hPa) were assessed at the four points surrounding the catchment. Precipitation time series from the reanalysis were not considered due to the presence of zeros, which may not always match with the prediction, leading to an undefined temporal repartition. A time lapse (-12h to +12h) 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, which was precipitable water, was identified by means of a correlation analyses (on non-zero values) with the 6-hourly precipitation time series. Once the best proxy had been selected, its temporal profile was used in order to disaggregate the daily precipitation time series.

These two simple methods did not result in valuable outputs (results not shown). Indeed, the performance improvement brought by the MTW was lost due to the poor quality of the reconstructedprecipitation archives. A slight improvement could be obtained for the second method relying on a proxy variable compared to the proportional distribution method, but it was still relatively small, and most of the benefit of the MTW was lost. Another reanalysis dataset with more accurate moisture variables might produce better proxies.

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

**5 Conclusions**

The 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 match optimally at the same hours of the day. It is probable that better matches may be found at another hour, which can change the selection of the analogue dates.

As Finet2008 had previously shown, the introduction of an MTW allows finding better analogue situations in terms of the atmospheric circulation. Using recent reanalysis datasets, MTWs with 12h, 6h, and 3h time steps were assessed. It has been demonstrated here that the improvement of the S1 criterion values was growing with the rank of the analogue. This was due to the accumulation of better analogues within the predicted distributions. The improvement of the S1 criterion increased for smaller MTW time steps.

A seasonal effect has been highlighted, as the MTW was more profitable for winter 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, a large part of the analogue situations were selected at a different time in winter.

The MTW improved the precipitation prediction, as the CRPS globally decreased for all stations, both on the CP and the VP. The prediction skill on the CP was almost always improved further by reducing the time step of the MTW, but not of the same magnitude. Results on the VP showed globally the same trends, but with more variability. It was also shown that the reliability of the considered AMs improved for the prediction of different threshold exceedences, especially with a 3h-MTW.

Both improvements of the analogy criterion and the performance scores were found to be higher for MTWs with a smaller time step (3h). A 3h-MTW provides 8 times more candidate situations, even though not fully independent. As the MTW provides additional situations to the pool of possible analogues, it can be seen as an inflation of the archive.

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

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

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

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

The technique being very simple and easily applicable, it should be considered for several applications in different contexts, may it be for operational forecasting or climate-related studies.

## A Acronyms

[Sorry. Ignored \begin{labeling} ... \end{labeling}]

The authors declare that they have no conflict of interest.

[Sorry. Ignored \begin{acknowledgements} ... \end{acknowledgements}]

References

[12010Ben Daoud] Ben Daoud, A.: Améliorations et développements d’une méthode de prévision probabiliste des pluies par analogie., Ph.D. thesis, Université de Grenoble, 2010.

[22008Ben Daoud et al.Ben Daoud, Sauquet, Lang, Obled, and Bontron] 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.

[32016Ben Daoud et al.Ben Daoud, Sauquet, Bontron, Obled, and Lang] 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, 10.1016/j.atmosres.2015.09.015, 2016.

[42010Bliefernicht] Bliefernicht, J.: Probability forecasts of daily areal precipitation for small river basins, Ph.D. thesis, Universität Stuttgart, 2010.

[52004Bontron] 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.

[62005Bontron and Obled] Bontron, G. and Obled, C.: L’adaptation probabiliste des prévisions météorologiques pour la prévision hydrologique, La Houille Blanche, 1, 23–28, 10.1051/lhb:200501002, 2005.

[71974Brown] Brown, T.: Admissible Scoring Systems for Continuous Distributions., Tech. rep., URL http://eric.ed.gov/?id=ED135799, 1974.

[82016Caillouet et al.Caillouet, Vidal, Sauquet, and Graff] 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, 10.5194/cp-12-635-2016, 2016.

[92014Chardon et al.Chardon, Hingray, Favre, Autin, Gailhard, Zin, and Obled] 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, 10.1175/JCLI-D-13-00464.1, 2014.

[102015Dayon et al.Dayon, Boé, and Martin] 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, 10.1002/2014JD022236, 2015.

[112008Desaint et al.Desaint, Nogues, Perret, and Garçon] 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., 10.1051/lhb/2009054, 2008.

[121970Duband] 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.

[132008Finet et al.Finet, Marty, Zin, and Obled] 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.

[142009García Hernández et al.García Hernández, Horton, Tobin, and Boillat] 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.

[152006Hamill and Whitaker] Hamill, T. and Whitaker, J.: Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application, Mon. Weather Rev., 134, 3209–3229, 10.1175/mwr3237.1, 2006.

[162006Hamill et al.Hamill, Whitaker, and Mullen] 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, 10.1175/BAMS-87-1-33, 2006.

[172015Hamill et al.Hamill, Scheuerer, and Bates] 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, 10.1175/MWR-D-15-0004.1, 2015.

[182000Hersbach] Hersbach, H.: Decomposition of the continuous ranked probability score for ensemble prediction systems, Weather Forecast., 15, 559–570, 10.1175/1520-0434(2000)015<0559:dotcrp>2.0.co;2, 2000.

[192012Horton] 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, URL https://tel.archives-ouvertes.fr/tel-01441762, 2012.

[202012Horton et al.Horton, Jaboyedoff, Metzger, Obled, and Marty] 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, 10.5194/nhess-12-777-2012, 2012.

[212017Horton et al.Horton, Jaboyedoff, and Obled] Horton, P., Jaboyedoff, M., and Obled, C.: Global Optimization of an Analog Method by Means of Genetic Algorithms, Mon. Weather Rev., 145, 1275–1294, 10.1175/MWR-D-16-0093.1, 2017.

[222015Junk et al.Junk, Delle Monache, Alessandrini, Cervone, and von Bremen] 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, 10.1127/metz/2015/0659, 2015.

[231996Kalnay et al.Kalnay, Kanamitsu, Kistler, Collins, Deaven, Gandin, Iredell, Saha, White, and Woollen] 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, 10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2, 1996.

[241956Lorenz] 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.

[251969Lorenz] Lorenz, E.: Atmospheric predictability as revealed by naturally occurring analogues, J. Atmos. Sci., 26, 636–646, 10.1175/1520-0469(1969)26<636:aparbn>2.0.co;2, 1969.

[262010Maraun et al.Maraun, Wetterhall, Chandler, Kendon, Widmann, Brienen, Rust, Sauter, Themeßl, Venema, Chun, Goodess, Jones, Onof, Vrac, and Thiele-Eich] Maraun, D., Wetterhall, F., Chandler, R. E., Kendon, E. J., Widmann, M., Brienen, S., Rust, H. W., Sauter, T., Themeßl, M., Venema, V. K. C., Chun, K. P., Goodess, C. M., Jones, R. G., Onof, C., Vrac, M., and Thiele-Eich, I.: Precipitation downscaling under climate change: Recent developements to bridge the gap between dynamical models and the end user, Rev. Geophys., 48, 1–34, 10.1029/2009RG000314, 2010.

[272010Marty] Marty, R.: Prévision hydrologique d’ensemble adaptée aux bassins à crue rapide. Elaboration de prévisions probabilistes de précipitations à 12 et 24 h. Désagrégation horaire conditionnelle pour la modélisation hydrologique. Application à des bassins de la région Cév, Ph.D. thesis, Université de Grenoble, 2010.

[282012Marty et al.Marty, Zin, Obled, Bontron, and Djerboua] Marty, R., Zin, I., Obled, C., Bontron, G., and Djerboua, A.: Toward real-time daily PQPF by an analog sorting approach: Application to flash-flood catchments, J. Appl. Meteorol. Clim., 51, 505–520, 10.1175/JAMC-D-11-011.1, 2012.

[291976Matheson and Winkler] Matheson, J. and Winkler, R.: Scoring rules for continuous probability distributions, Manage. Sci., 22, 1087–1096, 10.1287/mnsc.22.10.1087, 1976.

[302014Obled] Obled, C.: Daniel Duband - cinquante ans de contributions scientifiques à l’hydrologie (1962-2011), La Houille Blanche, pp. 55–68, 10.1051/lhb/2014017, 2014.

