Elsevier Editorial System(tm) for Journal of Hydrology
Manuscript Draft

Manuscript Number: HYDROL21431R1

Title: Using Genetic Algorithms to Optimize the Analogue Method for Precipitation Prediction in the Swiss Alps

Article Type: SI:Measuring & Modeling Rain

Keywords: Precipitation prediction; Precipitation downscaling; Analogue method; Optimization; Genetic algorithms; Alpine climate

Corresponding Author: Dr. Pascal Horton,

Corresponding Author's Institution: Terranum

First Author: Pascal Horton

Order of Authors: Pascal Horton; Michel Jaboyedoff; Charles Obled

Abstract: Analogue methods provide a statistical precipitation prediction based on synoptic predictors supplied by general circulation models or numerical weather prediction models. The method samples a selection of days in the archives that are similar to the target day to be predicted, and consider their set of corresponding observed precipitation (the predictand) as the conditional distribution for the target day. The relationship between the predictors and predictands relies on some parameters that characterize how and where the similarity between two atmospheric situations is defined.

This relationship is usually established by a semi-automatic sequential procedure that has strong limitations: (i) it cannot automatically choose the pressure levels and temporal windows for a given meteorological variable, (ii) it cannot handle dependencies between parameters, and (iii) it cannot easily handle new degrees of freedom. In this work, a global optimization approach relying on genetic algorithms was able to optimize all parameters jointly and automatically. It allowed consideration of parameter inter-dependencies, and objective selection of some parameters that were manually selected beforehand, which obviates the need to assess a large number of combinations of pressure levels and temporal windows of predictor variables.

The global optimization was applied to some variants of the analogue method for the Rhône catchment in the Swiss Alps. The performance scores increased compared to reference methods, especially for days with high precipitation totals. The resulting parameters were found to be relevant and coherent between the different subregions of the catchment. Moreover, they were obtained automatically and objectively, which reduces the effort that needs to be invested in exploration attempts when adapting the method to a new region or for a new predictand. In addition, the approach allowed for new degrees of freedom, such as a possible weighting between pressure levels, and non-overlapping spatial windows.

*Revision Notes

Response to Reviewer #1

The manuscript „Using genetic algorithms to optimize the Analogue Method for precipitation downscaling in the Swiss Alps" by Horton et al. addresses the problem of precipitation prediction by analogue methods (AM). The main goal of the paper is to present and test a new optimization method based on Genetic Algorithms (GA) which has in my opinion two main advantages to the traditional sequential procedure: (a) it optimizes all predictors jointly and independently; and (b) it allows for a free selection of the spatial integration window (and time) and can give weights to different pressure levels. The paper shows improvement compared to a reference method optimized by the sequential procedure. The analysis is conducted over the Rhone Basin in Switzerland divided into 10 different regions. Predictors come from NCEP-NCAR reanalysis, predicted precipitation comes from Meteoswiss stations averaged over 500 km2 grids.

My main impression is that this is solid work for people interested in improving AM for predicting precipitation. As far as I can tell methodologically the paper is correct and in my opinion the results are well explained. But I do have a concern about the interest for a general hydrological audience and suggest that the authors in their revision try to make the paper shorter and more focussed on the message and expand some discussion on the actual usefulness and application of the method to precipitation prediction. Also a clearer statement on the novelty of this paper compared to the recommendations of Horton et al. (2016b) is necessary (line 40-41). The paper can be published in JH after revision. My main questions/issues which the authors may wish to consider and/or expand on are listed below.

Thank you for your positive feedback.

We removed many figures and tables and some secondary analyses in order to focus on the main message. We also added applications examples for hydrology mainly in the introduction. Please note that this paper is part of the special issue "Precipitation measurement and modeling: uncertainty, variability, observations, ensemble simulation and downscaling", and so the focus on precipitation is justifiable in our view.

Horton et al. (2016b) focused on the parametrization of GAs in order to optimize AMs successfully, not on their application on more elaborated AMs. This has been clarified.

1. Problem of non-uniqueness. It is true that the global optimization method provides more freedom in the combination of parameters (predictors) and therefore performance. However this increase in the degrees of freedom is to me not necessarily a win-win situation. As the authors point out several different combinations of parameters may now lead to similar model performance, especially in the two analogue version. I do not see a constraint on parameters as necessarily negative. I have a feeling the authors could engage with this question more in their work.

We agree and added some sentences in the discussion (section 5): l. 575-580

2. The use for prediction. I am not very familiar with applications of AM for prediction of precipitation, but as I understood it from the paper only observed precipitation amounts and patterns can be reproduced (from the historical archive), even though N1 analogues can be selected and a "probabilistic forecast" from them derived. Do these N1 cases cover the entire uncertainty, including high precipitation (extremes) which have not been observed? In other words, what is the advantage of these methods against stochastic predictive models conditioned on similar (circulation and moisture) predictors? Do the authors have an opinion on this?

The following has been added in the discussion about extremes: l. 635-645

The following has been added in the introduction about advantages of the AMs: l. 46-51

3. RCM is not an option? Although I agree that GCM precipitation performance is poor, in all fairness I think it should be mentioned in the Introduction (lines 11-20) that limited area RCMs are an alternative at space and time scales that are getting pretty close to station data, or rather to the spatial averaging scale for precipitation used in this study.

Some sentences were added in the introduction (l. 37-45).

4. Methodology. I appreciate the approach in Chapter 5 starting from the two fixed pressure level reference and then releasing constraints, i.e. adding degrees of freedom, and eventually a second analogue (moisture). Although I agree that the added level of performance is not incredible, there is measurable improvement in the calibration. There is also consistency in the results, in that some pressure levels, etc., are consistently chosen over others. Finally it was concluded that 4 predictors are optimal because more dropped the performance in validation and the authors considered the model then over-parameterized. I would like to read some physical explanation of this feature in the paper connected to the fields of pressure and moisture and precipitation formation in the Alps. Why should it be that including more parameters, if they are responsible for rain formation, becomes eventually counterproductive?

About the optimal number of 4 predictors: the text was maybe not so clear, so it has been reworked (l. 319-334) and illustrated in figure 2. The general term “predictor” might be ambiguous here, because only geopotential heights were considered at this stage. So, there is no moisture predictors at this stage, only atmospheric circulation data. As shown in figure 2, with 4 geopotential heights, most of the predictive information seem to be integrated. Adding other thermodynamic predictors will increase the prediction skill, as shown with the moisture index.

5. Performance dependent on precipitation depth. The authors present that the performance of their GA-AM method is best for days with heavy rain. I am wondering what is the reason for this, especially considering that I would also expect the predictand derived from station observations averaged over 500 km2 grids to be most uncertain then. Meaning similar pressure/moisture analogues may lead to more different precipitation on the ground during heavy rain which is concentrated in space, but this would not reflect in the grid-averaged precipitation because the spatial density of gauges is not high enough to capture the large spatial variability all of the time.

As mentioned to reviewer #2, larger precipitation values contribute more to the performance score. Thus, GAs will optimize these days more thoroughly than the sequential calibration as they are more powerful and can handle supplementary degrees of freedom. The following sentence was added in section 3.3: l. 390-393

6. Cross-compatibility. This analysis (lines 263-265, 431-438) is really interesting. Is there any spatial pattern evident in the performance drop in space that the authors find physically explainable? There is an attempt starting on line 308 which I do not fully follow.

Yes, it is related to different climate properties of the subregions, mainly for heavy precipitation events. We developed a bit further this point in section 3.3 (l. 394-410) and we hope it is now more understandable.

7. Please explain in a sentence what accuracy and sharpness are (line 295) so that readers don't have to go to Bontron (2004).

For the sake of conciseness, the analysis of the CRPS decomposition into sharpness and accuracy has been removed as it was not extensively analyzed

8. Adding moisture analogue. The performance after adding the moisture analogue improved further, although 2 parameters were sufficient for this analogue. The transferability however starts to be a problem. This is an expected result because I expect atmospheric moisture to vary more significantly in space (one of many reasons?). It is however not super clear to me what the authors recommend: should moisture be used as a predictor or not?

A paragraph was added with recommendations (last one of section 4.2, l. 525-534). Another part was added to discuss the spatial variability of moisture variables (as the main difference between the 2 less transferable subregions with other regions is the spatial windows on which the moisture is considered).

9. Optimization window. The preselection period of 4 months was also tested, without much effect on the results. I understood this was +- 2 months around the target date. Would the results improve if seasons are instead considered, for example the window is fixed for selected summer months to capture convective events, etc, regardless of the target date in that season?

Maybe the sentence was not clear. We use the 4-month preselection window anyway. It was the optimization of the length of this window that was unsuccessful. The 4-month preselection window was introduced some decades ago as an improvement of a fixed seasonal preselection. It is expected to be better than a fixed seasonal window.

10. Finally, I find the paper very long and cumbersome to read. It has 17 figures and 9 tables. I wonder if some of the figures/tables can be passed to Supplementary materials to help the reader focus on the main messages. This is only my subjective opinion.

The number of figures was decreased to 10 and the number of tables to 6. We removed the following elements:

- The analysis of the CRPS decomposition into sharpness and accuracy, as it was not extensively analyzed, along with Fig. 4 and 10.
- Tables 3 and 7 (containing the resulting parameters for the Chablais subregion) were removed as the same information can be found in figures for all subregions.
- Table 4 was removed (values of the CRPSS score) and figure 3 and 8 were changed in order to represent the values of the CRPSS score instead of the relative difference.
- Figure 9 (CRPSS of the 4Zo-4Mio method) was removed as the results are very similar to those of the 4Zo-2Mio method, and thus not so interesting.
- Figures 5 and 6 were merged into a single figure, as well as figures 13 and 14.
- Figure 12 (relationship between the different number of analogues) was removed as it is a bit redundant with figure 11.
- Figure 15 (Optimized weighting for the pressure levels of the 4Zo method) was removed and standard deviations were added to figure 16 in order to show the variability.

11. Some editorial issues

> abstract: parameter inter-dependencies, not parameters inter-dependencies (check also elsewhere in text)

Corrected, thanks.

> line 38: presents

Corrected, thanks.

> typo in the subscript in Equation (1)

We didn't find any mistake here… The score name is written S1.

> line 138: Figure 17 where?

Some clarifications were added

> line 173: what are left side valleys?

This was changed for "southern valleys"

Response to Reviewer #2

Summary

The authors present a methodology to optimize the parameters of an analogue-based precipitation downscaling system using genetic algorithms (GA). The GA is not only an optimization technique but allows discovering parameter inter-dependencies and possibly give a better understanding of the dynamics that lead to high precipitation accumulations in Canton Valais, Switzerland.

The paper is well written and a pleasure to read. I believe that the use of genetic algorithms within analogue-based forecasting techniques is an interesting idea. In fact, it increases the objectivity of current "rule-of-thumb" decisions that are done to drive the selection of analogue situations. Consequently, I recommend the publication of the paper after having addressed the remarks that I list hereafter.

Thank you for your very nice feedback.

Major comments

Page 4, Line 66 - I would rather put the equation just after mentioning the Teweless-Wobus criterion S1.

We changed that, thanks.

Page 5, line 97 - Here I would also mention that the skill of analogue forecasts that include as predictor variable the moisture index depends on the skill of the NWP model in predicting moisture fields (when used in real-time).

Good point. We added a note on that.

Page 5, line 8-9 - Does the AM perform well also when looking for analogues for a single rain gauge? What is the consequence of computing a local average given the high spatial variability and intermittency of precipitation, e.g. for convective cases? There is no need to do analysis to answer this question

There is a slight increase in performance due to the smoothing of local variability when working on local averages. The difference is however rather low. A comment has been added.

Page 6, line 122 - Is the climatological distribution of precipitation over a single day sufficiently stable as reference to account for seasonality? Have you tried to include a temporal smoothing or pool the data over days before and after the given day? A harder reference to beat could be the Eulerian persistence forecast (the precipitation observed on the previous day).

The climatological distribution of precipitation usually considered is the one from the entire archive. It could be discussed that a climatological distribution built on the +-2 months' window (thus seasonal) might be more relevant. However, most applications use the distribution of the full archive. This does not play a major role in this paper, as we are interested in the improvement relatively to reference methods.

Page 9, line 200-202 - What is the overlapping constraint? The expression "what the sequential calibration cannot do" is not clear to me.

No overlapping constraint of the spatial windows means that they can differ from one pressure level to another (this has been specified in the article).

"what the sequential calibration cannot do" was changed for "which cannot be achieved with the sequential calibration technique".

Page 10, line 226 - It would be very interesting to show a plot with the CP and VP error as a function of number of predictors to illustrate that the VP error reaches an optimum around 4 predictors while the CP error keeps decreasing for increasing number of predictors (overfitting).

The assessment of the optimal number of predictors has been performed again on 3 subregions in order to consider the weighting between the predictors in that process. Consequently, this has slightly changed the results (not anymore a clear decrease on the VP), but it has not changed the conclusion that 4 seems to be an optimal number of predictors. A figure (Fig. 2) has been added to illustrate this aspect.

Page 13, line 316 - You could add that there are multiple local optima in very different regions of the parameter space that provide sufficiently good performance. Instead of using only one single optimal solution for the selection of analogues, you could use an ensemble of optimal solutions. This way you could both account for the parameter uncertainty of the analogue technique and increase the number of samples contributing to the empirical distribution of precipitation at the rain gauge (ensemble size). This could be considered for future studies.

This is a good idea. This point was added in the discussion section (l.597-606).

Page 14, lines 351- 356 - When optimizing an error function depending on precipitation totals, the large precipitation values (and errors) will contribute more to the total error. Thus, using GA allows to minimize the forecast error in particular for days with high precipitation accumulations. Therefore, it is quite reasonable that you beat the reference method, which has no optimization of an error function.

There are two different elements here:

- Indeed, larger precipitation values contribute more to the error function. The following sentence was thus added in section 3.3: l. 390-393
- However, the reference methods are calibrated by means of the sequential procedure, which also aims at reducing the error function. GAs can reduce the errors to a greater extent than the sequential procedure thanks to more efficient techniques and more degrees of freedom. We realized that the establishment of the references is not clear enough and thus we added a section on the sequential calibration (2.5).

Page 17, line 435-438 - Could the over-parametrization of the regions be due to the larger spatial variability of moisture fields? Pressure fields are known to be smoother and could be expected to generalize more to close regions than moisture fields.

It is likely to play a role. Indeed, the main difference with other regions is the spatial windows on which the moisture is considered. We added a comment on that in the end of section 4.2 (l. 518-524).

Page 17, line 450 - It would be interesting to mention that there is an interdependence between the location (or size) of the spatial window and the temporal window. In fact, if we follow Taylor's hypothesis, space and time could be easily related if we consider a moving precipitation system (or other) that has no significant growth and decay processes. More we go backwards in time more we have to move upstream the analogy window.

We now mention this point: l. 548-551

Page 20, line 536 - I wonder whether it would be useful to compute and show a correlation matrix between the different predictors.

We consider this idea interesting, but not feasible, because of the unlimited possible combinations. Indeed, the predictors are considered at different pressure levels, temporal and spatial windows.

Figures and Tables - The number of figures and tables in the paper is quite high, but I do not know which ones could be removed, perhaps those that are not discussed in detail in text or that are giving redundant conclusions.

We removed 7 figures and 3 tables. See the last comment to referee #1 about what was removed.

Minor comments

Abstract, line 2 - "provided by global models" is a bit too general. I would rather use general circulation models or numerical weather prediction models.

Corrected, thanks.

Abstract, par 2, line 2 - "strong limitations". You could complete the sentence by listing a couple of them.

Some information were added.

Page 2, line 9 - "Other predictands are also often considered". Here I would also add which ones, e.g. …

Examples and references were added.

Page 2, Line 15 - "get down" → resolve, compute, forecast. I would use a more appropriate term.

Corrected, thanks.

Page 2, Line 21 - "made" → "designed"?

Corrected, thanks.

Page 3, Line 26 - "criterion itself" or "criteria themselves"

Corrected, thanks.

Page 3, lines 27-29 - Here you could also mention that ad-hoc techniques for the selection of predictors were also used by Panziera et al. (2011) and Foresti et al. (2015) for ensemble radar rainfall nowcasting. The GA technique could also be adapted for these applications.

A sentence referring to this work was added at the end of the paragraph.

Page 3, Line 30 - I would find a better term for "reconsidering"

Corrected, thanks.

Page 3, Line 31 - "pressure levels" → "optimal pressure levels"

Corrected, thanks.

Page 3, Line 44 - "on precipitation predicting" → "for precipitation prediction"

Corrected, thanks.

Page 4, Line 74 - "of the geopotential height". I would add ", which represent better the upper level flow direction"

Added to the text, thanks.

Page 5, Line 81 - "both North and East directions"

Added to the text, thanks.

Page 5, line 102 - "Predictors are generally extracted from reanalysis datasets"

Added to the text, thanks.

Page 6, line 111 - It would be interesting to mention that you are trying to verify the performance of an ensemble-probabilistic forecast technique.

A sentence has been added.

Page 7, line 133 - "complex surface". You could add that you are trying to find the global optimum of a complex high-dimensional error function having multiple local optima.

A sentence has been added.

Page 8, line 165 - Here you could add that the high spatial variability of precipitation is due to complex orography.

Added to the text, thanks.

Page 9, line 196 - "are not provided in this paper"

Added to the text, thanks.

Page 9, line 204 - "respectively, w.r.t. the reference method based on Z500 and Z1000"

Added to the text, thanks.

Page 9, 205 - "tremendous" → "very significant", "large"

Corrected, thanks.

Page 9, line 209 - "other parameters (…)"

Information added.

Page 9, line 211 - “and may” → “but may”

We meant “and”, because both points are negative consequences.

Page 10, line 225 - “but always more to a smaller extent” could be rephrased

Corrected, thanks.

Page 10, line 229 - “another region than Valais” to clarify that it is not another region within your domain.

Clarification added.

Page 10, line 239 - “name” → “named”

Corrected, thanks.

Page 11, line 263 - “cross-compatibility and spatial coherence of the optimized parameters”

Information added.

Page 11, line 276 - “significant preference in the AM” is not clear.

The sentence has been removed.

Page 14, lines 360-365 - Does this mean that the two levels of analogy bring complementary information (not independent)? This is a good finding.

Thanks. A note has been added.

Page 15, line 382 - “spatial shift”?

No, a vertical shift (clarification added)

Page 18, line 481 - “does not”

Corrected, thanks.

Page 20, line 522 - “what the sequential calibration” is a strange expression to me.

Corrected, thanks.

Page 21, line 544 - “dependence in the selected parameters”

Corrected, thanks.

Page 21, line 555 - “significantly more improved” → “improved further” or other

Corrected, thanks.

Figure 2 and 7 - Would it be better to put the actual pressure levels (Z500, Z1000, etc) instead of the four levels (Z1, …, Z4)?

Corrected, thanks.

Table 1-3 - It is not clear if the provided hour (12h, 24h) is for the day before the target day that we want to forecast.

It has been clarified. It is the hour within the target day. In a forecasting application, the AM is applied to the NWP model outputs, thus letting the temporal extrapolation to the NWP model. It is then more an “adaptation technique”.

Table 6 - In the caption I would make clear whether the improvement is w.r.t. climatology or the reference method.

Clarification added.

Highlights

- Parameters of the Analogue Method are optimized jointly and entirely objectively
- Parameters that were manually selected can now be optimized automatically
- Some parameter inter-dependencies have been demonstrated
- New degrees of freedom could be added to the Analogue Method
- The performance scores of the prediction were increased significantly

Using Genetic Algorithms to Optimize the Analogue Method for Precipitation ~~Downscaling~~ Prediction in the Swiss Alps

Pascal Horton[a,b,1,*], Michel Jaboyedoff[a], Charles Obled[c]

[a] *University of Lausanne, Institute of Earth Sciences, Lausanne, Switzerland*
[b] *University of Bern, Oeschger Centre for Climate Change Research, Institute of Geography, Bern, Switzerland*
[c] *Université de Grenoble-Alpes, LTHE, Grenoble, France*

Abstract

~~The Analogue Method aims at predicting precipitation based on predictors variables provided by global models~~ Analogue methods provide a statistical precipitation prediction based on synoptic predictors supplied by general circulation models or numerical weather prediction models. The method samples a selection of days in the archives that are similar to the target day to ~~predict~~be predicted, and consider ~~the set of their~~ their set of corresponding observed precipitation (the predictand) as the conditional distribution for the target day. The relationship between the predictors and predictands relies on some parameters that characterize how and where the ~~similitude~~ similarity between two atmospheric situations is defined.

This relationship is usually established by a semi-automatic sequential procedure that has strong limitations~~. A new~~ : (i) it cannot automatically choose the pressure levels and temporal windows for a given meteorological variable, (ii) it cannot handle dependencies between parameters, and (iii) it cannot easily handle new degrees of freedom. In this work, a global optimization approach relying on ~~Genetic Algorithms can~~ genetic algorithms was able to optimize all parameters jointly and automatically~~, which is a breakthrough in the way~~

*Corresponding author
Email address: pascal.horton@alumnil.unil.ch (Pascal Horton)

~~the Analogue Method was calibrated until now. It allows taking into account parameters~~ . It allowed consideration of parameter inter-dependencies, and ~~selecting objectively~~ objective selection of some parameters that were manually selected beforehand~~(such as the~~ , which obviates the need to assess a large number of combinations of pressure levels and ~~the~~ temporal windows of ~~the predictor variables)~~predictor variables.

~~In this work, the global optimization is applied to the~~ The global optimization was applied to some variants of the analogue method for the Rhône catchment ~~,~~ in the Swiss Alps. The performance scores ~~are significantly~~ increased compared to ~~a reference method, and this even to a greater extent~~ reference methods, especially for days with high precipitation totals. The resulting parameters were found to be relevant and coherent between the different subregions of the catchment. Moreover, they ~~are~~ were obtained automatically and objectively, which reduces ~~efforts~~ the effort that needs to be invested in exploration attempts when adapting the method to a new region or for a new predictand. In addition, the approach ~~allows~~ allowed for new degrees of freedom, such as a ~~weighting between the~~ possible weighting between pressure levels, and ~~non overlapping~~ non-overlapping spatial windows.

Keywords: ~~Precipitation~~ precipitation prediction, ~~Precipitation downscaling~~~~Analogue~~ analogue method, ~~Optimization~~optimization, ~~Genetic~~ genetic algorithms, Alpine climate

---

1. Introduction

The analogue method (AM) is a downscaling technique based on the idea expressed by Lorenz (1956, 1969) ~~,~~that similar situations in terms of atmospheric circulation are likely to lead to similar local weather ~~(Duband, 1970; Bontron and Obled, 2005). It aims at forecasting a predictand, often the daily precipitation (eg. Guilbaud, 1997; Bontron and Obled, 2005; on the basis of~~ (Duband, 1970). It uses predictor variables describing the synoptic atmospheric circulation ~~. —~~in order to predict local-scale predictands of interest. It is often used to predict daily precipitation, either in an operational forecasting context (e.g. Guilbaud, 1997; Bontron and Obled, 2005; Hamill and Whitaker, 2006; Bliefernicht, 2 a climate downscaling context (e.g. Radanovics et al., 2013; Chardon et al., 2014; Dayon et al., 2015; Raynaud Other predictands are also ~~often considered(see Horton et al., 2016a, for a non-exhaustive list)~~considered, such as precipitation radar images (Panziera et al., 2011; Foresti et al., 2015), temperature (Radinovic, 1975; Woodcock, 1980; Kruizinga and Murphy, 1983; Delle Monache et al., 2013; Ca wind (Gordon, 1987; Delle Monache et al., 2013, 2011; Vanvyve et al., 2015; Alessandrini et al., 2015b; Junk e solar power (Alessandrini et al., 2015a; Bessa et al., 2015), snow avalanches (Obled and Good, 1980; Bolognesi and radiation (Bois et al., 1981; Raynaud et al., 2016).

In ~~operational~~ real-time forecasting, it ~~has been~~ is used mainly by practitioners, notably hydropower companies or flood forecasting services, ~~while in~~ that need to anticipate water yields or issue early flood warnings several days in advance. The classical forecasting chain consists of using limited area models (e.g. AROME, or COSMO) forced by global NWP (numerical weather prediction) models with a lower resolution. However, their use requires very important processing capacities, and the resulting forecast still presents large uncertainties and biases. Although these outputs are essential, they can be supplemented by other sources of forecasts providing useful information. In contrast to local NWP models, AMs can transform at low cost the synoptic-scale information provided by the global NWP model into precipitation forecasts, by using the natural local behaviour in response to synoptic-scale influences stored in the archive of observed precipitation. Running an AM approach is fast enough that it can search for analogues for each day, up to ten days ahead, eventually for the different traces of an ensemble forecast and/or those issued by different NWP models (e.g. NOAA-GFS or ECMWF-IFS), in a matter of minutes.

In climate studies, ~~it is~~ AMs are used to downscale the ~~results of global climate~~ outputs of a general circulation model (GCM) ~~simulation runs. In this last case, although the GCM represents the large or regional scaleevolutions of~~ or regional climate model (RCM) simulation runs (Dayon et al., 2015) or to reconstruct past weather conditions (Caillouet et al., 2016). In future climate studies, RCMs are often used to dynamically downscale precipitation to a local scale. However, even though the relevance of RCMs' outputs increases, a bias correction of the outputs is often still required, particularly in complex terrain. Moreover, their application is computer-intensive, which makes it difficult to cover all combinations of climate scenarios and GCMs. Therefore, the ~~atmosphere, it would be far too computer intensive to get down to variables such as precipitation at a rather small and representative scale. Thereforethe~~ idea is to ~~rely on past observations to bypass these unaffordable small scale~~ bypass the small-scale simulations and to go from the ~~large scale situation proposed by the GCM~~ large-scale situation to the end variables ~~like precipitation by searching analogues in the archives.~~ such as precipitation by statistical downscaling (Maraun et al., 2010).

