##### The Analogue Method for Precipitation Forecasting: Finding Better Analogue Situations at a Sub-Daily Time Step

Fri Dec 18 14:58:37 2015

The Analogue Method (AM) aims at forecasting local weather variables (predictands), such as precipitations, 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, in order to sample the days with a similar atmospheric circulation.

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, the same for both the target day and the candidate analogues, the constraint being the use of daily time series, due to the length of available archives they provide. However, it is unlikely that the best analogy happens at the very same hour, but it may occur at a different time of the day. In order to assess the potential of finding better analogues at a different hour, a sliding time window (STW) has been introduced.

The STW resulted in a better analogy in terms of the atmospheric circulation, with improved values of the analogy criteria on the whole distribution of analogue dates. The improvement was found to grow with the analogue ranks due to an accumulation of more similar situations 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, which favors predictors at the same hour for target and analogues.

The impact of the STW on the prediction performance 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 after recalibrating the AM parameters.

However, attempts to reconstruct longer precipitation series of running 24 h-totals by means of simple methods failed. It emphasized the need to use time series with an appropriate chronology. These should be available in a near future, either by means of growing observed archives, or by the establishment of precipitation reanalyses through regional modeling. Then, the use of a STW in the AM should be considered for any application, especially when the prediction quality of extreme events is important.

The Analogue Method (AM) is based on the principle that two relatively similar synoptic situations may produce relatively similar local effects [Lorenz(1956), Lorenz(1969)]. Multiple variations of the methods exist and listings can be found in [Horton et~al.(2016)Horton, Jaboyedoff, and Obled] or [Ben Daoud et~al.(2015)Ben Daoud, Sauquet, Bontron, Obled, and Lang]. The versions that will be considered here are often used as references or benchmarks for various improvements.

The first method is based on the analogy of the atmospheric circulation only [?, Table :method\_2Z,]]Obled2002, Bontron2005, Marty2012. Searching analogue situations for a target day starts by a preselection step of the potential candidate for analogy. Hereit 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 S1 criteria [?, Eq.:S1,]]Teweles1954, Drosdowsky2003, which is a comparison of gradients, over a certain spatial window and at certain hours. In fact, the target variable or predictandis not an instantanous value but often an accumulation over a certain time period (e.g. 24 h for rainfall, taken from 6-30 h UTC). Therefore, the meteorological situation generating this accumulated rainfall is not instantanous either, and is a mix of the instantantanous fieldsobserved at different times and heightsaround and within the predictand accumulation period. . In our sample case, the considered pressure levels are 1000 hPa (Z1000) at 12 h UTC and 500 hPa (Z500)at 24 h UTC).

Then  dates with the lowest values of S1 are considered as analogues to the target day,  being a parameter to calibrate. Then, the daily observed precipitation amount of the corresponding dates provide the empirical conditional distribution considered as the probabilistic forecast for the target day. Following the nomenclature proposed by [Horton et~al.(2016)Horton, Jaboyedoff, and Obled], this method will be named 2Z.

 (1)

where  is the forecast geopotential height difference between the *i*th pair of adjacent points of the gridded data describing the target situation, and  is the corresponding observed geopotential height difference in the candidate situation. The differences are processed separately in both directions. The smaller the S1 values, the more similar the pressure fields.

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

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

The second reference method adds a subsequent level of analogy on moisture variables (method 2Z-2MI, Table 2). This second predictor is a moisture index made 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,  dates are subsampled in the  analogues on the atmospheric circulation, to end up with 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.

Table 2: Parameters of the reference method with moisture variables (2Z-2MI). Same conventions as Table 1

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

To calibrate the method, the commonly used technique is a semi-automatic sequential procedure elaborated by [Bontron(2004)]. AtmoSwing [Horton et~al.(2016)Horton, Jaboyedoff, and Obled], 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 CRPS [?, Continuous Ranked Probability Score,]]Brown1974, Matheson1976, Hersbach2000. It allows evaluating the predicted cumulative distribution functions , for example of the precipitation values  from analogue situations, compared to the observed value . The better the forecast, the smaller the score. The mean CRPS of a forecast series of length  can be written:

 (2)

where  is the Heaviside function that is null when , and has the value 1 otherwise. The mean CRPS is averaged on the calibration, respectively the validation periods, on all days. [Bontron(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 the climatological distribution as the reference. The CRPSS (*Continuous Ranked Probability Skill Score*) is thus defined as following:

 (3)

where  is the CRPS value for the reference and  would be the one for a perfect forecast (which implies ). A better forecast is characterized by an increase in CRPSS.

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 do 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 sliding time window (STW) was introduced to allow searching for candidates at different hours of the day.