[312002Obled et al.Obled, Bontron, and Garçon] 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, 10.1016/S0169-8095(02)00038-8, 2002.

[322016Poli et al.Poli, Hersbach, Dee, Berrisford, Simmons, Vitart, Laloyaux, Tan, Peubey, Thépaut, Trémolet, Hólm, Bonavita, Isaksen, and Fisher] Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G. H., Peubey, C., Thépaut, J. N., Trémolet, Y., Hólm, E. V., Bonavita, M., Isaksen, L., and Fisher, M.: ERA-20C: An atmospheric reanalysis of the twentieth century, J. Climate, 29, 4083–4097, 10.1175/JCLI-D-15-0556.1, 2016.

[332013Radanovics et al.Radanovics, Vidal, Sauquet, Ben Daoud, and Bontron] 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, 10.5194/hess-17-4189-2013, 2013.

[342016Raynaud et al.Raynaud, Hingray, Zin, Anquetin, Debionne, and Vautard] 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., 10.1002/joc.4844, 2016.

[351988Ruosteenoja] 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.

[361954Teweles and Wobus] Teweles, S. and Wobus, H. B.: Verification of prognostic charts, B. Am. Meteorol. Soc., 35, 455–463, 1954.

[371994Van Den Dool] Van Den Dool, H. M.: Searching for analogues, how long must we wait?, Tellus, 46A, 314–324, 1994.

[381999Zorita and Storch] Zorita, E. and Storch, H. V.: The analog method as a simple statistical downscaling technique: comparison with more complicated methods, J. Climate, 12, 2474–2489, 10.1175/1520-0442(1999)012<2474:TAMAAS>2.0.CO;2, 1999.

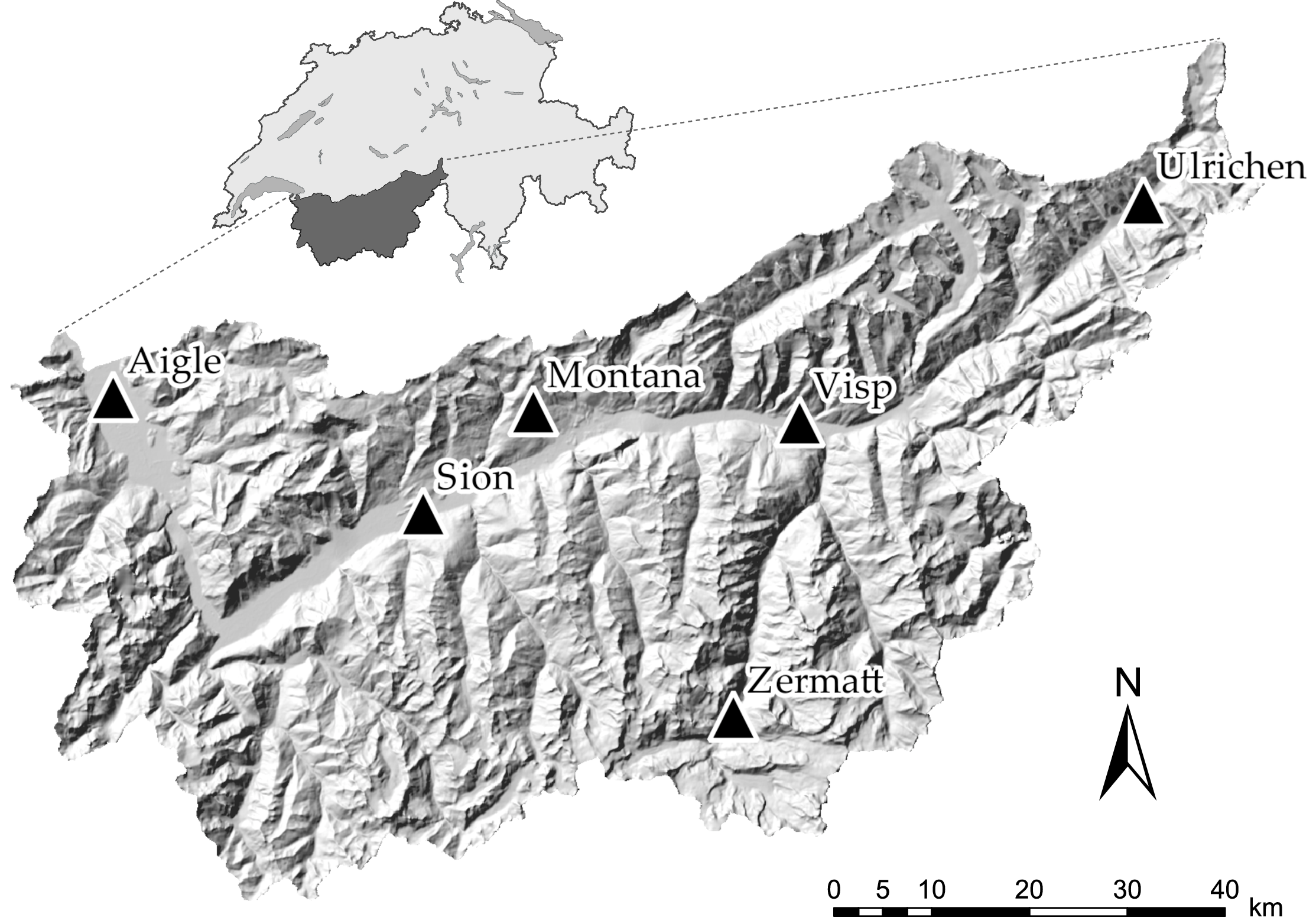


Figure 1: Position of the six weather stations of interest (Aigle, Montana, Sion, Ulrichen, Visp, and Zermatt) in the upper Rhône catchment in Switzerland.

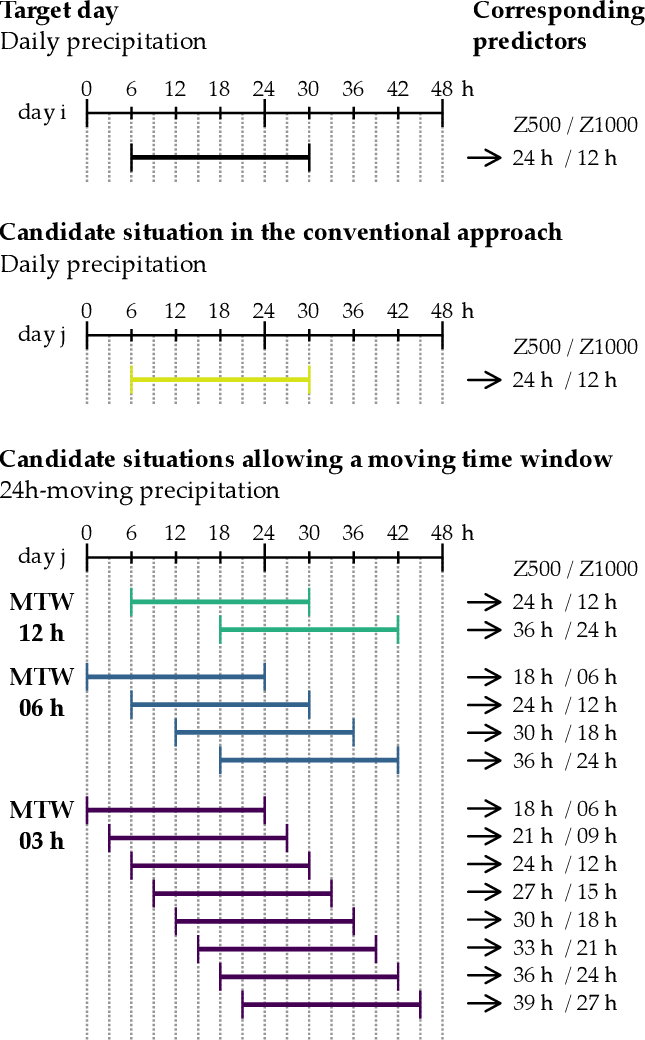


Figure 2: Illustration of the principle of a moving time window (MTW). The target situation is the same for the conventional approach and the MTW, while the candidate situations are 2, 4, or 8 times as many with the MTW. The horizontal bars represent the 24h precipitation totals, with their associated predictors being listed on the right hand side.