Beyond being computationally inexpensive, another big advantage of AMs is that they create realistic precipitation patterns for a region, provided that the analogue dates are the same, since they are based on observed situations with consistent spatial distribution (Radanovics et al., 2013; Chardon et al., 2014). For the same reason, they can also provide multivariate predictions that are physically consistent (Raynaud et al., 2016).

The method can be ~~made of~~ designed with multiple successive subsampling steps, or analogy levels, each of them relying on different meteorological variables. A certain number of parameters define the relationship between predictors and predictands, such as the choice of the predictor variable, its pressure level and temporal window to consider, the spatial domain to use for the comparison, as well as the analogy ~~criteria~~ criterion itself, and finally, the number of analogue situations to keep at each subsampling level. These parameters are usually calibrated by means of a semi-automatic sequential procedure ~~(see Bontron, 2004; Horton et al., 2016a, for the details)~~(Bontron, 2004; Radanovics et al., 2013), i.e. by optimizing each single parameter, one at a time, in an arbitrarily chosen order, with no or little ~~reconsidering~~reassessment. This sequential approach ~~has therefore~~ therefore has strong limitations: (i) it cannot automatically choose the optimal pressure levels and the temporal windows for a given meteorological variable, (ii) it cannot handle dependencies between the parameters within a level of analogy, and even less between them, and (iii) it ~~could not~~ cannot easily handle new degrees of freedom, such as a possible weighting between the pressure levels. Thus, even if the processing involved is relatively fast, the sequential approach requires laborious assessments of ~~predictors~~ predictor combinations (variables, pressure levels, temporal windows), and ~~present~~ presents a high risk of ending in a local optimum ~~due to~~ because of subjective initial choices and lack of consideration of ~~parameters~~ parameter inter-dependencies. Other calibration methods exist for specific applications, such as radar images (Panziera et al., 2011; Foresti et al., 2015).

~~With the perspective~~ Aiming to overcome these limitations, a global optimization by ~~Genetic Algorithms~~ genetic algorithms (GAs) was introduced~~by Horton et al. (2016b)~~. An intensive assessment ~~work~~ resulted in recommendations ~~of parametrization of~~ to parametrize GAs in order to optimize AMs successfully (Horton et al., 2016a). The present paper is based on these recommendations~~and illustrate them on precipitation predicting~~ , and applies them to precipitation prediction for the upper Rhône catchment in the Swiss Alps, using AMs of varying complexity. It aims at ~~proving~~ illustrating the relevance of a ~~fully-automatic~~fully automatic, objective, and global, optimization technique for AMs. The applications are indeed numerous, as ~~the AM has~~ AMs have to be adapted to every new location ~~it is~~ they are applied, or to any new predictand ~~it~~ they should predict.

~~A short overview of AMsis presented in section ??, as well as a summary of Genetic Algorithms in section 2.2. Section 2.1 describes the case study area~~The data, AMs, and optimization techniques (sequential and GAs) are presented in Section 2. The results are first ~~detailed~~ given for the optimization of the analogy ~~on the~~ of atmospheric circulation only (~~section~~ Section 3), before being extended to a method adding a second level of analogy on moisture variables (~~section~~ Section 4). General discussions (~~section~~ Section 5) and conclusions (~~section~~ Section 6) follow.

2. ~~The Analogue Method~~Data and methods

2.1. ~~References~~Case study description

The study area is the alpine upper Rhône catchment in Switzerland (Fig. 1). The altitude ranges from 372 to 4634 m.a.s.l. and the area is 5524 km^2. This region is the target of the MINERVE (Modélisation des Intempéries de Nature Extrême sur les Rivières Valaisannes et de leurs Effets) project, which aimed at real-time flood management on the upper Rhône catchment (García Hernández et al., 2009). Even though the region is rather small, the meteorological influences related to extreme weather conditions vary substantially within it (see Horton et al., 2012). Indeed, a high spatial variability of precipitation climatology exists, which is due to the complex orography of the region, and the mix of various meteorological influences. Based on different climatological analyses, the precipitation gauge stations in the catchment were clustered in ten subregions (Fig. 1):

1. Swiss Chablais
2. Trient Valley
3. West Bernese Alps
4. Lower Rhone Valley
5. Southern valleys
6. Southern ridges
7. Upper Rhone Valley
8. Southeast ridges
9. East Bernese Alps
10. Conches Valley

2.2. Data

AMs rely on two types of data: predictors, which are atmospheric variables describing the state of the atmosphere at a synoptic scale, and the predictand, which is the local weather variable one wants to predict.

Predictors are generally extracted from reanalysis datasets. The NCEP-NCAR reanalysis I (6-hourly, 17 pressure levels at a resolution of 2.5°, see Kalnay et al., 1996) was used here, but it could have been any other reanalysis dataset.

The predictand (which is to be predicted) is here the daily precipitation (6 a.m. to 6 a.m. the next day) measured at the MeteoSwiss network stations, for the period 1961–2008. The time series from every available gauge station were averaged over the ten subregions (Fig. 1), which were approximately 500 km^2 each, in order to smooth local effects (Obled et al., 2002; Marty et al., 2012). This helps account for local variability, mainly when convective processes are involved, which slightly increases the prediction skill.

It must be stressed that the predictand here is a temporally cumulated variable, compared to the meteorological predictors, which may be considered instantaneous. Depending on the duration of the accumulation period (here 24 h, but could have been 6 h, 12 h, or more than 24 h), the choice of predictors will vary.

The 48-yr precipitation dataset was divided into a calibration period (CP) and a validation period (VP). Using data independent of the CP to validate the results is very important in order to assess the robustness of the proposed improvements and to avoid over-parametrization of the method.

In order to reduce potential biases related to trends linked to climate change or to the evolution in measurement techniques, the selection of the VP was evenly distributed over the entire series (Ben Daoud, 2010). Thus, one out of every six years was selected for validation, which represents a total of 8 years for the VP and 40 for the CP. This choice of sequence was made in order to have similar statistical characteristics between the CP and VP.

2.3. The analogue method

Multiple variations of the ~~methods exist, and most of them will not be detailed hereafter (see Horton et al., 2016a; ?, for more comprehensive listings)~~ analogue method exist, most of which are not detailed here (see Ben Daoud et al., 2016, for a more comprehensive listing However, there are mainly ~~2~~ two parameterizations that are most often used for precipitation ~~forecasting and that will be~~ prediction and that are considered as reference: one that relies on an analogy of the atmospheric circulation, and another that adds a second level of analogy on moisture variables (Obled et al., 2002; Bontron and Obled, 2005; Marty et al., 2012).

The method based on the analogy of ~~the~~ synoptic circulation consists ~~in~~ of the following steps (Table 1): the similarity of the atmospheric circulation of a target date with every day of the archive is assessed by processing the S1 ~~criteria~~ criterion (Eq. 1, Teweles and Wobus, 1954; Drosdowsky and Zhang, 2003), which is a comparison of gradients, over a certain spatial window~~. Bontron and Obled (2005) showed~~ :

$$S1 = 100 \frac{\sum_i |\Delta\hat{z}_i - \Delta z_i|}{\sum_i max\left\{|\Delta\hat{z}_i|, |\Delta z_i|\right\}} \tag{1}$$

where $\Delta\hat{z}_i$ is the difference in geopotential height between the i-th pair of adjacent points of gridded data describing the target situation, and Δz_i is the corresponding observed geopotential height difference in the candidate situation. The differences are processed separately in both North and East directions over the selected spatial domain. The smaller the S1 values, the more similar the pressure fields.

Bontron and Obled (2005) show that the geopotential height at 500 hPa (Z500) and 1000 hPa (Z1000) are the best first predictors of the NCEP/NCAR reanalysis I dataset, and that the S1 ~~criteria~~ criterion performs better than scores based on absolute distances. The reason for such better results is that the S1 ~~criteria allows comparing the circulation pattern~~criterion allows comparison of the circulation patterns, by means of the gradients, rather than the absolute value of the geopotential height, which better represent the flow direction. To cope with seasonal effects, candidate dates are extracted within a period of ~~4 months centered~~ four months centred around the target date, for every year of the archive. ~~Following the nomenclature proposed by Horton et al. (2016a), this~~ This method using two geopotential heights ~~will be named~~ is named here

2Z.

The N_1 dates with the lowest values of S1 are considered as analogues to the target day. The number of analogues, N_1, is a parameter to calibrate. Then, the daily observed precipitation amount ~~of~~ for the N_1 resulting dates provide the empirical conditional distribution, considered as the probabilistic ~~forecast~~ prediction for the target day.

The other most ~~know~~ well-known parametrization adds a second level of analogy on the moisture variables (method 2Z-2MI, Table 2). The predictor that Bontron (2004) found optimal for ~~the France territory~~ France is a moisture index made of the product of the ~~precipitable water~~ total precipitable water (TPW) with the relative humidity at 850 hPa (RH850). Horton (2012) ~~confirmed~~ confirms that this index is also better for the Swiss Alps than any other variable from the NCEP/NCAR reanalysis I considered independently. When adding a second level of analogy, N_2 dates are subsampled ~~in~~ within the N_1 analogues ~~on~~ of the atmospheric circulation, to end up with a smaller number of analogue situations. When ~~a~~ this second level of analogy is added, a higher number of analogues N_1 is kept on the first level.

2.4. ~~*Data*~~

~~The AM relies on two types of data: predictors, that are atmospheric variablesdescribing the state of the atmosphere at a synoptic scale, and the predictand, which is the local weather time series one wants to predict~~Moisture fields are not as well-predicted by NWP models as pressure variables. This implies that the 2Z-2MI method, when used in real-time forecasting, is very dependent on the skill of the NWP model in predicting moisture fields, and thus its use is often restricted to the first lead times.

~~Predictors are generally reanalysis datasets. NCEP-NCAR reanalysis (6-hourly, 17 pressure levels at a resol~~ ~~used here, but it could be any other reanalysis dataset.~~

~~The predictand (which is to be predicted) is here the daily precipitation (6 a .m. to 6 a.m. the next day) measured at the MeteoSwiss' stations network, for the period 1961-2008. The time series from every available gauging station were~~

~~averaged over subregions of approximately 500 km^2 in order to smooth local effects (Obled et al., 2002; Marty et al., 2012).~~

2.4. Performance assessment

The performance assessment in the present context consists of verifying the prediction of an ensemble probabilistic technique. The set of precipitation values collected with each analogue can be considered as a sample drawn from the conditional distribution associated with the current circulation. The score that is most often used to assess an AM performance is the CRPS (Continuous Ranked Probability Score, Brown, 1974; Matheson and Winkler, 1976; Hersbach, 2000). It allows evaluating the predicted cumulative distribution functions $F(y)$, for example, of the precipitation values y from analogue situations, compared to the observed value y^0. The better the prediction, the smaller the score. The mean CRPS of a prediction series of length n can be written as:

$$CRPS = \frac{1}{n}\sum_{i=1}^{n}\left(\int_{-\infty}^{+\infty}\left[F_i(y) - H_i(y - y_i^0)\right]^2 dy\right) \tag{2}$$

where $H(y - y_i^0)$ is the Heaviside function that is null when $y - y_i^0 < 0$, and has the value 1 otherwise. ~~The mean CRPS is averaged on the calibration, respectively the validation periods, on all days.~~

In order to compare the value of the score ~~in regard~~ relative to a reference, one often considers its skill score expression, and ~~use~~ uses the climatological distribution of ~~daily precipitation~~ precipitation from the entire archive as the reference. The CRPSS (~~*Continuous Ranked Probability Skill Score*~~Continuous Ranked Probability Skill Score) is thus defined as ~~following~~follows:

$$CRPSS = \frac{CRPS - CRPS_r}{CRPS_p - CRPS_r} = 1 - \frac{CRPS}{CRPS_r} \tag{3}$$

where $CRPS_r$ is the CRPS value for the reference and $CRPS_p$ would be the one for a perfect prediction (which implies $CRPS_p = 0$). A better prediction is characterized by an increase in CRPSS.

3. ~~Genetic Algorithms~~

Note, however, that the choice of reference does not matter so much when assessing potential improvements of the method, since we consider more its relative increase or decrease rather than the CRPSS absolute value.

2.1. Sequential calibration

AMs are usually calibrated by a semi-automatic sequential procedure, as elaborated by Bontron (2004) (see also Radanovics et al., 2013; Ben Daoud et al., 2016). The calibration technique optimizes the spatial windows in which the predictors are compared and the number of analogues for every level of analogy, by maximizing the performance score (CRPSS). However, the different analogy levels are calibrated sequentially, and the meteorological variables, pressure levels, and temporal windows are chosen manually. The procedure, as defined by Bontron (2004), consists of the following steps:

1. Manual selection of the following parameters:
    (a) Meteorological variable
    (b) Pressure level
    (c) Temporal window (hour of the day)
    (d) Initial analogue numbers
2. For every level of analogy:
    (a) Identification, for the analogy level considered, of the most skilled unitary cell of all predictors jointly, over a large domain, by a full scanning of the grid.
    (b) From this most skilled cell, the spatial window is expanded by successive iterations in the direction of greater performance gain until no improvement is reached.
    (c) The number of analogue situations N_1 is then reconsidered and optimized for the current level of analogy.

3. A new level of analogy can then be added, based on other variables (such as the moisture index) with some chosen pressure levels, temporal windows, and initial number of analogues N_2. The procedure starts again from step 2 (calibration of the spatial window and the number of analogues) for the new level. The parameters calibrated on the previous analogy levels are fixed and do not change.
4. Finally, the numbers of analogues N_1 and N_2 for the different levels of analogy are reassessed by systematic increments.

The calibration is done in successive steps with a limited number of parameters. Previously calibrated parameters are generally not reassessed (except for the number of analogues).

This procedure was used to calibrate the methods that were here considered as references to further assess the ability of genetic algorithms to outperform the classic approach.

2.2. Genetic algorithms

Genetic ~~Algorithms~~ algorithms (GAs) were developed by Holland (1992) and Goldberg (1989). They are part of the Evolutionary Algorithms (Bäck and Schwefel, 1993; Schwefel, 1993), which ~~get inspiration from some mechanisms of~~ were inspired by some mechanisms in biological evolution, such as reproduction, genetic mutations, chromosomal crossovers, and natural selection. GAs seek the global optimum on a complex surface, theoretically without restriction, which is of interest for AMs, which are characterized by a complex high-dimensional error function having multiple local optima. Practically, GAs allow rapidly approaching satisfactory solutions, but they ~~do not~~ are not guaranteed to provide the optimum solution ~~for sure~~ (Zitzler et al., 2004). It is indeed mainly a matter of time. When the optimizer gets closer to the global optimum, any new improvement takes more time to appear~~(see for example Figure 12)~~, and the final adjustment of the parameters ~~is~~ can be very time consuming (Bäck, 1993). For problems that require a significant amount of time ~~in order~~ to evaluate the objective function, as in the case of AMs (because it needs to make a prediction for every day of the CP), the number of generations has to be limited in order to ~~get~~ ensure a reasonable processing time. Thus, different acceptable solutions can result from one or more ~~optimizations~~ optimization runs (Holland, 1992). This is both a strength and a weakness of GAs: they are very good at exploring complex parameter spaces in order to identify the most promising areas, but they will not necessarily always find the best solution with the optimal values ~~of~~ for all parameters (Holland, 1992).

The optimizations here were performed based on the recommended ~~GAs~~ GA parametrization for AMs as described in ~~Horton et al. (2016b)~~Horton et al. (2016a). As the optimization is mostly sensitive to the mutation operator (that randomly changes some values in the ~~parameters~~ parameter sets), parallel optimizations are considered with variants of this operator, according to ~~Horton et al. (2016b)~~Horton et al. (2016a):

- 3x non-uniform mutation (Michalewicz, 1996) with varying parameters ~~,~~
- 1x multi-scale mutation ~~(Horton et al., 2016b),~~ (Horton et al., 2016a)
- 2x chromosome of adaptive search radius ~~(Horton et al., 2016b)~~ (Horton et al., 2016a).

A population size of 500 individuals (i.e. parameter sets of the AM to be detailed ~~hereunder~~below) was considered, and the optimization was stopped when the best individual (with the highest CRPSS performance score) did not evolve for 20 generations (cycles of ~~the~~ optimization).

3. ~~Case study description~~

~~The study area is the alpine upper Rhône catchment in Switzerland (Fig. 1). The altitude ranges from 372 to 4634 m.a.s.l. and the area is 5524 km^2. This region is the target of the MINERVE (Modélisation des Intempéries de Nature Extrême sur les Rivières Valaisannes et de leurs Effets) project that aims at allowing real-time flood management on the upper Rhône catchment (García Hernández et al., 2009). Even though the region is rather small, the~~

~~meteorological influences related to extreme weather conditions varies substantially within it (see Horton et al., 2012). Based on different climatological analyses, the gauging stations in the catchment were clustered in 10 subregions (Fig. 1) :~~

1. ~~Swiss Chablais~~
2. ~~Trient Valley~~
3. ~~West Bernese Alps~~
4. ~~Lower Rhone Valley~~
5. ~~Left side valleys~~
6. ~~Southern ridges~~
7. ~~Upper Rhone Valley~~
8. ~~Southeast ridges~~
9. ~~East Bernese Alps~~
10. ~~Conches Valley~~

~~The 48 years precipitation dataset (see section 1.2.2) was divided into a calibration period (CP) and a validation period (VP). Using data independent from the CP to validate the results is very important in order to assess the robustness of the improvements and to avoid over-parametrization of the method. Parameters determined on the CP are then applied to the VP in order to obtain a performance score for the independent period.~~

~~In order to reduce potential bias related to trends linked to climate change or to the evolution in measurement techniques, the selection of the VP is evenly distributed over the entire series (Ben Daoud, 2010). Thus, one year every six years were selected for validation, which represents a total of 8 years for the VP and 40 for CP. The choice of the sequence was made in order to have similar statistical characteristics between the CP and the VP.~~

3. Optimization of the circulation analogy

The analogy of the atmospheric circulation was optimized for the ~~10 subregions~~ ten subregions (Section 2.1) independently. We started from the ~~most simple~~

simplest AM, and increased the complexity in order to identify the degrees of freedom that are of particular interest. Thus, the tested parametrization evolved iteratively in complexity. The detailed results of the intermediate stages are not provided in this paper (see Horton, 2012, for the details).

The reference method for the analogy of the atmospheric circulation (2Z, Table 1), based on Z500 and Z1000, was first considered. The optimizer had to choose simultaneously the number of analogues, both spatial windows with no overlapping constraint (i.e. they can differ from one pressure level to another), as well as the temporal windows (hours of observation of the geopotential), ~~what~~ which cannot be achieved with the sequential calibration ~~cannot do. With these~~ technique. The performance score (CRPSS) was slightly improved, with these limited degrees of freedom, ~~a relative CRPSS improvement of 3.97% and 2.45% in average was obtained for the CP and the VP respectively~~relative to the 2Z reference method calibrated with the sequential procedure. Some tests showed that most of the gains ~~are~~ were due to the non-overlapping spatial windows. This ~~is not a tremendous improvement, but it~~ demonstrated that the optimizer was able to ~~get relevant parameters~~ obtain relevant parameters for a simple method.

Then, an additional degree of freedom was provided to the GAs by letting them choose the pressure levels along with the other parameters ~~,~~ (analogue numbers, spatial and temporal windows), which is also a non-automated process in the sequential calibration. This degree of freedom increased the optimization time~~and may~~ , and might decrease the number of simulations that converge to a single solution. However, most solutions were very close in terms of ~~score. The averaged relative improvement of the CRPSS is 5.63% for the CP and 3.82% for the VP. The pressure levels that were chosen are~~ the performance score, which was further improved. The selected pressure levels were Z500 or Z700 for the upper level, and Z925 or Z1000 (most often) for the lower level.

Parallel analyses showed that the analogy of circulation is incomplete, and that ~~the geopotential still contains~~ geopotential heights still contain relevant information that can improve the statistical relationship. Therefore, a third~~predictor~~

~~was~~ , followed by a fourth circulation predictor were added (still ~~on the geopotential height)that the optimizer could use along with the previous parameters~~only geopotential heights). There was no constraint on the predictors, so that the same pressure level could be selected more than once. ~~Some~~ Further improvements were found ~~on the~~ in the performance score, both for the CP and the VP, confirming that this additional information ~~is~~ was beneficial for the quality of the prediction. ~~We then tried with 4 predictors, and so on, up to 8. Every time a new predictor to optimize was added, the score on the CP increased, but always more to a smaller extent. However, the score value on the VP dropped after 4 predictors, revealing an over-parametrization of the method, and thus a lack of robustness. Considering 4 predictors is optimal for this case study, since the gain in CRPSS is significant. It cannot be excluded that another number would prevail in another region under other meteorological conditions.~~

Finally, a weighting of the analogy criteria values per pressure level was proposed, again optimized by GAs. The weighting operates in the combination of the S1 criteria processed on every level, which were previously averaged with equal weights. The role of this new degree of freedom is to give more weight to the levels with greater predictive capacity, and to consider the ~~geopotential variability changes~~ differences in the geopotential height variability with altitude.

3.1. ~~*Which parameters are optimized ?*~~

The number of circulation predictors (still only geopotential heights) was then successively increased up to ten, considering the weighting of the analogy criteria values. The addition of circulation predictors globally improved the prediction skill (for both the CP and the VP) only up to four predictors (Figure 2). Afterwards, the score on the VP was more variable, eventually even showing a decrease, which revealed an over-parametrization of the method, and thus a lack of robustness. After four predictors, the score for the CP did not increase substantially, and even presented a local decrease due to increasing difficulty for the optimizer to converge. Selecting four circulation predictors (geopotential heights) was considered optimal for this case study, since the gain in CRPSS was significant, and the model remained relatively simple. It cannot be ruled out that another number would prevail in a region other than the upper Rhône catchment, under other meteorological conditions, or with another reanalysis dataset.

3.1. *Which parameters are optimized?*

The chosen method for the atmospheric circulation analogy, based on ~~4 levels of the geopotential , and that will be name~~ four circulation predictors (geopotential heights), and which is here named 4Zo ~~, is made of~~ (o for optimized), was based on the following degrees of freedom:

- ~~the selection of the~~ selection of pressure levels (4 degrees)
- ~~the~~ temporal windows (4 degrees)
- ~~the~~ spatial windows (~~4x4~~ 4 x 4 degrees)
- ~~the~~ weights (4 degrees)
- ~~the~~ number of analogues (1 degree).

This ~~sums~~ adds up to 29 degrees of freedom that ~~are~~ were optimized simultaneously.

3.2. *Results for the 4Zo method*

The resulting optimized parameters for 4Zo vary from one subregion to another. ~~An example of the detailed parameters is provided for the Swiss Chablais in Table ??.~~ The optimized spatial windows are given for every subregion in Figure 3, and the selected pressure levels in Table 3.

The resulting CRPSS scores are provided in ~~Table ?? and are in~~ Figure 4 and were on average 35.8% for the CP and 35.5% for the VP~~. The improvement of the CRPSS score relatively to~~ , compared to 31.1% and 32.3%, respectively, for the reference method 2Z on the atmospheric circulation (optimized by the sequential procedure)~~is illustrated in Figure ?? and is in average 15.3 % for the CP and 9.9 % for the VP~~. The score was also calculated for three precipitation thresholds: $P \geq 1$ mm, $P \geq 0.1 \cdot P10$, and $P \geq 0.5 \cdot P10$, P10 being the daily precipitation with a 10 year return period (Table 4). The gain in score ~~increases~~ increased with the precipitation threshold: the relative improvement of the CRPSS ~~is in average, respectively~~ was, on average, for the different thresholds, 13.3%, 15.4%, and 29.1% for the CP and 7.9%, 11.1%, and 34.5% for the VP. The optimization ~~improves thus even more the prediction~~ thus improved the prediction even more for days with significant precipitation than for the usual days.

To assess the parameters cross-compatibility and the spatial coherence of the resulting parameters, those optimized for one subregion were applied to the others. The resulting losses or gains of the CRPSS are displayed in Figure ~~?? for the CP and in Figure ?? for the VP~~5.

3.3. Analysis

The automatic selections of ~~the~~ pressure levels (Table 3) and ~~the~~ temporal windows (not shown) for the analogy of circulation ~~show~~ showed a great homogeneity and ~~are~~ were spatially consistent. First of all, the level Z1000 ~~is~~ was always selected twice (the first time at 6 or 12 h, and the second always at 30 h) and Z700 ~~is~~ was selected once for every subregion (always at 24 h). The level ~~which varies~~ that varied from one subregion to another~~is~~, albeit in a spatially consistent way, was the upper level (~~however~~ always at 12 h), which ~~is~~ was Z300 for the ~~North-West~~ north-west part of the catchment, Z500 for most of the other subregions, and Z600 for the Conches Valley. ~~Its spatial distribution is however homogeneous.~~ The optimizer thus provided consistent selections of pressure levels and temporal windows~~, which depicts a significant preference in the AM, and the success of GAs to provide consistent results~~. The automatic selection of ~~the~~ pressure levels is a big advantage in ~~favor of a~~ favour of global optimization.

The resulting spatial windows (Figure 3) may look very diverse first, but there are significant similarities for subregions located within the same vicinity. The first ~~4 subregions are~~ four subregions were characterized by a large spatial window on the upper level, whereas it ~~is~~ was smaller elsewhere. For most subregions, the second level (Z700) ~~is represented by~~ was compared on thin and longitudinally extended ~~domains~~spatial windows. The third level (Z1000 at 6 or 12 h) also ~~contains~~ had longitudinally extended domains, ~~but a bit~~ which were slightly larger. The last one (Z1000 at 30 h) ~~is~~ had rather large and squared windows. Subregions number 5 (~~Left side~~ southern valleys) and 6 (~~Southern ridges) have~~ southern ridges) had exactly the same spatial windows, which ~~suggest~~ suggests that they behave in a similar way and thus could have been merged. This similarity is a good sign for the accuracy of the optimized parameters.

