

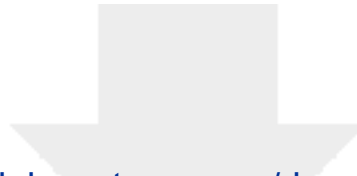
# Journal of Climate

## Global Optimization of the Analogue Method by Means of Genetic Algorithms

--Manuscript Draft--

<b>Manuscript Number:</b>	
<b>Full Title:</b>	Global Optimization of the Analogue Method by Means of Genetic Algorithms
<b>Article Type:</b>	Article
<b>Corresponding Author:</b>	Pascal Horton, Ph.D. University of Bern Bern, SWITZERLAND
<b>Corresponding Author's Institution:</b>	University of Bern
<b>First Author:</b>	Pascal Horton, Ph.D.
<b>Order of Authors:</b>	Pascal Horton, Ph.D. Michel Jaboyedoff Charles Obled
<b>Abstract:</b>	<p>The Analogue Method is based on a statistical relationship between synoptic atmospheric variables (predictors) and local weather (predictand), which we aim at forecasting. This relationship is expressed through many parameters that are usually calibrated by means of a semi-automatic sequential procedure. This calibration approach has strong limitations: it is made of successive steps and thus cannot handle parameters dependencies, and it cannot automatically optimize some parameters, such as the selection of the pressure levels and the temporal windows on which the predictors are compared.</p> <p>In order to surpass these limitations, a global optimization technique was assessed, namely Genetic Algorithms, which can optimize jointly all parameters of the method and get closer to a global optimum by taking into account the parameters dependencies. Moreover, it can choose objectively parameters that were previously manually assessed, and can take into account new degrees of freedom that were unthinkable before.</p> <p>These kind of optimization techniques need however to be tailored to the problem to solve. Multiple combinations of algorithms had to be assessed, and even new operators were developed, such as the chromosome of adaptive search radius which was found to be very robust, in order to make recommendations for the use of Genetic Algorithms to optimize the Analogue Method. These recommendations are the main outcome of this work. It opens new perspective for the improvement of the Analogue Method, and its application to new regions or to new predictands.</p>
<b>Suggested Reviewers:</b>	<p>Rémy Garçon EDF remy.garcon@edf.fr He knows the Analogue Method for years, as it has been developed at EDF.</p> <p>Alain Gautheron Alain.Gautheron@developpement-durable.gouv.fr He works with the Analogue Method for operational forecasting.</p> <p>Anne-Catherine Favre Université Joseph Fourier - Grenoble anne-catherine.favre@ujf-grenoble.fr She is an expert in hydrology statistics and knows the analogue method.</p> <p>Benoît Hingray Université Joseph Fourier - Grenoble benoit.hingray@ujf-grenoble.fr He knows the Analogue Method and is very conscientious when he writes or review.</p> <p>Guillaume Bontron CNR G.Bontron@cnr.tm.fr</p>

	He established the references we are comparing to.
	Renaud Marty, Ph.D. renaud.marty@developpement-durable.gouv.fr He did a Ph.D. on the analogue method.



[Click here to access/download](#)

**Cost Estimation and Agreement Worksheet**  
**Cost Estimation and Agreement Worksheet.pdf**



# Global Optimization of the Analogue Method by Means of Genetic

## Algorithms

Pascal Horton<sup>\*†</sup>

*University of Lausanne, Lausanne, Switzerland*

Michel Jaboyedoff

*University of Lausanne, Lausanne, Switzerland*

Charles Obled

*Université de Grenoble-Alpes, LTHE, Grenoble, France*

<sup>\*</sup> *Corresponding author address:* Pascal Horton, Terranum, Rue de l'industrie 35 bis, 1030 Bussigny, Switzerland.

E-mail: pascal.horton@alumni.unil.ch

<sup>†</sup> *Current affiliation:* Terranum, Bussigny, Switzerland

## ABSTRACT

13 The Analogue Method is based on a statistical relationship between synoptic  
14 atmospheric variables (predictors) and local weather (predictand), which we  
15 aim at forecasting. This relationship is expressed through many parameters  
16 that are usually calibrated by means of a semi-automatic sequential procedure.  
17 This calibration approach has strong limitations: it is made of successive steps  
18 and thus cannot handle parameters dependencies, and it cannot automatically  
19 optimize some parameters, such as the selection of the pressure levels and  
20 the temporal windows on which the predictors are compared. In order to sur-  
21 pass these limitations, a global optimization technique was assessed, namely  
22 Genetic Algorithms, which can optimize jointly all parameters of the method  
23 and get closer to a global optimum by taking into account the parameters  
24 dependencies. Moreover, it can choose objectively parameters that were pre-  
25 viously manually assessed, and can take into account new degrees of freedom  
26 that were unthinkable before. These kind of optimization techniques need  
27 however to be tailored to the problem to solve. Multiple combinations of al-  
28 gorithms had to be assessed, and even new operators were developed, such as  
29 the *chromosome of adaptive search radius* which was found to be very robust,  
30 in order to make recommendations for the use of Genetic Algorithms to op-  
31 timize the Analogue Method. These recommendations are the main outcome  
32 of this work. It opens new perspective for the improvement of the Analogue  
33 Method, and its application to new regions or to new predictands.

## 34 **1. Introduction**

35 The Analogue Method (AM) relies on the hypothesis that similar situations in terms of atmo-  
36 spheric circulation are likely to lead to similar local weather (Lorenz 1956, 1969; Duband 1970;  
37 Bontron and Obled 2005). The principle consists in sampling a certain number of past situations  
38 based on different atmospheric variables (predictors) in order to build a probabilistic forecast for a  
39 local weather variable of interest (predictand). The most common usage of the AM is for precip-  
40 itation forecasting (eg. Guilbaud 1997; Bontron and Obled 2005; Bliefernicht 2010; Marty et al.  
41 2012; Horton et al. 2012; Radanovics et al. 2013; Ben Daoud et al. 2015), but AMs, or equiva-  
42 lent, were also used for short to medium term forecasting of daily temperatures (Radinovic 1975;  
43 Woodcock 1980; Kruizinga and Murphy 1983), wind (Gordon 1987), snow avalanches (Obled and  
44 Good 1980; Bolognesi 1993), insolation (Bois et al. 1981), and the trajectory of tropical cyclones  
45 (Keenan and Woodcock 1981; Sievers et al. 2000; Fraedrich et al. 2003). Applications for monthly  
46 forecasts in many countries also exist, including Canada (Shabbar and Knox 1986), Hungary (Toth  
47 1989), the Netherlands (Nap et al. 1981), and England (Murray 1974), as well as seasonal fore-  
48 casts: Barnett and Preisendorfer (1978), Bergen and Harnack (1982) and Livezey and Barnston  
49 (1988).

50 Even though the method is rather simple compared to numerical weather prediction (NWP)  
51 models, it contains a certain number of parameters one needs to determine, such as the choice of  
52 the predictor variable, its pressure level and temporal window to consider, the spatial domain to  
53 use for the comparison, as well as the analogy criteria itself, and finally the number of analogue  
54 situations to keep at each subsampling level.

55 The AM needs to be adapted to every new region considered, because the leading meteorological  
56 influences may be location specific. Even the selection of the pressure levels and the temporal  
57 windows should be reconsidered, when not the predictor variable itself.

58 Up to now, the optimization of the method has been undertaken by semi-automatic sequen-  
59 tial calibration procedures (see Bontron 2004; Horton et al. 2016a, for its definition). However,  
60 the selection of predictor variables, pressure levels and temporal windows had still to be made  
61 manually before optimizing the spatial windows and the number of analogues. Testing multiple  
62 combinations of these is very combinatorial and becomes quickly cumbersome, especially when  
63 considering multiple predictors within the same level of analogy. Thus, optimizing the method  
64 with the sequential technique is laborious, as many combinations of predictors (variables, pres-  
65 sure levels, temporal windows) have to be assessed. Moreover, proceeding to the optimization  
66 sequentially ignores potential dependencies between the parameters of the method, may they be  
67 within a single level of analogy or between them, which could lead to another parametrization if  
68 the parameters were calibrated together. Simultaneous calibration of all parameters has never been  
69 undertaken so far. Thus, due to the sequential approach, the risk of ending in a local optimum is  
70 high and can not be avoided. Indeed, during the AM calibration it has been found that the resulting  
71 parameters may vary with initial choices (such as the number of analogues).

72 When creating the sequential calibration procedure, Bontron (2004) was aware of the problem  
73 of dependencies between parameters and wrote: *"We perceive here the combinatorial aspect of*  
74 *our problem: variables and spatial windows are not independent. We will present our results*  
75 *by first searching the best variable [note: e.g. selection of the pressure level and the temporal*  
76 *window for the geopotential height] on a chosen spatial window, and next, the best window for*  
77 *the chosen variable. However, even by repeating the process, are we sure to obtain the optimal*  
78 *combination?"*. And later in his work: *"Our approach, which is again to vary the parameters one*

79 *by one – the others being fixed in a more or less arbitrary manner – may therefore not exactly*  
80 *lead us to the optimal solution”*. Bliefernicht (2010) has also faced the combinatorial issue of  
81 the parameters of the AM and concludes that one needs to be an expert to have a sense of their  
82 respective influence, sensitivity and nonlinear interactions. Ben Daoud (2010), when calibrating  
83 the AM, also stated that *”the combinatory aspect related to the calibration was found to be too*  
84 *high for all the parameters to be calibrated simultaneously”*.

85 In order to overcome these limitations, two optimization techniques were assessed. First, Horton  
86 (2012) assessed the ability of the Nelder and Mead (1965) method based on a simplex approach.  
87 This technique did not provide satisfying results and failed at converging toward a unique solu-  
88 tion. The parameter space of the AM is very complex and is inappropriate for a linear optimization  
89 technique. The conclusion was that global optimization techniques were necessary in order to cal-  
90 ibrate AMs, as it is the only way to optimize all parameters of all analogy levels simultaneously.  
91 In addition, it can overcome the systematic manual assessments of all pressure levels and tempo-  
92 ral windows. Finally, it can open new perspectives by allowing the addition of new degrees of  
93 freedom, such as a weighting of the criteria values between the pressure levels (see Horton et al.  
94 2016b), and the consideration of differentiated spatial windows between the pressure levels. The  
95 relevance of Genetic Algorithms (GAs) is presented here, which does not exclude that other global  
96 optimization techniques could eventually be successful.