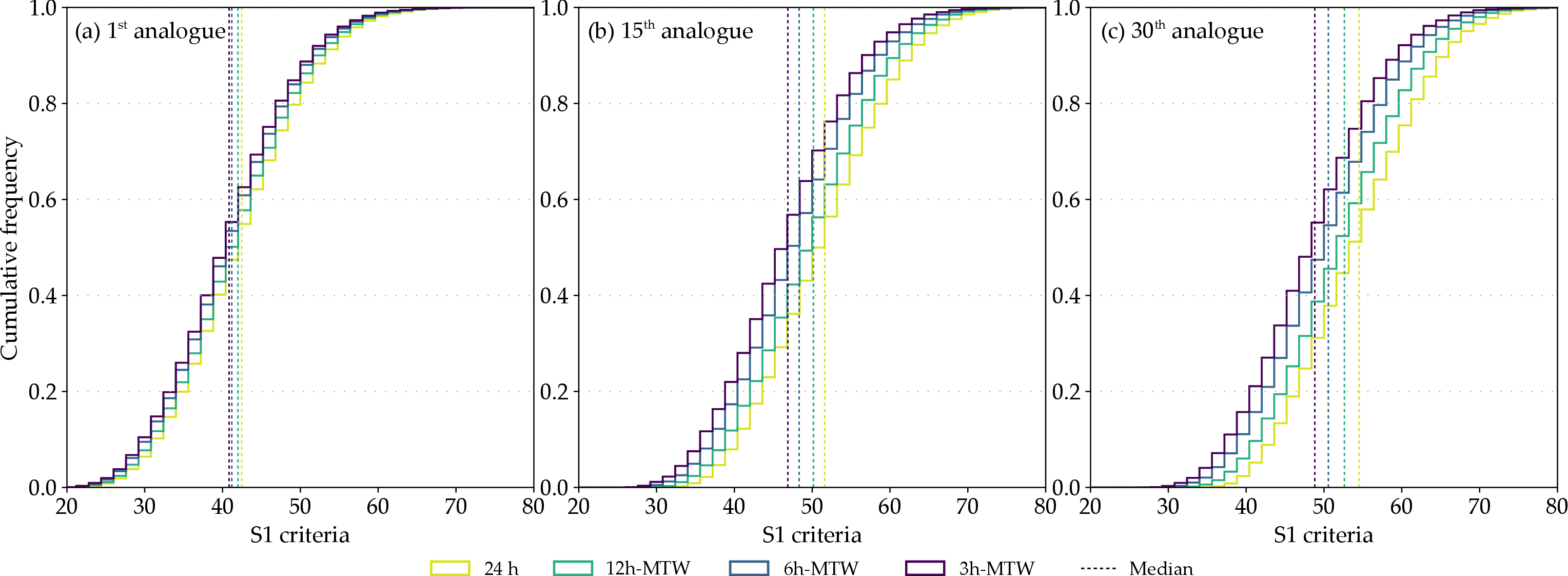


Figure 3: Changes in the S1 criterion distributions due to the introduction of an MTW with a 12h, 6h or 3h time step. Distributions are provided for (a) the , (b) , and (c) analogue ranks for all days of the CP at the Ulrichen station.

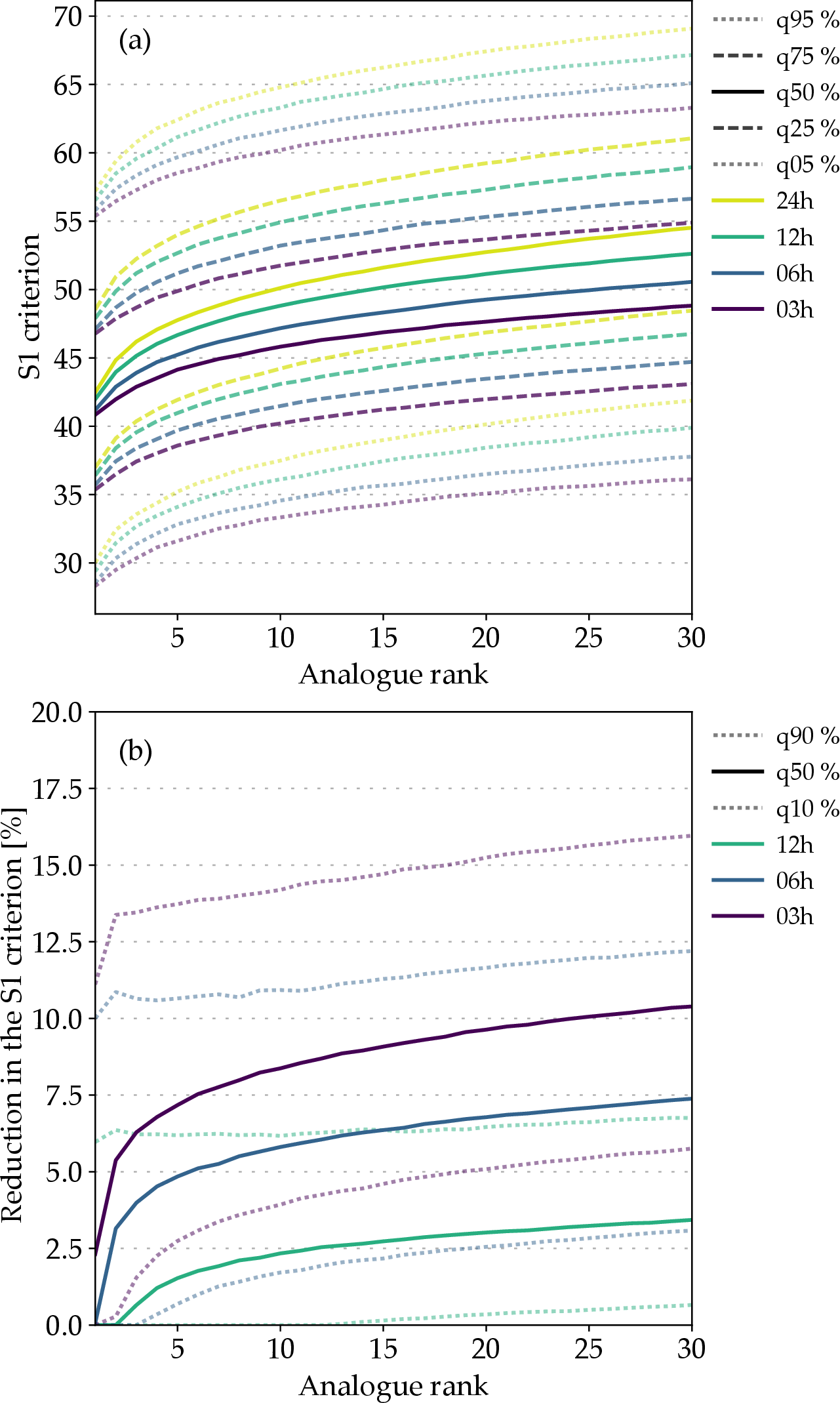


Figure 4: Synthesis of the changes in the S1 criterion due to the MTW (with a 12h, 6h, and 3h time step) for the Ulrichen station depending on the rank of the analogue. (a) Quantiles of the S1 distributions corresponding to the MTW with different time steps (24h is the conventional approach without MTW).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