Previous tests showed the benefit, in terms of analogy criteria values, of searching for analogue synoptic situations at a finer time step, but without assessing the impact on the forecast skills [Finet et~al.(2008)Finet, Marty, Zin, and Obled]. 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 (section 1). It was then aggregated in the form of 24 h-totals, but starting at different 6-hourly time step, by means of a moving average (section 2).

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), 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 24h daily totals.

The changes due to the introduction of a STW was assessed first in regards to the analogy criteria improvement (section 4), and then in terms of precipitation prediction skill (section 5). Finally, attempts were made to reconstruct longer 24 h-total precipitations series (section 6).

## 1 Study area and data

The study area is the upper Rhône catchment in Switzerland [?, see also]]Horton2012a. 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 the period 1982-2007) come from 6 automatic weather stations, namely Ulrichen, Zermatt, Visp, Montana, Sion and Aigle (Figure 1), which are subject to various meteorological influences [Horton et~al.(2012)Horton, Jaboyedoff, Metzger, Obled, and Marty]. The results will thereafter be presented arbitrarily for the Ulrichen station, but they are equivalent for all stations.

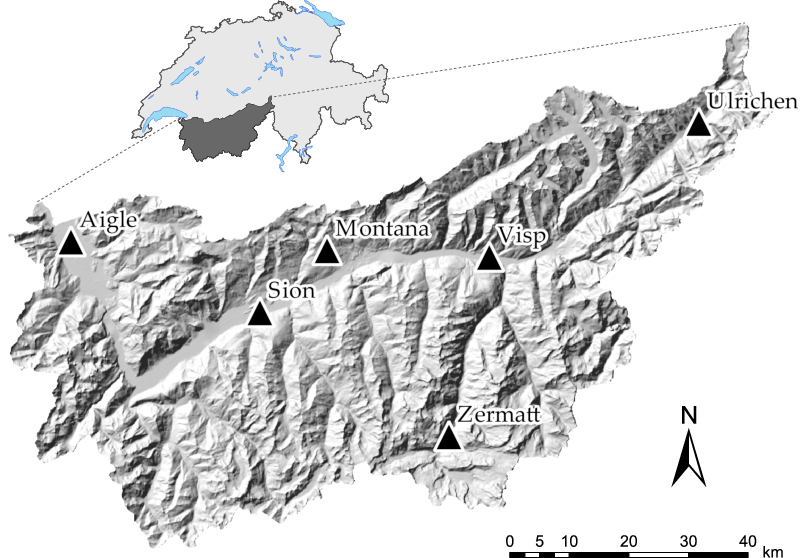


Figure 1: Position of the 6 weather stations of interest (Ulrichen, Zermatt, Visp, Montana, Sion and Aigle) in the study area.

Predictors are extracted from the NCEP/NCAR reanalysis I [Kalnay et~al.(1996)Kalnay, Kanamitsu, Kistler, Collins, Deaven, Gandin, Iredell, Saha, White, and Woollen] dataset with a 6-hourly temporal resolution, 17 pressure levels, and a spatial resolution of 2.5.

## 2 The sliding time window (STW) 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 (section 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.

The target situations and their corresponding observed precipitation values (used for validation) do not change, because the forecast is always established for a fixed period of the target day (6-30 h), as before. The difference is that the candidates are 4 times more numerous (Figure 2).

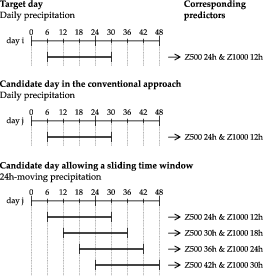


Figure 2: Illustration of the principle of a sliding time window.

It could have been expected that, when a given time window is selected, the neighbouring ones could also be selected even with a lower ranking. ThereforeLogic an exclusion of the same dates was tested in two different ways: (1) by excluding the worse analogues in a period of  24 h, and (2) by excluding the worse analogues with the same date. After several tests, it turned out that no restriction led to better skills, and this on a significant and recurring basis. When an STW emerges, it proves superior to the neighbouring ones, at least at 6h timestep.Thus, no constraint was further used.

## 3 Consequences of the archive reduction

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 reduction of a 47 years archive (1961 to 2008) to 25 years (1982 to 2007). This change is assessed with the original method, without STW.

Both 2Z (Table 1) and 2Z-2MI (Table 2) methods were considered. The AM parameters were calibrated using AtmoSwing [Horton et~al.(2016)Horton, Jaboyedoff, and Obled] on the original archive (Tables 3 and 4) and will be used thereafter.