97 This article is not about discussing the details of the results of an optimization with GAs, but  
98 describing how GAs are to be used in order to successfully optimize AMs. Indeed, GAs variants  
99 are numerous and always need to be tailored to the problem addressed. This requires intensive  
100 and systematic comparisons of operators and options in order to identify the key factors leading  
101 the optimization and the respective sensitivity of the options. Such analyses are presented here  
102 by application to the upper Rhône catchment in Switzerland, and will result in recommendations



103 for the use of GAs when applied to AMs. The demonstration of the benefit brought by such an  
104 approach on a specific case study is the topic of Horton et al. (2016b).

105 We will begin by presenting the basics of the AM (Sect. 2) and the concepts of GAs as well  
106 as the assessed operators (operations performed on the AM parameters; see Sect. 3). The com-  
107 parative analyses of the operators and their results are presented in Sect. 4, which lead us to the  
108 recommendations in Sect. 5.

## 109 **2. The Analogue Method**

110 The purpose of this paper is not to present in details the AM and its different parameterizations,  
111 which can be found in Horton et al. (2016a) or Ben Daoud et al. (2015). The concepts one needs to  
112 understand here are the following: the AM consists in searching a certain number of past situations  
113 in a meteorological archive (typically reanalysis datasets) that are the most similar, according to an  
114 analogy criteria, and to extract the observed values of the local weather variable of interest from  
115 another archive (typically local or basin averaged time series) in order to build the conditional  
116 empirical distribution considered as the probabilistic forecast for the target day. Predictors can  
117 be varied: for example the geopotential height at different pressure levels on different temporal  
118 windows (time of observation). The method is usually made of several levels of analogy, leading  
119 to successive subsampling on predictors of different nature (eg. atmospheric circulation, moisture  
120 variables, vertical motion, and air temperature).

121 There are mainly 2 parameterizations that are most often used for precipitation forecasting: one  
122 that relies on an analogy of the atmospheric circulation, and another that adds a second level of  
123 analogy on moisture variables (Obled et al. 2002; Bontron and Obled 2005; Marty et al. 2012).

124 The method based on the analogy of the synoptic circulation consists in the following steps: the  
125 similarity of the atmospheric circulation of a target date with every day of the archive is assessed

by processing the S1 criteria (Teweles and Wobus 1954; Drosowsky and Zhang 2003), which is a comparison of gradients, on 2 different the geopotential heights (e.g. 500 and 1000 hPa) and over a certain spatial window. To cope with seasonal effects, candidate dates are extracted within a period of 4 months centered around the target date, for every year of the archive. The  $n_1$  (parameter to calibrate) dates with the lowest values of S1 are considered as analogues to the target day. Then, the daily observed precipitation amount of the  $n_1$  resulting dates provide the empirical conditional distribution considered as the probabilistic forecast for the target day.

The other current parametrization adds a second level of analogy on moisture variables, more specifically on a moisture index made of the product of the precipitable water with the relative humidity at a certain pressure level (e.g. 850 hPa, see Bontron 2004). When adding a second level of analogy,  $n_2$  dates are subsampled in the  $n_1$  analogues on the atmospheric circulation, to end up with a smaller number of analogue situations.

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  $l$  can be written:

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

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.

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

$$CRPSS = \frac{CRPS - CRPS_r}{CRPS_p - CRPS_r} = 1 - \frac{CRPS}{CRPS_r} \quad (2)$$

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

### 152 3. Assessed Genetic Algorithms variants

153 Genetic Algorithms (GAs, Holland 1992; Goldberg 1989) come from the world of stochastic  
 154 optimization, more specifically from metaheuristic approaches. These are stochastic iterative al-  
 155 gorithms that behave like search algorithms by exploiting the characteristics of a problem and are  
 156 particularly suitable for complex parameter spaces.

157 GAs are part of the family of Evolutionary Algorithms (Bäck and Schwefel 1993; Schwefel  
 158 1993), inspired by some mechanisms of biological evolution, such as reproduction, genetic mu-  
 159 tations, chromosomal crossovers, and natural selection. Unlike a linear or local optimization,  
 160 GAs seek the global optimum on a complex surface, theoretically without restriction, but with no  
 161 guarantee to reach it.

#### 162 a. Basic concepts of the Genetic Algorithms

163 GAs mimic the evolution of a population of individuals in a new environment, by applying rules  
 164 based on natural processes, such as DNA mutation, chromosomes crossover, natural selection, etc.

Generation after generation, the DNA mixes and the best suited, or best adapted, genes cumulate in some individuals (Beasley and Chu 1996).

Applications of GAs are diversified and can handle problems of various types (Joines et al. 1996), even with very complex cost surfaces (Haupt and Haupt 2004). The objective function to optimize (often named fitness function in this context) can be of different types, but GAs must be adapted in order to perform optimally.

Two conditions guarantee in theory the convergence to the global optimum (Zitzler et al. 2004): (1) Parameters mutations that can allow to explore the entire parameter space, thereby ensures that any value can be achieved with a non-zero probability. (2) A rule of elitism ensuring that an optimal solution cannot be lost or damaged.

A key element of the parametrization of GAs is finding the right balance between exploration and exploitation (Bäck 1992a; Smith and Fogarty 1997). Exploration is characterized by a relatively high probability to assess the regions of the parameter space that have not yet been visited. This probability must be sufficiently large at the beginning of the optimization, so that the algorithm is capable of identifying the region where the global optimum is likely located. Exploitation is characterized by a local search in an area of interest, and generally makes small movements. The latter is interesting to refine the results at the end of the optimization.

#### *b. Structure and operators*

The GAs optimize a population of  $N$  individuals (parameter sets). Each individual contains a chromosome (set of parameters of the AM in this case). Genes are the individual parameters constituting the chromosome. They can be either categorical (e.g. geopotential, or temperature), discrete (e.g. number of analogues to select, from 1 to 40), or continuous. For example, a chromosome may look like:

$$[p_{i,j}, h_{i,j}, \lambda w_{i,j}, \lambda e_{i,j}, \varphi s_{i,j}, \varphi n_{i,j}, p_{i,j+1}, h_{i,j+1}, \dots, n_i, \dots, p_{i+1,j}, h_{i+1,j}, \dots, n_{i+1}, \dots] \quad (3)$$

where, for the level of analogy  $i$  and the predictor number  $j$ ,  $p_{i,j}$  is the pressure level,  $h_{i,j}$  is the temporal window,  $\lambda w_{i,j}$  and  $\lambda e_{i,j}$  are the western and eastern longitudes of the spatial window,  $\varphi s_{i,j}$  and  $\varphi n_{i,j}$  are its southerly and northerly latitudes, and  $n_{i,j}$  is the number of analogues. A chromosome for the simple AM on the atmospheric circulation only (Sect. 2) may be:  $[500 \text{ hPa}, 12 \text{ h}, -2.5^\circ, 15^\circ, 40^\circ, 55^\circ, 1000 \text{ hPa}, 24 \text{ h}, 0^\circ, 15^\circ, 42.5^\circ, 47.5^\circ, 28]$

The floating-point representation (instead of the canonical binary, see Goldberg 1989, 1990; Holland 1992) of the genes was considered, as it was found more suited in multiple applications (Michalewicz 1996; Herrera et al. 1998; Haupt and Haupt 2004; Bäck and Schütz 1996; Gaffney et al. 2010).

There are numerous implementation variants of GAs, often optimal for a given problem (Hart and Belew 1991; Schraudolph and Belew 1992). The divergences are in the operators implementation, through significantly different algorithms, which has an important effect on the results (Gaffney et al. 2010). Here, operators are defined as the mechanisms that modify the values of the genes to try bringing individuals (or chromosomes) closer to an optimum of the fitness function. The structure of the method (Fig. 1) resulting from the work of Holland (1992) is common to most applications (Bäck and Schwefel 1993), and consists in the following steps:

1. A population of  $N$  individuals (parameter sets of the AM) is randomly generated, which constitutes the initial population.
2. The fitness (performance score or objective function) of every individual is assessed.
3. A natural selection is applied, after which only the best individuals remain, which constitutes the intermediate generation (IG).

209 4. From this IG, couples are formed according to given rules.

210 5. These couples proceed to reproduction, or chromosome crossover, to mix their genes (param-  
211 eters) according to the selected operator version. New children are generated in order to refill  
212 the IG back to  $N$  individuals.

213 6. Parents and children are then subject to mutation, where some genes get affected, meaning  
214 some parameter values are randomly changed.

215 7. The new formed generation is then re-assessed.

216 8. According to the ending criteria, the optimization ends or starts again for another iteration.

217 All considered operators and their options, applied to floating-point coding, are described in the  
218 following sections. Many other operators exist, but only the ones evaluated are presented.

## 219 1) GENESIS OF THE POPULATION

220 The first step of the optimization is to generate an initial population. A population is a set of  $N$   
221 individuals (each of which represents, in this application, a parameter set of the AM and thus a  
222 point in the space of potential solutions) that are going to evolve. A generation is the population  
223 considered at a given time.

224 A random initialization based on a uniform sampling is the most current version. The size  $N$  of  
225 the population is often a compromise between the computation time and the quality of the solution.  
226  $N$  must allow sufficient sampling of the solutions field (Beasley and Chu 1996), and should thus  
227 vary as a function of chromosome size (ie the number of genes or parameters to be optimized).

## 228 2) NATURAL SELECTION

229 Natural selection is performed on the basis of the values of the objective function. The selection  
230 allows to only keep a certain part of the population, usually half ( $N/2$ ), which can access the IG  
231 (with  $N_{IG}$  members). If  $N_{IG}$  is too high, the reproduction rate is too low, whereas if it is too small,  
232 the strong traits of individuals will not have the ability to accumulate in the same chromosome  
233 (Haupt and Haupt 2004). Several techniques exist, such as:

- 234 •  $N_{IG}$ -elitism (Michalewicz 1996): the population is ranked according to the value of the ob-  
235 jective function and only the better half is preserved.
- 236 • Tournament selection (Michalewicz 1996; Zitzler et al. 2004): two individuals are randomly  
237 selected and fight. The one with the highest performance score is chosen, but with a certain  
238 probability, in order to reduce the selection pressure. This procedure is repeated until the IG  
239 is full. Individuals can be selected several times, and thus be represented several times in the  
240 IG.