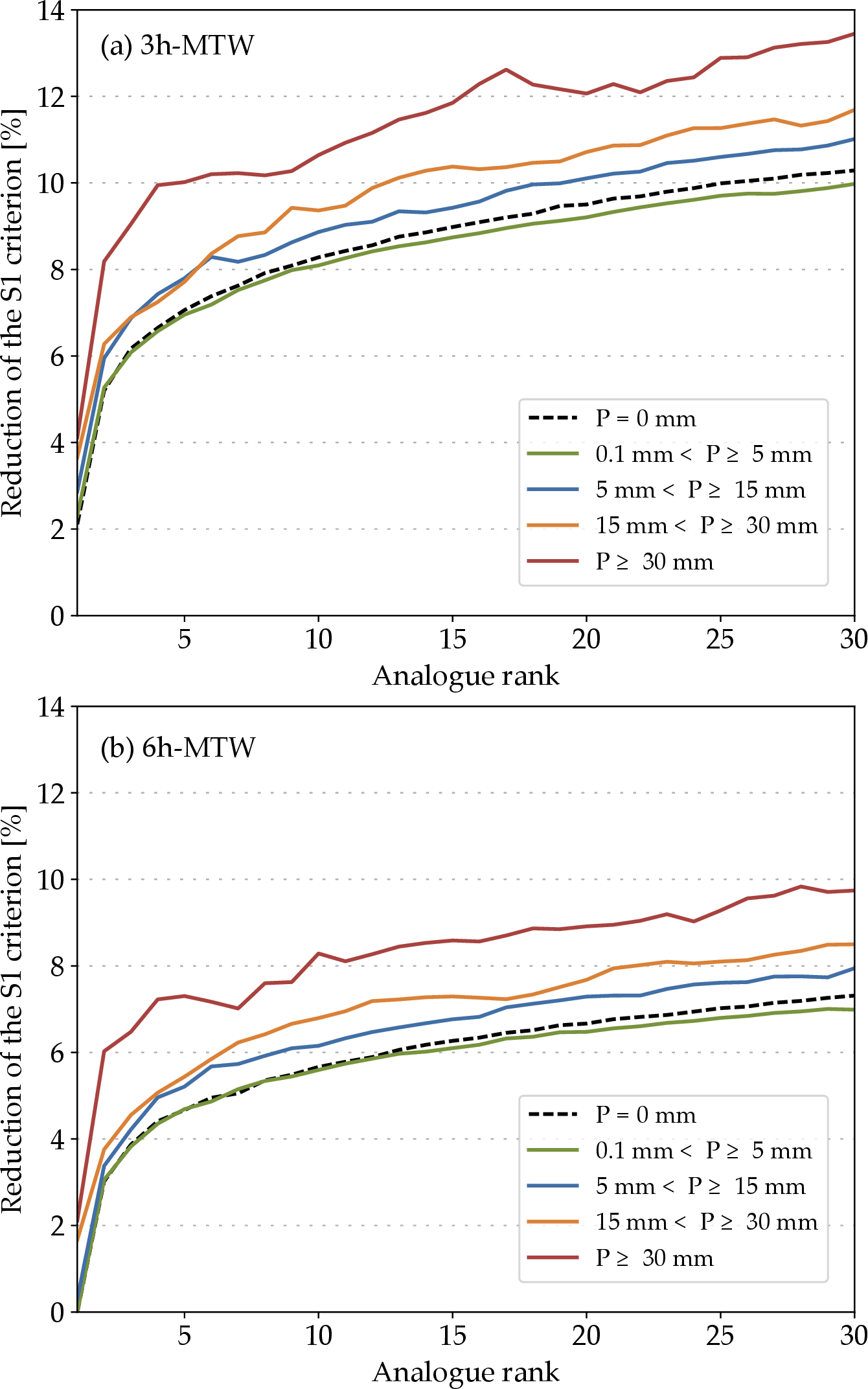


Figure 5: Distribution of the median improvements of the S1 criterion due to the (a) 3h-MTW and (b) 6h-MTW, depending on daily precipitation thresholds at the Ulrichen station.

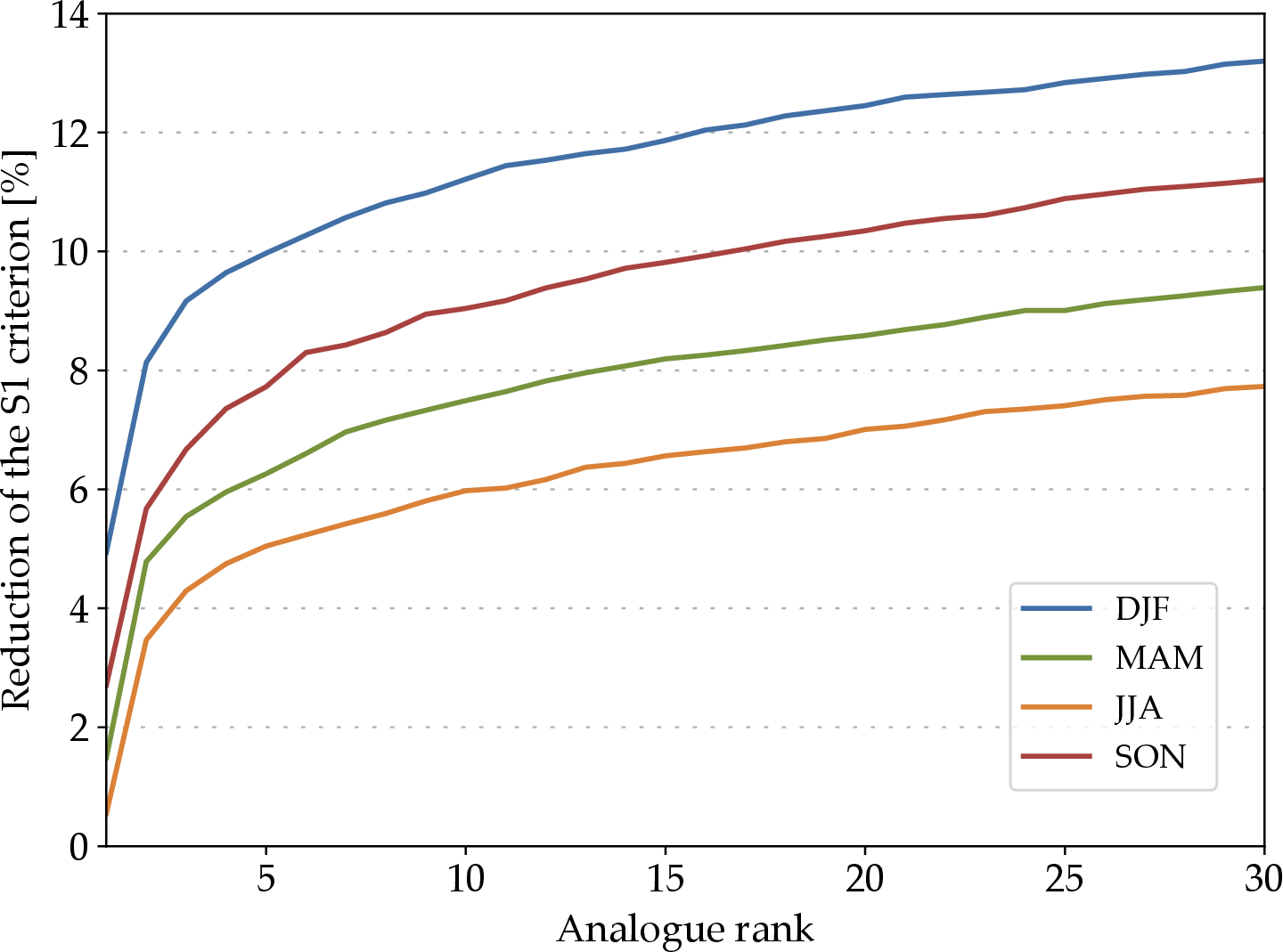


Figure 6: Seasonal effect on the median reduction of the S1 criterion for the Ulrichen station due to the MTW. DJF: winter, MAM: spring, JJA: summer, and SON: fall.