Table 3: Calibrated parameters (spatial windows and number of analogues) for the analogy on the geopotential at 500 hPa and 1000 hPa (method 2Z) and skill score (%) of the method on the full archive.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Station | Longitudes | Latitudes | N | CRPSS |
| Ulrichen | 0  17.5 | 42.5  47.5 | 40 | 30.73 |
| Zermatt | 0  20 | 37.5  50 | 35 | 23.87 |
| Visp | -2.5  20 | 40  47.5 | 30 | 25.11 |
| Montana | -2.5  17.5 | 40  47.5 | 40 | 32.55 |
| Sion | -2.5  17.5 | 40  47.5 | 40 | 26.23 |
| Aigle | -5  17.5 | 40  50 | 50 | 30.59 |

Table 4: Parameters of the moisture variables from the 2Z-2MI method and the corresponding skill score (%) on the complete archive. The parameters for the atmospheric circulation are the same as in Table 3, except the number of analogues of the first analogy level (N), which are here different. The  column is the number of analogues for the second level (N).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Station | Longitudes | Latitudes | N | N | CRPSS |
| Ulrichen | 5  10 | 45  47.5 | 60 | 25 | 34.31 |
| Zermatt | 5  10 | 45  47.5 | 55 | 25 | 28.28 |
| Visp | 5  10 | 45  47.5 | 45 | 25 | 28.85 |
| Montana | 5  7.5 | 45  47.5 | 55 | 30 | 36.11 |
| Sion | 5  10 | 45  47.5 | 90 | 30 | 31.16 |
| Aigle | 7.5 | 45  47.5 | 100 | 35 | 35.82 |

Table 5: Influence of the archive reduction on the CRPSS. CRPSS values are provided for both considered methods and the differences are expressed in absolute value.

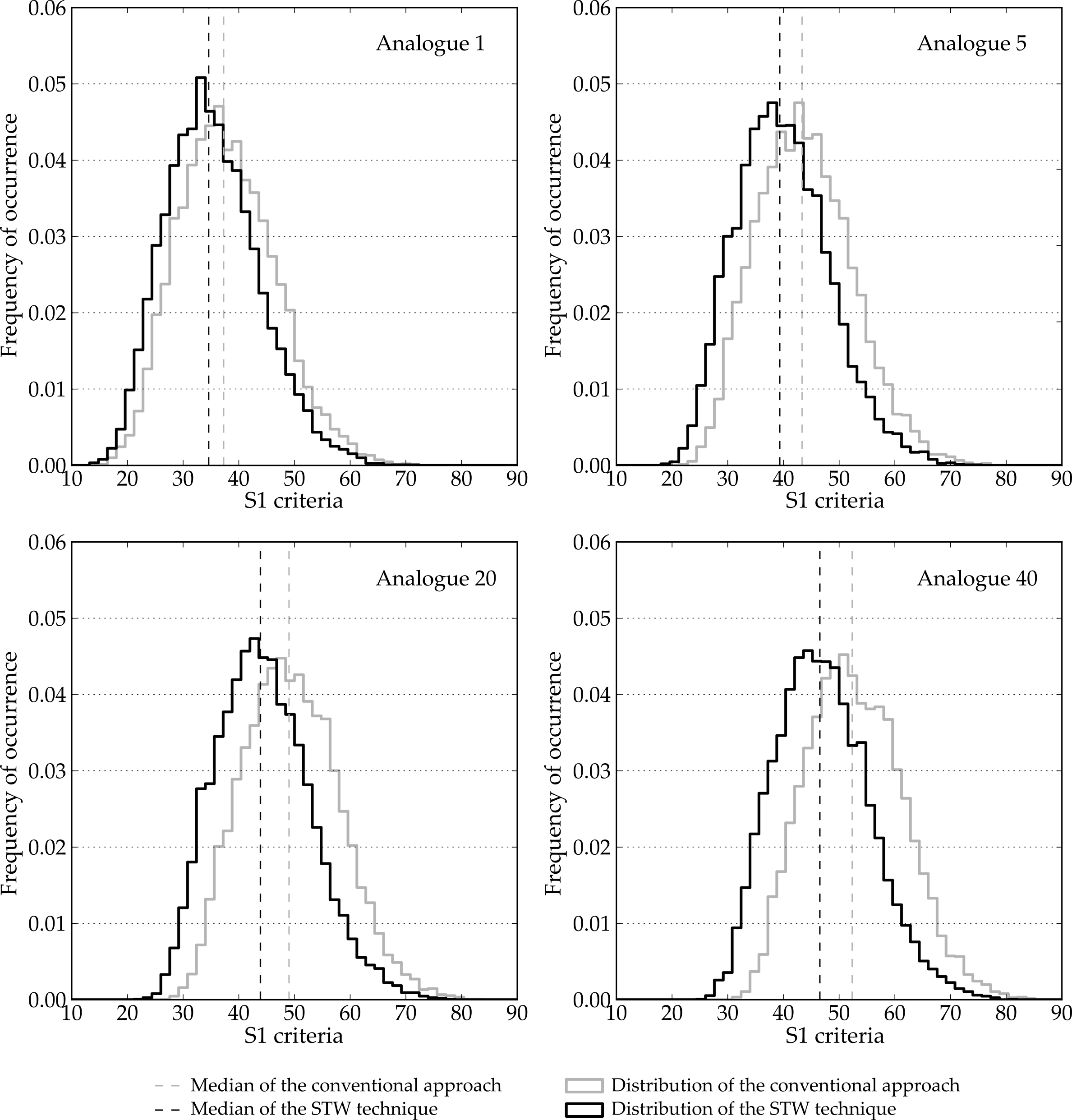
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2\*Station | 2Z | | 2Z-2MI | |
|  | 82-07 |  | 82-07 |  |
| Ulrichen | 29.37 | -1.36 | 33.24 | -1.08 |
| Zermatt | 22.20 | -1.67 | 26.95 | -1.32 |
| Visp | 23.23 | -1.89 | 27.77 | -1.08 |
| Montana | 30.79 | -1.76 | 34.77 | -1.34 |
| Sion | 24.78 | -1.45 | 29.36 | -1.80 |
| Aigle | 30.57 | -0.01 | 35.95 | 0.13 |