The ~~scores show significant~~ performance scores showed non-negligible improvements for both the CP and ~~the~~ VP (Figure ~~??~~4) compared to the 2Z reference method optimized by the sequential procedure. Even more interestingly, the results for higher precipitation thresholds (Table 4) ~~show~~ showed the largest improvements. This is of particular interest in the framework of flood forecasting.

~~The CRPS score was also discretized into its accuracy and sharpness components, as suggested by Bontron (2004). The changes in each of these components has been illustrated in Figure ?? relatively to the total CRPS value of the 2Z reference method for the different regions and both the CP and the VP. One can see that the accuracy part is always improved to a greater extent than the sharpness, which can occasionally deteriorate. It means that the medians of the predicted precipitations distributions are closer to the observed value, whereas the spread of the distribution vary, but is in general a bit narrower~~The further improvement of days with higher precipitation totals is likely related to the fact that larger values contribute more to the CRPS score, which means that better predicting these days results in significant increase in the global performance score.

The analysis of the parameters cross-compatibility ~~shows that obviously, the parameters are~~ showed that the parameters were obviously optimal on the CP ~~when they are optimized for a given subregion (Figure ??~~for the subregion for which they were optimized (Figure 5 top). However, the losses in CRPSS when exchanging the parameters ~~are~~ were not of the same magnitude ~~between~~ among the different subregions. Indeed, the Upper Rhone Valley (7) and~~moreover the Southeast~~ , moreover, the southeast ridges (8) ~~seem~~ seemed to behave significantly differently~~, likely due to their particular sensitivity to southerly flows(Horton et al., 2012).~~ . These two regions have different climatic properties than the others, as they are particularly sensitive to southerly flows. Indeed, almost all heavy precipitation events occurred under a southerly regime, such as in the Liguria, Piedmont, and Aosta regions in Italy, whereas the other subregions of the catchment had extreme events mainly under a westerly regime (Horton et al., 2012). Thus, as the performance score is significantly influenced by heavy precipitation values, the parameters for the different subregions are likely optimized to better predict these days. It can then be expected that the optimal parameters differ between these two subregions and the others. This points at the importance of taking into account leading meteorological influences during ~~discretization, that~~ precipitation station clustering, which are not always best represented by ~~the physical distance.~~ geographical distance.

Globally, the same ~~pattern can~~ cross-compatibility structure could be observed for the VP (Figure ~~??~~5 bottom), but in this case, minor improvements ~~may occur~~ were occasionally observed when crossing the parameters, ~~due to~~ because of the presence of other events in the VP that ~~may~~ might be better predicted by a different parameter set. The relatively small ~~gaps in score between parameterizations indicate~~ differences in scores between parameterizations indicated that even though the parameters ~~may~~ might differ significantly, the performance ~~may~~ might not be drastically affected. Even a change in the pressure level ~~does~~ did not mean a radical drop ~~of~~ in the score value. A different parametrization may lead to a distinct selection of analogue days, and thus to an improvement of the prediction under certain weather conditions at the expense of others.

4. Optimization of the analogy with moisture information

It is known that moisture variables as a second level of analogy do provide improvements to the method (section ~~??~~2.3). The moisture index, which is a combination of the relative humidity and ~~the~~ precipitable water, has thus also to be optimized. In order to do so, a constraint ~~to~~ on the optimizer had to be introduced, so as to select the same temporal window (time of observation) for both variables.

Two methods were assessed: one with ~~2~~ two moisture predictors (moisture index on ~~2~~ two pressure levels or at ~~2~~ two different hours), named 4Zo-2MIo, and one with ~~4 moisture predictors~~four moisture predictors, named 4Zo-4MIo. When introducing ~~2~~ two predictors for the moisture analogy, the number of degrees of freedom ~~raises~~ increased to 42, and to 54 with ~~4 predictors.~~ four predictors. However, there was no substantial difference in the performance scores between both 4Zo-2MIo and 4Zo-4MIo methods, which suggests that considering four moisture predictors is not necessary. For this reason, only the results of 4Zo-2MIo are presented.

The optimization was processed on both levels of analogy simultaneously. This implies that the analogy of the atmospheric circulation ~~may change due to~~ could change because of the new moisture information.

4.1. Results for the 4Zo-2MIo ~~and 4Zo-4MIo methods~~method

~~Due to significant similarities between the results from 4Zo-2MIo and 4Zo-4MIo, the latter will only be partly shown in order to improve readability.~~

~~As~~ As seen previously, the optimized parameters ~~differ~~ differed from one subregion to another, ~~and this even to a~~ but to an even greater extent. ~~Detailed examples are again provided for the Swiss Chablais subregion in Table **??** for 4Zo-2MIo.~~ The resulting spatial windows are displayed in Figure 6 for 4Zo-2MIo, along with the selected pressure levels for both the circulation and ~~the~~ moisture analogy (Table **??**).

The CRPSS scores of the optimized ~~methods~~ 4Zo-2MIo method are provided in ~~Table **??** and amounts to slightly more than~~ Figure 7 and amounted on average to 40% ~~in average for both methods and for both periods. This results in a relative improvements for 4Zo-2MIo that is in average 14.0 % for the CPand 11.5 % for the VP (Figure **??**~~(CP) and 40.3% (VP), compared to 35.2% (CP) and 36.2% (VP) for the reference method 2Z-2MI on the moisture analogy optimized with the sequential procedure. ~~For 4Zo-4MIo, the average improvement is 15.6 % for the CP and 12.1 % for the VP (Figure **??**).~~

The parameters cross-compatibilities are shown in Figure 9. As for 4Zo, the 4Zo-2MIo ~~and 4Zo-4MIo methods present~~ method presented larger improvements in the prediction ~~for~~ of significant rainfall (~~thresholds P≥1 mm, P≥0.1·P10 and P≥0.5·P10). The improvement are relatively similar for 4Zo-2MIo (~~Table **??**)~~and 4Zo-4MIo (not shown), with slightly superior scores for the latter on small precipitation (P≥1 mm : 17.7% and 13.0%) and extremes (P≥0.5·P10 : 23.7% and 29.2%)~~.

~~The parameters cross-compatibility has also been assessed for the methods with moisture variables, and are shown in Figures **??** and **??** for the method 4Zo-2MIo (not shown for 4Zo-4MIo, but similar).~~

4.2. Analysis

When optimizing a method ~~made of 2~~ consisting of two levels of analogy, the introduction of moisture variables in the second level has an influence on the parameter values of the first level. This means that the two levels of analogy bring complementary information, and are thus not independent. This is first visible ~~on~~ in the number N_1 of analogues to be selected on the first level, and ~~on~~ in the selection of the pressure levels for the circulation analogy. If the change in the optimal value of N_1 was already known, a change in the optimal pressure levels for the circulation analogy has never been identified before.

As for the sequential procedure, the optimal value of N_1 ~~increases~~ increased when adding a second level of analogy (Figure ~~**??**~~10). One can also ~~notice~~ see that the optimal number of analogues N_2 for the second level of analogy of 4Zo-2MIo ~~is~~ was slightly inferior to N_1 from 4Zo, but very close. There is a ~~trend~~ globally common tendency between the optimal analogue number values of both methods: ~~the higher~~ N_1 of the 4Zo method, ~~the higher~~ and N_1 and N_2 of 4Zo-2MIo ~~is (Figure **??**). This relationship is not to be considered perfectly robust and should not be transposed to another case study, but in this case, it relates the magnitudes of the N_1 and N_2 on the various levels and methods~~tend to be higher or lower together for a given region.

The optimal final ~~number of analogues do~~ numbers of analogues did not vary much: $23 \leq N_1 \leq 33$ for 4Zo and $21 \leq N_2 \leq 28$ for 4Zo-2MIo. However, the optimal number of the N_1 analogues of the first level of 4Zo-2MIo ~~varies~~ varied to a greater extent: $48 \leq N_1 \leq 84$. In this latter method, it may be problematic to consider a fixed and unique value for all regions.

As for the pressure levels, Z1000~~that~~, which was previously systematically selected twice (Table 3) ~~is now less~~ was here less often chosen (once or even not at all) for ~~both~~ 4Zo-2MIo (Table **??**)~~and 4Zo-4MIo (not shown). There is indeed a shift of~~. There was indeed a vertical shift in the previously selected Z1000 for higher levels ~~, that is~~ that was even slightly stronger with ~~4~~ four moisture predictors than with ~~2.~~ two. This change is likely due to the fact that when considering only the circulation analogy, the method ~~tries~~ tried to take into account information that can serve as a proxy for moisture assessment, whereas it ~~does~~ did not need it with the moisture index. This ~~aspect has never been demonstrated before, as sequential calibration tools do not allow it. It~~ can only be assessed by a global optimization technique that can ~~tune jointly~~ work jointly on both levels of analogy.

The selected pressure levels for the analogy ~~on~~ of the moisture index ~~are strongly centered~~ were strongly centred around 700 hPa and 600 hPa. No other value ~~has been~~ was selected when considering ~~2 pressure levels~~ two moisture predictors (Table **??**)~~, and when considering 4 levels, 850 hPa and 500 hPa were sometimes also selected (not shown). However, even in this latter method, the 700 hPa and 600 hPa levels still hold 78 % of the selection. It is thus~~. It was sometimes more efficient, in terms of prediction performance, to consider ~~one of this level several times at different hours~~ the moisture at 700 hPa twice, but at different hours, rather than selecting another pressure level. Besides, the optimizer never chose the same pressure level at the same hour for any variable, even though it was allowed to do so. The selected pressure levels for the ~~analogy on the moisture index differ from the parameters resulting from the reference method optimized by the sequential procedure~~ moisture analogy differed from the reference method (Tables 2 and **??**, last row). ~~It is possible that the level 850 hPa is more optimal for the region for which the reference parameterization was established. However, this selection is usually not reconsidered when applying the sequential procedure.~~

The selection of ~~the~~ temporal windows for ~~the atmospheric circulation is~~ atmospheric circulation was similar to the preceding optimization (in ~~the~~ order of increasing pressure: 12 h, 24/30 h, 12 h, 30 h), but sometimes with ~~a bit more~~ some variability. When it comes to the moisture analogy, there ~~is a clear trend~~ was a clear tendency to select 12 h and 24 h~~when considering 2 predictors.~~. However, it must be remembered that this holds for our predictand, the accumulated precipitation over 06–30 h UTC, and that it is expected to differ if the temporal window changes (e.g. 00–24 h UTC, or another accumulation duration).

The optimized spatial windows for the atmospheric circulation have also changed (Figure 6~~; results of 4Zo-4MIo are not shown but present the same trends~~). The very large domains on the upper level of the ~~4 first subregions are~~ first four subregions were not present anymore, and more variability ~~can~~ could be observed. The selected points for the ~~analogy on the moisture index are always located nearby~~ moisture analogy were always located near the catchment, including at least ~~1~~ one of the nearest points from the reanalysis dataset, and the spatial windows ~~are~~ were relatively small. Thus, for this case study, there is no need to look for distant moisture information, and the search could be reduced to a smaller domain.

The CRPSS scores were improved by considering the moisture information (~~Table **??**~~Figure 7 to be compared with Figure 4). The optimized ~~methods also perform~~ method also performed significantly better than the ~~reference method based on the moisture index (Figures **??** and **??**)~~ 2Z-2MI reference method optimized by the sequential procedure. ~~However, there is no drastic difference between both 4Zo-2MIo and 4Zo-4MIo methods, which suggests that considering 4 moisture predictors is not necessary.~~ When it comes to improvements for days with precipitation above the ~~3~~ three thresholds (P≥1 mm, P ≥ 0.1·P10, and P ≥ 0.5·P10), the conclusion is the same as before, that is, a significant improvement ~~of~~ in the prediction compared to the reference method, mainly for heavy rainfall.

~~As previously, the CRPS scores were discretized into their sharpness and accuracy components. The changes in each of these components, expressed relatively to the total CRPS values, is once again in favor of the accuracy over the sharpness (Figure **??** for 4Zo-2MIo; 4Zo-4MIo not shown, but similar).~~

The analysis of the parameters cross-compatibility (~~Figures ?? and ??)~~ ~~is~~ Figure 9) was also very similar to the one ~~on~~ of the circulation analogy only~~(results for 4Zo-4MIo not shown as very similar)~~. The same pattern ~~can~~ could be observed, with a drop of performance for the subregions submitted to different meteorological influences. However, the losses ~~of performance are~~ in performance were globally more important than before, suggesting that more complex methods with moisture variables are less transposable to another subregion (consistent with the observations of Chardon et al. (2014)), even though both ~~are~~ were located within the same grid cell of the reanalysis dataset. Moisture fields have greater variability than pressure fields, and thus a change in the spatial windows can have a greater impact on the method performance. Indeed, the two regions with the lowest cross-compatibility with the others were the upper Rhone Valley (7) and the southeast ridges (8), which had similar optimal pressure levels and temporal windows to other regions, but had rather different spatial windows on the moisture predictor.

Predictors based on moisture variables do significantly increase the prediction skill, and are thus recommended, as long as they are reliable. In real-time forecasting, their reliability depends on the lead time: for lead times superior to 3–4 days, the uncertainties related to moisture variables from NWP models become fairly high, which reduces the relevance of methods relying on this information. In climate downscaling studies, it mainly depends on the coherence of the climatologies between the archive and the GCM model outputs. One should, however, not establish an AM with moisture variables for too large a region, as the transferability is reduced (see Chardon et al., 2014, for alternative approaches).

5. Discussion

The optimization of the AM by means of GAs has been undertaken in successive stages by releasing progressively new degrees of freedom. This approach allowed us to differentiate the contributions to performance gains, as well as to identify possible over-parametrization. The main improvements ~~for~~ obtained in the present case study are due to the following ~~factors~~elements:

- Using ~~4~~ four pressure levels for the ~~analogy of circulation . It seems to be the~~ circulation analogy seemed to be an optimal number for the studied region~~. Beyond that value~~ , length of archive available, and target predictand considered. Beyond four, the validation score ~~drops~~was more variable, revealing a loss in robustness due to over-parametrization.

- The automatic and joint optimization of all parameters: the ~~analogues number, the selection of the~~ analogue numbers, selection of pressure levels and ~~the~~ temporal windows, and ~~the~~ spatial windows. These parameters are highly interdependent, so one needs to optimize them jointly in order to identify optimal combinations. ~~Traditional calibration procedures based on a systematic assessment of every combination is not possible anymore when considering more than 2 pressure levels.~~Indeed, there is a strong interdependence between space and time in the atmospheric circulation, so that, e.g. the spatial window should move upstream the main atmospheric flow for earlier temporal windows.

- The introduction of distinct spatial windows between pressure levels. The synoptic circulation is characterized by features with very different scales depending on the height, and important information for predicting ~~rainfall~~

precipitation is not necessarily located in the same area from one level to another. ~~However, the optimized spatial windows are consistent in between the subregions.~~

- The weighting of the analogy criteria between different pressure levels. ~~It~~ This can be influenced by the variability of the geopotential height with altitude, ~~and the change of some levels significance with the targeted~~ or the levels of significance in regards to the meteorological processes specific to a region. There is a trend ~~for~~ in the weighting of circulation predictors to decrease with the increase in pressure, as one can see in Figure ~~?? for the method 4Zo, and in Figure~~ 11 for the ~~averages over the 3~~ three optimized methods. However, the values ~~stay around equity~~remained approximately equal. This may not be the most influencing factor, and we may suggest ~~to remove~~ removing it first when trying to reduce the ~~number of~~ degrees of freedom.

- The joint optimization of the circulation and moisture analogy levels, ~~that~~ which are usually calibrated successively. We ~~have been~~ were able to demonstrate that there is a dependency between the analogy levels, and that in order to approach the optimal parameters, one must consider them jointly.

GAs have proved very useful to optimize complex variants of the AM, and to assess new degrees of freedom that were not available ~~so~~ thus far. However, it can be dangerous to add too many parameters to optimize. Indeed, the optimizer will probably use them to successfully improve the calibration score, ~~but~~ so the validation control remains very important in order to determine if one is actually improving the method, or if it is being over-parametrized. Moreover, it might not always be desirable to increase the degrees of freedom, and some constraints (e.g. same weighting of the analogy criteria between different pressure levels) can be justified. However, one should first assess the consequence of a constraint before establishing it. In this sense, even though not all degrees of freedom are useful, GAs allow us to assess their influence. Finally, GAs could be used to identify, among other things, the best pairs of spatial and temporal windows, in order to later create a simpler regional method.

The convergence of parallel optimizations ~~decreases when the method to optimize becomes~~ decreased when the AM to optimize became more and more complex. The optimizer ~~do~~ did not always converge to the exact global optimum, but to its surroundings. This is related to the fact that the optimization slows down when it gets closer to the global optimum, and that one has to stop it before the end, ~~due to~~ because of the required processing time (see ~~Figure 12 as example~~for example the slow-down of the improvements over generations in Figure 12). The resulting parameters ~~may sometimes present non negligible~~ might sometimes present non-negligible differences, even though the score is ~~almost~~ similar. Through ~~some~~ Monte-Carlo analyses of the parameter space properties of the AM, Horton (2012) showed that some parameters of the method have a wide range of acceptable values~~(see also Horton et al., 2016a)~~. The spatial windows, for example, can be larger than the optimal size without much impact on the score, while they cannot be smaller (see also Bontron, 2004). We ~~could also observe~~ also observed that the selection of ~~the~~ pressure level is not a parameter as discrete as we would have thought, and that choosing another level may have reduced the consequences on the performance. This is particularly true for higher pressure levels, but can be more critical for lower layers. It was thus interesting to sometimes ~~get~~ obtain several sets of near-optimal parameters, but with some nuances, in order to get an idea of the sensitivity of the parameters for a given region, and to compare the score ~~on~~ for the VP. In this regard, a cross-validation technique may be advisable. However, as solutions identified at different regions of the parameter space might provide sufficiently good performance, an ensemble of these could be used, instead of a unique solution. These could account for the parameters uncertainty in the AM, and could also increase the sample size contributing to the empirical distribution of precipitation values. An approach that can also be recommended is to first explore a wide range of the parameter space with some optimizations, and to narrow it according to the results for more targeted optimizations that are likely to go faster and to perform better.

We tried to optimize the length of the preselection period (i.e. the ~~4-months~~ seasonal stratification, which is a 4-month window) jointly with the other parameters, but no improvement was observed. Optimizing the moisture flux, which is composed of the moisture index multiplied ~~with~~ by the wind flux, was also assessed. However, the results were not better than when considering the moisture index alone. This may be related to the fact that the optimizer tries to provide the best analogy of the atmospheric circulation in the first place, which makes the wind information less relevant in the second level of analogy.

As ~~it~~ has been observed, methods with ~~a~~ higher complexity that integrate moisture predictors are less transposable than simpler ones. It was also noticed in another unpublished work, that it is by far better to optimize for ~~2 subregions jointly rather~~ two subregions jointly than to optimize on one and to apply its parametrization to the other. Finally, the discretization in subregions is an important process and should be handled with care. Indeed, the ~~physical~~ geographical distance is not always the leading factor to define a subregion. For example, the ~~Southeast ridges subregion do~~ southeast ridges subregion does not behave like its ~~surrounding and differ~~ surroundings, and differs in its parametrization ~~, due to~~ because of different leading meteorological influences.

GAs are relatively heavy on processing and require an IT infrastructure capable of performing thousands of hours of calculations. However, they automatically optimize all parameters of the AM, ~~what~~ which is not possible with the sequential calibration~~does not allow~~. Therefore, much human time~~is saved, that was~~ , previously required to ~~assess successively~~ successively assess numerous combinations of parameters (particularly the selection of ~~the~~ pressure levels and ~~the~~ temporal windows), is saved. The ability to ~~optimize jointly~~ jointly optimize all parameters is important given the strong dependencies between them and between the levels of analogy.

Furthermore, AMs optimized with GAs showed an improvement in predictions for days with heavier precipitation, including extremes. Even though no new extreme value was added to the existing precipitation archive, the distribution of analogue precipitation values for a target situation can move towards the targeted extreme by sampling better candidate situations. Then, the subset of precipitation values collected on the analogue dates can be considered as a sample of the conditional distribution of precipitation associated with this situation. A truncated exponential or a gamma distribution model can be fit and extrapolated to extreme values not contained in the sample or even in the whole precipitation archive (Obled et al., 2002). Another possible approach is to combine AMs with other methods (e.g. Chardon et al., 2014).

6. Conclusions and perspectives

The parameters resulting from the optimization by GAs ~~are~~ were very consistent in terms of ~~selection of the pressure levels , and the~~ the selection of pressure levels and temporal and spatial windows. There ~~are~~ were clear trends or even identical results for subregions under similar meteorological influences, which confirm that the optimized parameters ~~are~~ were coherent, despite an eventual first impression of ~~a~~ great variability in the spatial windows. When adding moisture variables, the results ~~show~~ showed a higher variability, but ~~remains~~ remained highly acceptable and coherent.

Strong dependencies between the parameters of the AM ~~could be~~ were observed. Thus, the sequential calibration, which optimizes the parameters successively, may not lead to the optimal combination. Moreover, it contains several manual systematic ~~assessment~~assessments, such as the selection of ~~the~~ pressure levels and ~~the~~ temporal windows. GAs, however, can ~~select the~~ automatically select pressure levels and ~~the time windowsautomatically~~temporal windows, which can save a considerable amount of human time. A great advantage of a global optimization is its ability to approach or reach optimal parameter values when they are considered jointly.

A ~~parametric dependence between the analogy of circulation and the~~ dependence in the selected parameters between the circulation analogy and moisture variables was identified. When the two analogy levels are considered together, the optimal parameters of the circulation analogy ~~are different~~changed. This complexity can only be exploited in a suitable manner by global optimization methods.

For the present case study, there ~~seems~~ seemed to be an optimum number of pressure levels to consider for the ~~analogy of circulation~~ circulation analogy, which is four, before losing ~~performance in validation. The analogy of circulation has also been~~ consistency of the real gains. The circulation analogy was improved by introducing a weighting between pressure levels, and considering independent spatial windows between pressure levels.

GAs ~~provide~~ provided parameterizations of AMs that ~~exceed~~ exceeded the performance of the sequential calibration. In addition, it has been observed that the prediction for days with strong precipitation were ~~significantly more improved~~ improved to a greater extent, which is clearly interesting in the ~~framework~~ context of flood forecasting.

This work is by ~~far not exhaustiveand means~~ no means exhaustive, and is meant to open a door ~~for~~ to new explorations of AMs with GAs. It is even possible to let the optimizer chose the meteorological variable to be used as a predictor, as well as the analogy criteria~~. It is already possible to undertake such approach with our code~~, which is the topic of work in progress. Moreover, the AM has been explored for decades for precipitation ~~forecasting, but very few works analyze its potential~~ prediction, but not as intensively for other predictands. A global optimizer, such as ~~GAs~~a GA, can speed up this assessment significantly.

Acknowledgments

~~Thanks to~~ We thank Hamid Hussain-Khan of the University of Lausanne for his help and availability, and for the intensive use of the cluster he ~~is in charge of. Thanks to~~ administers. We also thank Renaud Marty for ~~his~~ fruitful collaboration over the years~~. Thanks to~~, and Dominique Bérod for his support.

~~Thanks to~~ We thank the Roads and Water courses Service, the Energy and Water Power Service of the Wallis Canton, the Water, Land and Sanitation Service of the Vaud Canton, and the Swiss Federal Office for Environment (FOEV) ~~who financed~~ for financing the MINERVE (Modélisation des Intempéries de Nature Extrême des Rivières Valaisannes et de leurs Effets) project which started this research. ~~The fruitful~~ 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~~ were provided by NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, ~~from their Web site~~ via their website at http://www.esrl.noaa.gov/psd/. ~~Precipitation time series~~ The precipitation time series were provided by MeteoSwiss.

The authors would also like to acknowledge the work of Loris Foresti and another anonymous reviewer, which contributed to improving this paper.

References

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

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

Bäck, T., 1993. Optimal Mutation Rates in Genetic Search, in: Proceedings of the 5th International Conference on Genetic Algorithms, pp. 2–8.

Bäck, T., Schwefel, H.P., 1993. An Overview of Evolutionary Algorithms for Parameter Optimization. Evol. Comput. 1, 1–23. doi:`10.1162/evco.1993.1.1.1`.

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

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

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

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

Bois, P., Obled, C., Thalamy, J., 1981. Etude des liaisons entre champs de pression et températures sur l'Europe avec la durée d'insolation: Application à la possibilité de la prévision d'insolation, in: C.R. du colloque Météorologie de l'énergie solaire, PIRDES-CNRS, Toulouse, pp. 145–177.

Bolognesi, R., 1993. Premiers développements d'un modèle hybride pour le diagnostic spatial des risques d'avalanches. La Houille Blanche 8, 551–553. doi:10.1051/lhb/1993045.

Bontron, G., 2004. 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.

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

Brown, T., 1974. Admissible Scoring Systems for Continuous Distributions. Technical Report.

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

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

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

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

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

Drosdowsky, W., Zhang, H., 2003. Verification of Spatial Fields, in: Jolliffe, I.T., Stephenson, D.B. (Eds.), Forecast verification: a practitioner's guide in atmospheric science. Wiley. chapter 6, pp. 121–136.

Duband, D., 1970. Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations. Ph.D. thesis. Faculté des sciences de Paris.

Foresti, L., Panziera, L., Mandapaka, P.V., Germann, U., Seed, A., 2015. Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall. Meteorol. Appl. 22, 141–155. doi:10.1002/met.1416.

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

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman Publishing Co., Inc.

Gordon, N.D., 1987. Statistical very short-range forecasting via analogues, in: Proceedings of the Symposium on mesoscale analysis and forecasting, ESA SP-282, Vancouver, Canada. pp. 307–312.