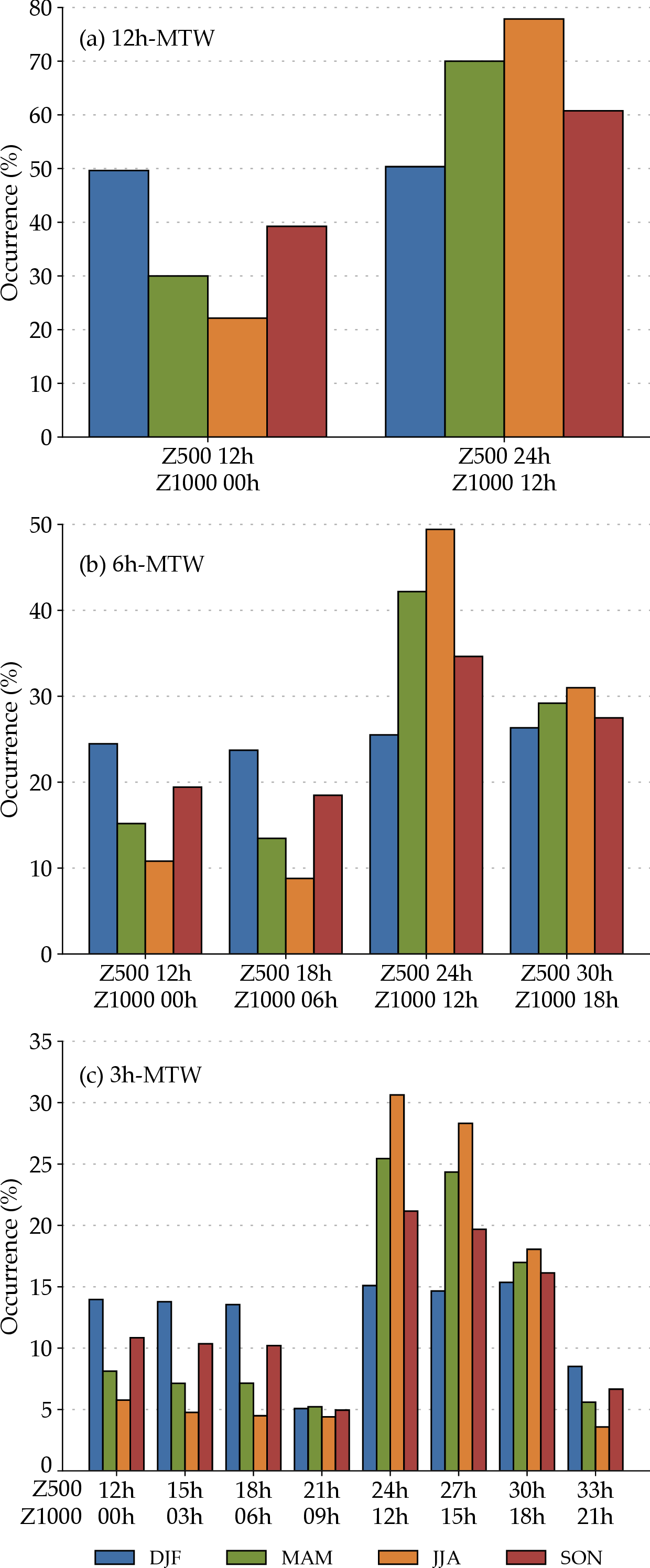


Figure 7: Distribution of the predictor hours on the selected analogue dates for the Ulrichen station when using a (a) 12h-MTW, (b) 6h-MTW, and (c) 3h-MTW, depending on the season.

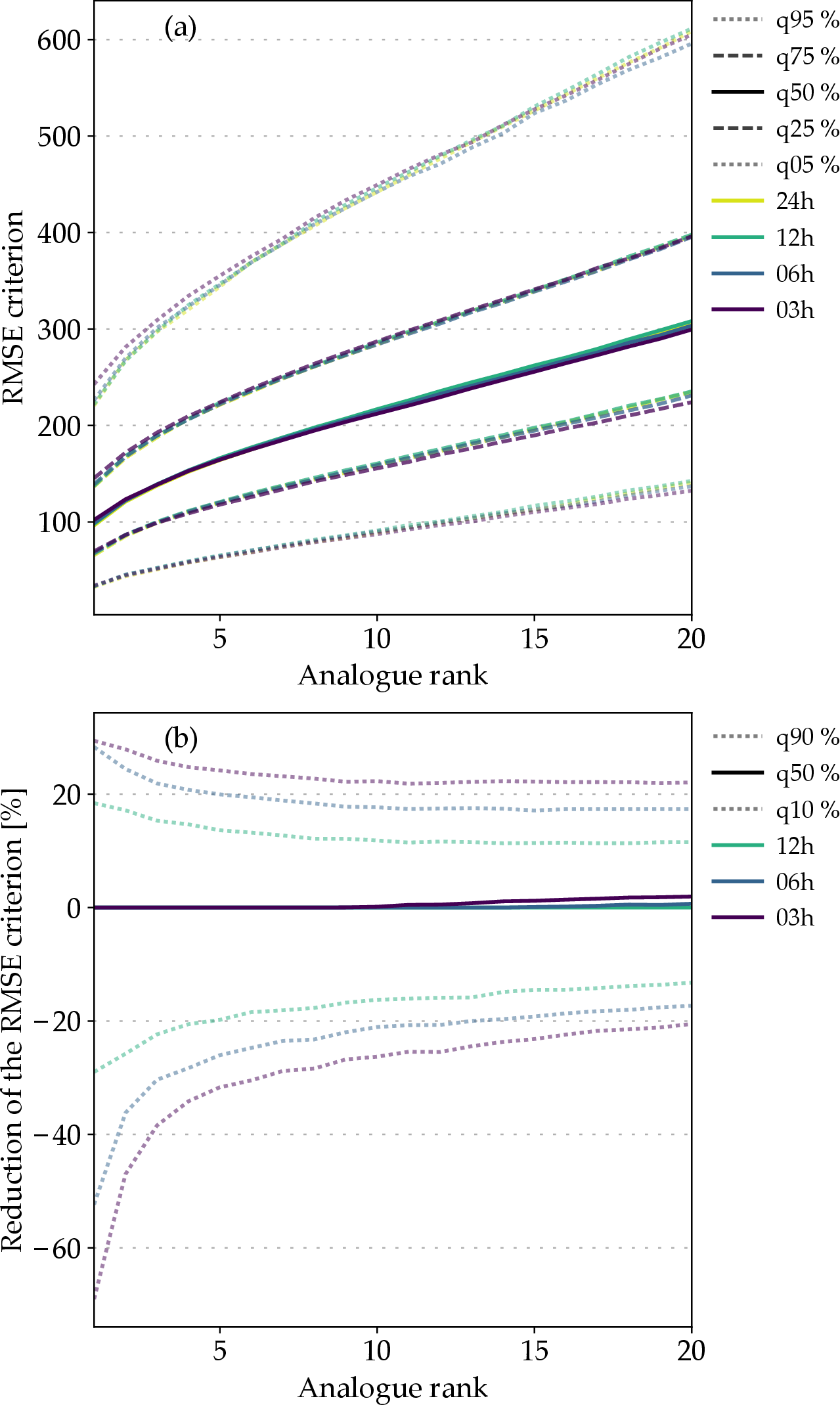


Figure 8: Synthesis of the changes in the RMSE criterion due to the MTW for the Ulrichen station depending on the rank of the analogue. (a) Quantiles of the RMSE distributions with and without the MTW. (b) Quantiles of the relative improvements of the RMSE criterion when using the MTW.