## 241 3) SELECTION OF THE COUPLES

242 Individuals of the IG can reproduce. It begins with the selection of pairs (the parents). The  
243 techniques implemented in this work are the following:

- 244 • Rank pairing: individuals are gathered in pairs according to their rank (based on the perfor-  
245 mance scores). Consecutive ranks are put together (odd rows are associated with even rows).
- 246 • Random pairing: two individuals are randomly selected to form a couple, according to a  
247 uniform law.
- 248 • Roulette wheel weighting: this technique refers to gambling, but with a biased roulette. Each  
249 individual is associated with a sector of the wheel with a certain opening angle, which is

its probability of selection. The probability ( $p_n$ ) assigned to the individuals is proportional to their fitness (objective function), so that the most adapted individuals have the greatest probability of reproduction. There are two techniques for weighting the individuals of the IG:

*Roulette wheel weighting on rank*: the probability of each individual depends on its rank  $n$ :

$$p_n = \frac{N_{IG} - n + 1}{\sum_{n=1}^{N_{IG}} n} \quad (4)$$

*Roulette wheel weighting on fitness*: the selection probability is calculated based on the value of the objective function:

$$p_n = \frac{score_n - score_{N_{IG}}}{\sum_{n=1}^{N_{IG}} (score_n - score_{N_{IG}})} \quad (5)$$

In this application, the last individual ( $N_{IG}$ ) has zero probability of being selected.

- **Tournament selection**: This operator is similar to the one used in natural selection, but is applied here for the successive selection of each parent. To select a parent, a number of individuals (2 or 3) are randomly picked and the best is kept. This operation is performed twice, once for each partner and imitates the breeding competition in nature (Haupt and Haupt 2004).

#### 4) CHROMOSOME Crossover

Once the two parents are selected for breeding, they combine their chromosomes and produce two children, bringing the number of individuals in the population back to  $N$  (the parents also return back in the total population in order to complement the next generation). The combination of chromosomes is carried out using a crossover operator, thereby generating two offspring having characteristics derived from both parents. It allows a mixing of genes and a potential accumulation of positive mutations.

The evaluated crossover operators are the following:



- Single-point crossover: within the chromosome, a crossover point is randomly chosen for the pair. The genes (AM parameters) located after that point are exchanged in between the two chromosomes.
- Two-point crossover: works like the single-point crossover, but there are two intersections defining the segments to be exchanged. This approach, which significantly extends the search space for the children, is considered more efficient than the previous (Beasley et al. 1993).
- Multiple-point crossover (De Jong 1975): a generalization of the previous, with a number of crossover points up to the number of genes.
- Uniform crossover (Syswerda 1989): for each gene of the chromosome, it is randomly chosen to exchange or not the values between the parents.
- Binary-like crossover (Haupt and Haupt 2004): chromosome crossover on a binary coding can generate new values for variables located at intersection points, since the crossovers are applied at the bit level, thus often within a gene. This is not the case for the floating-point representation, since the crossover is performed between the genes. To reproduce the behaviour present in the original algorithms, which introduces new information, Haupt and Haupt (2004) propose an operator that combines standard crossover with an interpolation approach. The genes located after a crossover point are exchanged, but the gene located at the intersection is modified according to:

$$\begin{cases} g_{o1,n} = g_{p1,n} - \beta(g_{p1,n} - g_{p2,n}) \\ g_{o2,n} = g_{p2,n} + \beta(g_{p1,n} - g_{p2,n}) \end{cases} \quad (6)$$

where  $g_{o1,n}$  and  $g_{o2,n}$  are the  $n$ -th gene of the two new offspring, and  $g_{p1,n}$  and  $g_{p2,n}$  are those of the two parents.  $\beta$  is a random value between 0 and 1.

- Blending method (Radcliffe 1991): in this approach, instead of exchanging the genes in between the chromosomes after one or multiple crossover points, these are combined by linear combination (Eq. 7). The genes of the parents are blended together using a random value ( $\beta$ ) that can be unique for the whole chromosome, or that can change for every gene. The genes of the offspring are bounded by the genes of the parents, no value can be out of their range.

$$\begin{cases} g_{o1,n} = \beta g_{p1,n} + (1 - \beta) g_{p2,n} \\ g_{o2,n} = (1 - \beta) g_{p1,n} + \beta g_{p2,n} \end{cases} \quad (7)$$

- Linear crossover (Wright 1991): in order to allow the genes to take values outside the interval defined by the parents, a method of extrapolation is necessary. Linear crossover introduces such an approach, and produces three children from two parents, following Eq. 8. Less couples are required in order to fill up the population.

$$\begin{cases} g_{o1,n} = 0.5 g_{p1,n} + 0.5 g_{p2,n} \\ g_{o2,n} = 1.5 g_{p1,n} - 0.5 g_{p2,n} \\ g_{e3,n} = -0.5 g_{p1,n} + 1.5 g_{p2,n} \end{cases} \quad (8)$$

- Heuristic crossover (Michalewicz 1996): it is a variation of the latter methods that relies on the following equation:

$$\begin{cases} g_{o1,n} = \beta (g_{p1,n} - g_{p2,n}) + g_{p1,n} \\ g_{o2,n} = \beta (g_{p2,n} - g_{p1,n}) + g_{p2,n} \end{cases} \quad (9)$$

- Linear interpolation: unlike previous techniques, this technique does not rely on crossover points, but on a linear interpolation on every gene of the couple (Eq. 10).

$$\begin{cases} c_{o1} = c_{p1} - \beta (c_{p1} - c_{p2}) \\ c_{o2} = c_{p2} + \beta (c_{p1} - c_{p2}) \end{cases} \quad (10)$$

where  $c_{o1}$  and  $c_{o2}$  are the full chromosomes of the offspring, and  $c_{p1}$  and  $c_{p2}$  are the ones of the parents. As before,  $\beta$  is a random value between 0 and 1, and is here the same for every gene.

- Free interpolation: this technique performs interpolation on each gene, like the previous one; but in this case, the weighting factor changes for each gene:

$$\begin{cases} c_{o1} = c_{p1} - [\beta_1(g_{p1,1} - g_{p2,1}), \beta_2(g_{p1,2} - g_{p2,2}), \dots, \beta_{N_g}(g_{p1,N_g} - g_{p2,N_g})] \\ c_{o2} = c_{p2} + [\beta_1(g_{p1,1} - g_{p2,1}), \beta_2(g_{p1,2} - g_{p2,2}), \dots, \beta_{N_g}(g_{p1,N_g} - g_{p2,N_g})] \end{cases} \quad (11)$$

where  $N_g$  is the number of genes, and  $\beta$  is here independent between the genes.

## 5) MUTATION

The combination of strong genes by the operator of chromosomes crossover is theoretically the most important operating mechanism in the conventional GAs (Holland 1992; Bäck and Schwefel 1993). However, many studies identify the mutation process as main operator, and crossovers as secondary (see Bäck 1992a, 1996; Bäck and Schütz 1996; Smith and Fogarty 1997; Deb and Beyer 1999; Costa et al. 2005, 2007).

The mutation operator is a direct modification of genes. In real coding, it is done by changing the gene values. Mutations add diversity to the population and prevent a freeze of the evolution, or a genetic drift to a local optimum. Thus, it makes the convergence to the global optimum theoretically possible (Beasley et al. 1993), as they allow exploring beyond the current region of the parameter space by bringing new characteristics that were not present in the original population (Haupt and Haupt 2004).

321 The evaluated and developed mutation operators are listed hereafter. They apply to genes made  
 322 of continuous or discrete variables, but not categorical (eg meteorological variable or analogy  
 323 criterion). In the latter case, the random choice of a new value is always based on a uniform  
 324 distribution, without notion of distance in the parameters space.

- 325 • Uniform mutation: The mutation rate is constant and equal for every gene of each individual;  
 326 they all have the same probability to mutate. When a gene is selected for mutation, a new  
 327 random value is assigned, according to a uniform law.
- 328 • Variable uniform mutation (Fogarty 1989): a variable mutation rate over the generations was  
 329 first suggested by Holland (1992) and evaluated by Fogarty (1989). In most applications, the  
 330 mutation rate decreases with the generations, in a deterministic and global (for all individuals)  
 331 manner (Bäck 1992b). Its optimum configuration depends on the size of the chromosomes,  
 332 of the properties of the objective function, and of the population size (Bäck 1992b). This  
 333 operator was implemented according to:

$$p_{n,G} = p_{G_0} + \left( \frac{p_{G_0} - p_{G_{m,p}}}{G_{m,p}} \right) \min \{ G, G_{m,p} \} \quad (12)$$

334 where  $p_{n,G}$  is the mutation rate (probability) of the gene  $n$  for generation number  $G$ ,  $G_{m,p}$  is  
 335 the maximum number of generations during which the mutation rate varies.  $p_{G_0}$  is the initial  
 336 mutation probability, and  $p_{G_{m,p}}$  is the final one.  $p_{G_0}$ ,  $p_{G_{m,p}}$  and  $G_{m,p}$  are the three controlling  
 337 parameters of the operator. The evolution of the mutation rate is linear.

- 338 • Constant normal mutation: many applications use normal distributions to generate new val-  
 339 ues. The gene  $g$  that mutate becomes:

$$g' = N(g, \sigma^2) \quad (13)$$

340 where  $\sigma$  is the standard deviation of the distribution, which has to be estimated.

- Variable normal mutation (new): with the same logic as the variable uniform mutation, a mutation operator was tested using a normal distribution with a variable mutation rate and standard deviation. The mutation rate is calculated with Eq. 12. On the same principle, the standard deviation decreases linearly over the generations:

$$\sigma_{n,G} = \sigma_{G_0} + \left( \frac{\sigma_{G_0} - \sigma_{G_{m,\sigma}}}{G_{m,\sigma}} \right) \min \{G, G_{m,\sigma}\} \quad (14)$$

where  $\sigma_{n,G}$  is the standard deviation of gene  $n$  and generation number  $G$ ,  $\sigma_{G_0}$  is the initial standard deviation,  $\sigma_{G_{m,\sigma}}$  is the final standard deviation,  $G_{m,\sigma}$  is the maximum number of generations during which the standard deviation varies.  $p_{G_0}$ ,  $p_{G_{m,p}}$ ,  $G_{m,p}$ ,  $\sigma_{G_0}$ ,  $\sigma_{G_{m,\sigma}}$  and  $G_{m,\sigma}$  are the six parameters of the method.