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

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

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

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

Holland, J.H., 1992. Genetic Algorithms. Sci. Am. 267, 66–72. doi:10.1038/scientificamerican0792-66.

Horton, P., 2012. 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.

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

Horton, P., Jaboyedoff, M., Obled, C., 2016a. Global Optimization of an Analog Method by Means of Genetic Algorithms. Mon. Weather Rev. Accepted. doi:10.1175/MWR-D-16-0093.1.

Horton, P., Obled, C., Jaboyedoff, M., 2016b. The Analogue Method for Precipitation Forecasting: Finding Better Analogue Situations at a Sub-Daily Time Step. Hydrol. Earth Syst. Sci. Discuss. , in reviewdoi:10.5194/hess-2016-246.

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

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

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

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

Lorenz, E., 1956. Empirical orthogonal functions and statistical weather prediction. Technical Report. Massachusetts Institute of Technology, Department of Meteorology, Massachusetts Institute of Technology, Dept. of Meteorology.

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

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., Thiele-Eich, I., 2010. Precipitation downscaling under climate change: Recent developements to bridge the gap between dynamical models and the end user. Rev. Geophys. 48, 1–34. doi:10.1029/2009RG000314.

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

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

Michalewicz, Z., 1996. Genetic Algorithms + Data Structures = Evolution Programs. 3rd editio ed., Springer-Verlag.

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

Obled, C., Good, W., 1980. Recent developments of avalanche forecasting by discriminant analysis techniques: a methodological review and some applications to the Parsenn area (Davos, Switzerland). J. Glaciol. 25, 315–346. doi:10.3198/1980JoG25-92-315-346.

Panziera, L., Germann, U., Gabella, M., Mandapaka, P.V., 2011. NORA-Nowcasting of Orographic Rainfall by means of Analogues. Q. J. R. Meteorol. Soc. 137, 2106–2123. doi:10.1002/qj.878.

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

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

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

Schwefel, H.P., 1993. Evolution and Optimum Seeking. John Wiley & Sons, Inc., New York, NY, USA.

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

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

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

Zitzler, E., Laumanns, M., Bleuler, S., 2004. A Tutorial on Evolutionary Multiobjective Optimization, in: Metaheuristics for Multiobjective Optimisation, pp. 3–37. doi:10.1007/978-3-642-17144-4_1.

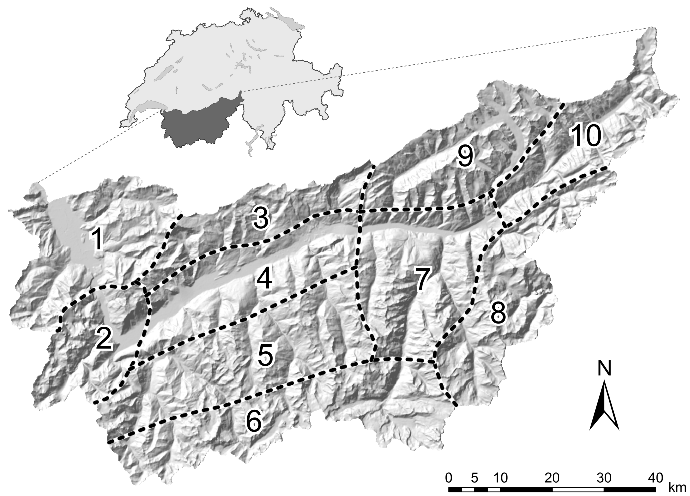


Figure 1: Location of the alpine Rhône catchment in Switzerland. (source: Swisstopo)

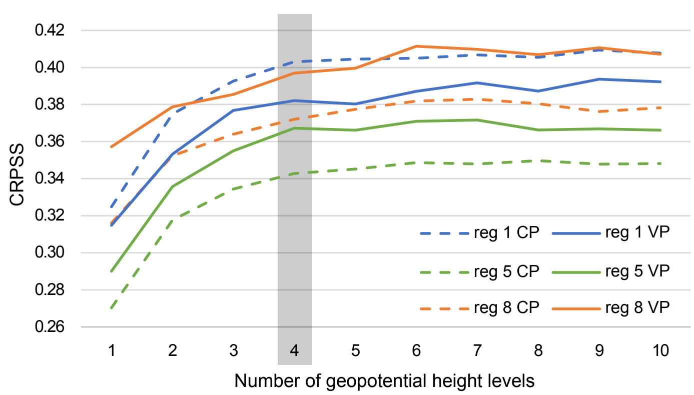


Figure 2: Performance score (CRPSS) for CP and VP for three stations (1 - Swiss Chablais; 5 - Southern valleys; 8 - Southeast ridges) when varying the number of geopotential height levels available to the optimizer.

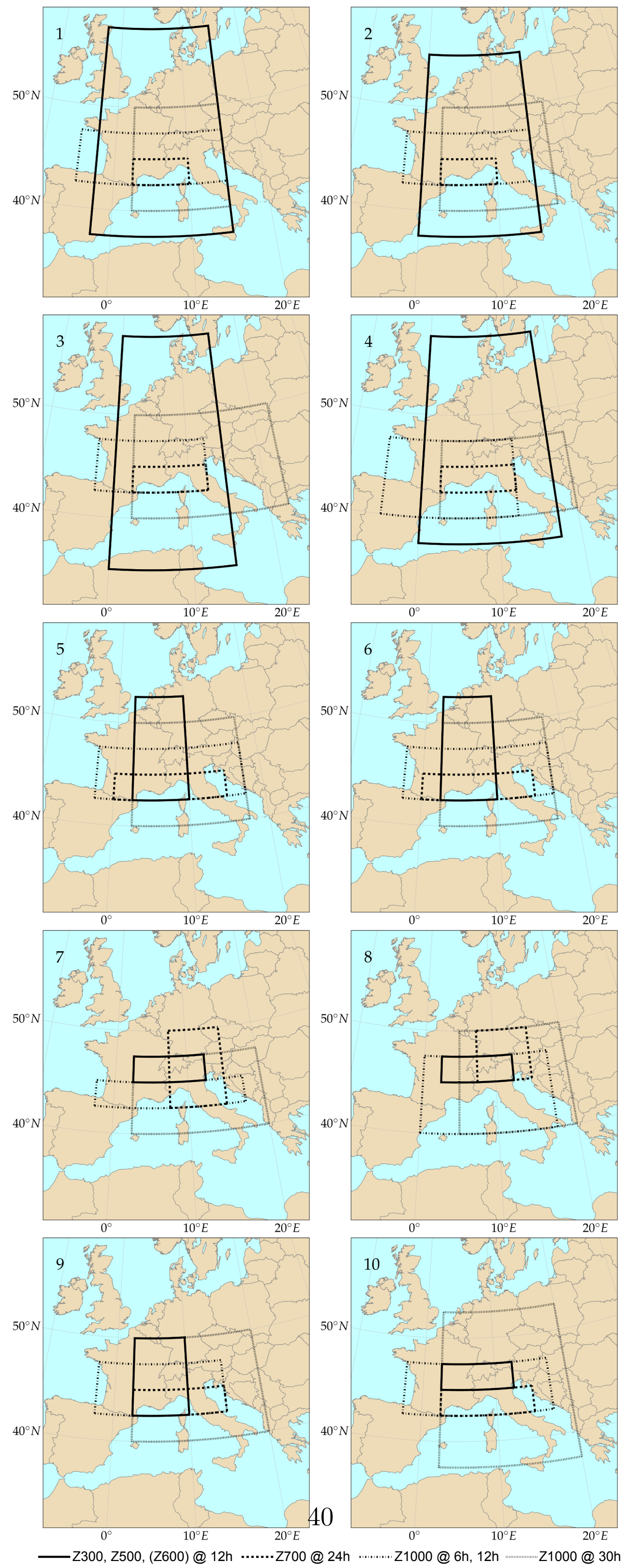


Figure 3: Optimized spatial windows for the 4Zo method (analogy of the atmospheric circulation on ~~4~~ four pressure levels).~~The pressure levels are ordered by increasing pressure and increasing time for the same levels.~~

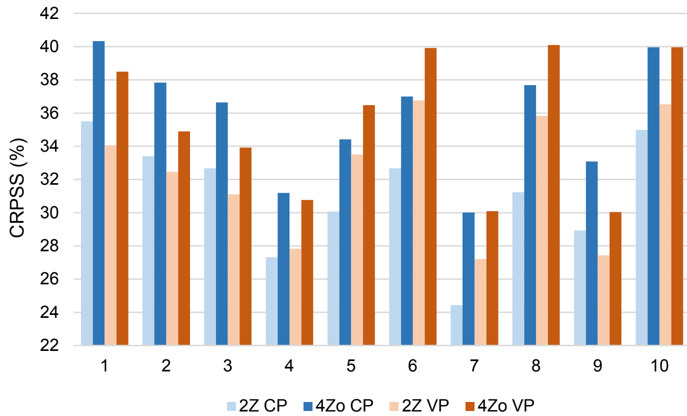


Figure 4: Performance score (CRPSS) of the reference method 2Z (Table 1) and the optimized 4Zo method for CP and VP for every subregion.

Groupment Ids (parameters used)

Groupment Ids (targets)

CP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.3	-0.8	-0.7	-0.7	-0.7	-3.1	-4.4	-1.3	-2.0
2	-0.2	0.0	-0.6	-0.3	-0.6	-0.6	-2.2	-3.6	-0.9	-1.1
3	-0.4	-0.2	0.0	-0.5	-0.4	-0.3	-1.8	-2.9	-0.1	-0.5
4	-0.3	-0.2	-0.4	0.0	-0.3	-0.2	-1.2	-2.6	-0.4	-0.6
5	-0.6	-0.5	-0.8	-0.5	0.0	0.0	-1.2	-2.3	-0.5	-0.7
6	-1.2	-0.8	-1.1	-0.8	0.0	0.0	-0.8	-1.6	-0.3	-0.4
7	-3.6	-3.1	-2.6	-2.5	-1.5	-1.5	0.0	-0.7	-1.1	-1.1
8	-6.4	-5.3	-5.0	-4.6	-2.8	-2.8	-0.2	0.0	-2.7	-1.6
9	-0.9	-0.7	-0.5	-0.6	-0.4	-0.3	-0.8	-1.7	0.0	-0.3
10	-1.9	-1.4	-1.2	-1.6	-0.6	-0.5	-0.9	-1.1	-0.5	0.0

Groupment Ids (targets)

VP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.3	-0.5	0.1	-0.8	-0.9	-3.2	-4.7	-0.9	-1.4
2	-0.1	0.0	0.0	0.2	-0.6	-0.6	-2.1	-3.8	-0.3	-0.7
3	-0.4	-0.6	0.0	-0.3	-1.1	-1.1	-2.1	-4.4	-0.3	-1.1
4	-0.2	0.0	-0.2	0.0	-0.5	-0.6	-2.0	-3.6	-0.7	-1.0
5	0.2	0.2	0.2	0.3	0.0	0.0	-1.8	-3.2	0.0	-0.5
6	-0.5	-0.4	-0.6	0.0	0.0	0.0	-0.9	-1.9	-0.5	-0.4
7	-1.4	-1.1	-0.6	-0.6	-0.6	-0.6	0.0	-1.0	-0.5	0.0
8	-3.6	-3.0	-2.2	-1.9	-1.3	-1.3	0.1	0.0	-1.4	0.2
9	0.2	0.0	0.2	0.4	-0.2	-0.2	-0.7	-2.2	0.0	0.1
10	-0.4	-0.3	-0.1	-0.2	0.1	0.1	-0.8	-1.1	-0.1	0.0

Figure 5: Losses or gains (in %) ~~of the~~ in CRPSS ~~by~~ from applying ~~the~~ optimized parameters for the series in ~~column~~ columns to those in ~~line~~rows. Method 4Zo, calibration ~~period~~and validation periods.

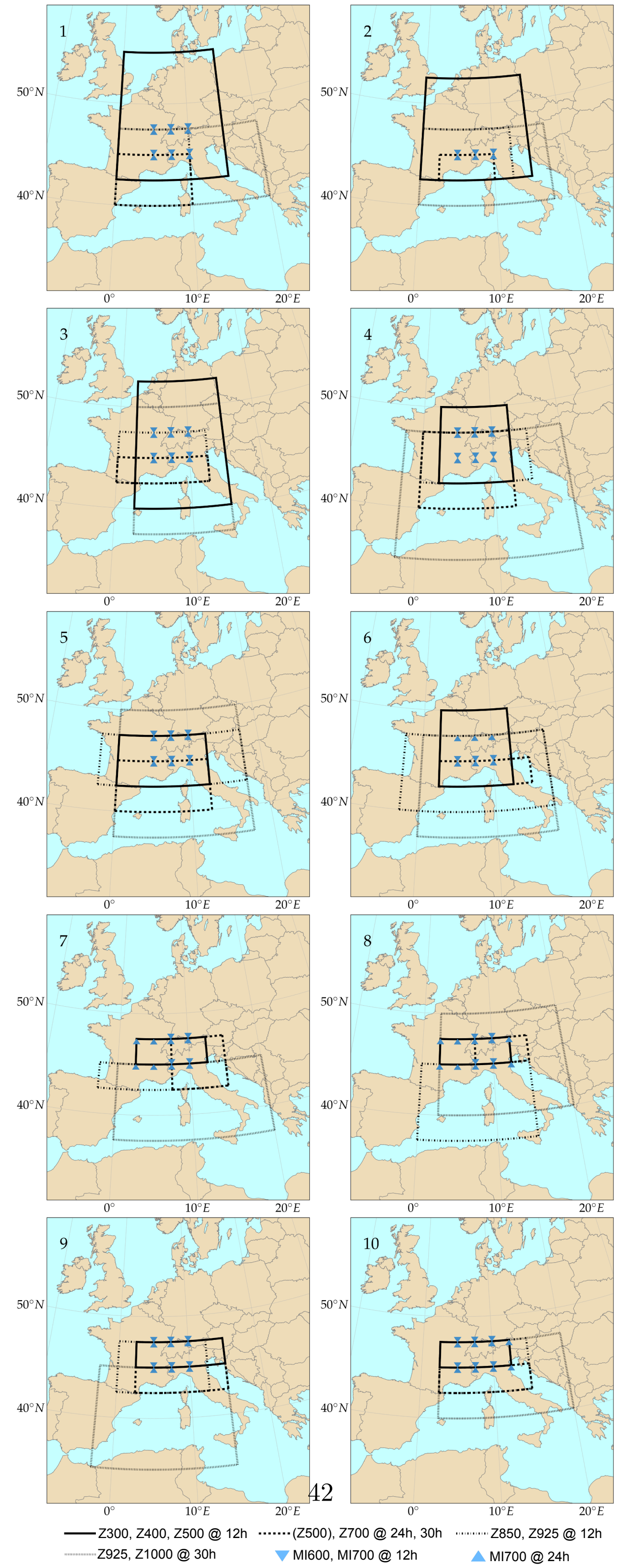


Figure 6: Optimized spatial windows for the 4Zo-2MIo method (analogy of atmospheric circulation on four pressure levels and analogy on the moisture index on two pressure levels).

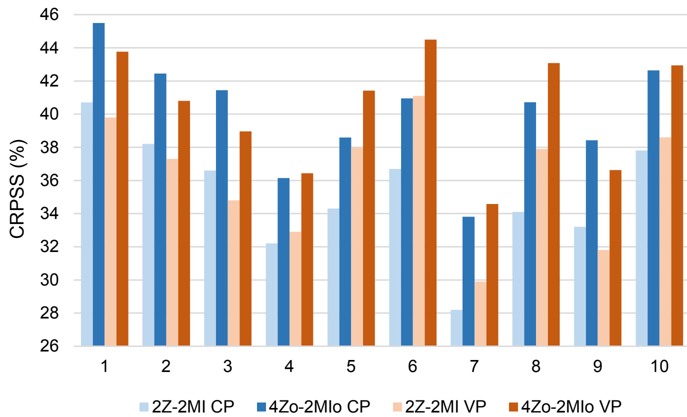


Figure 7: Performance score (CRPSS) of the reference method 2Z-2MI (Table 2) and the optimized 4Zo-2MIo method for CP and VP for every subregion.

Groupment Ids (parameters used)

CP — Groupment Ids (targets)	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.6	-0.9	-1.4	-0.9	-1.8	-3.9	-5.4	-2.2	-2.7
2	-0.9	0.0	-0.5	-0.9	-0.6	-1.2	-2.5	-4.1	-1.4	-2.2
3	-1.3	-0.7	0.0	-0.6	-0.8	-1.3	-2.0	-3.7	-0.6	-1.7
4	-1.2	-0.3	-0.3	0.0	-0.2	-0.8	-1.3	-2.8	-0.3	-1.6
5	-0.8	-0.5	-0.4	-0.3	0.0	-0.5	-1.0	-2.5	-0.4	-1.1
6	-1.5	-1.0	-1.0	-0.7	-0.6	0.0	-0.8	-1.7	-0.7	-0.9
7	-4.6	-3.3	-3.3	-2.1	-2.7	-1.3	0.0	-1.2	-1.3	-1.5
8	-7.5	-5.9	-5.7	-4.2	-5.1	-2.5	-0.7	0.0	-2.8	-1.8
9	-2.0	-1.4	-0.9	-0.6	-1.0	-1.2	-1.0	-2.5	0.0	-1.5
10	-2.5	-1.9	-1.4	-1.6	-1.6	-0.6	-0.7	-1.3	-0.9	0.0

VP — Groupment Ids (targets)	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.2	0.1	-1.0	-0.3	-1.4	-3.3	-6.1	-1.4	-2.6
2	-0.9	0.0	-0.3	-1.0	-0.9	-1.2	-2.9	-5.9	-1.2	-2.5
3	-1.3	-0.7	0.0	-1.3	-0.6	-1.1	-2.5	-6.0	-0.8	-2.0
4	-0.9	0.2	-0.2	0.0	-0.3	-0.1	-1.8	-5.1	0.1	-1.6
5	-0.5	0.1	-0.1	-0.4	0.0	0.0	-1.1	-3.9	0.0	-1.1
6	-1.2	-0.7	-0.9	-1.2	-0.7	0.0	-0.9	-3.0	-0.9	-1.0
7	-3.4	-2.1	-2.7	-1.5	-1.7	-0.5	0.0	-2.4	-0.4	-0.8
8	-5.3	-3.3	-3.9	-2.7	-3.0	-0.6	-0.1	0.0	-1.1	-0.7
9	-2.0	-1.3	-0.7	-0.8	-1.2	-1.3	-1.6	-4.9	0.0	-2.2
10	-1.8	-0.9	-0.5	-1.3	-0.6	0.2	-0.5	-2.4	-0.6	0.0

Figure 8: Losses or gains (in %) of the CRPSS from applying optimized parameters for the series in columns to those in rows. Method 4Zo-2MIo, calibration and validation periods.

Groupment Ids (parameters used)

CP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.6	-0.9	-1.4	-0.9	-1.8	-3.9	-5.4	-2.2	-2.7
2	-0.9	0.0	-0.5	-0.9	-0.6	-1.2	-2.5	-4.1	-1.4	-2.2
3	-1.3	-0.7	0.0	-0.6	-0.8	-1.3	-2.0	-3.7	-0.6	-1.7
4	-1.2	-0.3	-0.3	0.0	-0.2	-0.8	-1.3	-2.8	-0.3	-1.6
5	-0.8	-0.5	-0.4	-0.3	0.0	-0.5	-1.0	-2.5	-0.4	-1.1
6	-1.5	-1.0	-1.0	-0.7	-0.6	0.0	-0.8	-1.7	-0.7	-0.9
7	-4.6	-3.3	-3.3	-2.1	-2.7	-1.3	0.0	-1.2	-1.3	-1.5
8	-7.5	-5.9	-5.7	-4.2	-5.1	-2.5	-0.7	0.0	-2.8	-1.8
9	-2.0	-1.4	-0.9	-0.6	-1.0	-1.2	-1.0	-2.5	0.0	-1.5
10	-2.5	-1.9	-1.4	-1.6	-1.6	-0.6	-0.7	-1.3	-0.9	0.0
VP	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0.0	-0.2	0.1	-1.0	-0.3	-1.4	-3.3	-6.1	-1.4	-2.6
2	-0.9	0.0	-0.3	-1.0	-0.9	-1.2	-2.9	-5.9	-1.2	-2.5
3	-1.3	-0.7	0.0	-1.3	-0.6	-1.1	-2.5	-6.0	-0.8	-2.0
4	-0.9	0.2	-0.2	0.0	-0.3	-0.1	-1.8	-5.1	0.1	-1.6
5	-0.5	0.1	-0.1	-0.4	0.0	0.0	-1.1	-3.9	0.0	-1.1
6	-1.2	-0.7	-0.9	-1.2	-0.7	0.0	-0.9	-3.0	-0.9	-1.0
7	-3.4	-2.1	-2.7	-1.5	-1.7	-0.5	0.0	-2.4	-0.4	-0.8
8	-5.3	-3.3	-3.9	-2.7	-3.0	-0.6	-0.1	0.0	-1.1	-0.7
9	-2.0	-1.3	-0.7	-0.8	-1.2	-1.3	-1.6	-4.9	0.0	-2.2
10	-1.8	-0.9	-0.5	-1.3	-0.6	0.2	-0.5	-2.4	-0.6	0.0

Groupment Ids (targets)

Figure 9: ~~Changes~~ Losses or gains (in ~~the CRPS components~~ %) of ~~4Zo-2MIo relatively to~~ the ~~total CRPS of the reference methods 2Z-2MI (Table 2)~~ CRPSS from applying optimized parameters for the ~~CP~~ series in columns to those in rows. Method 4Zo-2MIo, calibration and ~~the VP~~validation periods.

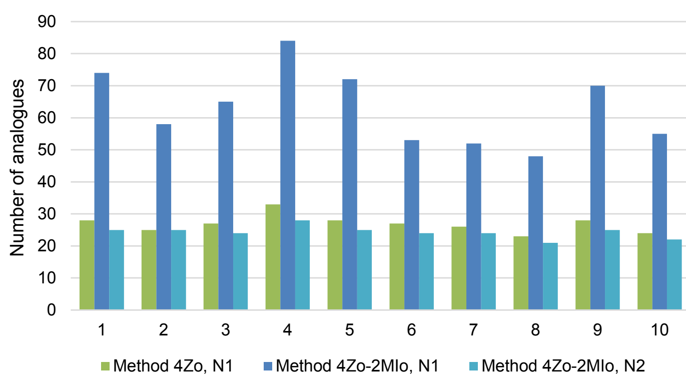


Figure 10: ~~Number~~ Optimal numbers of analogues for the different regions and the ~~various~~ two methods, resulting from ~~the~~ optimization. Method 4Zo is made of a single level of analogy with N_1 analogues, whereas 4Zo-2MIo has two levels of analogy with ~~respectively~~ N_1 and N_2 analogues.

~~Relationships between the optimal number of analogues for both levels of the 4Zo-2MIo method and the corresponding ones for the unique level of the 4Zo method.~~

~~Losses or gains (in %) of the CRPSS by applying the optimized parameters for the series in column to those in line. Method 4Zo-2MIo, calibration period.~~

~~Same as Figure **??** but for the validation period.~~

~~Optimized weighting for the pressure levels of the 4Zo method.~~

~~Averaged weighting for the pressure levels of the circulation analogy of the three methods.~~

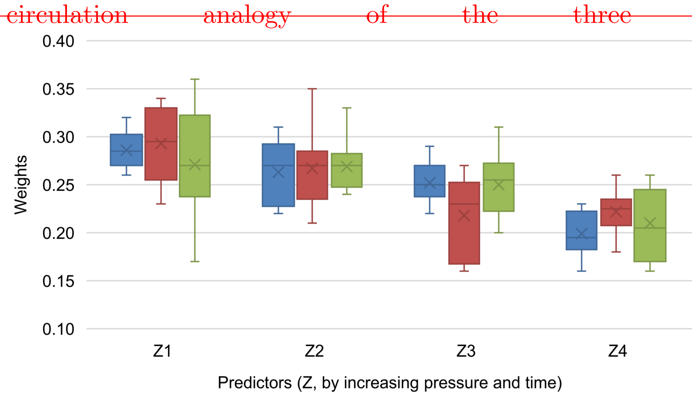


Figure 11: Distribution of optimal weights for the predictors of the first level of analogy (geopotential heights) of (blue) 4Zo, (red) 4Zo-2MIo, and (green) 4Zo-4MIo methods. Results are for the ten subregions. Geopotential heights are sorted by increasing pressure and hour.

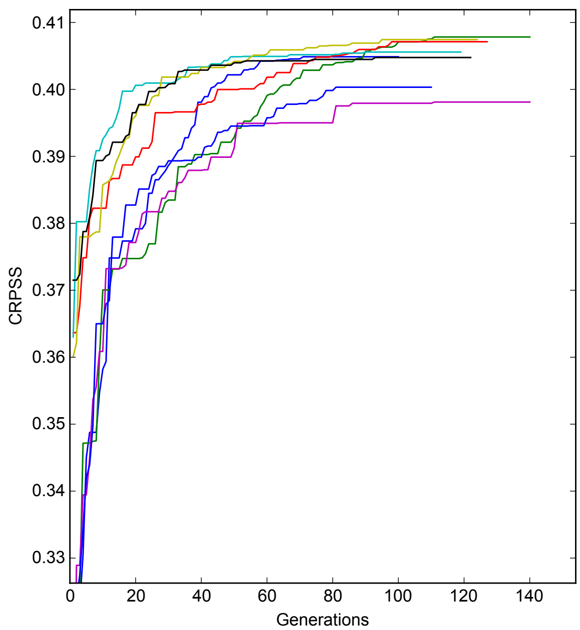


Figure 12: Example of ~~the~~ evolution of the performance score of the best individual over ~~8~~ eight independent optimizations.