The impact of the change in the archive length is summarized in Table 5 for both 2Z and 2Z-2MI methods. As expected, a loss of performance can be observed for each station, except for that of Aigle, which seems relatively indifferent to this change. This loss is globally significant, with up to -1.89 points for Visp and the 2Z method.

## 4 Influence on the analogy criteria

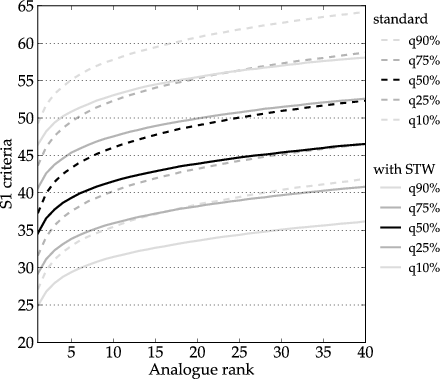
### 4.1 Changes in the atmospheric circulation analogy

When searching for analogues on the geopotential heights, as in the 2Z method, there are now 4 times more candidates than before, which obviously allows to find better matches.



Changes in the S1 criteria distributions of the , ,  and  analogue ranks for the Ulrichen station, due to the STW.

Figure 4.1 presents the changes in the distributions of the S1 criterion for the , ,  and  analogues for the Ulrichen station on the whole calibration period, due to introduction of the STW. The precipitation target remains at before, that is centered on 18 h UTC (6 h UTC to 6 h UTC the next day). The shapes of the distributions of the conventional approach and the STW are similar, but the values of the analogy criteria are now reduced, shifted to the left, and therefore better. An increase in the difference between a fixed window and a sliding window is identifiable, 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.



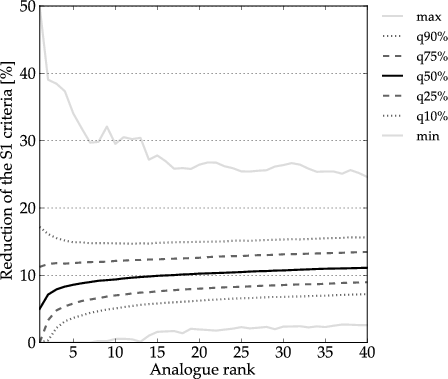


Figure 3: Synthesis of the changes in the S1 criteria, due to the STW, for the Ulrichen station, depending on the ranks of the analogue. (top) Quantiles of the S1 distributions with and without the STW. (bottom) Quantiles of the relative improvements of the S1 criteria when using the STW.

The improvements of the S1 criteria are summarized in Figure 3, which shows (top) quantiles of the S1 criteria according to the analogue rank for the conventional method and the STW, and (bottom) quantiles of the relative reduction. This confirms that all quantile seem similarly reduced (S1 distributions keep their shape), and that this improvement is constantly increasing from the first to the last analogue (Figure 3 bottom).

The median of the S1 values reduction (Figure 3 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 S1 criteria, both in terms of distribution shape and amplitude.

### 4.2 Influence of the weather situation

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 STW. 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 STW, since 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. The results of this analysis are summarized in Figure 4 by the median reduction of S1 for days with precipitation organised into classes between two thresholds. The number of cases per class being reduced, the curves are not as smooth as in previous analyzes. It is nevertheless clear that the improvement tends to increase on days with higher precipitation. This is true for all our stations and confirms our intuition.

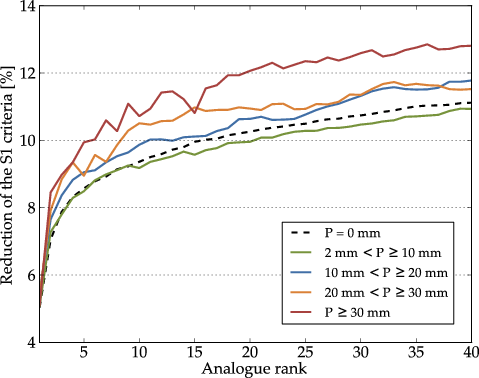


Figure 4: Distribution of the median improvements of the S1 criteria, due to the STW, depending on precipitation thresholds at the Ulrichen station.