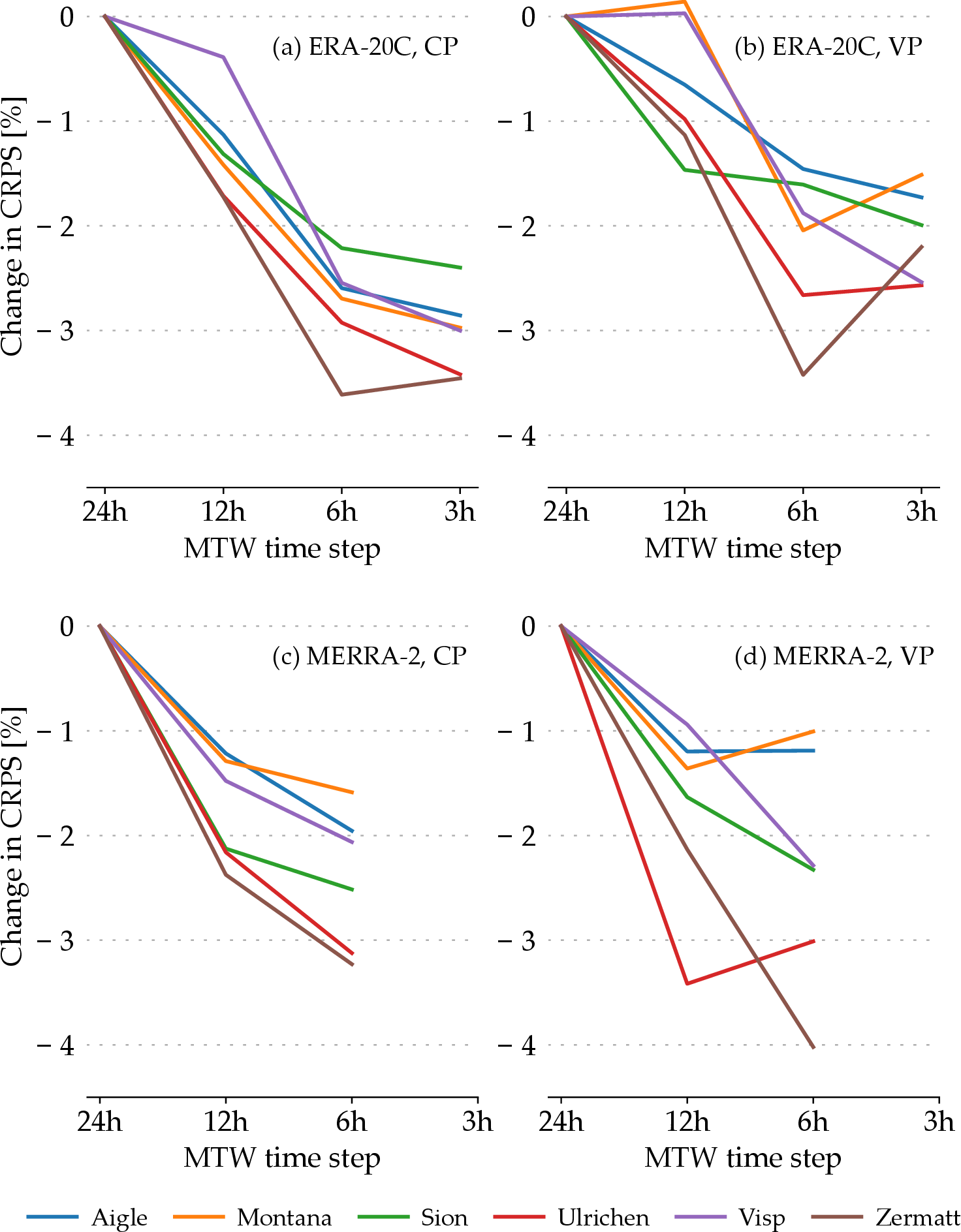


Figure 9: Change in performance score (CRPS – lower is better) of the 2Z method at the different stations for MTW of varying time steps relatively to the conventional approach (24h). Results are provided for ERA-20C on (a) the CP and (b) the VP, as well as for MERRA-2 on (c) the CP and (d) the VP.

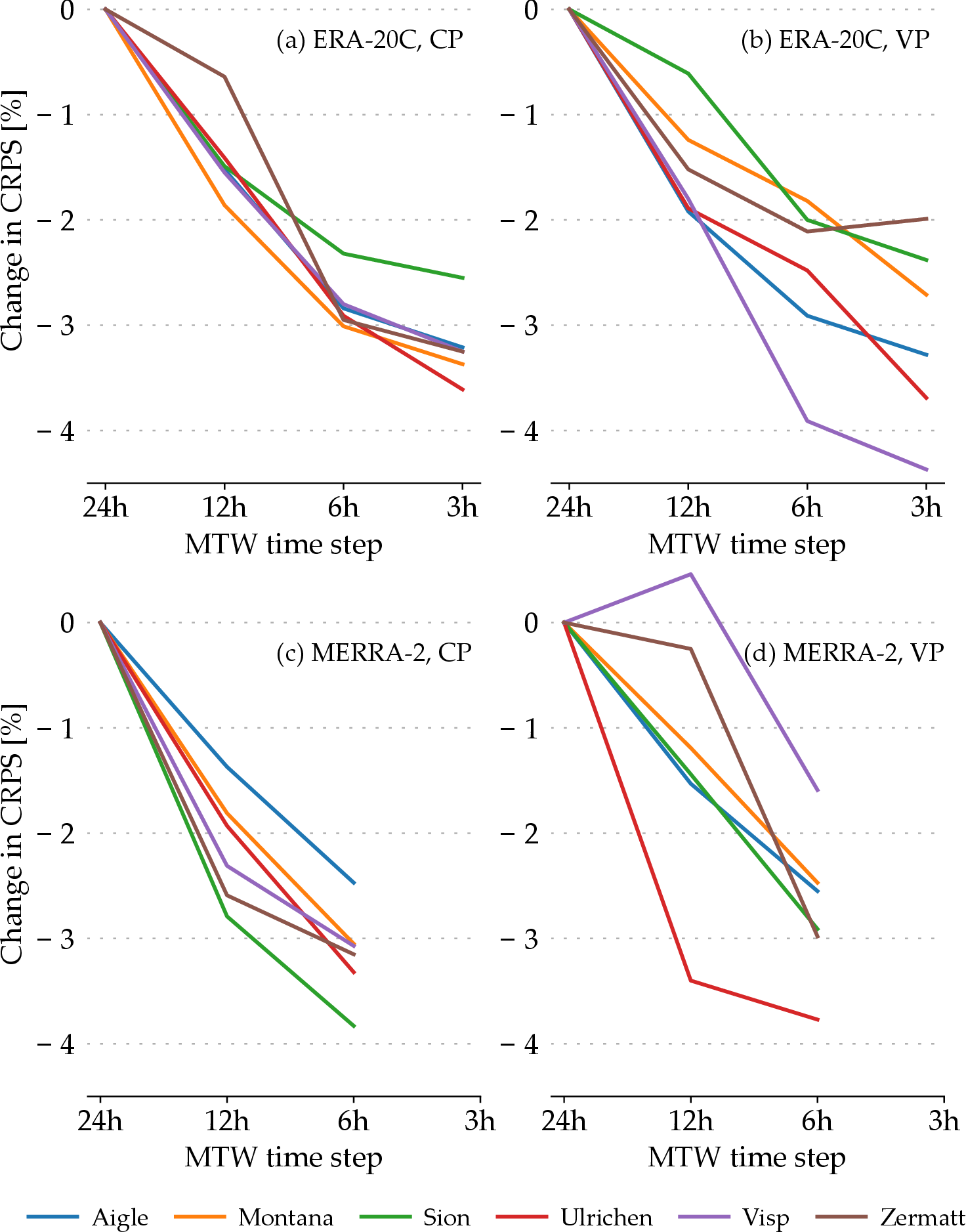


Figure 10: Same as Figure 9, but for the 2Z-2MI method.