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

Level	Variable	Hour	~~Criteria~~ Criterion	Nb
0	±60 days around the target date			
1	Z1000	12 h	S1	50
	Z500	24 h		

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

Level	Variable	Hour	~~Criteria~~ Criterion	Nb
0	±60 days around the target date			
1	Z1000	12 h	S1	70
	Z500	24 h		
2	TPW * RH850	12 h	RMSE	30
	TPW * RH850	24 h		

Table 3: ~~Parameters of~~ Pressure levels (∼) automatically selected for the 4Zo method for different subregions (~~analogy on 4 levels of the atmospheric circulation~~ID)~~optimized for the Chablais subregion~~. ~~The columns are~~ R represents the ~~following: L = level of analogy, V = meteorological variable, H = hour of observation or temporal window, D = domain or spatial window, W = weighting of the score for the selected pressure levels, C = criteria, N = number of analogues~~2Z reference method (Table 1).

~~L V H D W C N 0 -2.5 - 15.0 E 37.5 - 57.5 N 2.5 - 10.0 E 42.5 - 45.0 N -5.0 - 15.0 E 42.5 - 47.5 N 2.5 - 15.0 E 40.0 - 50.0 N~~

~~Pressure levels (∼) automatically selected for the 4Zo method at the different subregions. R represents the reference method (Table 1).~~

ID	300	400	500	600	700	850	925	1000
1	∼				∼			∼∼
2	∼				∼			∼∼
3	∼				∼			∼∼
4	∼				∼			∼∼
5			∼		∼			∼∼
6			∼		∼			∼∼
7			∼		∼			∼∼
8			∼		∼			∼∼
9			∼		∼			∼∼
10				∼	∼			∼∼
R			∼					∼

Table 4: ~~CRPSS score~~ Relative improvement (%) ~~of~~ in CRPSS for different precipitation thresholds for the ~~three~~ optimized ~~methods (~~4Zo ~~is the atmospheric circulation analogy detailed in the present section; 4Zo-2MIo and 4Zo-4MIo add a second level of analogy on moisture indexes~~method, ~~as explained in section 4)~~compared to the reference method.

ID	P≥1 mm		P≥0.1·P10		P≥0.5·P10	
	CP	VP	CP	VP	CP	VP
1	10.2	9.4	8.5	7.9	17.0	14.2
2	9.9	3.4	10.2	7.3	19.3	13.7
3	13.3	10.5	13.3	10.9	19.7	9.7
4	11.0	7.4	12.9	10.0	23.2	23.8
5	8.6	4.2	10.9	6.2	25.2	23.8
6	10.5	5.1	11.1	7.1	21.2	41.1
7	24.3	12.4	33.1	26.0	71.2	104.3
8	19.0	12.7	26.2	19.2	39.4	34.9
9	12.4	6.8	13.8	9.9	24.9	48.1
10	13.6	6.8	14.4	6.9	29.9	31.5
av.	13.3	7.9	15.4	11.1	29.1	34.5

Table 5: ~~Parameters of~~ Atmospheric levels automatically selected for the ~~4Zo-2MIo method~~ ~~(~~analogy of atmospheric circulation ~~on 4 levels~~ (∼) and moisture ~~index on 2 levels~~analogy (•) of the 4Zo-2MIo method, ~~optimized~~ for ~~the Chablais subregion~~different subregions (ID). ~~Same conventions as~~ R represents the 2Z-2MI reference method (Table ~~??.~~2)

~~L V H D W C N 0 0.0 - 15.0 E 42.5 - 55.0 N 0.0 - 10.0 E 40.0 - 45.0 N 0.0 - 10.0 E 42.5 - 47.5 N 0.0 - 20.0 E 40.0 - 47.5 N TPW 5.0 - 10.0 E *RH700 45.0 - 47.5 N TPW 5.0 - 10.0 E *RH700 45.0 - 47.5 N~~

~~Atmospheric levels automatically selected for the analogy of the atmospheric circulation (∼) and the analogy of moisture (•) of the 4Zo-2MIo method, at the different subregions. R represents the reference method (Table 2)~~

ID	300	400	500	600	700	850	925	1000
1	∼		∼		••	∼		∼
2	∼				∼ ••	∼		∼
3	∼				∼ ••	∼	∼	
4			∼	•	∼ •	∼		∼
5		∼			∼ ••		∼∼	
6		∼		•	∼ •	∼		∼
7		∼		•	∼ •	∼		∼
8			∼	•	∼ •		∼∼	
9		∼		•	∼ •	∼	∼	
10		∼		•	∼ •	∼		∼
R			∼			••		∼

Table 6: ~~Improvement~~ Relative improvement (%) ~~of the~~ in CRPSS for different precipitations thresholds for the optimized 4Zo-2MIo method, compared to the reference method.

ID	P≥1 mm		P≥0.1·P10		P≥0.5·P10	
	CP	VP	CP	VP	CP	VP
1	12.6	9.3	12.4	9.7	15.8	11.0
2	10.4	7.7	11.2	10.5	18.9	16.6
3	14.5	11.6	14.1	11.4	18.7	14.6
4	11.4	9.4	11.5	11.6	14.9	22.7
5	11.8	8.0	12.2	8.9	12.0	12.8
6	11.3	7.1	11.2	8.0	15.3	29.1
7	20.5	15.5	25.2	24.0	43.0	79.5
8	19.3	15.7	23.1	18.6	25.2	31.7
9	17.0	15.4	17.4	16.5	23.7	39.4
10	12.9	9.6	13.8	11.1	28.5	32.1
av.	14.2	10.9	15.2	13.0	21.6	28.9

Using Genetic Algorithms to Optimize the Analogue Method for Precipitation Prediction in the Swiss Alps

Pascal Horton[a,b,1,*], Michel Jaboyedoff[a], Charles Obled[c]

[a] *University of Lausanne, Institute of Earth Sciences, Lausanne, Switzerland*
[b] *University of Bern, Oeschger Centre for Climate Change Research, Institute of Geography, Bern, Switzerland*
[c] *Université de Grenoble-Alpes, LTHE, Grenoble, France*

Abstract

Analogue methods provide a statistical precipitation prediction based on synoptic predictors supplied by general circulation models or numerical weather prediction models. The method samples a selection of days in the archives that are similar to the target day to be predicted, and consider their set of corresponding observed precipitation (the predictand) as the conditional distribution for the target day. The relationship between the predictors and predictands relies on some parameters that characterize how and where the similarity between two atmospheric situations is defined.

This relationship is usually established by a semi-automatic sequential procedure that has strong limitations: (i) it cannot automatically choose the pressure levels and temporal windows for a given meteorological variable, (ii) it cannot handle dependencies between parameters, and (iii) it cannot easily handle new degrees of freedom. In this work, a global optimization approach relying on genetic algorithms was able to optimize all parameters jointly and automatically. It allowed consideration of parameter inter-dependencies, and objective selection of some parameters that were manually selected beforehand, which obviates the need to assess a large number of combinations of pressure levels and temporal windows of predictor variables.

*Corresponding author
Email address: pascal.horton@alumnil.unil.ch (Pascal Horton)

The global optimization was applied to some variants of the analogue method for the Rhône catchment in the Swiss Alps. The performance scores increased compared to reference methods, especially for days with high precipitation totals. The resulting parameters were found to be relevant and coherent between the different subregions of the catchment. Moreover, they were obtained automatically and objectively, which reduces the effort that needs to be invested in exploration attempts when adapting the method to a new region or for a new predictand. In addition, the approach allowed for new degrees of freedom, such as a possible weighting between pressure levels, and non-overlapping spatial windows.

Keywords: precipitation prediction, analogue method, optimization, genetic algorithms, Alpine climate

---

1. Introduction

The analogue method (AM) is a downscaling technique based on the idea expressed by Lorenz (1956, 1969) that similar situations in terms of atmospheric circulation are likely to lead to similar local weather (Duband, 1970). It uses predictor variables describing the synoptic atmospheric circulation in order to predict local-scale predictands of interest. It is often used to predict daily precipitation, either in an operational forecasting context (e.g. Guilbaud, 1997; Bontron and Obled, 2005; Hamill and Whitaker, 2006; Bliefernicht, 2010; Marty et al., 2012; Horton et al., 2012, 2016b; Hamill et al., 2015; Ben Daoud et al., 2016) or a climate downscaling context (e.g. Radanovics et al., 2013; Chardon et al., 2014; Dayon et al., 2015; Raynaud et al., 2016). Other predictands are also considered, such as precipitation radar images (Panziera et al., 2011; Foresti et al., 2015), temperature (Radinovic, 1975; Woodcock, 1980; Kruizinga and Murphy, 1983; Delle Monache et al., 2013; Caillouet et al., 2016; Raynaud et al., 2016), wind (Gordon, 1987; Delle Monache et al., 2013, 2011; Vanvyve et al., 2015; Alessandrini et al., 2015b; Junk et al., 2015b,a), solar power (Alessandrini et al., 2015a; Bessa et al., 2015), snow avalanches (Obled and Good, 1980; Bolognesi, 1993), and radiation (Bois et al., 1981; Raynaud et al., 2016).

In real-time forecasting, it is used mainly by practitioners, notably hydropower companies or flood forecasting services, that need to anticipate water yields or issue early flood warnings several days in advance. The classical forecasting chain consists of using limited area models (e.g. AROME, or COSMO) forced by global NWP (numerical weather prediction) models with a lower resolution. However, their use requires very important processing capacities, and the resulting forecast still presents large uncertainties and biases. Although these outputs are essential, they can be supplemented by other sources of forecasts providing useful information. In contrast to local NWP models, AMs can transform at low cost the synoptic-scale information provided by the global NWP model into precipitation forecasts, by using the natural local behaviour in response to synoptic-scale influences stored in the archive of observed precipitation. Running an AM approach is fast enough that it can search for analogues for each day, up to ten days ahead, eventually for the different traces of an ensemble forecast and/or those issued by different NWP models (e.g. NOAA-GFS or ECMWF-IFS), in a matter of minutes.

In climate studies, AMs are used to downscale the outputs of a general circulation model (GCM) or regional climate model (RCM) simulation runs (Dayon et al., 2015) or to reconstruct past weather conditions (Caillouet et al., 2016). In future climate studies, RCMs are often used to dynamically downscale precipitation to a local scale. However, even though the relevance of RCMs' outputs increases, a bias correction of the outputs is often still required, particularly in complex terrain. Moreover, their application is computer-intensive, which makes it difficult to cover all combinations of climate scenarios and GCMs. Therefore, the idea is to bypass the small-scale simulations and to go from the large-scale situation to the end variables such as precipitation by statistical downscaling (Maraun et al., 2010).

Beyond being computationally inexpensive, another big advantage of AMs is that they create realistic precipitation patterns for a region, provided that the analogue dates are the same, since they are based on observed situations with consistent spatial distribution (Radanovics et al., 2013; Chardon et al., 2014). For the same reason, they can also provide multivariate predictions that are physically consistent (Raynaud et al., 2016).

The method can be designed with multiple successive subsampling steps, or analogy levels, each of them relying on different meteorological variables. A certain number of parameters define the relationship between predictors and predictands, such as the choice of the predictor variable, its pressure level and temporal window to consider, the spatial domain to use for the comparison, as well as the analogy criterion itself, and finally, the number of analogue situations to keep at each subsampling level. These parameters are usually calibrated by means of a semi-automatic sequential procedure (Bontron, 2004; Radanovics et al., 2013), i.e. by optimizing each single parameter, one at a time, in an arbitrarily chosen order, with no or little reassessment. This sequential approach therefore has strong limitations: (i) it cannot automatically choose the optimal pressure levels and the temporal windows for a given meteorological variable, (ii) it cannot handle dependencies between the parameters within a level of analogy, and even less between them, and (iii) it cannot easily handle new degrees of freedom, such as a possible weighting between the pressure levels. Thus, even if the processing involved is relatively fast, the sequential approach requires laborious assessments of predictor combinations (variables, pressure levels, temporal windows), and presents a high risk of ending in a local optimum because of subjective initial choices and lack of consideration of parameter interdependencies. Other calibration methods exist for specific applications, such as radar images (Panziera et al., 2011; Foresti et al., 2015).

Aiming to overcome these limitations, a global optimization by genetic algorithms (GAs) was introduced. An intensive assessment resulted in recommendations to parametrize GAs in order to optimize AMs successfully (Horton et al., 2016a). The present paper is based on these recommendations, and applies them to precipitation prediction for the upper Rhône catchment in the Swiss Alps, using AMs of varying complexity. It aims at illustrating the relevance of a fully automatic, objective, and global, optimization technique for AMs. The applications are indeed numerous, as AMs have to be adapted to every new location they are applied, or to any new predictand they should predict.

The data, AMs, and optimization techniques (sequential and GAs) are presented in Section 2. The results are first given for the optimization of the analogy of atmospheric circulation only (Section 3), before being extended to a method adding a second level of analogy on moisture variables (Section 4). General discussions (Section 5) and conclusions (Section 6) follow.

2. Data and methods

2.1. Case study description

The study area is the alpine upper Rhône catchment in Switzerland (Fig. 1). The altitude ranges from 372 to 4634 m.a.s.l. and the area is 5524 km^2. This region is the target of the MINERVE (Modélisation des Intempéries de Nature Extrême sur les Rivières Valaisannes et de leurs Effets) project, which aimed at real-time flood management on the upper Rhône catchment (García Hernández et al., 2009). Even though the region is rather small, the meteorological influences related to extreme weather conditions vary substantially within it (see Horton et al., 2012). Indeed, a high spatial variability of precipitation climatology exists, which is due to the complex orography of the region, and the mix of various meteorological influences. Based on different climatological analyses, the precipitation gauge stations in the catchment were clustered in ten subregions (Fig. 1):

1. Swiss Chablais
2. Trient Valley
3. West Bernese Alps
4. Lower Rhone Valley
5. Southern valleys
6. Southern ridges
7. Upper Rhone Valley

8. Southeast ridges
9. East Bernese Alps
10. Conches Valley

2.2. Data

AMs rely on two types of data: predictors, which are atmospheric variables describing the state of the atmosphere at a synoptic scale, and the predictand, which is the local weather variable one wants to predict.

Predictors are generally extracted from reanalysis datasets. The NCEP-NCAR reanalysis I (6-hourly, 17 pressure levels at a resolution of 2.5°, see Kalnay et al., 1996) was used here, but it could have been any other reanalysis dataset.

The predictand (which is to be predicted) is here the daily precipitation (6 a.m. to 6 a.m. the next day) measured at the MeteoSwiss network stations, for the period 1961–2008. The time series from every available gauge station were averaged over the ten subregions (Fig. 1), which were approximately 500 km^2 each, in order to smooth local effects (Obled et al., 2002; Marty et al., 2012). This helps account for local variability, mainly when convective processes are involved, which slightly increases the prediction skill.

It must be stressed that the predictand here is a temporally cumulated variable, compared to the meteorological predictors, which may be considered instantaneous. Depending on the duration of the accumulation period (here 24 h, but could have been 6 h, 12 h, or more than 24 h), the choice of predictors will vary.

The 48-yr precipitation dataset was divided into a calibration period (CP) and a validation period (VP). Using data independent of the CP to validate the results is very important in order to assess the robustness of the proposed improvements and to avoid over-parametrization of the method.

In order to reduce potential biases related to trends linked to climate change or to the evolution in measurement techniques, the selection of the VP was evenly distributed over the entire series (Ben Daoud, 2010). Thus, one out of every six years was selected for validation, which represents a total of 8 years for the VP and 40 for the CP. This choice of sequence was made in order to have similar statistical characteristics between the CP and VP.

2.3. The analogue method

Multiple variations of the analogue method exist, most of which are not detailed here (see Ben Daoud et al., 2016, for a more comprehensive listing). However, there are mainly two parameterizations that are most often used for precipitation prediction and that are considered as reference: one that relies on an analogy of the atmospheric circulation, and another that adds a second level of analogy on moisture variables (Obled et al., 2002; Bontron and Obled, 2005; Marty et al., 2012).

The method based on the analogy of synoptic circulation consists of the following steps (Table 1): the similarity of the atmospheric circulation of a target date with every day of the archive is assessed by processing the S1 criterion (Eq. 1, Teweles and Wobus, 1954; Drosdowsky and Zhang, 2003), which is a comparison of gradients, over a certain spatial window:

$$S1 = 100 \frac{\sum_i |\Delta \hat{z}_i - \Delta z_i|}{\sum_i max\left\{|\Delta \hat{z}_i|, |\Delta z_i|\right\}} \tag{1}$$

where $\Delta \hat{z}_i$ is the difference in geopotential height between the i-th pair of adjacent points of gridded data describing the target situation, and Δz_i is the corresponding observed geopotential height difference in the candidate situation. The differences are processed separately in both North and East directions over the selected spatial domain. The smaller the S1 values, the more similar the pressure fields.

Bontron and Obled (2005) show that the geopotential height at 500 hPa (Z500) and 1000 hPa (Z1000) are the best first predictors of the NCEP/NCAR reanalysis I dataset, and that the S1 criterion performs better than scores based on absolute distances. The reason for such better results is that the S1 criterion allows comparison of the circulation patterns, by means of the gradients, rather than the absolute value of the geopotential height, which better represent the flow direction. To cope with seasonal effects, candidate dates are extracted within a period of four months centred around the target date, for every year of the archive. This method using two geopotential heights is named here 2Z.

The N_1 dates with the lowest values of S1 are considered as analogues to the target day. The number of analogues, N_1, is a parameter to calibrate. Then, the daily observed precipitation amount for the N_1 resulting dates provide the empirical conditional distribution, considered as the probabilistic prediction for the target day.

The other most well-known parametrization adds a second level of analogy on the moisture variables (method 2Z-2MI, Table 2). The predictor that Bontron (2004) found optimal for France is a moisture index made of the product of the total precipitable water (TPW) with the relative humidity at 850 hPa (RH850). Horton (2012) confirms that this index is also better for the Swiss Alps than any other variable from the NCEP/NCAR reanalysis I considered independently. When adding a second level of analogy, N_2 dates are subsampled within the N_1 analogues of the atmospheric circulation, to end up with a smaller number of analogue situations. When this second level of analogy is added, a higher number of analogues N_1 is kept on the first level. Moisture fields are not as well-predicted by NWP models as pressure variables. This implies that the 2Z-2MI method, when used in real-time forecasting, is very dependent on the skill of the NWP model in predicting moisture fields, and thus its use is often restricted to the first lead times.

2.4. Performance assessment

The performance assessment in the present context consists of verifying the prediction of an ensemble probabilistic technique. The set of precipitation values collected with each analogue can be considered as a sample drawn from the conditional distribution associated with the current circulation. The score that is most often used to assess an AM performance is the CRPS (Continuous

Ranked Probability Score, Brown, 1974; Matheson and Winkler, 1976; Hersbach, 2000). It allows evaluating the predicted cumulative distribution functions $F(y)$, for example, of the precipitation values y from analogue situations, compared to the observed value y^0. The better the prediction, the smaller the score. The mean CRPS of a prediction series of length n can be written as:

$$CRPS = \frac{1}{n} \sum_{i=1}^{n} \left(\int_{-\infty}^{+\infty} \left[F_i(y) - H_i(y - y_i^0) \right]^2 dy \right) \tag{2}$$

where $H(y - y_i^0)$ is the Heaviside function that is null when $y - y_i^0 < 0$, and has the value 1 otherwise.

In order to compare the value of the score relative to a reference, one often considers its skill score expression, and uses the climatological distribution of precipitation from the entire archive as the reference. The CRPSS (Continuous Ranked Probability Skill Score) is thus defined as follows:

$$CRPSS = \frac{CRPS - CRPS_r}{CRPS_p - CRPS_r} = 1 - \frac{CRPS}{CRPS_r} \tag{3}$$

where $CRPS_r$ is the CRPS value for the reference and $CRPS_p$ would be the one for a perfect prediction (which implies $CRPS_p = 0$). A better prediction is characterized by an increase in CRPSS.

Note, however, that the choice of reference does not matter so much when assessing potential improvements of the method, since we consider more its relative increase or decrease rather than the CRPSS absolute value.

2.5. Sequential calibration

AMs are usually calibrated by a semi-automatic sequential procedure, as elaborated by Bontron (2004) (see also Radanovics et al., 2013; Ben Daoud et al., 2016). The calibration technique optimizes the spatial windows in which the predictors are compared and the number of analogues for every level of analogy, by maximizing the performance score (CRPSS). However, the different analogy levels are calibrated sequentially, and the meteorological variables, pressure levels, and temporal windows are chosen manually. The procedure, as defined by Bontron (2004), consists of the following steps:

1. Manual selection of the following parameters:
   (a) Meteorological variable
   (b) Pressure level
   (c) Temporal window (hour of the day)
   (d) Initial analogue numbers
2. For every level of analogy:
   (a) Identification, for the analogy level considered, of the most skilled unitary cell of all predictors jointly, over a large domain, by a full scanning of the grid.
   (b) From this most skilled cell, the spatial window is expanded by successive iterations in the direction of greater performance gain until no improvement is reached.
   (c) The number of analogue situations N_1 is then reconsidered and optimized for the current level of analogy.
3. A new level of analogy can then be added, based on other variables (such as the moisture index) with some chosen pressure levels, temporal windows, and initial number of analogues N_2. The procedure starts again from step 2 (calibration of the spatial window and the number of analogues) for the new level. The parameters calibrated on the previous analogy levels are fixed and do not change.
4. Finally, the numbers of analogues N_1 and N_2 for the different levels of analogy are reassessed by systematic increments.

The calibration is done in successive steps with a limited number of parameters. Previously calibrated parameters are generally not reassessed (except for the number of analogues).

This procedure was used to calibrate the methods that were here considered as references to further assess the ability of genetic algorithms to outperform the classic approach.

2.6. Genetic algorithms

Genetic algorithms (GAs) were developed by Holland (1992) and Goldberg (1989). They are part of the Evolutionary Algorithms (Bäck and Schwefel, 1993; Schwefel, 1993), which were inspired by some mechanisms in biological evolution, such as reproduction, genetic mutations, chromosomal crossovers, and natural selection. GAs seek the global optimum on a complex surface, theoretically without restriction, which is of interest for AMs, which are characterized by a complex high-dimensional error function having multiple local optima. Practically, GAs allow rapidly approaching satisfactory solutions, but they are not guaranteed to provide the optimum solution (Zitzler et al., 2004). It is indeed mainly a matter of time. When the optimizer gets closer to the global optimum, any new improvement takes more time to appear, and the final adjustment of the parameters can be very time consuming (Bäck, 1993). For problems that require a significant amount of time to evaluate the objective function, as in the case of AMs (because it needs to make a prediction for every day of the CP), the number of generations has to be limited in order to ensure a reasonable processing time. Thus, different acceptable solutions can result from one or more optimization runs (Holland, 1992). This is both a strength and a weakness of GAs: they are very good at exploring complex parameter spaces in order to identify the most promising areas, but they will not necessarily always find the best solution with the optimal values for all parameters (Holland, 1992).

The optimizations here were performed based on the recommended GA parametrization for AMs as described in Horton et al. (2016a). As the optimization is mostly sensitive to the mutation operator (that randomly changes some values in the parameter sets), parallel optimizations are considered with variants of this operator, according to Horton et al. (2016a):

- 3x non-uniform mutation (Michalewicz, 1996) with varying parameters
- 1x multi-scale mutation (Horton et al., 2016a)
- 2x chromosome of adaptive search radius (Horton et al., 2016a).

A population size of 500 individuals (i.e. parameter sets of the AM to be detailed below) was considered, and the optimization was stopped when the best individual (with the highest CRPSS performance score) did not evolve for 20 generations (cycles of optimization).

3. Optimization of the circulation analogy

The analogy of the atmospheric circulation was optimized for the ten subregions (Section 2.1) independently. We started from the simplest AM, and increased the complexity in order to identify the degrees of freedom that are of particular interest. Thus, the tested parametrization evolved iteratively in complexity. The detailed results of the intermediate stages are not provided in this paper (see Horton, 2012, for the details).

The reference method for the analogy of the atmospheric circulation (2Z, Table 1), based on Z500 and Z1000, was first considered. The optimizer had to choose simultaneously the number of analogues, both spatial windows with no overlapping constraint (i.e. they can differ from one pressure level to another), as well as the temporal windows (hours of observation of the geopotential), which cannot be achieved with the sequential calibration technique. The performance score (CRPSS) was slightly improved, with these limited degrees of freedom, relative to the 2Z reference method calibrated with the sequential procedure. Some tests showed that most of the gains were due to the non-overlapping spatial windows. This demonstrated that the optimizer was able to obtain relevant parameters for a simple method.

Then, an additional degree of freedom was provided to the GAs by letting them choose the pressure levels along with the other parameters (analogue numbers, spatial and temporal windows), which is also a non-automated process in the sequential calibration. This degree of freedom increased the optimization time, and might decrease the number of simulations that converge to a single solution. However, most solutions were very close in terms of the performance score, which was further improved. The selected pressure levels were Z500 or

Z700 for the upper level, and Z925 or Z1000 (most often) for the lower level.

Parallel analyses showed that the analogy of circulation is incomplete, and that geopotential heights still contain relevant information that can improve the statistical relationship. Therefore, a third, followed by a fourth circulation predictor were added (still only geopotential heights). There was no constraint on the predictors, so that the same pressure level could be selected more than once. Further improvements were found in the performance score, both for the CP and the VP, confirming that this additional information was beneficial for the quality of the prediction.

Finally, a weighting of the analogy criteria values per pressure level was proposed, again optimized by GAs. The weighting operates in the combination of the S1 criteria processed on every level, which were previously averaged with equal weights. The role of this new degree of freedom is to give more weight to the levels with greater predictive capacity, and to consider the differences in the geopotential height variability with altitude.