- Non-uniform mutation (Michalewicz 1996): two random numbers are picked based on a uniform law:  $r_1$ , which determines the direction of the change, and  $r_2$ , which determines its magnitude. The new value of the gene is given by the following equation, according to a predefined number of generations:

$$g'_n = \begin{cases} g_n + (b_n - g_n) r_2 \left(1 - \frac{G}{G_m}\right)^2 & \text{if } r_1 < 0.5 \\ g_n - (g_n - a_n) r_2 \left(1 - \frac{G}{G_m}\right)^2 & \text{if } r_1 \geq 0.5 \end{cases} \quad (15)$$

where  $a_n$  is the lower bound of the  $n$ -th gene,  $b_n$  its upper bound,  $G$  the present generation, and  $G_m$  the maximum number of generations.

This operator was adapted for this application, which is not based on a predefined number of generations:

$$g'_n = \begin{cases} g_n + (b_n - g_n) r_2 \varphi^2 & \text{if } r_1 < 0.5 \\ g_n - (g_n - a_n) r_2 \varphi^2 & \text{if } r_1 \geq 0.5 \end{cases} \quad (16)$$

with

$$\varphi = 1 - \min \left\{ \frac{G}{G_{m,r}}, 1 \right\} (1 - \omega) \quad (17)$$

where  $G_{m,r}$  is the maximum number of generations during which the magnitude of the search varies, and  $\omega$  is a threshold chosen by the user to maintain a minimum search radius when  $G > G_{m,r}$ . During the first generations, the exploration extent covers the entire parameter space. However, this area is reduced over generations, allowing exploitation of local solutions.

- Individual adaptive mutation rate (Bäck 1992a): based on the ideas of Evolution Strategies (see Rechenberg 1973; Schwefel 1981), Bäck (1992a) introduced a concept of self-adaptive GAs. The idea is to distribute control parameters within individuals themselves, which partially decentralize control of the evolution. It allows reducing the parametrization of GAs and introducing a notion of self-management. The first approach is the introduction of a mutation rate per individual, that mutates itself under its own probability (Bäck 1992a). Then, the eventual new rate is used to mutate the genes of the individual. Thus, as this rate decreases, it will have less probability of being itself mutated. Mutations are performed according to a constant uniform distribution. The initial mutation rates are randomly chosen (Bäck 1992a) and the method has no parameter. Other approaches exist to introduce a self-adaptation (see Smith and Fogarty 1997; Deb and Beyer 1999, 2001).
- Individual adaptive search radius (new): based on the ideas of the non-uniform mutation, a search radius was introduced in the approach of individual adaptive mutation rates. This search radius  $r_a$ , bounded between 0 and 1 (relatively to the parameters ranges), is also adaptive and behaves similarly to the adaptive mutation rates. In order to separate its evolution from the one of the mutation rate, its own value is considered initially as a self-mutation rate

to eventually mutate before being used as a normalized search radius. The value of a mutated gene is given by the following equation, which is a simplification of the non-uniform mutation:

$$g'_n = \begin{cases} g_n + (b_n - g_n) r_2 r_a & \text{if } r_1 < 0.5 \\ g_n - (g_n - a_n) r_2 r_a & \text{if } r_1 \geq 0.5 \end{cases} \quad (18)$$

where  $r_1$  and  $r_2$  are randomly selected, in the same way as for the non-uniform mutation. No external parameter is therefore necessary.

- Chromosome of adaptive mutation rate (or *n adaptive mutation rate*, Bäck 1992a): analogously to the individual adaptive mutation rate, this approach leaves the control of the evolution rate to the individuals themselves. The difference here is that each gene has a specific mutation rate. The main advantage is that the tuning of the mutation can be much more precise (Smith and Fogarty 1997). A second chromosome containing the mutation rate for each gene of the first chromosome was therefore considered. The operations of mutation and self-mutation are similar to the case of the individual adaptive mutation rate, but in a distributed way, within the chromosome. Another difference is that the same crossover operations are applied as those applied to the first chromosome, and this for the same crossing points. Thus, during an exchange of genes, children also inherit the mutation rates specific for each of these genes.
- Chromosome of adaptive search radius (new): this operator combines the operations of the chromosome of adaptive mutation rate to the adaptive search radius approach. Similarly, an individual has 3 chromosomes: the first containing the values to be optimized, the second the distributed mutation rate, and the last one, the distributed search radius. Again, no external parameters are required.

- Multi-scale mutation (new): finally, another approach was developed that is also based on the search radius concept. However, the latter is not decreasing with time. Methods based on a reduction of the mutation rate or radius simulate a transition from the exploration phase to the exploitation one. The idea was to test an approach that combines both exploration and exploitation during the whole optimization. Thus, the search radius  $r_a$  of Eq. 18 was considered as a random value for each individual, but restricted to 4 equiprobable values: 1, 0.5, 0.1, 0.02, which range from full exploration to fine exploitation. The only external parameter is the mutation rate which is fixed.

## 6) ELITISM

A process of elitism was introduced on the natural selection as well as on mutations. This ensures the survival of the best individual so that a better solution is never lost. After the natural selection operator, if the previously best individual has not been selected, it is copied to the IG instead of an individual randomly picked. After mutation, if the previously best individual has mutated and if its new version has a lower performance score than the original, the latter is also reinserted in the IG instead of an individual randomly chosen.

## 7) ENDING THE OPTIMIZATION

The convergence check determines whether the solution is acceptable and if the algorithm may stop. The stopping criteria are not often well documented in GAs case studies. The optimization is here stopped if the best individual does not change for  $x$  generations. This value should not be too low to allow the algorithm to escape from a local optima. In addition, the rate of improvement decreases with the progression of the optimization. It is thus common that the best individual does



not evolve over several generations when getting closer to the global solution. A value of  $x = 20$  generations was chosen.

### *c. Implementation and constraints*

Some constraints need to be taken into account. For example, when a crossover or a mutation operation results in a parameter value standing out of the authorized bounds, it has to be brought back within the limits. Moreover, the parameters are of different nature: some are continuous, some are discrete, and finally, some are categorical, i.e. independent elements in an array, such as the selection of the meteorological variable. New values resulting from the optimizer need to respect the type of data it represents.

Other constraints exist in between the parameters, such as the temporal window of the moisture index (see Horton et al. 2016a) that has to be consistent in between the relative humidity and the precipitable water.

GAs are very computationally intensive because they require many evaluations of the objective function. These assessments are very long in this application, as they require calculating and assessing a forecast for every day of the calibration period, thus over several decades. In order to reduce the computation time, recalculating the performance score of an individual who has previously been evaluated and that has not changed was avoided. Thus, the score of each individual living in the selection was kept until it mutates.

As the assessment (calculation of the objective function) of each member of the population of a generation is completely independent, it was performed in parallel on different processors of a computer. In order to perform optimizations for multiple time series, the use of a cluster is a necessity, which our code allows.

#### 4. Assessment process and results

The GAs parametrization, i.e. the a priori choices such as the mutation rate, population size, natural selection options, etc, appears difficult given the high number of existing variants, each developed for a specific problem (Haupt and Haupt 2004; Costa et al. 2007). The parametrization depends on the objective function, implementation variants, the range of the parameters to be optimized, and performance indicators. Thus, different studies suggest very different parametrizations.

De Jong (1975) and Grefenstette (1986) compared different implementations and parametrizations of GAs on functions of varying complexity. They observed that a small population size improves the initial performance, while a large population improves latter performance. They also observed that the proportion of the population to keep for the IG is around 50% (45% to 60%).

Values of the mutation rate ( $p$ ) varies broadly between studies: from 0.001 (De Jong 1975) to 0.2 (Haupt and Haupt 2004). Bäck and Schütz (1996) showed that mutation rates higher than the usual ranges are more optimal at the beginning of the optimization, allowing further exploration. Varying mutation rates are certainly more optimal but more complex to implement (Bäck 1996; Bäck and Schütz 1996).

##### *a. Comparison process*

One of our goals being to make recommendations of parametrization in view of optimizing the AM, a systematic procedure was adopted. The results are summarized hereafter (see Horton 2012, for the details). Concepts were taken from the factorial design approach (see eg. Costa et al. 2005, 2007; Mariano et al. 2010), which is sometimes used for comparative analysis of different parametrizations of GAs. It allows isolating the effect of a parameter under different combinations of the other options. A procedure by stages was adopted, analyzing in details and in a systematic

way every variants of the implemented operators, in combination with multiple other options and parameters in order to take into account eventual co-dependencies. The goal here is not to focus on the performance score obtained through optimization, neither the values of the new optimized parameters (covered in Horton et al. 2016b), but to explain how to use GAs to optimize AMs in an efficient way.

In order to evaluate a combination of operators/options, 10 optimizations per parametrization of GAs were processed. Such assessment is not possible on the whole archive length, and had to be performed on a reduced period. The performances were characterized by four indicators: (i) mean performance score: average of the final scores of the 10 optimizations, (ii) convergence: the number of optimizations that converged to a supposed global optimum, (iii) number of generations: characterization of the convergence speed, and (iv) number of evaluations of the objective function: characterization of the required calculation time.

#### *b. Success of the approach*

After a first overview of the results, GAs have quickly proved successful at optimizing the AM. The performance score of the sequential approach has been quickly exceeded without adding any new parameter to the method (Fig. 2). Even GAs parameterizations that will be considered later as inadequate (Fig. 3) did significantly better than the sequential approach. The amplitude of the improvement is not the main outcome here. The most important point is that GAs proved successful at optimizing AMs automatically, globally, and objectively.

#### *c. Results of the comparison*

The results illustrate the effect of an operator when its contribution is isolated from the other operators. It means that we analyze the effect of a given operator for equivalent conditions (same

488 settings of other operators). Multiple combinations with other operators are assessed. This con-  
489 tribution is then summarized as a percentage of gain/loss regarding the mean of all variants, for  
490 equivalent external conditions. For example, to evaluate the performance of the uniform crossover  
491 operator, its performance is compared to the average of all crossover operators while retaining the  
492 same population size, the same mutation operators, natural selection, and selection of couples.

493 From the very beginning of the assessments, the importance of the mutation operator was obvi-  
494 ous (see Horton 2012, for the details), and its leading influence on the optimization performance  
495 was evident. Its role is analyzed later on.

## 496 1) BREEDING OPERATORS

497 Every combination of 6 options for the couples selection (Table 1) and 21 for the chromosome  
498 crossover operators (Table 2) were evaluated, along with variants of the other operators. This  
499 resulted in 1,008 combinations, requiring 10,080 optimizations.