### 4.3 Seasonal effect

Atmospheric dynamics varies greatly from one season to another, which reflects on the performance of the AM that is generally lower between June and August [Bliefernicht(2010)]. It therefore makes sense to verify the effect of the STW separately per season.

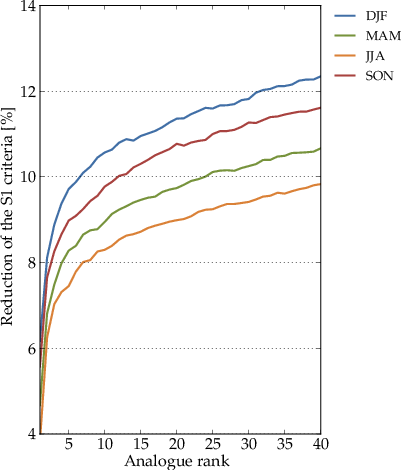


Figure 5: Seasonal effect on the median reduction of the S1 criteria for the Ulrichen station due to the STW. DJF: winter, MAM: spring, JJA: summer SON: fall.

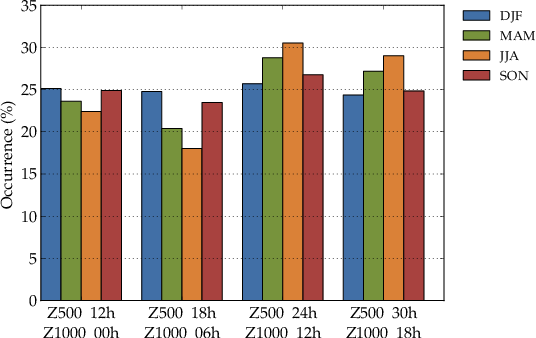


Figure 6: Distribution of the predictors hours in the selected analogue dates depending on the season, for the Ulrichen station.

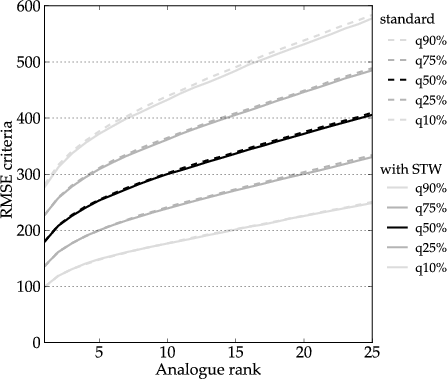
A seasonal effect can be observed on the reduction of the S1 criteria due to the STW (Figure 5). The improvements are greater for winter than summer. One hypothesis is that the diurnal effects of the summer months have an influence on the atmospheric circulation, at least in the lower layers. This effect is based on the daily cycle and good analogues are essentially found for the same hours or time windows.

An analysis of the selected hours for the geopotential predictor seems to confirm this assumption (Figure 6). It was found that the new choice of the temporal window in winter, when using the STW 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 24h & Z1000 12h), 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 larger amplitude than for Ulrichen.

### 4.4 Changes in the moisture analogy

When considering the second level of analogy of the 2Z-2MI method (Table 2), the N2 candidates situations are less numerous, as they are subsampled in the previously  selected analogues, but their dates may have changed. In contrast to earlier, both a reduction or an increase of the RMSE analogy criterion values are possible. In fact, Figure 7 shows an almost insignificant improvement of the RMSE values. Unlike for the first level of analogy, the relative improvement of the RMSE values are distributed relatively symmetrically around zero, with improvements and losses of the same amplitude. Once again, the results for the other stations are similar.



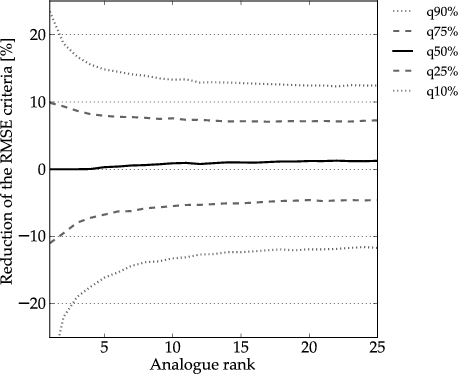


Figure 7: Synthesis of the changes in the RMSE criteria, due to the STW, for the Ulrichen station, depending on the ranks of the analogue. (top) Quantiles of the RMSE distributions with and without the STW. (bottom) Quantiles of the relative improvements of the RMSE criteria when using the STW.