The number of circulation predictors (still only geopotential heights) was then successively increased up to ten, considering the weighting of the analogy criteria values. The addition of circulation predictors globally improved the prediction skill (for both the CP and the VP) only up to four predictors (Figure 2). Afterwards, the score on the VP was more variable, eventually even showing a decrease, which revealed an over-parametrization of the method, and thus a lack of robustness. After four predictors, the score for the CP did not increase substantially, and even presented a local decrease due to increasing difficulty for the optimizer to converge. Selecting four circulation predictors (geopotential heights) was considered optimal for this case study, since the gain in CRPSS was significant, and the model remained relatively simple. It cannot be ruled out that another number would prevail in a region other than the upper Rhône catchment, under other meteorological conditions, or with another reanalysis dataset.

3.1. Which parameters are optimized?

The chosen method for the atmospheric circulation analogy, based on four circulation predictors (geopotential heights), and which is here named 4Zo (o for optimized), was based on the following degrees of freedom:

- selection of pressure levels (4 degrees)
- temporal windows (4 degrees)
- spatial windows (4 x 4 degrees)
- weights (4 degrees)
- number of analogues (1 degree).

This adds up to 29 degrees of freedom that were optimized simultaneously.

3.2. Results for the 4Zo method

The resulting optimized parameters for 4Zo vary from one subregion to another. The optimized spatial windows are given for every subregion in Figure 3, and the selected pressure levels in Table 3.

The resulting CRPSS scores are provided in Figure 4 and were on average 35.8% for the CP and 35.5% for the VP, compared to 31.1% and 32.3%, respectively, for the reference method 2Z on the atmospheric circulation (optimized by the sequential procedure). The score was also calculated for three precipitation thresholds: $P \geq 1$ mm, $P \geq 0.1 \cdot P10$, and $P \geq 0.5 \cdot P10$, P10 being the daily precipitation with a 10 year return period (Table 4). The gain in score increased with the precipitation threshold: the relative improvement of the CRPSS was, on average, for the different thresholds, 13.3%, 15.4%, and 29.1% for the CP and 7.9%, 11.1%, and 34.5% for the VP. The optimization thus improved the prediction even more for days with significant precipitation than for the usual days.

To assess the parameters cross-compatibility and the spatial coherence of the resulting parameters, those optimized for one subregion were applied to the others. The resulting losses or gains of the CRPSS are displayed in Figure 5.

3.3. Analysis

The automatic selections of pressure levels (Table 3) and temporal windows (not shown) for the analogy of circulation showed a great homogeneity and were spatially consistent. First of all, the level Z1000 was always selected twice (the first time at 6 or 12 h, and the second always at 30 h) and Z700 was selected once for every subregion (always at 24 h). The level that varied from one subregion to another, albeit in a spatially consistent way, was the upper level (always at 12 h), which was Z300 for the north-west part of the catchment, Z500 for most of the other subregions, and Z600 for the Conches Valley. The optimizer thus provided consistent selections of pressure levels and temporal windows. The automatic selection of pressure levels is a big advantage in favour of global optimization.

The resulting spatial windows (Figure 3) may look very diverse first, but there are significant similarities for subregions located within the same vicinity. The first four subregions were characterized by a large spatial window on the upper level, whereas it was smaller elsewhere. For most subregions, the second level (Z700) was compared on thin and longitudinally extended spatial windows. The third level (Z1000 at 6 or 12 h) also had longitudinally extended domains, which were slightly larger. The last one (Z1000 at 30 h) had rather large and squared windows. Subregions number 5 (southern valleys) and 6 (southern ridges) had exactly the same spatial windows, which suggests that they behave in a similar way and thus could have been merged. This similarity is a good sign for the accuracy of the optimized parameters.

The performance scores showed non-negligible improvements for both the CP and VP (Figure 4) compared to the 2Z reference method optimized by the sequential procedure. Even more interestingly, the results for higher precipitation thresholds (Table 4) showed the largest improvements. This is of particular interest in the framework of flood forecasting. The further improvement of days with higher precipitation totals is likely related to the fact that larger values contribute more to the CRPS score, which means that better predicting these days results in significant increase in the global performance score.

The analysis of the parameters cross-compatibility showed that the parameters were obviously optimal on the CP for the subregion for which they were optimized (Figure 5 top). However, the losses in CRPSS when exchanging the parameters were not of the same magnitude among the different subregions. Indeed, the Upper Rhone Valley (7) and, moreover, the southeast ridges (8) seemed to behave significantly differently. These two regions have different climatic properties than the others, as they are particularly sensitive to southerly flows. Indeed, almost all heavy precipitation events occurred under a southerly regime, such as in the Liguria, Piedmont, and Aosta regions in Italy, whereas the other subregions of the catchment had extreme events mainly under a westerly regime (Horton et al., 2012). Thus, as the performance score is significantly influenced by heavy precipitation values, the parameters for the different subregions are likely optimized to better predict these days. It can then be expected that the optimal parameters differ between these two subregions and the others. This points at the importance of taking into account leading meteorological influences during precipitation station clustering, which are not always best represented by geographical distance.

Globally, the same cross-compatibility structure could be observed for the VP (Figure 5 bottom), but in this case, minor improvements were occasionally observed when crossing the parameters, because of the presence of other events in the VP that might be better predicted by a different parameter set. The relatively small differences in scores between parameterizations indicated that even though the parameters might differ significantly, the performance might not be drastically affected. Even a change in the pressure level did not mean a radical drop in the score value. A different parametrization may lead to a distinct selection of analogue days, and thus to an improvement of the prediction under certain weather conditions at the expense of others.

4. Optimization of the analogy with moisture information

It is known that moisture variables as a second level of analogy do provide improvements to the method (section 2.3). The moisture index, which is a combination of the relative humidity and precipitable water, has thus also to be optimized. In order to do so, a constraint on the optimizer had to be introduced, so as to select the same temporal window (time of observation) for both variables.

Two methods were assessed: one with two moisture predictors (moisture index on two pressure levels or at two different hours), named 4Zo-2MIo, and one with four moisture predictors, named 4Zo-4MIo. When introducing two predictors for the moisture analogy, the number of degrees of freedom increased to 42, and to 54 with four predictors. However, there was no substantial difference in the performance scores between both 4Zo-2MIo and 4Zo-4MIo methods, which suggests that considering four moisture predictors is not necessary. For this reason, only the results of 4Zo-2MIo are presented.

The optimization was processed on both levels of analogy simultaneously. This implies that the analogy of the atmospheric circulation could change because of the new moisture information.

4.1. Results for the 4Zo-2MIo method

As seen previously, the optimized parameters differed from one subregion to another, but to an even greater extent. The resulting spatial windows are displayed in Figure 6 for 4Zo-2MIo, along with the selected pressure levels for both the circulation and moisture analogy (Table 5).

The CRPSS scores of the optimized 4Zo-2MIo method are provided in Figure 7 and amounted on average to 40% (CP) and 40.3% (VP), compared to 35.2% (CP) and 36.2% (VP) for the reference method 2Z-2MI on the moisture analogy optimized with the sequential procedure. The parameters cross-compatibilities are shown in Figure 8. As for 4Zo, the 4Zo-2MIo method presented larger improvements in the prediction of significant rainfall (Table 6).

4.2. Analysis

When optimizing a method consisting of two levels of analogy, the introduction of moisture variables in the second level has an influence on the parameter values of the first level. This means that the two levels of analogy bring complementary information, and are thus not independent. This is first visible in the number N_1 of analogues to be selected on the first level, and in the selection of the pressure levels for the circulation analogy. If the change in the optimal value of N_1 was already known, a change in the optimal pressure levels for the circulation analogy has never been identified before.

As for the sequential procedure, the optimal value of N_1 increased when adding a second level of analogy (Figure 9). One can also see that the optimal number of analogues N_2 for the second level of analogy of 4Zo-2MIo was slightly inferior to N_1 from 4Zo, but very close. There is a globally common tendency between the optimal analogue number values of both methods: N_1 of the 4Zo method, and N_1 and N_2 of 4Zo-2MIo tend to be higher or lower together for a given region.

The optimal final numbers of analogues did not vary much: $23 \leq N_1 \leq 33$ for 4Zo and $21 \leq N_2 \leq 28$ for 4Zo-2MIo. However, the optimal number of the N_1 analogues of the first level of 4Zo-2MIo varied to a greater extent: $48 \leq N_1 \leq 84$. In this latter method, it may be problematic to consider a fixed and unique value for all regions.

As for the pressure levels, Z1000, which was previously systematically selected twice (Table 3) was here less often chosen (once or even not at all) for 4Zo-2MIo (Table 5). There was indeed a vertical shift in the previously selected Z1000 for higher levels that was even slightly stronger with four moisture predictors than with two. This change is likely due to the fact that when considering only the circulation analogy, the method tried to take into account information that can serve as a proxy for moisture assessment, whereas it did not need it with the moisture index. This can only be assessed by a global optimization technique that can work jointly on both levels of analogy.

The selected pressure levels for the analogy of the moisture index were strongly centred around 700 hPa and 600 hPa. No other value was selected when considering two moisture predictors (Table 5). It was sometimes more efficient, in terms of prediction performance, to consider the moisture at 700 hPa twice, but at different hours, rather than selecting another pressure level. Besides, the optimizer never chose the same pressure level at the same hour for any variable, even though it was allowed to do so. The selected pressure levels for the moisture analogy differed from the reference method (Tables 2 and 5, last row).

The selection of temporal windows for atmospheric circulation was similar to the preceding optimization (in order of increasing pressure: 12 h, 24/30 h, 12 h, 30 h), but sometimes with some variability. When it comes to the moisture analogy, there was a clear tendency to select 12 h and 24 h. However, it must be remembered that this holds for our predictand, the accumulated precipitation over 06–30 h UTC, and that it is expected to differ if the temporal window changes (e.g. 00–24 h UTC, or another accumulation duration).

The optimized spatial windows for the atmospheric circulation have also changed (Figure 6). The very large domains on the upper level of the first four subregions were not present anymore, and more variability could be observed. The selected points for the moisture analogy were always located near the catchment, including at least one of the nearest points from the reanalysis dataset, and the spatial windows were relatively small. Thus, for this case study, there is no need to look for distant moisture information, and the search could be reduced to a smaller domain.

The CRPSS scores were improved by considering the moisture information (Figure 7 to be compared with Figure 4). The optimized method also performed significantly better than the 2Z-2MI reference method optimized by the sequential procedure. When it comes to improvements for days with precipitation above the three thresholds (P≥1 mm, P $\geq$ 0.1·P10, and P $\geq$ 0.5·P10), the conclusion is the same as before, that is, a significant improvement in the prediction compared to the reference method, mainly for heavy rainfall.

The analysis of the parameters cross-compatibility (Figure 8) was also very similar to the one of the circulation analogy only. The same pattern could be observed, with a drop of performance for the subregions submitted to different meteorological influences. However, the losses in performance were globally more important than before, suggesting that more complex methods with moisture variables are less transposable to another subregion (consistent with the observations of Chardon et al. (2014)), even though both were located within the same grid cell of the reanalysis dataset. Moisture fields have greater variability than pressure fields, and thus a change in the spatial windows can have a greater impact on the method performance. Indeed, the two regions with the lowest cross-compatibility with the others were the upper Rhone Valley (7) and the southeast ridges (8), which had similar optimal pressure levels and temporal windows to other regions, but had rather different spatial windows on the moisture predictor.

Predictors based on moisture variables do significantly increase the prediction skill, and are thus recommended, as long as they are reliable. In real-time forecasting, their reliability depends on the lead time: for lead times superior to 3–4 days, the uncertainties related to moisture variables from NWP models become fairly high, which reduces the relevance of methods relying on this information. In climate downscaling studies, it mainly depends on the coherence of the climatologies between the archive and the GCM model outputs. One should, however, not establish an AM with moisture variables for too large a region, as the transferability is reduced (see Chardon et al., 2014, for alternative approaches).

5. Discussion

The optimization of the AM by means of GAs has been undertaken in successive stages by releasing progressively new degrees of freedom. This approach allowed us to differentiate the contributions to performance gains, as well as to identify possible over-parametrization. The main improvements obtained in the present case study are due to the following elements:

- Using four pressure levels for the circulation analogy seemed to be an optimal number for the studied region, length of archive available, and target predictand considered. Beyond four, the validation score was more variable, revealing a loss in robustness due to over-parametrization.

- The automatic and joint optimization of all parameters: the analogue numbers, selection of pressure levels and temporal windows, and spatial windows. These parameters are highly interdependent, so one needs to optimize them jointly in order to identify optimal combinations. Indeed, there is a strong interdependence between space and time in the atmospheric circulation, so that, e.g. the spatial window should move upstream the main atmospheric flow for earlier temporal windows.

- The introduction of distinct spatial windows between pressure levels. The synoptic circulation is characterized by features with very different scales depending on the height, and important information for predicting precipitation is not necessarily located in the same area from one level to another.

- The weighting of the analogy criteria between different pressure levels. This can be influenced by the variability of the geopotential height with altitude, or the levels of significance in regards to the meteorological processes specific to a region. There is a trend in the weighting of circulation predictors to decrease with the increase in pressure, as one can see in Figure 10 for the three optimized methods. However, the values remained approximately equal. This may not be the most influencing factor, and we may suggest removing it first when trying to reduce the degrees of freedom.

- The joint optimization of the circulation and moisture analogy levels, which are usually calibrated successively. We were able to demonstrate that there is a dependency between the analogy levels, and that in order to approach the optimal parameters, one must consider them jointly.

GAs have proved very useful to optimize complex variants of the AM, and to assess new degrees of freedom that were not available thus far. However, it can be dangerous to add too many parameters to optimize. Indeed, the optimizer will probably use them to successfully improve the calibration score, so the validation control remains very important in order to determine if one is actually improving the method, or if it is being over-parametrized. Moreover, it might not always be desirable to increase the degrees of freedom, and some constraints (e.g. same weighting of the analogy criteria between different pressure levels) can be justified. However, one should first assess the consequence of a constraint before establishing it. In this sense, even though not all degrees of freedom are useful, GAs allow us to assess their influence. Finally, GAs could be used to identify, among other things, the best pairs of spatial and temporal windows, in order to later create a simpler regional method.

The convergence of parallel optimizations decreased when the AM to optimize became more and more complex. The optimizer did not always converge to the exact global optimum, but to its surroundings. This is related to the fact that the optimization slows down when it gets closer to the global optimum, and that one has to stop it before the end, because of the required processing time (see for example the slow-down of the improvements over generations in Figure 11). The resulting parameters might sometimes present non-negligible differences, even though the score is similar. Through Monte-Carlo analyses of the parameter space properties of the AM, Horton (2012) showed that some parameters of the method have a wide range of acceptable values. The spatial windows, for example, can be larger than the optimal size without much impact on the score, while they cannot be smaller (see also Bontron, 2004). We also observed that the selection of pressure level is not a parameter as discrete as we would have thought, and that choosing another level may have reduced the consequences on the performance. This is particularly true for higher pressure levels, but can be more critical for lower layers. It was thus interesting to sometimes obtain several sets of near-optimal parameters, but with some nuances, in order to get an idea of the sensitivity of the parameters for a given region, and to compare the score for the VP. In this regard, a cross-validation technique may be advisable. However, as solutions identified at different regions of the parameter space might provide sufficiently good performance, an ensemble of these could be used, instead of a unique solution. These could account for the parameters uncertainty in the AM, and could also increase the sample size contributing to the empirical distribution of precipitation values. An approach that can also be recommended is to first explore a wide range of the parameter space with some optimizations, and to narrow it according to the results for more targeted optimizations that are likely to go faster and to perform better.

We tried to optimize the length of the preselection period (i.e. the seasonal stratification, which is a 4-month window) jointly with the other parameters, but no improvement was observed. Optimizing the moisture flux, which is composed of the moisture index multiplied by the wind flux, was also assessed. However, the results were not better than when considering the moisture index alone. This may be related to the fact that the optimizer tries to provide the best analogy of the atmospheric circulation in the first place, which makes the wind information less relevant in the second level of analogy.

As has been observed, methods with higher complexity that integrate moisture predictors are less transposable than simpler ones. It was also noticed in another unpublished work, that it is by far better to optimize for two subregions jointly than to optimize on one and to apply its parametrization to the other. Finally, the discretization in subregions is an important process and should be handled with care. Indeed, the geographical distance is not always the leading factor to define a subregion. For example, the southeast ridges subregion does not behave like its surroundings, and differs in its parametrization because of different leading meteorological influences.

GAs are relatively heavy on processing and require an IT infrastructure capable of performing thousands of hours of calculations. However, they automatically optimize all parameters of the AM, which is not possible with the sequential calibration. Therefore, much human time, previously required to successively assess numerous combinations of parameters (particularly the selection of pressure levels and temporal windows), is saved. The ability to jointly optimize all parameters is important given the strong dependencies between them and between the levels of analogy.

Furthermore, AMs optimized with GAs showed an improvement in predictions for days with heavier precipitation, including extremes. Even though no new extreme value was added to the existing precipitation archive, the distribution of analogue precipitation values for a target situation can move towards the targeted extreme by sampling better candidate situations. Then, the subset of precipitation values collected on the analogue dates can be considered as a sample of the conditional distribution of precipitation associated with this situation. A truncated exponential or a gamma distribution model can be fit and extrapolated to extreme values not contained in the sample or even in the whole precipitation archive (Obled et al., 2002). Another possible approach is to combine AMs with other methods (e.g. Chardon et al., 2014).

6. Conclusions and perspectives

The parameters resulting from the optimization by GAs were very consistent in terms of the selection of pressure levels and temporal and spatial windows. There were clear trends or even identical results for subregions under similar meteorological influences, which confirm that the optimized parameters were coherent, despite an eventual first impression of great variability in the spatial windows. When adding moisture variables, the results showed a higher variability, but remained highly acceptable and coherent.

Strong dependencies between the parameters of the AM were observed. Thus, the sequential calibration, which optimizes the parameters successively, may not lead to the optimal combination. Moreover, it contains several manual systematic assessments, such as the selection of pressure levels and temporal windows. GAs, however, can automatically select pressure levels and temporal windows, which can save a considerable amount of human time. A great advantage of a global optimization is its ability to approach or reach optimal parameter values when they are considered jointly.

A dependence in the selected parameters between the circulation analogy and moisture variables was identified. When the two analogy levels are considered together, the optimal parameters of the circulation analogy changed. This complexity can only be exploited in a suitable manner by global optimization methods.

For the present case study, there seemed to be an optimum number of pressure levels to consider for the circulation analogy, which is four, before losing consistency of the real gains. The circulation analogy was improved by introducing a weighting between pressure levels, and considering independent spatial windows between pressure levels.

GAs provided parameterizations of AMs that exceeded the performance of the sequential calibration. In addition, it has been observed that the prediction for days with strong precipitation were improved to a greater extent, which is clearly interesting in the context of flood forecasting.

This work is by no means exhaustive, and is meant to open a door to new explorations of AMs with GAs. It is even possible to let the optimizer chose the meteorological variable to be used as a predictor, as well as the analogy criteria, which is the topic of work in progress. Moreover, the AM has been explored for decades for precipitation prediction, but not as intensively for other predictands. A global optimizer, such as a GA, can speed up this assessment significantly.

Acknowledgments

We thank Hamid Hussain-Khan of the University of Lausanne for his help and availability, and for the intensive use of the cluster he administers. We also thank Renaud Marty for fruitful collaboration over the years, and Dominique Bérod for his support.

We thank the Roads and Water courses Service, the Energy and Water Power Service of the Wallis Canton, the Water, Land and Sanitation Service of the Vaud Canton, and the Swiss Federal Office for Environment (FOEV) for financing the MINERVE (Modélisation des Intempéries de Nature Extrême des Rivières Valaisannes et de leurs Effets) project which started this research. 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 were provided by NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, via their website at http://www.esrl.noaa.gov/psd/. The precipitation time series were provided by MeteoSwiss.

The authors would also like to acknowledge the work of Loris Foresti and another anonymous reviewer, which contributed to improving this paper.

References

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

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

Bäck, T., 1993. Optimal Mutation Rates in Genetic Search, in: Proceedings of the 5th International Conference on Genetic Algorithms, pp. 2–8.

Bäck, T., Schwefel, H.P., 1993. An Overview of Evolutionary Algorithms for Parameter Optimization. Evol. Comput. 1, 1–23. doi:10.1162/evco.1993.1.1.1.

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

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

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

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

Bois, P., Obled, C., Thalamy, J., 1981. Etude des liaisons entre champs de pression et températures sur l'Europe avec la durée d'insolation: Application à la possibilité de la prévision d'insolation, in: C.R. du colloque Météorologie de l'énergie solaire, PIRDES-CNRS, Toulouse, pp. 145–177.

Bolognesi, R., 1993. Premiers développements d'un modèle hybride pour le diagnostic spatial des risques d'avalanches. La Houille Blanche 8, 551–553. doi:10.1051/lhb/1993045.

Bontron, G., 2004. 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.

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

Brown, T., 1974. Admissible Scoring Systems for Continuous Distributions. Technical Report.

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

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

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

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

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

Drosdowsky, W., Zhang, H., 2003. Verification of Spatial Fields, in: Jolliffe, I.T., Stephenson, D.B. (Eds.), Forecast verification: a practitioner's guide in atmospheric science. Wiley. chapter 6, pp. 121–136.

Duband, D., 1970. Reconnaissance dynamique de la forme des situations météorologiques. Application à la prévision quantitative des précipitations. Ph.D. thesis. Faculté des sciences de Paris.

Foresti, L., Panziera, L., Mandapaka, P.V., Germann, U., Seed, A., 2015. Retrieval of analogue radar images for ensemble nowcasting of orographic rainfall. Meteorol. Appl. 22, 141–155. doi:10.1002/met.1416.

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

Goldberg, D.E., 1989. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman Publishing Co., Inc.

Gordon, N.D., 1987. Statistical very short-range forecasting via analogues, in: Proceedings of the Symposium on mesoscale analysis and forecasting, ESA SP-282, Vancouver, Canada. pp. 307–312.

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

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

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

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

Holland, J.H., 1992. Genetic Algorithms. Sci. Am. 267, 66–72. doi:10.1038/scientificamerican0792-66.

Horton, P., 2012. 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.

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

Horton, P., Jaboyedoff, M., Obled, C., 2016a. Global Optimization of an Analog Method by Means of Genetic Algorithms. Mon. Weather Rev. Accepted. doi:10.1175/MWR-D-16-0093.1.

Horton, P., Obled, C., Jaboyedoff, M., 2016b. The Analogue Method for Precipitation Forecasting: Finding Better Analogue Situations at a Sub-Daily Time Step. Hydrol. Earth Syst. Sci. Discuss. , in reviewdoi:10.5194/hess-2016-246.

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

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

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

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

Lorenz, E., 1956. Empirical orthogonal functions and statistical weather prediction. Technical Report. Massachusetts Institute of Technology, Department of Meteorology, Massachusetts Institute of Technology, Dept. of Meteorology.

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

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., Thiele-Eich, I., 2010. Precipitation downscaling under climate change: Recent developements to bridge the gap between dynamical models and the end user. Rev. Geophys. 48, 1–34. doi:10.1029/2009RG000314.

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

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

Michalewicz, Z., 1996. Genetic Algorithms + Data Structures = Evolution Programs. 3rd editio ed., Springer-Verlag.

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

Obled, C., Good, W., 1980. Recent developments of avalanche forecasting by discriminant analysis techniques: a methodological review and some applications to the Parsenn area (Davos, Switzerland). J. Glaciol. 25, 315–346. doi:10.3198/1980JoG25-92-315-346.

Panziera, L., Germann, U., Gabella, M., Mandapaka, P.V., 2011. NORA-Nowcasting of Orographic Rainfall by means of Analogues. Q. J. R. Meteorol. Soc. 137, 2106–2123. doi:10.1002/qj.878.

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

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

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

Schwefel, H.P., 1993. Evolution and Optimum Seeking. John Wiley & Sons, Inc., New York, NY, USA.

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

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

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

Zitzler, E., Laumanns, M., Bleuler, S., 2004. A Tutorial on Evolutionary Multiobjective Optimization, in: Metaheuristics for Multiobjective Optimisation, pp. 3–37. doi:10.1007/978-3-642-17144-4_1.

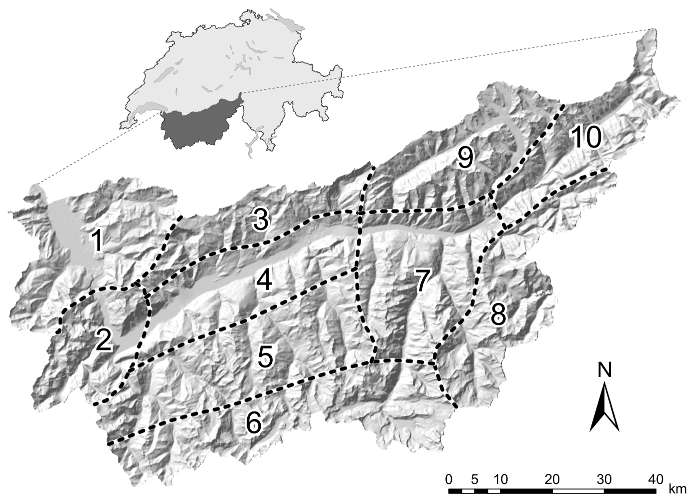


Figure 1: Location of the alpine Rhône catchment in Switzerland. (source: Swisstopo)

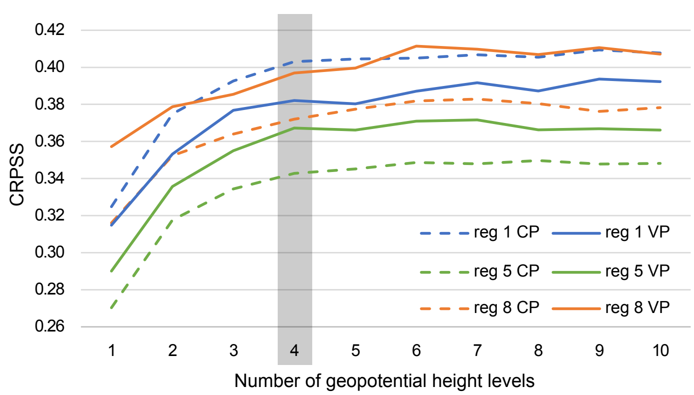


Figure 2: Performance score (CRPSS) for CP and VP for three stations (1 - Swiss Chablais; 5 - Southern valleys; 8 - Southeast ridges) when varying the number of geopotential height levels available to the optimizer.