500 The performance of the couples selection operator are relatively close (Fig. 4). Overall, the  
501 tournament selection with 3 candidates is slightly superior to others, along with the roulette wheel  
502 weighting. This last one is however a bit less effective in terms of convergence and number of  
503 evaluations (not shown). The couples selection operator has not a significant role in this applica-  
504 tion.

505 Analysis of crossover operators (Fig. 5) reveal some slightly superior options, some inappropri-  
506 ate, and many average. Binary-like crossover (especially with 2 points of intersection, whether  $\beta$   
507 is shared or not) are significantly better than the others, especially in terms of convergence (not  
508 shown). Others operators can also be considered usable.

## 509 2) MUTATION OPERATOR

510 Having identified the leading role of the mutation operator, the next sensitivity analysis focused  
511 on it. Each of the 10 different implementations (see Sect. 5) was tested, with different param-  
512 eters for those who require some (Table 3), bringing the number of variations up to 109. Some  
513 optimizations without any mutation were also performed as a reference. Along with variants of  
514 the other operators (see Horton 2012, for the details), this resulted in 660 combinations (so 6,600  
515 optimizations).

516 Figure 6 show the results of this analysis and illustrates the important role of the mutation on  
517 the performance of the optimizations. Those without mutation (last box on the figure) are inferior  
518 to most mutation operators, and the scale of the influence of this operator is significantly more  
519 important than those for the other options. The details of the analysis (see Horton 2012) show that  
520 the other reproduction operators seem of secondary importance. This observation is in line with  
521 the work of Bäck (1996), who argues for the importance of mutation over reproduction, as well as  
522 other authors (see Sect. 3.5).

523 The mutation operators based on a variable normal or variable uniform laws work very poorly  
524 and are difficult to configure. Many operators present more or less the same performance scores  
525 and require a variable amount of assessments. The convergence analysis (see Horton 2012) allows  
526 to highlight three best operators: non-uniform mutation, chromosome of adaptive search radius,  
527 and multi-scale mutation. Thus, different optimizations were further performed using variants of  
528 these 3 operators (Table 4).

529 The first analysis was the optimization of the precipitation forecasting over a subcatchment  
530 (Binn-Simplon region) in the Swiss Alps (Fig. 7). The optimizer could choose the 2 pressure lev-  
531 els of the atmospheric circulation analogy (method with a single level of analogy). The resulting

CRPSS performance score (Sect. 2) is obviously superior to the one obtained by the sequential calibration (red line on Fig. 7). For most options, it is also slightly better than the results from the optimization without selection of the pressure levels (blue line). A clear breakthrough of the performances was not expected, as the former selection of pressure levels results already from intensive comparative work (Bontron 2004). This application however demonstrates that, when correctly parametrized, GAs can automatically and successfully choose the pressure levels. However, some less relevant optimizations do not match the previous results. Through different applications, the automatic selection of the pressure level was shown to significantly increase the difficulty for GAs to converge to a unique solution, ideally the global optimum. A difficulty is that the pressure levels are considered within the optimization without continuity between values, and thus the approaches relying on distance in the parameters space, such as the search radius, cannot fully exploit the properties that make them efficient. However, even though the results show a certain variability, most of them present very good performance scores, despite different parameters of the AM.

Then, the same optimization was performed, but for another region, sensitive to other meteorological influences (Fig. 8), in order to assess eventual dependencies of the operator with the predictand. Even though differences can be observed with Fig. 7, it is globally the same options that perform better.

Next, a second level of analogy was proposed (Fig. 9) based on moisture variables (see Bontron 2004; Horton et al. 2016a). GAs had to optimize both levels of analogy simultaneously. Once again, despite the difficulty to do so, the results were better than the sequential calibration (red line on Fig. 9). And finally, a preselection on air temperature was added instead of the fixed calendar window, as proposed by Ben Daoud et al. (2015). The results show generally higher scores (Fig. 10), demonstrating the success of the optimizer to take advantage of this new degree of freedom,

556 and its capacity to handle optimization of 3 analogy levels jointly. Again, the most relevant options  
557 are generally the same.

558 After analysis of the most relevant mutation operators, the following advice can be raised (de-  
559 tailed parametrization are provided in Sect. 5):

- 560 • *Non-uniform mutation* (Michalewicz 1996): this operator is good in terms of convergence,  
561 mainly when the number of parameters to optimize is rather low. The number of required  
562 evaluations, however, can be quite substantial. The main disadvantage of the non-uniform  
563 mutation is the number of parameters it requires, that are difficult to estimate a priori. The  
564 mutation rate was found to be more important than the others. The difficulty is that its optimal  
565 value may be case-related.
- 566 • *Chromosome of adaptive search radius* (new): unlike the previous one, this new operator is  
567 very robust, as it requires no option and is auto-adapting. It is interesting to notice that the  
568 insertion of an extra chromosome representing the search radius gives better performance than  
569 other self-adaptive operators (such as, for example, the chromosome of adaptive mutation  
570 rate). If one had to choose a single option for the mutation operator, we would recommend  
571 this one, as it was proven effective and needs no parameter.
- 572 • *Multi-scale mutation* (new): finally, the multi-scale mutation, which also performs pretty  
573 well, requires one parameter, the mutation rate. However, it can also be difficult to estimate a  
574 correct value a priori.

575 In this application, the mutation operator has a leading effect and should be chosen with care. It  
576 may be wise to perform multiple optimizations and to consider these three operators in parallel in  
577 order to obtain results from options that are sometimes either more efficient or more robust. It is  
578 interesting to note that the three best techniques incorporate a notion of search distance. It is likely

579 that this notion is the key to these algorithms, for this application, and allows them to initially  
580 explore the parameter domain, and then to converge. The search radius in fact directly represents  
581 the notion of transition between exploration and exploitation, in our opinion more than a possible  
582 evolution of mutation rates.

### 583 3) OTHER OPTIONS

584 The analysis of the natural selection operator (Fig. 11) reveals a slight preference for the ratio-  
585 elitism compared to the tournament selection, but not so significant. This operator, or at least the  
586 two assessed versions, do not appear to significantly influence the optimization performances.

587 The size of the population ( $N$ ), i.e. the number of sets of AM parameters considered, has an  
588 effect on the performance of the optimization (Fig. 12). A bigger population leads to better  
589 results, but also to significantly longer optimizations. Indeed, the required number of evaluation,  
590 and thus the required time, is approximately proportional to the population size. The optimal size  
591 seems to depend on the complexity of the AM to optimize: a more complex AM (ie. with more  
592 degrees of freedom) requires a bigger population size. A rule of thumb based on a limited number  
593 of case studies (not shown here) is provided hereafter:

- 594 •  $N \approx 100$  for very simple implementations of the AM (1 level of analogy with 2 pressure  
595 levels),
- 596 •  $N \approx 200$  for a slightly more complex AM (1 level of analogy with 4 pressure levels, or 2 level  
597 of analogy with less pressure levels),
- 598 •  $N \approx 500$  for significantly more complex AMs (2-3 levels of analogy with 4 pressure levels  
599 for the atmospheric circulation, and 2 to 4 levels for the moisture analogy).



600 The influence of the size of the IG (proportion of the total population) selected for mating was  
601 also assessed (Fig. 13). It does not appear that this parameter is critical to the quality of the  
602 optimizations, provided it is not too big. A value of 50 % seems a wise choice.

## 603 **5. Recommended parametrization of GAs**

604 Optimizations by GAs for AMs of varying complexities were performed with a large number  
605 of combinations of operators in order to make recommendations for optimizing the AM. The  
606 conclusions are:

- 607 • The population size should be in accordance with the complexity of the AM to optimize:  
608 from 100 for the simple ones, up to 500 for the most complex AMs.
- 609 • The value of the ratio for the IG is not so important, and value of 50% seems quite appropriate.
- 610 • Ratio-elitism is slightly better than tournaments for the natural selection operator, but it is not  
611 decisive.
- 612 • The performance of the operators for the couples selection perform relatively similarly. The  
613 roulette wheel weighting and the tournament selection are more efficient in terms of conver-  
614 gence and required number of evaluations.
- 615 • Most crossover operators have relatively similar performance. Binary-like crossover with two  
616 points of intersection are better than others, especially for convergence.
- 617 • Mutation has a clearly dominant influence. Three mutation operators stand out: the non-  
618 uniform mutation, the multi-scale mutation, and the chromosome of adaptive search radius.  
619 The latter is the most robust as it has no controlling parameter.

620 The optimization does not systematically converge to the global optimum (but still often nearby),  
621 which is why it is recommended to do several optimizations in parallel in order to compare the  
622 results, analyze the convergence, and keep the best. It may be wise to consider the three mutation  
623 operators in parallel. In order to be confident in the optimized AMs, we propose using a set of the  
624 following mutation operators:

- 625 • 1x non-uniform,  $p_{mut} = 0.05$ ,  $G_m = 50$ ,  $\omega = 0.1$
- 626 • 1x non-uniform,  $p_{mut} = 0.05$ ,  $G_m = 100$ ,  $\omega = 0.1$
- 627 • 1x non-uniform,  $p_{mut} = 0.1$ ,  $G_m = 100$ ,  $\omega = 0.1$
- 628 • 1x multi-scale,  $p_{mut} = 0.1$
- 629 • 2x chromosome of adaptive search radius

## 630 6. Conclusions

631 In order to automatically optimize the AM and to get rid of the limitations of the usual sequen-  
632 tial calibration, GAs were evaluated. Given the large number of existing operators and options,  
633 multiple variants were assessed systematically in order to identify which operators are important,  
634 and which variants work best for the AM. The mutation operator was identified as a key element  
635 for this application, and new variants that proved efficient were provided, such as the chromo-  
636 some of adaptive search radius that is very robust (no control parameter). Recommendations were  
637 established for a relevant use of GAs for the optimization of AMs.

638 It is not excluded that another global optimization method or other operators of GAs may per-  
639 form still better. However, the relevance of such an approach has now been proved as it results in  
640 parametrization of AMs that are relevant and fully automatically, globally and objectively estab-  
641 lished. A global optimization is the only way to take into account all the dependencies between  
642 parameters and levels of analogy.

643 The global optimization approach allows easily adapting the AM to new regions by potentially  
644 taking into account local meteorological influences, and has thus a great potential of use. More-  
645 over, it allows exploring automatically datasets in order to extract the most relevant variables. It  
646 is thus possible to try assessing other predictands, such as the temperature, the limit of snowfall,  
647 the occurrence of hail, or wind, while leaving the algorithms select the best variables and the  
648 associated parameters.