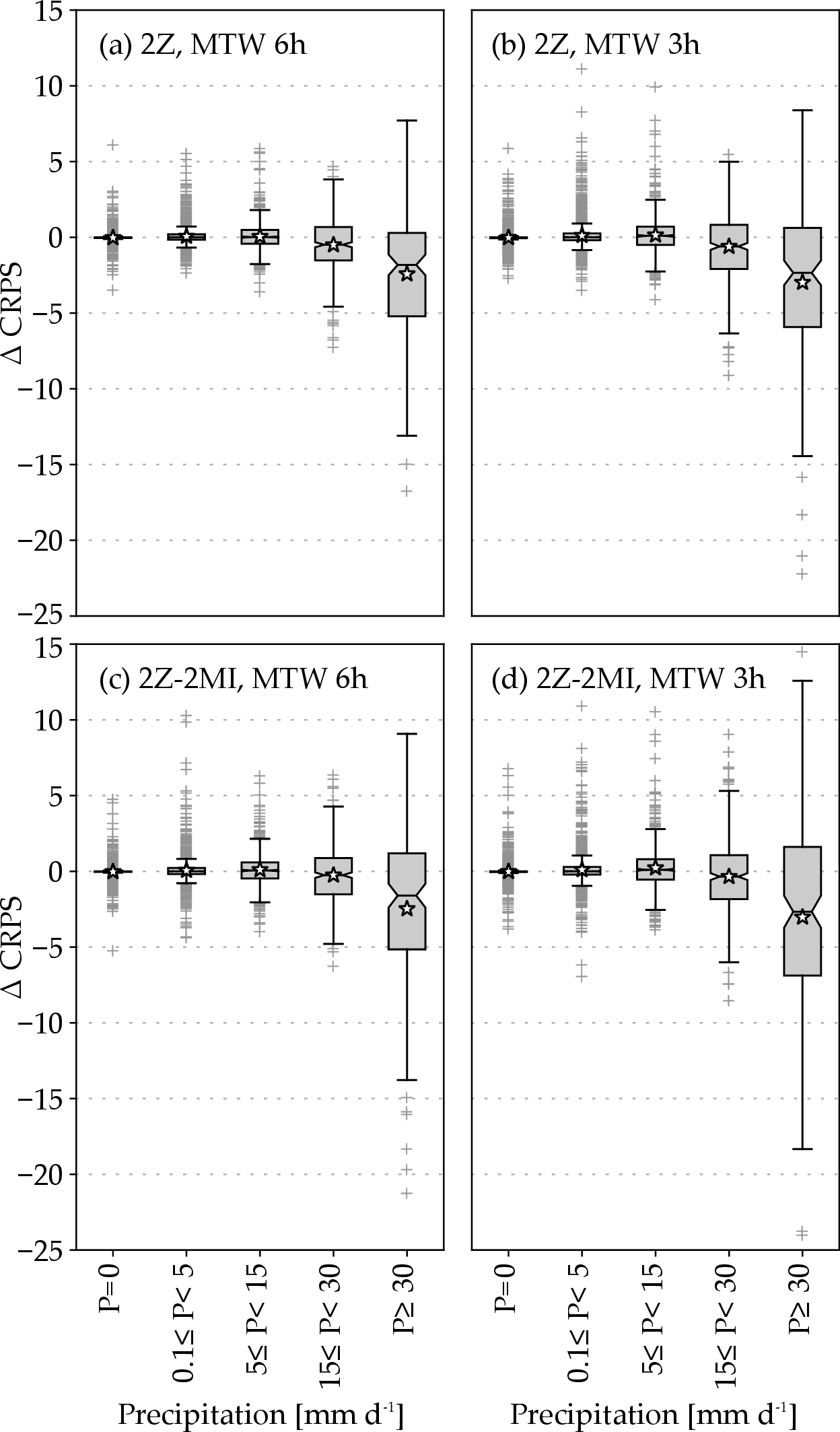


Figure 11: Differences of the CRPS performance score (lower is better) at the Ulrichen station due to the introduction of the MTW as a function of different daily precipitation thresholds (for the target date). The results are provided for (top) the 2Z method and (bottom) the 2Z-2MI method, and with a (left) 6h-MTW and a (right) 3h-MTW. The stars represent averages.

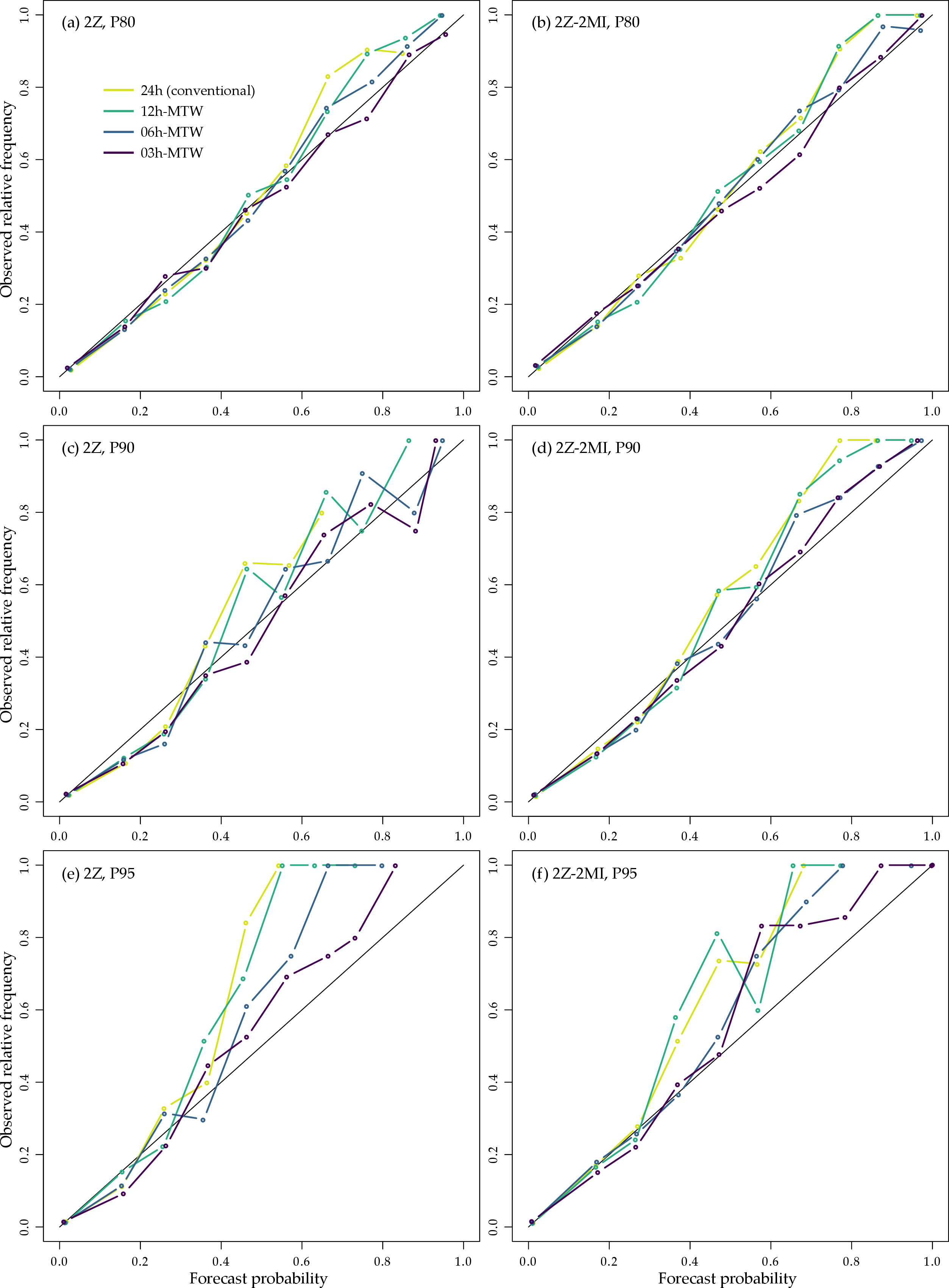


Figure 12: Reliability diagrams for (left) the 2Z (right) the 2Z-2MI methods and the prediction of the exceedence of percentiles (top) 80 %, (middle) 90 %, and (bottom) 95 % at the Ulrichen station on the VP. The conventional approach (24h) is provided as well as the 12h-MTW, 6h-MTW and 3h-MTW.