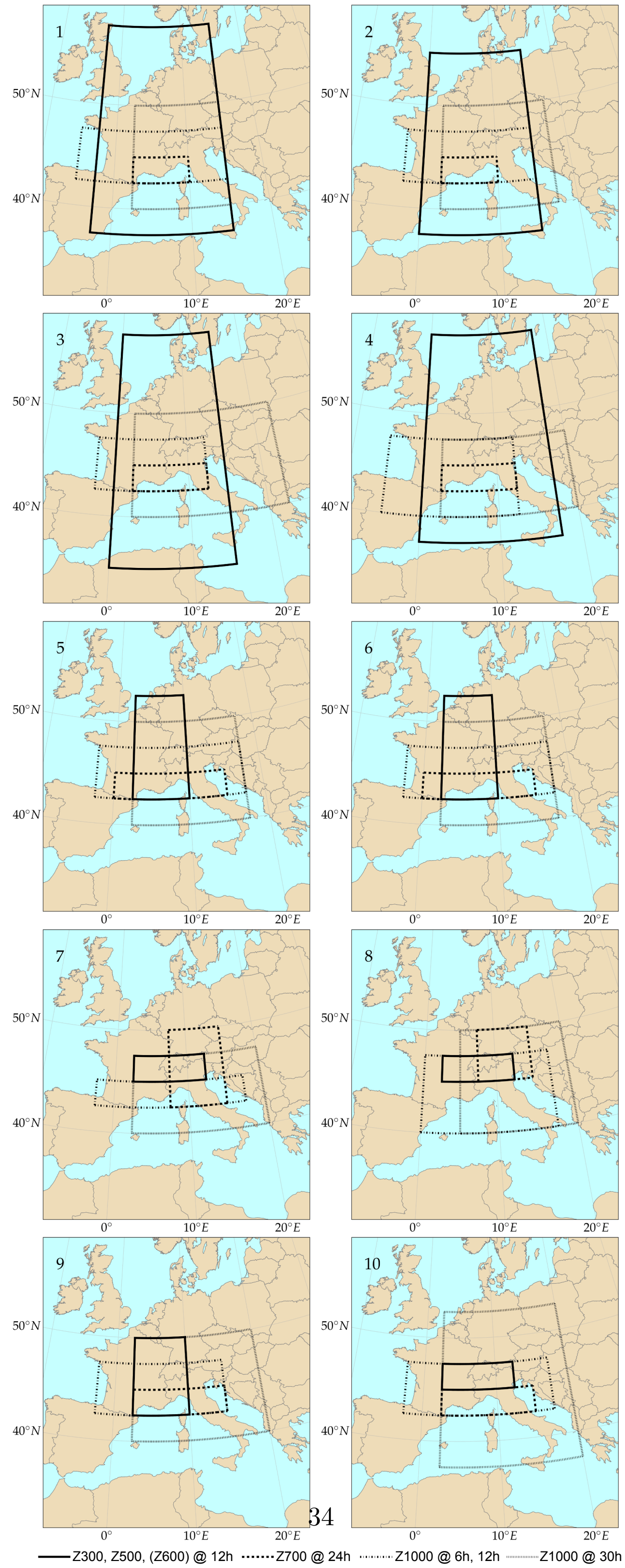


Figure 3: Optimized spatial windows for the 4Zo method (analogy of the atmospheric circulation on four pressure levels).

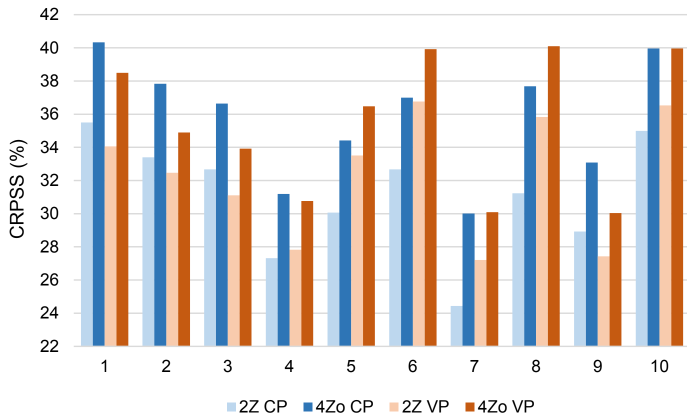


Figure 4: Performance score (CRPSS) of the reference method 2Z (Table 1) and the optimized 4Zo method for CP and VP for every subregion.

Groupment Ids (parameters used)

CP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.3	-0.8	-0.7	-0.7	-0.7	-3.1	-4.4	-1.3	-2.0
2	-0.2	0.0	-0.6	-0.3	-0.6	-0.6	-2.2	-3.6	-0.9	-1.1
3	-0.4	-0.2	0.0	-0.5	-0.4	-0.3	-1.8	-2.9	-0.1	-0.5
4	-0.3	-0.2	-0.4	0.0	-0.3	-0.2	-1.2	-2.6	-0.4	-0.6
5	-0.6	-0.5	-0.8	-0.5	0.0	0.0	-1.2	-2.3	-0.5	-0.7
6	-1.2	-0.8	-1.1	-0.8	0.0	0.0	-0.8	-1.6	-0.3	-0.4
7	-3.6	-3.1	-2.6	-2.5	-1.5	-1.5	0.0	-0.7	-1.1	-1.1
8	-6.4	-5.3	-5.0	-4.6	-2.8	-2.8	-0.2	0.0	-2.7	-1.6
9	-0.9	-0.7	-0.5	-0.6	-0.4	-0.3	-0.8	-1.7	0.0	-0.3
10	-1.9	-1.4	-1.2	-1.6	-0.6	-0.5	-0.9	-1.1	-0.5	0.0

Groupment Ids (targets)

VP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.3	-0.5	0.1	-0.8	-0.9	-3.2	-4.7	-0.9	-1.4
2	-0.1	0.0	0.0	0.2	-0.6	-0.6	-2.1	-3.8	-0.3	-0.7
3	-0.4	-0.6	0.0	-0.3	-1.1	-1.1	-2.1	-4.4	-0.3	-1.1
4	-0.2	0.0	-0.2	0.0	-0.5	-0.6	-2.0	-3.6	-0.7	-1.0
5	0.2	0.2	0.2	0.3	0.0	0.0	-1.8	-3.2	0.0	-0.5
6	-0.5	-0.4	-0.6	0.0	0.0	0.0	-0.9	-1.9	-0.5	-0.4
7	-1.4	-1.1	-0.6	-0.6	-0.6	-0.6	0.0	-1.0	-0.5	0.0
8	-3.6	-3.0	-2.2	-1.9	-1.3	-1.3	0.1	0.0	-1.4	0.2
9	0.2	0.0	0.2	0.4	-0.2	-0.2	-0.7	-2.2	0.0	0.1
10	-0.4	-0.3	-0.1	-0.2	0.1	0.1	-0.8	-1.1	-0.1	0.0

Groupment Ids (targets)

Figure 5: Losses or gains (in %) in CRPSS from applying optimized parameters for the series in columns to those in rows. Method 4Zo, calibration and validation periods.

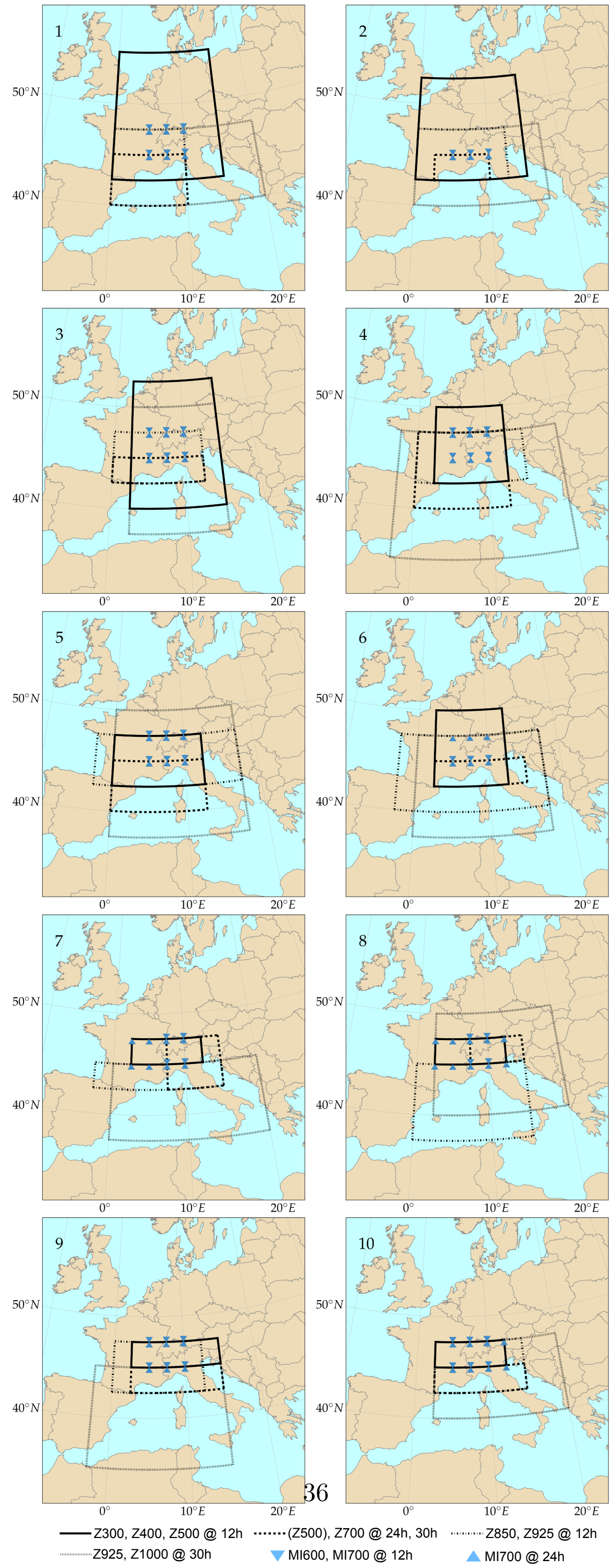


Figure 6: Optimized spatial windows for the 4Zo-2MIo method (analogy of atmospheric circulation on four pressure levels and analogy on the moisture index on two pressure levels).

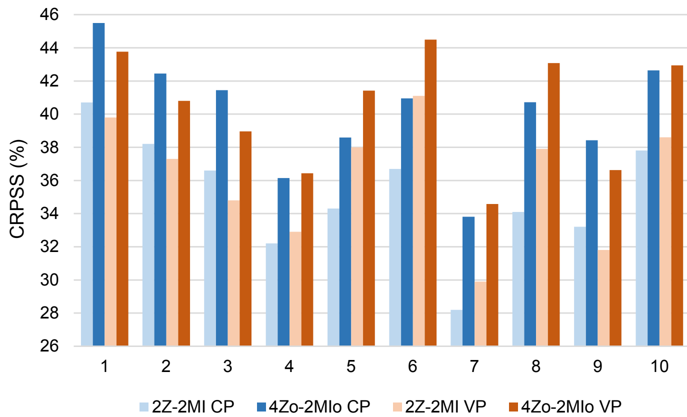


Figure 7: Performance score (CRPSS) of the reference method 2Z-2MI (Table 2) and the optimized 4Zo-2MIo method for CP and VP for every subregion.

Groupment Ids (parameters used)

Groupment Ids (targets)

CP	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0.0	-0.6	-0.9	-1.4	-0.9	-1.8	-3.9	-5.4	-2.2	-2.7
2	-0.9	0.0	-0.5	-0.9	-0.6	-1.2	-2.5	-4.1	-1.4	-2.2
3	-1.3	-0.7	0.0	-0.6	-0.8	-1.3	-2.0	-3.7	-0.6	-1.7
4	-1.2	-0.3	-0.3	0.0	-0.2	-0.8	-1.3	-2.8	-0.3	-1.6
5	-0.8	-0.5	-0.4	-0.3	0.0	-0.5	-1.0	-2.5	-0.4	-1.1
6	-1.5	-1.0	-1.0	-0.7	-0.6	0.0	-0.8	-1.7	-0.7	-0.9
7	-4.6	-3.3	-3.3	-2.1	-2.7	-1.3	0.0	-1.2	-1.3	-1.5
8	-7.5	-5.9	-5.7	-4.2	-5.1	-2.5	-0.7	0.0	-2.8	-1.8
9	-2.0	-1.4	-0.9	-0.6	-1.0	-1.2	-1.0	-2.5	0.0	-1.5
10	-2.5	-1.9	-1.4	-1.6	-1.6	-0.6	-0.7	-1.3	-0.9	0.0

VP	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0.0	-0.2	0.1	-1.0	-0.3	-1.4	-3.3	-6.1	-1.4	-2.6
2	-0.9	0.0	-0.3	-1.0	-0.9	-1.2	-2.9	-5.9	-1.2	-2.5
3	-1.3	-0.7	0.0	-1.3	-0.6	-1.1	-2.5	-6.0	-0.8	-2.0
4	-0.9	0.2	-0.2	0.0	-0.3	-0.1	-1.8	-5.1	0.1	-1.6
5	-0.5	0.1	-0.1	-0.4	0.0	0.0	-1.1	-3.9	0.0	-1.1
6	-1.2	-0.7	-0.9	-1.2	-0.7	0.0	-0.9	-3.0	-0.9	-1.0
7	-3.4	-2.1	-2.7	-1.5	-1.7	-0.5	0.0	-2.4	-0.4	-0.8
8	-5.3	-3.3	-3.9	-2.7	-3.0	-0.6	-0.1	0.0	-1.1	-0.7
9	-2.0	-1.3	-0.7	-0.8	-1.2	-1.3	-1.6	-4.9	0.0	-2.2
10	-1.8	-0.9	-0.5	-1.3	-0.6	0.2	-0.5	-2.4	-0.6	0.0

Figure 8: Losses or gains (in %) of the CRPSS from applying optimized parameters for the series in columns to those in rows. Method 4Zo-2MIo, calibration and validation periods.

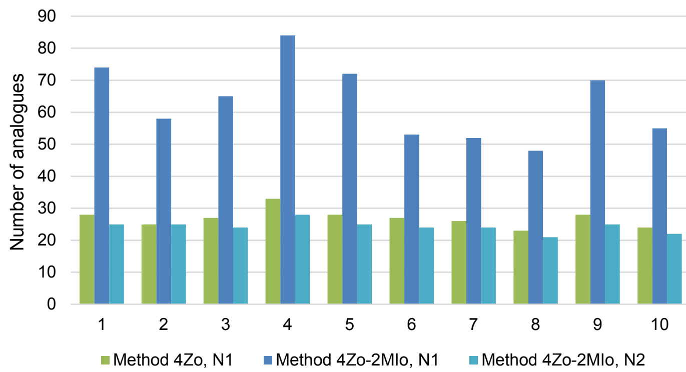


Figure 9: Optimal numbers of analogues for the different regions and the two methods, resulting from optimization. Method 4Zo is made of a single level of analogy with N_1 analogues, whereas 4Zo-2MIo has two levels of analogy with N_1 and N_2 analogues.

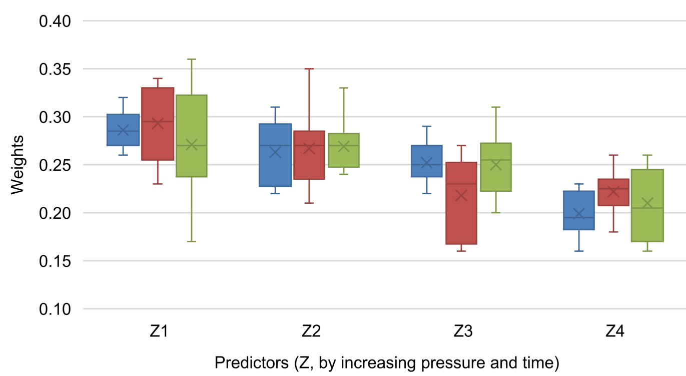


Figure 10: Distribution of optimal weights for the predictors of the first level of analogy (geopotential heights) of (blue) 4Zo, (red) 4Zo-2MIo, and (green) 4Zo-4MIo methods. Results are for the ten subregions. Geopotential heights are sorted by increasing pressure and hour.

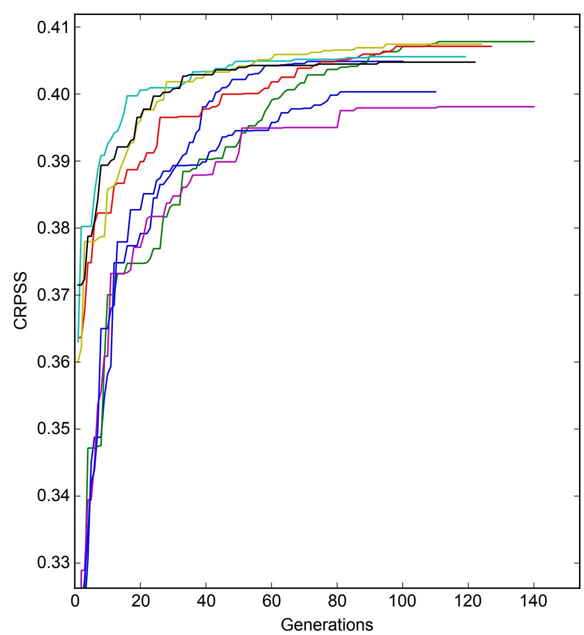


Figure 11: Example of evolution of the performance score of the best individual over eight independent optimizations.

Table 1: Parameters of the reference method on atmospheric circulation (2Z). First column is level of analogy (0 for preselection); subsequent columns list meteorological variable, its hour of observation within the target day (temporal window), criterion used for current level of analogy, and number of analogues.

Level	Variable	Hour	Criterion	Nb
0	±60 days around the target date			
1	Z1000	12 h	S1	50
	Z500	24 h		

Table 2: Parameters of the reference method with moisture variables (2Z-2MI). Conventions are the same as in Table 1

Level	Variable	Hour	Criterion	Nb
0	±60 days around the target date			
1	Z1000	12 h	S1	70
	Z500	24 h		
2	TPW * RH850	12 h	RMSE	30
	TPW * RH850	24 h		

Table 3: Pressure levels (∼) automatically selected for the 4Zo method for different subregions (ID). R represents the 2Z reference method (Table 1).

ID	300	400	500	600	700	850	925	1000
1	∼				∼			∼∼
2	∼				∼			∼∼
3	∼				∼			∼∼
4	∼				∼			∼∼
5			∼		∼			∼∼
6			∼		∼			∼∼
7			∼		∼			∼∼
8			∼		∼			∼∼
9			∼		∼			∼∼
10				∼	∼			∼∼
R			∼					∼

Table 4: Relative improvement (%) in CRPSS for different precipitation thresholds for the optimized 4Zo method, compared to the reference method.

ID	P≥1 mm		P≥0.1·P10		P≥0.5·P10	
	CP	VP	CP	VP	CP	VP
1	10.2	9.4	8.5	7.9	17.0	14.2
2	9.9	3.4	10.2	7.3	19.3	13.7
3	13.3	10.5	13.3	10.9	19.7	9.7
4	11.0	7.4	12.9	10.0	23.2	23.8
5	8.6	4.2	10.9	6.2	25.2	23.8
6	10.5	5.1	11.1	7.1	21.2	41.1
7	24.3	12.4	33.1	26.0	71.2	104.3
8	19.0	12.7	26.2	19.2	39.4	34.9
9	12.4	6.8	13.8	9.9	24.9	48.1
10	13.6	6.8	14.4	6.9	29.9	31.5
av.	13.3	7.9	15.4	11.1	29.1	34.5

Table 5: Atmospheric levels automatically selected for the analogy of atmospheric circulation ($\sim$) and moisture analogy ($\bullet$) of the 4Zo-2MIo method, for different subregions (ID). R represents the 2Z-2MI reference method (Table 2)

ID	300	400	500	600	700	850	925	1000
1	$\sim$		$\sim$		$\bullet\bullet$	$\sim$		$\sim$
2	$\sim$				$\sim\bullet\bullet$	$\sim$		$\sim$
3	$\sim$				$\sim\bullet\bullet$	$\sim$	$\sim$	
4			$\sim$	$\bullet$	$\sim\bullet$	$\sim$		$\sim$
5		$\sim$			$\sim\bullet\bullet$		$\sim\sim$	
6		$\sim$		$\bullet$	$\sim\bullet$	$\sim$		$\sim$
7		$\sim$		$\bullet$	$\sim\bullet$	$\sim$		$\sim$
8			$\sim$	$\bullet$	$\sim\bullet$		$\sim\sim$	
9		$\sim$		$\bullet$	$\sim\bullet$	$\sim$	$\sim$	
10		$\sim$		$\bullet$	$\sim\bullet$	$\sim$		$\sim$
R			$\sim$			$\bullet\bullet$		$\sim$

Table 6: Relative improvement (%) in CRPSS for different precipitations thresholds for the optimized 4Zo-2MIo method, compared to the reference method.

ID	P≥1 mm		P≥0.1·P10		P≥0.5·P10	
	CP	VP	CP	VP	CP	VP
1	12.6	9.3	12.4	9.7	15.8	11.0
2	10.4	7.7	11.2	10.5	18.9	16.6
3	14.5	11.6	14.1	11.4	18.7	14.6
4	11.4	9.4	11.5	11.6	14.9	22.7
5	11.8	8.0	12.2	8.9	12.0	12.8
6	11.3	7.1	11.2	8.0	15.3	29.1
7	20.5	15.5	25.2	24.0	43.0	79.5
8	19.3	15.7	23.1	18.6	25.2	31.7
9	17.0	15.4	17.4	16.5	23.7	39.4
10	12.9	9.6	13.8	11.1	28.5	32.1
av.	14.2	10.9	15.2	13.0	21.6	28.9

Figure 1

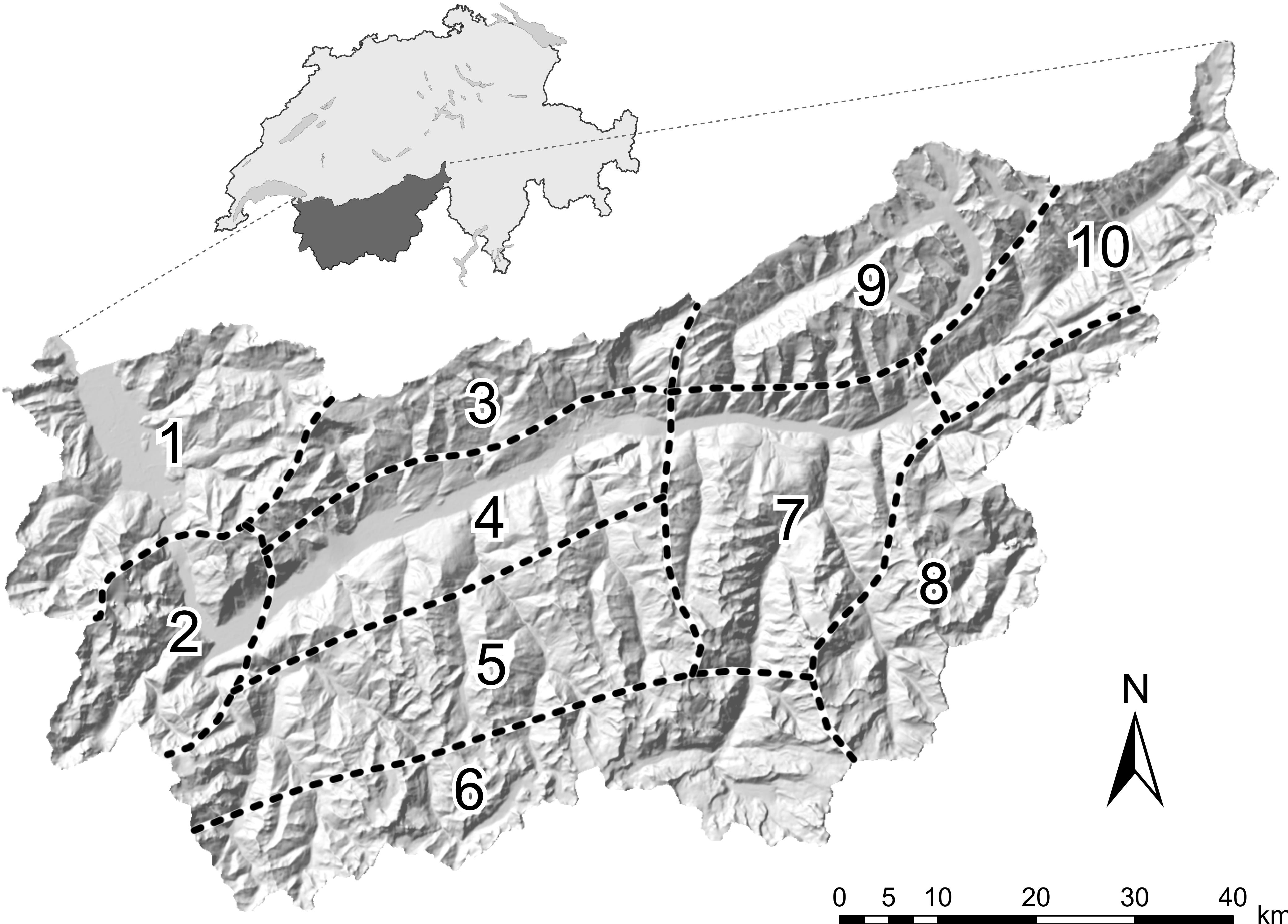


Figure 2

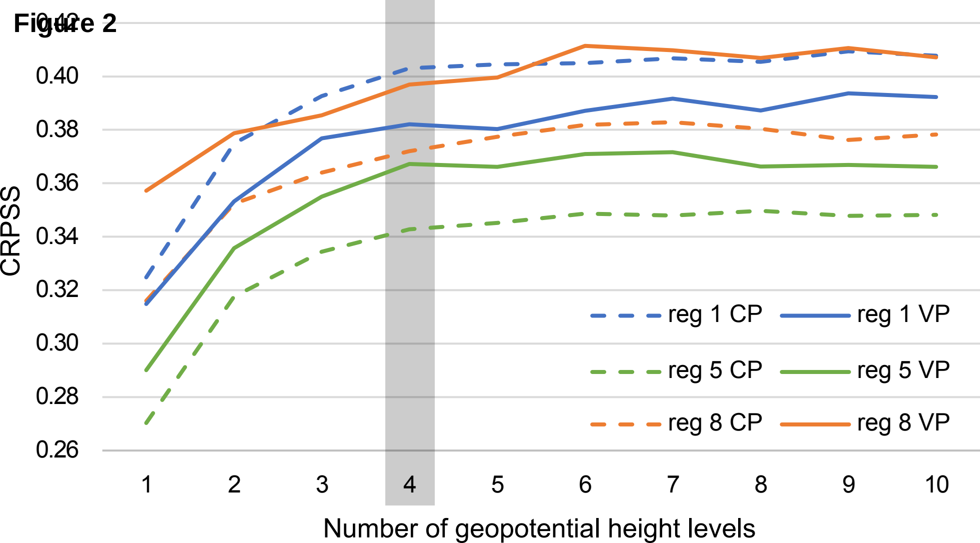

0.42
0.40
0.38
0.36
0.34
0.32
0.30
0.28
0.26
CRPSS
1
2
3
4
5
6
7
8
9
10
Number of geopotential height levels
reg 1 CP
reg 1 VP
reg 5 CP
reg 5 VP
reg 8 CP
reg 8 VP