This result of a globally null improvement of the RMSE values does not mean that the 2Z-2MI method cannot be improved by the STW. 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.

## 5 Impact on performance scores

A systematic improvement of the S1 values was previously observed. However, finding better analogue situations does not obligatory imply better skills to predict the precipitation. Therefore, the impact of the STW 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 section 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 6, along with the differences regarding the conventional method with the same archive length. The differences ranges from 0.57 to 2.14 points for the 2Z method and from 1.53 to 2.20 points for the 2Z-2MI method. The introduction of the STW 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 S1 criteria, neither with the season.

Table 6: New CRPSS performance scores for both 2Z and 2Z-2MI methods obtained by the STW approach. The differences are expressed regarding the conventional method with the same archive length.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2\*Station | 2Z | | 2Z-2MI | |
|  | STW |  | STW |  |
| Ulrichen | 31.12 | 1.74 | 35.44 | 2.20 |
| Zermatt | 24.34 | 2.14 | 28.92 | 1.97 |
| Visp | 24.39 | 1.16 | 29.42 | 1.64 |
| Montana | 31.59 | 0.80 | 36.30 | 1.53 |
| Sion | 25.35 | 0.57 | 31.07 | 1.71 |
| Aigle | 31.78 | 1.21 | 38.11 | 2.16 |

### 5.1 Improvement by precipitation threshold or classes ?

The S1 criteria was previously found to be improved to a greater extent for the most dynamic situations related with higher precipitation values (section 4.2). The changes in terms of performance scores will now be assessed regarding precipitation thresholds. Figure 8 synthesizes these differences for the Ulrichen station, 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 non-rainy days and small accumulations are not improved.

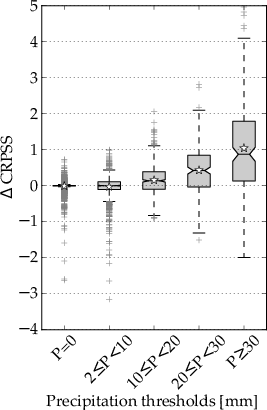


Figure 8: Differences of the CRPSS performance score, due to the introduction of the STW, as a function of precipitation thresholds at the Ulrichen station. The stars represent averages.

### 5.2 Recalibrating the 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 STW may change the optimum of some parameters. The calibration has then been reprocessed?

Table 7: Calibrated parameters (spatial windows and number of analogues) for the analogy on the geopotential at 500 hPa and 1000 hPa (method 2Z) with the STW approach.

|  |  |  |  |
| --- | --- | --- | --- |
| Station | Longitudes | Latitudes | N |
| Ulrichen | 0  17.5 | 42.5  50 | 50 |
| Zermatt | 0  17.5 | 40  50 | 55 |
| Visp | -2.5  20 | 40  50 | 55 |
| Montana | -2.5  15 | 42.5  47.5 | 55 |
| Sion | -2.5  15 | 37.5  50 | 55 |
| Aigle | -2.5  15 | 40  50 | 75 |

Table 8: Parameters of the moisture variables from the 2Z-2MI method with the STW approach. The parameters for the atmospheric circulation are the same as in Table 7, except the number of analogues of the first analogy level (N), which are here different. The  column is the number of analogues for the second level (N).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Station | Longitudes | Latitudes | N | N |
| Ulrichen | 5  10 | 45  47.5 | 110 | 35 |
| Zermatt | 7.5 | 45  47.5 | 80 | 30 |
| Visp | 7.5 | 45  47.5 | 135 | 35 |
| Montana | 5  7.5 | 45 | 110 | 40 |
| Sion | 5  10 | 45  47.5 | 140 | 50 |
| Aigle | 5  7.5 | 45 | 135 | 45 |

After recalibrating, some changes in the optimal parameters can be observed for both 2Z (Table 7) and 2Z-2MI methods (Table 8). Among these, the east-west dimension of the spatial windows of the circulation analogy tends to decrease. More significantly, the optimal number of analogues increases after introducing the STW, of a significant order of magnitude: 25 % to 83 % for the 2Z method and 20 % 67 % for the 2Z-2MI method. The number of analogues of 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 percive the increase in the archive length and diversity by allowing to extract more analogues.

As we saw in Figure 3, the improvement of the S1 criteria grows along with the rank of the analogue, which shows an accumulation of better analogue situations in the distributions. It seems thus profitable to widen 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 number of good analogues is thus globally increased.

This increase in number of analogues has a slight effect on the performance of the different precipitation classes. The same analysis as in section 5.1 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 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.