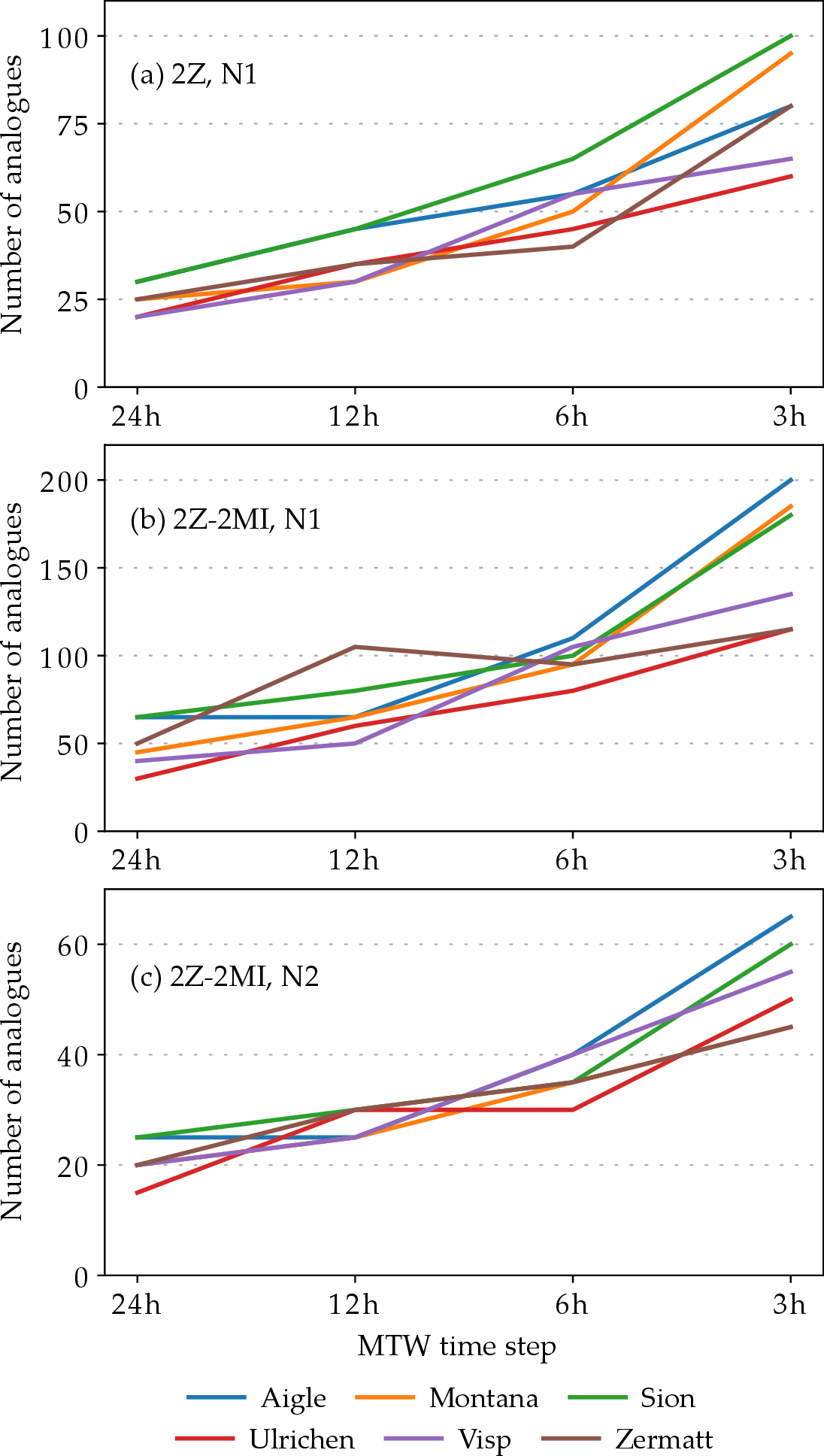


Figure 13: Optimal number of analogues (of the first and second level of analogy, respectively, and ) of the (a) 2Z and (b, c) 2Z-2MI methods after recalibration using the MTW.

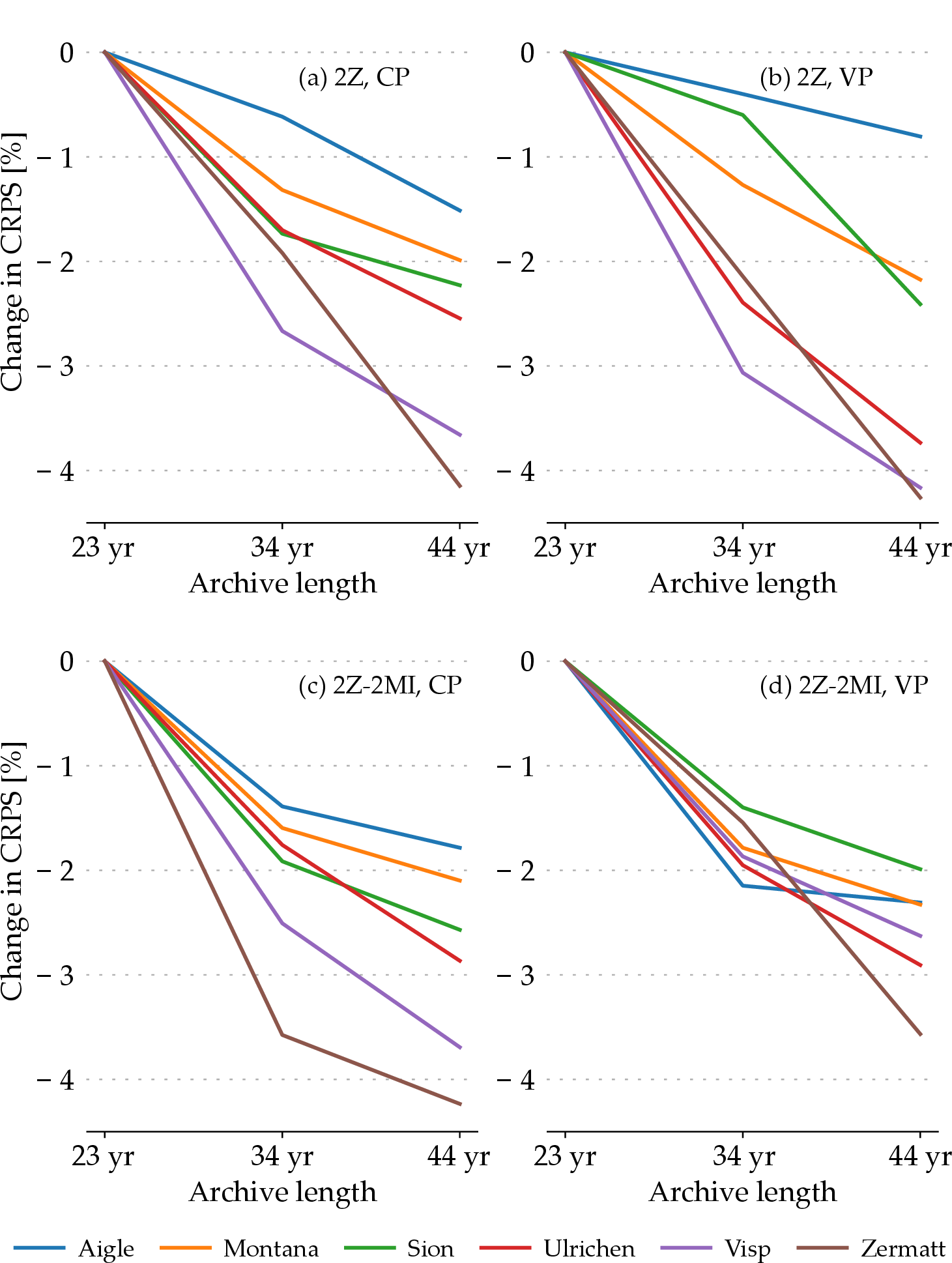


Figure 14: Change in performance score (CRPS – lower is better) of the (top) 2Z and (bottom) 2Z-2MI methods at the different stations on the (left) CP and (right) VP, for an increasing archive length with the conventional approach (24h windows).

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

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Level | Variable | Hour | Criterion | Nb |
| 0 | ±60 days around the target date | | | |
| 2\*1 | Z1000 | 12 h | 2\*S1 | 2\* |
|  | Z500 | 24 h |  |  |

Table 2: Parameters of the reference method with moisture variables (2Z-2MI). Same conventions as Table

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