Figure 3

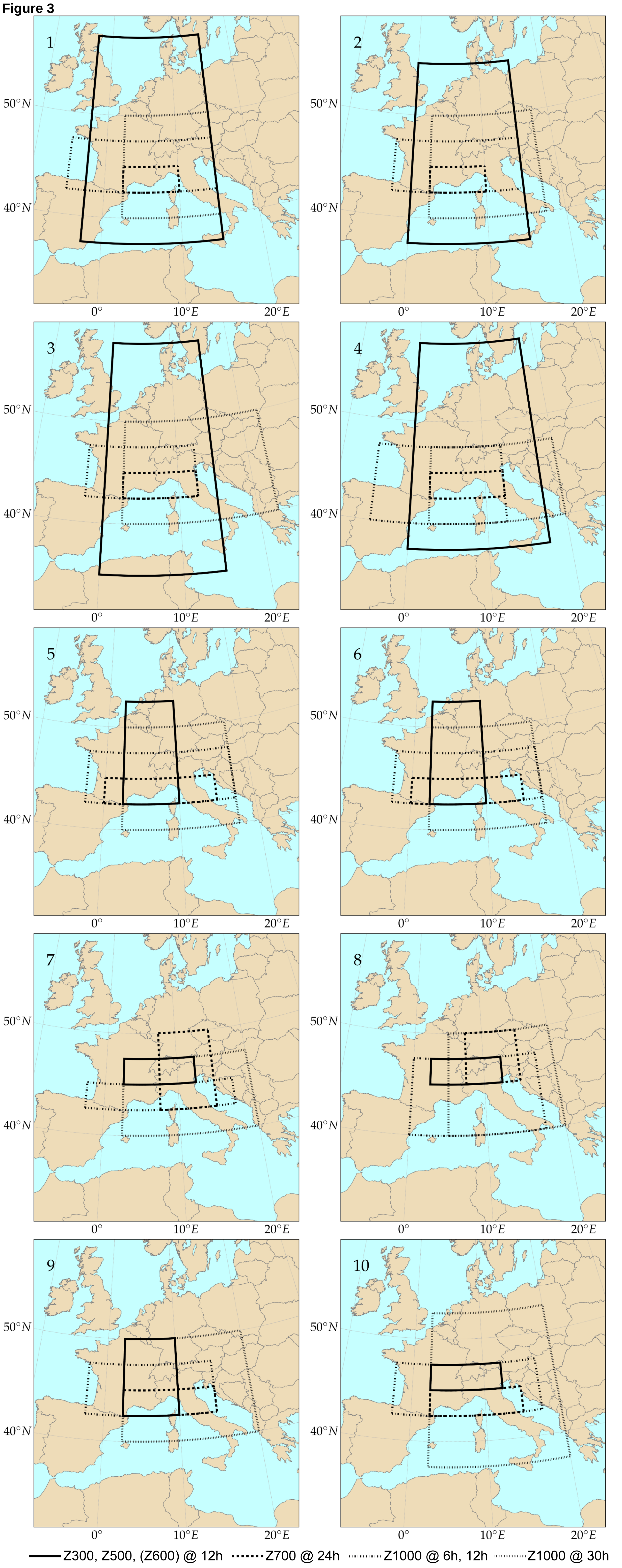
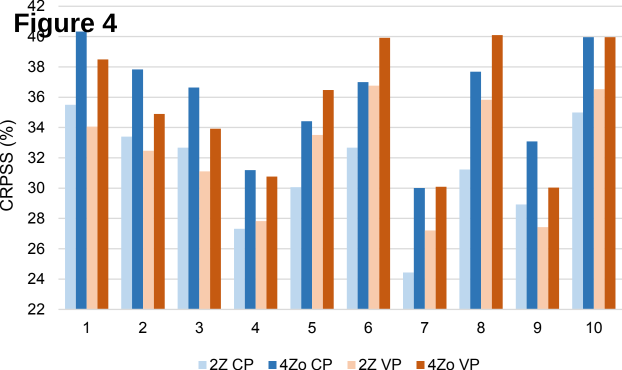

1
2
3
4
5
6
7
8
9
10
50°N
40°N
0°
10°E
20°E
Z300, Z500, (Z600) @ 12h
Z700 @ 24h
Z1000 @ 6h, 12h
Z1000 @ 30h

Figure 4



CRPSS (%)
42
40
38
36
34
32
30
28
26
24
22
1
2
3
4
5
6
7
8
9
10
2Z CP
4Zo CP
2Z VP
4Zo VP

Figure 5

Groupment Ids (parameters used)

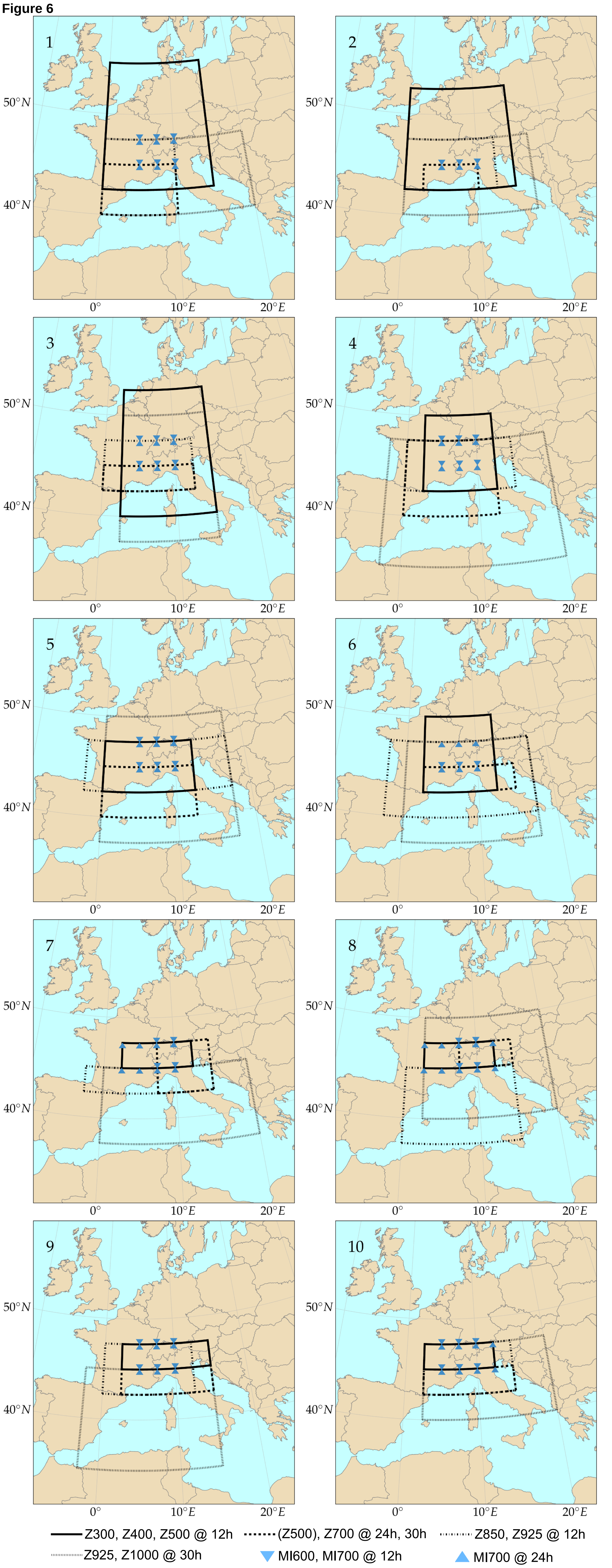
Rows: Groupment Ids (targets)

CP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.3	-0.8	-0.7	-0.7	-0.7	-3.1	-4.4	-1.3	-2.0
2	-0.2	0.0	-0.6	-0.3	-0.6	-0.6	-2.2	-3.6	-0.9	-1.1
3	-0.4	-0.2	0.0	-0.5	-0.4	-0.3	-1.8	-2.9	-0.1	-0.5
4	-0.3	-0.2	-0.4	0.0	-0.3	-0.2	-1.2	-2.6	-0.4	-0.6
5	-0.6	-0.5	-0.8	-0.5	0.0	0.0	-1.2	-2.3	-0.5	-0.7
6	-1.2	-0.8	-1.1	-0.8	0.0	0.0	-0.8	-1.6	-0.3	-0.4
7	-3.6	-3.1	-2.6	-2.5	-1.5	-1.5	0.0	-0.7	-1.1	-1.1
8	-6.4	-5.3	-5.0	-4.6	-2.8	-2.8	-0.2	0.0	-2.7	-1.6
9	-0.9	-0.7	-0.5	-0.6	-0.4	-0.3	-0.8	-1.7	0.0	-0.3
10	-1.9	-1.4	-1.2	-1.6	-0.6	-0.5	-0.9	-1.1	-0.5	0.0

Rows: Groupment Ids (targets)

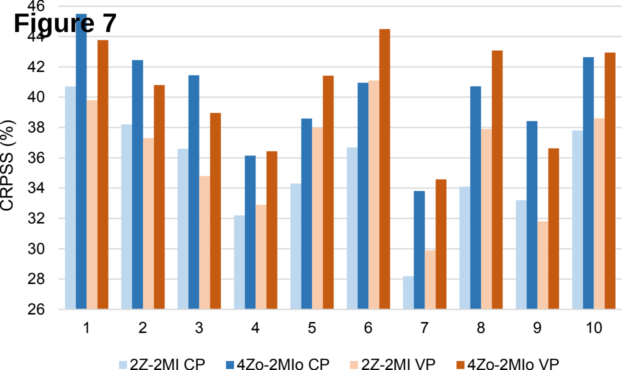
VP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.3	-0.5	0.1	-0.8	-0.9	-3.2	-4.7	-0.9	-1.4
2	-0.1	0.0	0.0	0.2	-0.6	-0.6	-2.1	-3.8	-0.3	-0.7
3	-0.4	-0.6	0.0	-0.3	-1.1	-1.1	-2.1	-4.4	-0.3	-1.1
4	-0.2	0.0	-0.2	0.0	-0.5	-0.6	-2.0	-3.6	-0.7	-1.0
5	0.2	0.2	0.2	0.3	0.0	0.0	-1.8	-3.2	0.0	-0.5
6	-0.5	-0.4	-0.6	0.0	0.0	0.0	-0.9	-1.9	-0.5	-0.4
7	-1.4	-1.1	-0.6	-0.6	-0.6	-0.6	0.0	-1.0	-0.5	0.0
8	-3.6	-3.0	-2.2	-1.9	-1.3	-1.3	0.1	0.0	-1.4	0.2
9	0.2	0.0	0.2	0.4	-0.2	-0.2	-0.7	-2.2	0.0	0.1
10	-0.4	-0.3	-0.1	-0.2	0.1	0.1	-0.8	-1.1	-0.1	0.0

Figure 6



1
2
3
4
5
6
7
8
9
10
50°N
40°N
0°
10°E
20°E
Z300, Z400, Z500 @ 12h
(Z500), Z700 @ 24h, 30h
Z850, Z925 @ 12h
Z925, Z1000 @ 30h
MI600, MI700 @ 12h
MI700 @ 24h

Figure 7



CRPSS (%)
26
28
30
32
34
36
38
40
42
44
46
1
2
3
4
5
6
7
8
9
10
2Z-2MI CP
4Zo-2MIo CP
2Z-2MI VP
4Zo-2MIo VP

Figure 8

Groupment Ids (parameters used)

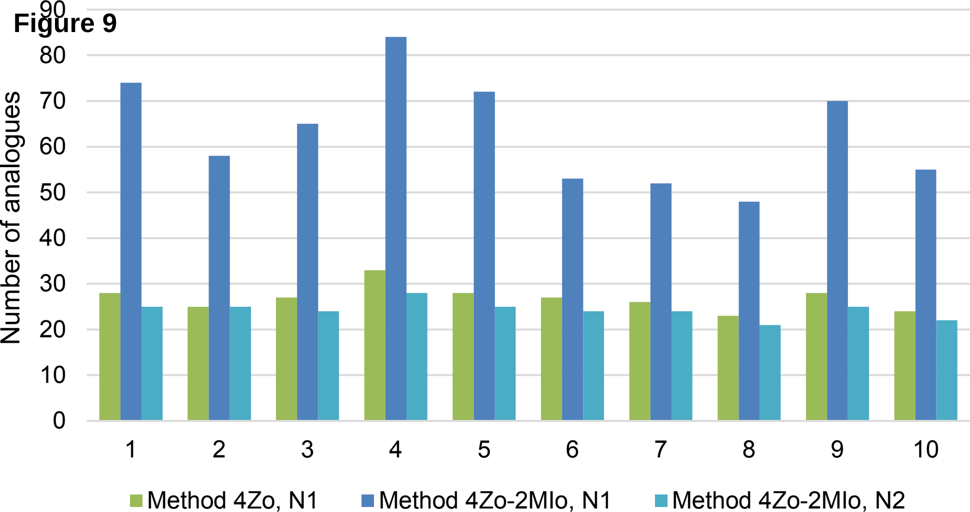
Groupment Ids (targets)

CP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.6	-0.9	-1.4	-0.9	-1.8	-3.9	-5.4	-2.2	-2.7
2	-0.9	0.0	-0.5	-0.9	-0.6	-1.2	-2.5	-4.1	-1.4	-2.2
3	-1.3	-0.7	0.0	-0.6	-0.8	-1.3	-2.0	-3.7	-0.6	-1.7
4	-1.2	-0.3	-0.3	0.0	-0.2	-0.8	-1.3	-2.8	-0.3	-1.6
5	-0.8	-0.5	-0.4	-0.3	0.0	-0.5	-1.0	-2.5	-0.4	-1.1
6	-1.5	-1.0	-1.0	-0.7	-0.6	0.0	-0.8	-1.7	-0.7	-0.9
7	-4.6	-3.3	-3.3	-2.1	-2.7	-1.3	0.0	-1.2	-1.3	-1.5
8	-7.5	-5.9	-5.7	-4.2	-5.1	-2.5	-0.7	0.0	-2.8	-1.8
9	-2.0	-1.4	-0.9	-0.6	-1.0	-1.2	-1.0	-2.5	0.0	-1.5
10	-2.5	-1.9	-1.4	-1.6	-1.6	-0.6	-0.7	-1.3	-0.9	0.0

Groupment Ids (targets)

VP	1	2	3	4	5	6	7	8	9	10
1	0.0	-0.2	0.1	-1.0	-0.3	-1.4	-3.3	-6.1	-1.4	-2.6
2	-0.9	0.0	-0.3	-1.0	-0.9	-1.2	-2.9	-5.9	-1.2	-2.5
3	-1.3	-0.7	0.0	-1.3	-0.6	-1.1	-2.5	-6.0	-0.8	-2.0
4	-0.9	0.2	-0.2	0.0	-0.3	-0.1	-1.8	-5.1	0.1	-1.6
5	-0.5	0.1	-0.1	-0.4	0.0	0.0	-1.1	-3.9	0.0	-1.1
6	-1.2	-0.7	-0.9	-1.2	-0.7	0.0	-0.9	-3.0	-0.9	-1.0
7	-3.4	-2.1	-2.7	-1.5	-1.7	-0.5	0.0	-2.4	-0.4	-0.8
8	-5.3	-3.3	-3.9	-2.7	-3.0	-0.6	-0.1	0.0	-1.1	-0.7
9	-2.0	-1.3	-0.7	-0.8	-1.2	-1.3	-1.6	-4.9	0.0	-2.2
10	-1.8	-0.9	-0.5	-1.3	-0.6	0.2	-0.5	-2.4	-0.6	0.0

Figure 9



90
80
70
60
50
40
30
20
10
0
Number of analogues
1
2
3
4
5
6
7
8
9
10
Method 4Zo, N1
Method 4Zo-2MIo, N1
Method 4Zo-2MIo, N2

Figure 10
Click here to download high resolution image

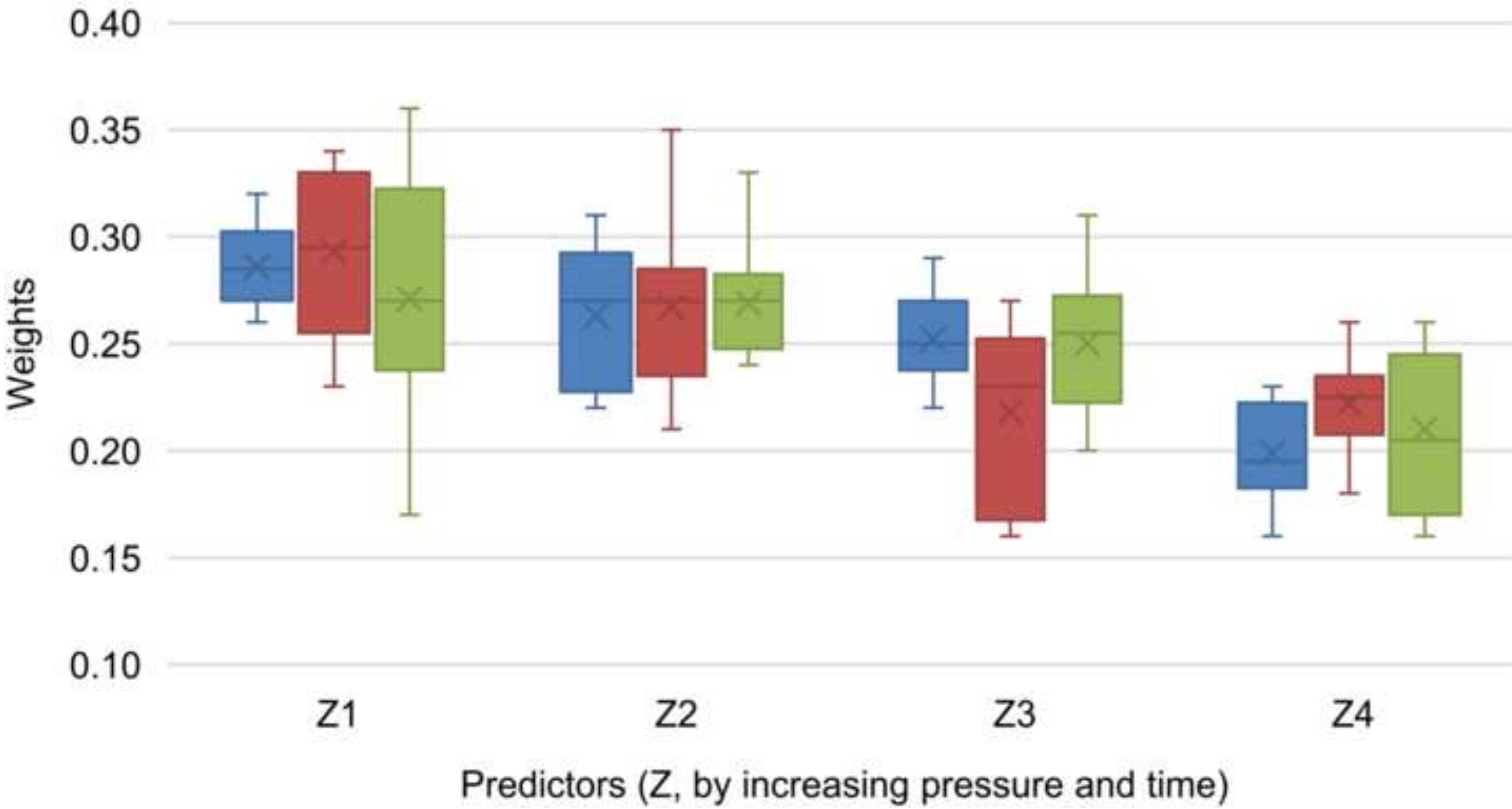
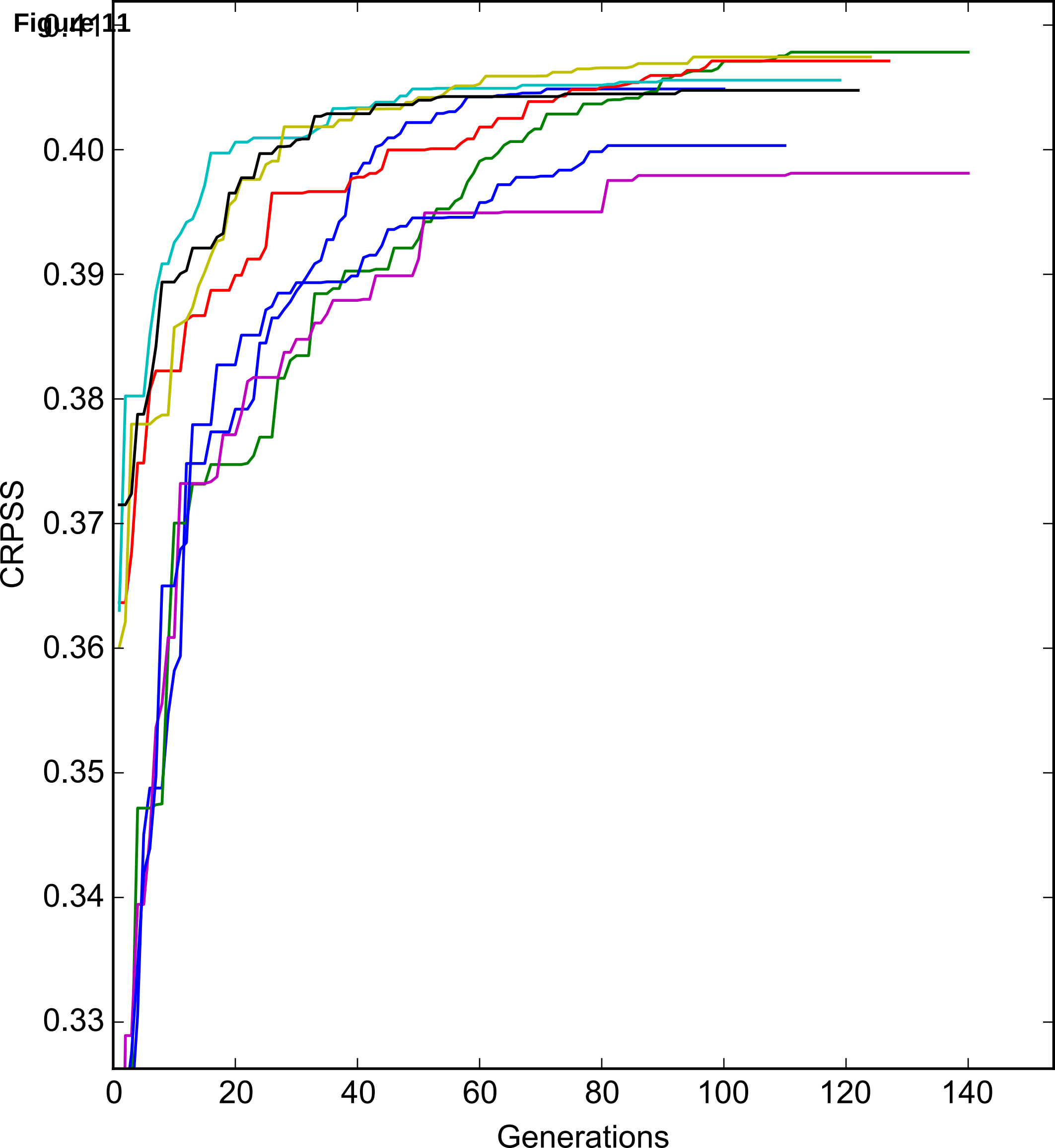


Figure 11


0.41
0.40
0.39
0.38
0.37
0.36
0.35
0.34
0.33
CRPSS
0
20
40
60
80
100
120
140
Generations

LaTeX Source Files
Click here to download LaTeX Source Files: src.zip