649 *Acknowledgments.* Thanks to Hamid Hussain-Khan of the University of Lausanne for his help  
650 and availability, and for the intensive use of the cluster he is in charge of. Thanks to Dominique  
651 Bérodt for his support and to Michel Bierlaire for his advices on optimization methods.

652 Thanks to the Roads and Water courses Service, Energy and Water Power Service of the Wal-  
653 lis Canton, the Water, Land and Sanitation Service of the Vaud Canton, and the Swiss Federal  
654 Office for Environment (FOEV) who financed the MINERVE (Modélisation des Intempéries de  
655 Nature Extrême des Rivières Valaisannes et de leurs Effets) project which started this research.  
656 The fruitful collaboration with the Laboratoire d’Etude des Transferts en Hydrologie et Environ-  
657 nement of the Grenoble Institute of Technology (G-INP) was made possible thanks to the Herbette  
658 Foundation. NCEP reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado,  
659 USA, from their Web site at <http://www.esrl.noaa.gov/psd/>. Precipitation time series provided by  
660 MeteoSwiss.

## 661 **References**

662 Bäck, T., 1992a: Self-adaptation in genetic algorithms. *Proceedings of the First European Confer-*  
663 *ence on Artificial Life*, 263–271.

- 664 Bäck, T., 1992b: The Interaction of Mutation Rate, Selection, and Self-Adaptation Within a Ge-  
665 netic Algorithm. *Parallel Problem Solving from Nature*, **2**, 85–94.
- 666 Bäck, T., 1996: Evolution strategies: An alternative evolutionary algorithm. *Artificial Evolution*,  
667 3–20.
- 668 Bäck, T., and M. Schütz, 1996: Intelligent mutation rate control in canonical genetic algorithms.  
669 *Foundation of Intelligent Systems, 9th International Symposium*, 158–167.
- 670 Bäck, T., and H.-P. Schwefel, 1993: An Overview of Evolutionary Algorithms for Parameter  
671 Optimization. *Evolutionary Computation*, **1** (1), 1–23, doi:10.1162/evco.1993.1.1.1.
- 672 Barnett, T., and R. Preisendorfer, 1978: Multifield analog prediction of short-term climate fluctu-  
673 ations using a climate state vector. *Journal of the Atmospheric Sciences*, **35**, 1771–1787.
- 674 Beasley, D., R. Martin, and D. Bull, 1993: An overview of genetic algorithms: Part 1. Fundamen-  
675 tals. *University computing*, **15** (2), 58–69.
- 676 Beasley, J., and P. Chu, 1996: A genetic algorithm for the set covering problem. *European Journal*  
677 *of Operational Research*, **94** (2), 392–404, doi:10.1016/0377-2217(95)00159-X.
- 678 Ben Daoud, A., 2010: Améliorations et développements d’une méthode de prévision probabiliste  
679 des pluies par analogie. Ph.D. thesis, Université de Grenoble.
- 680 Ben Daoud, A., E. Sauquet, G. Bontron, C. Obled, and M. Lang, 2015: Daily quantitative precipi-  
681 tation forecasts based on the analogue method: improvements and application to a French large  
682 river basin. *Atmospheric Research*, **169**, 147–159, doi:10.1016/j.atmosres.2015.09.015.
- 683 Bergen, R., and R. Harnack, 1982: Long-range temperature prediction using a simple analog  
684 approach. *Monthly Weather Review*, **110**, 1083–1099.

- Bliefernicht, J., 2010: Probability forecasts of daily areal precipitation for small river basins. Ph.D. thesis, Universität Stuttgart.
- Bois, P., C. Obled, and J. Thalamy, 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. *C.R. du colloque Météorologie de l'énergie solaire, PIRDES-CNRS, Toulouse*, 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.
- 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, 262 p. pp.
- Bontron, G., and C. Obled, 2005: L'adaptation probabiliste des prévisions météorologiques pour la prévision hydrologique. *La Houille Blanche*, **1**, 23–28.
- Brown, T., 1974: Admissible Scoring Systems for Continuous Distributions. Tech. rep., P-5235 pp.
- Costa, C. B., E. a.C. Rivera, M. C. A. F. Rezende, M. R. W. Maciel, and R. M. Filho, 2007: Prior detection of genetic algorithm significant parameters: Coupling factorial design technique to genetic algorithm. *Chemical Engineering Science*, **62 (17)**, 4780–4801, doi:10.1016/j.ces.2007.03.042.
- Costa, C. B., M. R. W. Maciel, and R. M. Filho, 2005: Factorial design technique applied to genetic algorithm parameters in a batch cooling crystallization optimisation. *Computers & Chemical Engineering*, **29 (10)**, 2229–2241, doi:10.1016/j.compchemeng.2005.08.005.

707 De Jong, K. A., 1975: Analysis of the behavior of a class of genetic adaptive systems. Ph.D. thesis,  
708 University of Michigan.

709 Deb, K., and H. Beyer, 1999: Self-adaptation in real-parameter genetic algorithms with simu-  
710 lated binary crossover. *Proceedings of the Genetic and Evolutionary Computation Conference*  
711 *(GECCO), Orlando, USA*, 172–179.

712 Deb, K., and H. Beyer, 2001: Self-adaptive genetic algorithms with simulated binary crossover.  
713 *Evolutionary Computation*, **9** (2), 197–221.

714 Drosowsky, W., and H. Zhang, 2003: Verification of Spatial Fields. *Forecast verification: a*  
715 *practitioner’s guide in atmospheric science*, I. T. Jolliffe, and D. B. Stephenson, Eds., Wiley,  
716 chap. 6, 121–136.

717 Duband, D., 1970: Reconnaissance dynamique de la forme des situations météorologiques. Ap-  
718 plication à la prévision quantitative des précipitations. Ph.D. thesis, Thèse de 3ème cycle de la  
719 faculté des sciences de Paris.

720 Fogarty, T. C., 1989: Varying the probability of mutation in the genetic algorithm. *Proceedings*  
721 *of the third international conference on Genetic algorithms*, Morgan Kaufmann Publishers Inc.,  
722 104–109.

723 Fraedrich, K., C. C. Raible, and F. Sielmann, 2003: Analog Ensemble Forecasts of Tropi-  
724 cal Cyclone Tracks in the Australian Region. *Weather and Forecasting*, **18** (1), 3–11, doi:  
725 10.1175/1520-0434(2003)018<0003:AEFOTC>2.0.CO;2.

726 Gaffney, J., C. Pearce, and D. Green, 2010: Binary versus real coding for genetic algorithms: A  
727 false dichotomy? *ANZIAM Journal*, **51**, 347–359.

- 728 Goldberg, D., 1990: Real-coded genetic algorithms, virtual alphabets, and blocking. *Urbana*, **51**,  
729 1–21.
- 730 Goldberg, D. E., 1989: *Genetic Algorithms in Search, Optimization and Machine Learning*.  
731 Addison-Wesley Longman Publishing Co., Inc.
- 732 Gordon, N. D., 1987: Statistical very short-range forecasting via analogues. *European Space*  
733 *Agency Special Publication*, 307–312.
- 734 Grefenstette, J. J., 1986: Optimization of control parameters for genetic algorithms. *IEEE Trans-*  
735 *actions on Systems, Man and Cybernetics*, **16 (1)**, 122–128.
- 736 Guilbaud, S., 1997: Pr vision quantitative des pr cipitations journali res par une m thode  
737 statistico-dynamique des recherche d’analogues - Application   des bassins du pourtour  
738 m diterran en. Ph.D. thesis, Institut National Polytechnique de Grenoble, 386 p. pp.
- 739 Hart, W. E., and R. K. Belew, 1991: Optimizing an arbitrary function is hard for the genetic  
740 algorithm. *Proceedings of the Fourth International Conference on Genetic Algorithms*, 190–  
741 195.
- 742 Haupt, R. L., and S. E. Haupt, 2004: *Practical genetic algorithms*. John Wiley & Sons, 272 p. pp.
- 743 Herrera, F., M. Lozano, and J. Verdegay, 1998: Tackling real-coded genetic algorithms: Operators  
744 and tools for behavioural analysis. *Artificial Intelligence Review*, **12 (4)**, 265–319.
- 745 Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble pre-  
746 diction systems. *Weather and Forecasting*, **15 (5)**, 559–570.
- 747 Holland, J., 1992: Genetic algorithms. *Scientific American*, **267 (1)**, 66–72.

748 Horton, P., 2012: Améliorations et optimisation globale de la méthode des analogues pour la  
749 prévision statistique des précipitations. Développement d'un outil de prévision et application  
750 opérationnelle au bassin du Rhône à l'amont du Léman. Thèse de doctorat, Université de Lau-  
751 sanne, 320 p. pp.

752 Horton, P., M. Jaboyedoff, R. Metzger, C. Obled, and R. Marty, 2012: Spatial relationship between  
753 the atmospheric circulation and the precipitation measured in the western Swiss Alps by means  
754 of the analogue method. *Natural Hazards and Earth System Sciences*, **12**, 777–784, doi:10.  
755 5194/nhess-12-777-2012.

756 Horton, P., M. Jaboyedoff, and C. Obled, 2016a: AtmoSwing (v1.4): Analog Technique model for  
757 Statistical weather forecastING. *Geoscientific Model Development*, **submitted**.

758 Horton, P., M. Jaboyedoff, and C. Obled, 2016b: Using Genetic Algorithms to Optimize the Ana-  
759 logue Method for Precipitation forecasting in the Swiss Alps. *Journal of Hydrology*, **submitted**.

760 Joines, J., C. Culbreth, and R. King, 1996: Manufacturing cell design: an integer programming  
761 model employing genetic algorithms. *IIE transactions*, **28 (1)**, 69–85.

762 Keenan, T. D., and F. Woodcock, 1981: Objective Tropical Cyclone Movement Forecasts Using  
763 Synoptic and Track Analogue Information. Tech. rep., Bureau of Meteorology.

764 Kruizinga, S., and A. Murphy, 1983: Use of an analogue procedure to formulate objective prob-  
765 abilistic temperature forecasts in the Netherlands. *Monthly Weather Review*, **111 (11)**, 2244–  
766 2254.

767 Livezey, R. E., and A. G. Barnston, 1988: An operational multifield analog/antianalog predic-  
768 tion system for United States seasonal temperatures. 1. system design and winter experiments.  
769 *Journal of Geophysical Research*, **93 (D9)**, 10 953–10 974.