Table 9: Values of the CRPSS (%) skill score for the newly calibrated parameters using the STW approach. The differences are expressed regarding the conventional method with the same archive length.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2\*Station | 2Z | | 2Z-2MI | |
|  | STW |  | STW |  |
| Ulrichen | 31.58 | 2.20 | 35.72 | 2.48 |
| Zermatt | 24.71 | 2.51 | 29.63 | 2.68 |
| Visp | 25.08 | 1.85 | 30.29 | 2.52 |
| Montana | 32.22 | 1.43 | 37.15 | 2.38 |
| Sion | 26.07 | 1.29 | 31.68 | 2.32 |
| Aigle | 32.21 | 1.64 | 38.50 | 2.55 |

The values of the CRPSS scores for both methods (Table 9) have significantly increased after recalibration. With the introduction of the STW, the performance loss related to the reduction of the archive length is retrieved or compensated?.

### 5.3 Changes in sharpness and accuracy

The CRPS score can be decomposed into two components, namely sharpness and accuracy. The impact of the STW has been analyzed on these components and the results can be found in Table 10 for the original parameters and in Table 11 for the re-calibrated methods. The changes are expressed relative to the total CRPS. Both components do not have the same value ranges: accuracy is in our case almost twice the sharpness values. Since the CRPS is considered here and not the CRPSS, improved prediction capacity results in a lower score. A decrease of the score is thus desirable.

Table 10: Changes in sharpness and accuracy relatively to the total CRPS, due to the introduction of the STW. The changes are presented for both 2Z and 2Z-2MI methods with the original parameters. A decrease of the score is desirable.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2\*Station | 2Z | | 2Z-2MI | |
|  | Sharpness | Accuracy | Sharpness | Accuracy |
| Ulrichen | 2.82 | -5.29 | 1.00 | -4.30 |
| Zermatt | 2.26 | -5.01 | 0.88 | -3.58 |
| Visp | 3.66 | -5.18 | 2.67 | -4.94 |
| Montana | 1.62 | -2.78 | 0.57 | -2.91 |
| Sion | 2.02 | -2.78 | 0.33 | -2.75 |
| Aigle | 0.52 | -2.26 | -1.20 | -2.17 |

Table 11: Same as Table 10 but with the re-calibrated parameters.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2\*Station | 2Z | | 2Z-2MI | |
|  | Sharpness | Accuracy | Sharpness | Accuracy |
| Ulrichen | 1.44 | -4.56 | -0.37 | -3.34 |
| Zermatt | 0.80 | -4.03 | 0.75 | -4.42 |
| Visp | 1.53 | -3.94 | 1.47 | -4.96 |
| Montana | -0.27 | -1.80 | 0.06 | -3.71 |
| Sion | 0.95 | -2.66 | -0.10 | -3.19 |
| Aigle | -0.35 | -2.01 | -2.08 | -1.90 |



Table 12: Influence (%) of the STW on the CRPS components (accuracy and sharpness) relatively to the total CRPS, 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.

It appears that the sharpness is a bit inferior with the STW in favor of accuracy (Table 10). When using the recalibrated parameters, the components are a bit more balanced, but still the accuracy prevails on the sharpness (Table 11). Figure 12 illustrates the changes (relative to the total CRPS) in sharpness and accuracy for different precipitation thresholds at the Ulrichen station. Be it for the 2Z or 2Z-2MI method, the original or the recalibrated parameters, the changes are similar: the improvements concern the accuracy over the sharpness, and this to a greater extent for significant precipitation events. The STW 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.

## 6 Application attempts on the full archive

The improvement provided by the STW 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.

### 6.1 Proportional distribution

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 (Figure 9). When using this reconstructed time series on the shorter period, the method performance was degraded (Table 13 to compare to Table 6) and was even below the conventional method without STW (Table 5). The benefit of a better selection of the analogue situations is lost due to a precipitation archive of poor quality.

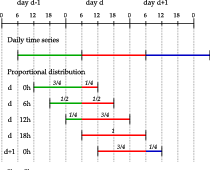


Figure 9: Illustration of the generation of 24 h-totals moving averages by means of a proportional distribution.

Table 13: Values of the CRPSS (%) skill score for the original and the recalibrated parameters using the STW approach on the disaggregated precipitation time series (short period) by means of the proportional distribution.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2\*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 |

### 6.2 Use of a proxy variable

As demonstrated in the previous section, the performance improvement brought by the STW 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 NCEP/NCAR reanalysis 1 (section 1) was used in this perspective, despite their rough resolution.

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 and 7.5 longitude and 45 and 47.5 latitude), 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 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.