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

773 Lorenz, E., 1969: Atmospheric predictability as revealed by naturally occurring analogues. *Jour-*  
 774 *nal of the Atmospheric Sciences*, **26**, 636–646.

775 Mariano, A., C. Costa, D. de Angelis, D. Pires Atala, F. Maugeri Filho, M. Wolf Maciel, and  
 776 R. Maciel Filho, 2010: Genetic Algorithms (Binary and Real Codes) for the Optimisation of a  
 777 Fermentation Process for Butanol Production. *International Journal of Chemical Reactor Engi-*  
 778 *neering*, **8** (8).

779 Marty, R., I. Zin, C. Obled, G. Bontron, and A. Djerboua, 2012: Toward Real-Time Daily PQPF  
 780 by an Analog Sorting Approach: Application to Flash-Flood Catchments. *Journal of Applied*  
 781 *Meteorology and Climatology*, **51** (3), 505–520, doi:10.1175/JAMC-D-11-011.1.

782 Matheson, J., and R. Winkler, 1976: Scoring rules for continuous probability distributions. *Man-*  
 783 *agement science*, **22** (10), 1087–1096.

784 Michalewicz, Z., 1996: *Genetic Algorithms + Data Structures = Evolution Programs*. 3rd ed.,  
 785 Springer-Verlag.

786 Murray, R., 1974: Indicators of monthly mean temperature and rainfall for England and Wales  
 787 based on antecedent monthly pressure anomalies over the Northern hemisphere. *The Meteoro-*  
 788 *logical magazine*, **103**, 70.

789 Nap, J., H. van den Dool, and J. Oerlemans, 1981: A verification of monthly weather forecasts in  
 790 the seventies. *Monthly Weather Review*, **109**, 306–312.

791 Nelder, J., and R. Mead, 1965: A simplex method for function minimization. *The Computer Jour-*  
792 *nal*, **7** (4), 308–313.

793 Obled, C., G. Bontron, and R. Garçon, 2002: Quantitative precipitation forecasts: a statistical  
794 adaptation of model outputs through an analogues sorting approach. *Atmospheric Research*,  
795 **63** (3-4), 303–324, doi:10.1016/S0169-8095(02)00038-8.

796 Obled, C., and W. Good, 1980: Recent developments of avalanche forecasting by discriminant  
797 analysis techniques: a methodological review and some applications to the Parsenn area (Davos,  
798 Switzerland). *Journal of Glaciology*, **25**, 315–346.

799 Radanovics, S., J.-P. Vidal, E. Sauquet, A. Ben Daoud, and G. Bontron, 2013: Optimising pre-  
800 dictor domains for spatially coherent precipitation downscaling. *Hydrology and Earth System*  
801 *Sciences*, **17** (10), 4189–4208, doi:10.5194/hess-17-4189-2013.

802 Radcliffe, N., 1991: *Proceedings of the Fourth International Conference on Genetic Algorithms*,  
803 222–229.

804 Radinovic, D., 1975: An analogue method for weather forecasting using the 500/1000 mb relative  
805 topography. *Monthly Weather Review*, **103** (7), 639–649.

806 Rechenberg, I., 1973: *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der*  
807 *biologischen Evolution*. Frommann-Holzboog.

808 Schraudolph, N., and R. Belew, 1992: Dynamic parameter encoding for genetic algorithms. *Ma-*  
809 *chine Learning*, **9** (1), 9–21.

810 Schwefel, H. P., 1981: *Numerical optimization of computer models*. John Wiley & Sons, Inc.

811 Schwefel, H.-P. P., 1993: *Evolution and Optimum Seeking*. John Wiley & Sons, Inc., New York,  
812 NY, USA.

- 813 Shabbar, A., and J. Knox, 1986: Monthly prediction by the analogue method. *Proceedings of the*  
814 *first WMO workshop on the diagnosis and prediction of monthly and seasonal atmospheric vari-*  
815 *ations over the globe, Long range forecasting Res. Rep. Ser. 6, Vol II, Tech. Doc. WMO/TD87,*  
816 *672–681.*
- 817 Sievers, O., K. Fraedrich, and C. C. Raible, 2000: Self-Adapting Analog Ensemble Predic-  
818 tions of Tropical Cyclone Tracks. *Weather and Forecasting*, **15** (5), 623–629, doi:10.1175/  
819 1520-0434(2000)015<0623:SAAEPO>2.0.CO;2.
- 820 Smith, J., and T. Fogarty, 1997: Operator and parameter adaptation in genetic algorithms. *Soft*  
821 *Computing - A Fusion of Foundations, Methodologies and Applications*, **1** (2), 81–87.
- 822 Syswerda, G., 1989: Uniform Crossover in Genetic Algorithms. *Proceedings of the 3rd Interna-*  
823 *tional Conference on Genetic Algorithms*, Morgan Kaufmann Publishers Inc., 2–9.
- 824 Teweles, S., and H. B. Wobus, 1954: Verification of prognostic charts. *Bulletin of the American*  
825 *Meteorological Society*, **35**, 455–463.
- 826 Toth, Z., 1989: Long-range weather forecasting using an analog approach. *Journal of Climate*, **2**,  
827 594–607.
- 828 Woodcock, F., 1980: On the use of analogues to improve regression forecasts. *Monthly Weather*  
829 *Review*, **108** (3), 292–297.
- 830 Wright, A. H., 1991: Genetic Algorithms for Real Parameter Optimization. *Foundations of genetic*  
831 *algorithms*, **1**, 205–218, doi:10.1016/B978-0-08-050684-5.50016-1.
- 832 Zitzler, E., M. Laumanns, and S. Bleuler, 2004: A tutorial on evolutionary multiobjective opti-  
833 mization. *Metaheuristics for Multiobjective Optimisation*, 3–37.

834	<b>LIST OF TABLES</b>	
835	<b>Table 1.</b> Assessed operator for couples selection. . . . .	41
836	<b>Table 2.</b> Assessed operators for chromosome crossover. . . . .	42
837	<b>Table 3.</b> Assessed mutation operators with the number of variants considered (combina-	
838	tion of parameters). . . . .	43
839	<b>Table 4.</b> Further assessments of mutation operators. . . . .	44

TABLE 1. Assessed operator for couples selection.

<b>Couples selection operators</b>	
A	Rank pairing
B	Random pairing
C	Roulette wheel weighting on rank
D	Roulette wheel weighting on fitness
E	Tournament selection (3 candidates)
F	Tournament selection (4 candidates)

TABLE 2. Assessed operators for chromosome crossover.

Chromosome crossover operators	
1	Single-point crossover
2	Two-point crossover
3	Multiple-point crossover (3 points)
4	Multiple-point crossover (5 points)
5	Uniform crossover
6	Blending method (2 points, unshared $\beta$ )
7	Blending method (4 points, unshared $\beta$ )
8	Blending method (2 points, shared $\beta$ )
9	Blending method (4 points, shared $\beta$ )
10	Linear crossover (2 points)
11	Linear crossover (4 points)
12	Heuristic crossover (2 points, unshared $\beta$ )
13	Heuristic crossover (4 points, unshared $\beta$ )
14	Heuristic crossover (2 points, shared $\beta$ )
15	Heuristic crossover (4 points, shared $\beta$ )
16	Binary-like crossover (2 points, unshared $\beta$ )
17	Binary-like crossover (4 points, unshared $\beta$ )
18	Binary-like crossover (2 points, shared $\beta$ )
19	Binary-like crossover (4 points, shared $\beta$ )
20	Linear interpolation
21	Free interpolation

TABLE 3. Assessed mutation operators with the number of variants considered (combination of parameters).

	<b>Mutation operator</b>	<b>Variants</b>
1	Uniform mutation	3
2	Variable uniform mutation	27
3	Constant normal mutation	9
4	Variable normal mutation	36
5	Non-uniform mutation	27
6	Individual adaptive mutation rate	1
7	Individual adaptive search radius	1
8	Chromosome of adaptive mutation rate	1
9	Chromosome of adaptive search radius	1
10	Mutli-scale mutation	3
11	No mutation	1

TABLE 4. Further assessments of mutation operators.

	<b>Mutation operator</b>	$p_{mut}$	$G_{max}$	$\omega$
1	Non-uniform mutation	0.01	50	0.1
2	Non-uniform mutation	0.05	50	0.1
3	Non-uniform mutation	0.1	50	0.1
4	Non-uniform mutation	0.2	50	0.1
5	Non-uniform mutation	0.4	50	0.1
6	Non-uniform mutation	0.01	100	0.1
7	Non-uniform mutation	0.05	100	0.1
8	Non-uniform mutation	0.1	100	0.1
9	Non-uniform mutation	0.2	100	0.1
10	Non-uniform mutation	0.4	100	0.1
11	Mutli-scale mutation	0.01		
12	Mutli-scale mutation	0.05		
13	Mutli-scale mutation	0.1		
14	Mutli-scale mutation	0.2		
15	Mutli-scale mutation	0.4		
16	Chromosome of adaptive search radius			



840	<b>LIST OF FIGURES</b>	
841	<b>Fig. 1.</b> Genetic Algorithms operational flowchart . . . . .	46
842	<b>Fig. 2.</b> Evolution of the score of the best individuals over generations for the 10 optimizations	
843	processed for a given parametrization. The continuous bottom line represents the score	
844	of the sequential approach and the dashed one (top), the supposed global optimum. . . . .	47
845	<b>Fig. 3.</b> Same as Fig. 2, but for a GAs parametrization considered as less relevant. . . . .	48
846	<b>Fig. 4.</b> Influence of the couples selection operators (Table 1) on the optimization performance (im-	
847	provement of the score). The box extends from the lower to upper quartile values of the data,	
848	with a line at the median. The whiskers extend from the box to 1.5 times the interquartile	
849	range. Flier points are those past the end of the whiskers. The star represents the median.	
850	The gray box highlights the best options. . . . .	49
851	<b>Fig. 5.</b> Influence of the chromosome crossover operators (Table 2) on the optimization performance	
852	(improvement of the score). Same conventions as Fig. 4. . . . .	50
853	<b>Fig. 6.</b> Influence of the mutation operators (Table 3) on the optimization performance. Same con-	
854	ventions as Fig. 4. . . . .	51
855	<b>Fig. 7.</b> Influence of the mutation operators (Table 4) on the optimization performance, leaving the	
856	optimizer choose the pressure level of the atmospheric circulation analogy (single level of	
857	analogy). The predictand is precipitation over a subcatchment in the Swiss Alps (Binn-	
858	Simplon region). The continuous bottom line represents the score of the sequential calibra-	
859	tion and the dashed superior line, the score of the optimization without automatic selection	
860	of the pressure levels. Same conventions as Fig. 4. . . . .	52
861	<b>Fig. 8.</b> Same as Fig. 7, but for another region in the Swiss Alps, with different atmospheric influ-	
862	ences. . . . .	53
863	<b>Fig. 9.</b> Same as Fig. 7, but with a second level of analogy on moisture variables. . . . .	54
864	<b>Fig. 10.</b> Same as Fig. 9, but with a preselection on air temperature rather than a fixed calendar	
865	window. . . . .	55
866	<b>Fig. 11.</b> Influence of the natural selection operators on the optimization performance. Same conven-	
867	tions as Fig. 4. . . . .	56
868	<b>Fig. 12.</b> Influence of the population size on the optimization performance. Same conventions as Fig.	
869	4. . . . .	57
870	<b>Fig. 13.</b> Influence of the intermediate population (IG) ratio on the optimization performance. Same	
871	conventions as Fig. 4. . . . .	58