Table 14: Value of the coefficient of determination between the reconstructed 6-hourly precipitation time series using the listed variables, and the accurate time series on the period 1982-2007. The grid points are the following (lon/lat): 1) 5/47.5, 2) 5/45, 3) 7.5/47.5, 4) 7.5/45. The highest coefficient of determination is indicated in bold.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 2\*Variable | 2\*Point | Time lapse | | | | |
|  |  | -12h | -6h | 0h | +6h | +12h |
| 4\*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 |
| 4\*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 |
| 4\*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 |
| 4\*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 | **0.721** | 0.696 | 0.643 | 0.636 |

Table 15: Values of the CRPSS (%) skill score for Zermatt, for the original and the recalibrated parameters using the STW approach on the disaggregated precipitation time series (short and long periods) by means of variables from the reanalysis.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 2\*Period | 2Z | | 2Z-2MI | |
|  | original | recalib. | original | recalib. |
| -2007 | 22.57 | 23.14 | 27.11 | 27.71 |
| -2008 | 23.81 | 24.38 | 28.42 | 28.86 |

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 14 for Zermatt (on the short period, 1982-2007). The best proxy is the precipitable water at 45 latitude and 7.5 longitude, 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.

Table 15 presents the CRPSS scores obtained by the reconstituted series using the proxy of precipitable water at the optimal point (Table 14). A slight improvement can be obtained compared to the proportional distribution method, but it is still relatively small, and most of the benefit of the STW is again lost.

These attempts to transpose the STW on the total archive highlight the importance of the actual rainfall chronology. The STW is profitable, provided that the precipitation series are close to the observed one. Without precipitation series with an accurate sub-daily chronology, the introduction of a STW remains fruitless.

The AM is 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 of the day. It is thus more likely that a better match can be found at another time, which can change the selection of the analogues dates.

As [Finet et~al.(2008)Finet, Marty, Zin, and Obled] had previously shown, the introduction of a STW allows finding better analogue situations in terms of the atmospheric circulation. It has been here demonstrated that the improvement of the S1 criteria values is growing with the rank of the analogue. This is 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, which are generally related to more dynamic atmospheric situations. This dynamism implies more specific circulation patterns that are evolving more rapidly. Thus, a STW is found to be of particular interest in this kind of situation, and benefits the prediction of extreme precipitation events.

A seasonal effect has been highlighted, as the STW is 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 75% of the analogue situations are selected at a different time in winter.

The impact of the STW on the prediction performance has never been assessed before, due to the shortcoming of long precipitation series at a sub-daily time step. The impact of the STW approach could be assessed on a 25 years time series with a high temporal resolution. After the introduction of the STW, 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 STW. Some parameters changed, 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 for the prediction performance. The number of good analogues is thus globally 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 to a loss of all the improvement brought by better analogue situations identified by the STW. 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 using meteorological variables from the NCEP reanalysis 1 dataset as proxy did not succeed. Other reanalysis datasets could eventually preform better.

Perspectives of using the STW rely 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, these archives of high temporal resolution precipitation data are getting longer. The second source is the establishment of precipitation reanalysis at a country scale, or the use of reanalysis-driven RCMs (Regional Climate Models) or LAMs (Limited Area Models) on a long period. When such precipitation data will be available, the use of a STW in the AM should be considered for any application, especially when the prediction quality of extreme events is important.

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

Thanks to the Swiss Federal Office for Environment (FOEV), the Roads and Water courses Service, Energy and Water Power Service of the Wallis Canton and the Water, Land and Sanitation Service of the Vaud Canton who financed the MINERVE (Modélisation des Intempéries de Nature Extrême des Rivières Valaisannes et de leurs Effets) project which started this research.

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

**References** urlstyle

[Ben Daoud et~al.(2015)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, Atmospheric Research, 169, 147–159, doi:10.1016/j.atmosres.2015.09.015 , URL http://linkinghub.elsevier.com/retrieve/pii/S0169809515002951 , 2015.

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

[Bontron(2004)] 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.

[Finet et~al.(2008)Finet, Marty, Zin, and Obled] Finet, T., Marty, R., Zin, I., and Obled, C.: Developing and transferring the ANALOG approach for PQPFâ€™s from French OHMCV catchments to quick responding catchments in Venetia and Romania., Tech. rep., LTHE & CNRS, Grenoble, France, 2008.

[Horton et~al.(2012)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, Natural Hazards and Earth System Sciences, 12, 777–784, doi:10.5194/nhess-12-777-2012 , 2012.

[Horton et~al.(2016)Horton, Jaboyedoff, and Obled] Horton, P., Jaboyedoff, M., and Obled, C.: AtmoSwing (v1.4): Analog Technique model for Statistical weather forecastING, Geoscientific Model Development, submitted, 2016.

[Kalnay et~al.(1996)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, Bulletin of the American Meteorological Society, 77, 437–471, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0477(1996)077%3C0437:TNYRP%3E2.0.CO;2 , 1996.

[Lorenz(1956)] 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, URL http://www.o3d.org/abracco/Atlantic/Lorenz1956.pdf , 1956.

[Lorenz(1969)] Lorenz, E.: Atmospheric predictability as revealed by naturally occurring analogues, Journal of the Atmospheric Sciences, 26, 636–646, URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1969)26%3C636:APARBN%3E2.0.CO;2 , 1969.