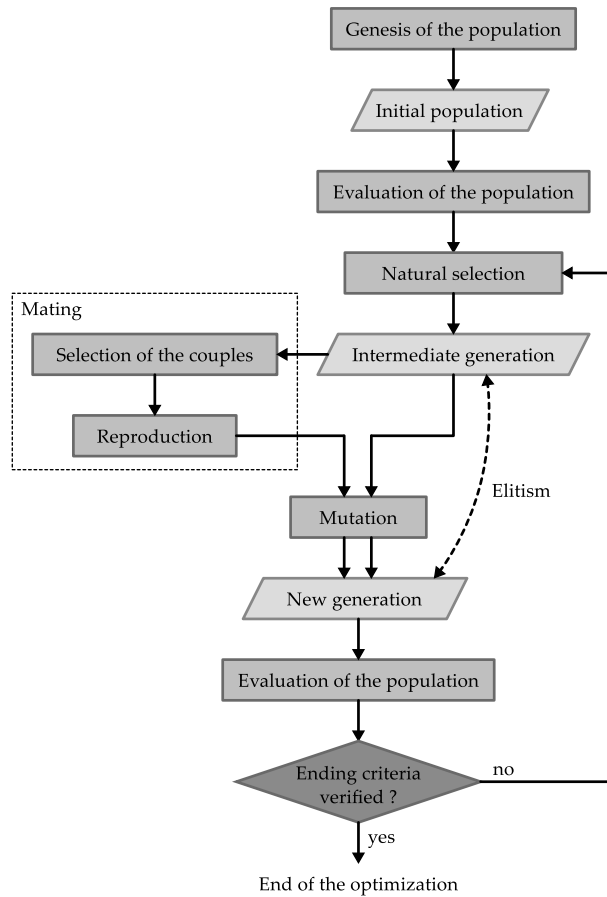


FIG. 1. Genetic Algorithms operational flowchart

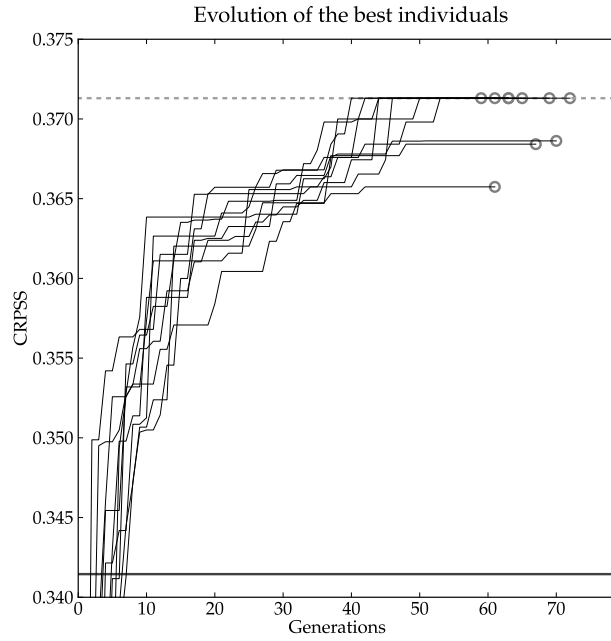


FIG. 2. Evolution of the score of the best individuals over generations for the 10 optimizations processed for a given parametrization. The continuous bottom line represents the score of the sequential approach and the dashed one (top), the supposed global optimum.

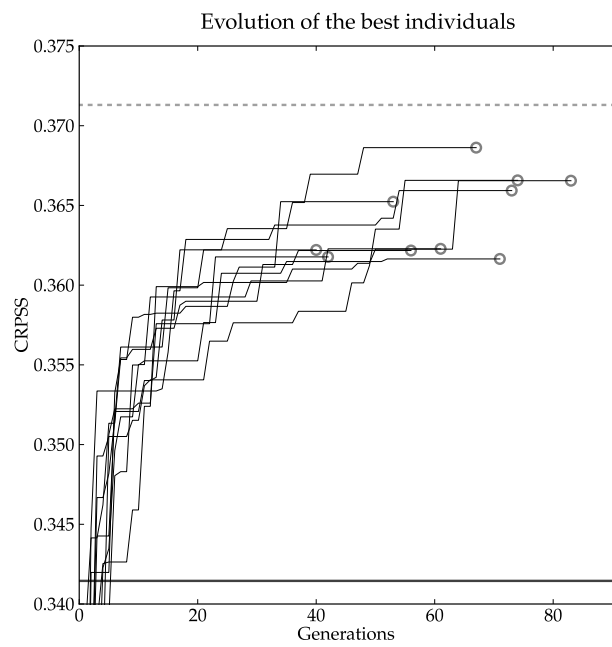


FIG. 3. Same as Fig. 2, but for a GAs parametrization considered as less relevant.

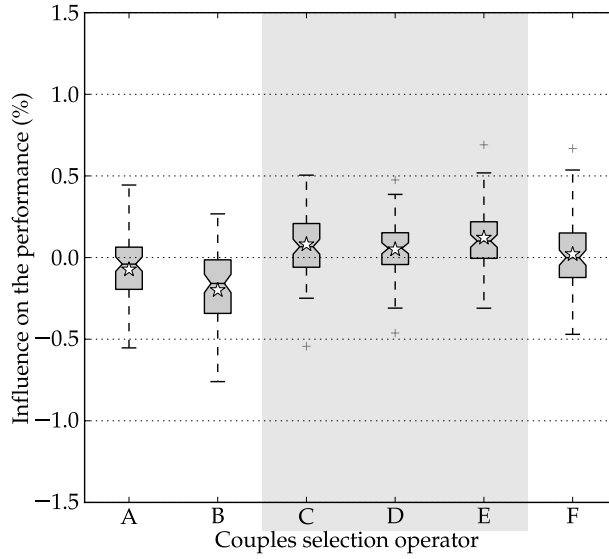


FIG. 4. Influence of the couples selection operators (Table 1) on the optimization performance (improvement of the score). The box extends from the lower to upper quartile values of the data, with a line at the median. The whiskers extend from the box to 1.5 times the interquartile range. Flier points are those past the end of the whiskers. The star represents the median. The gray box highlights the best options.

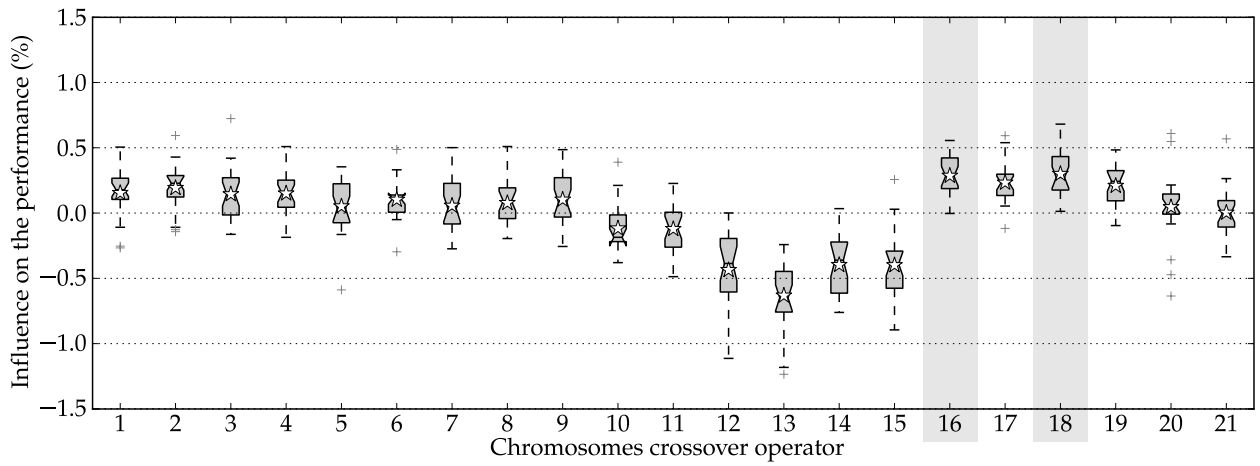
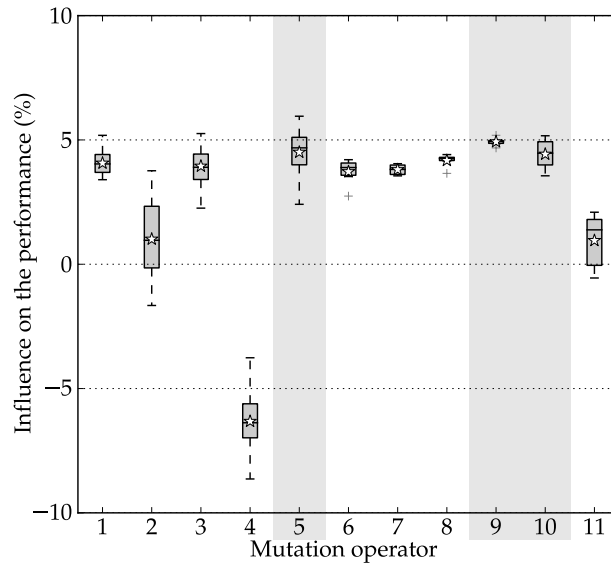
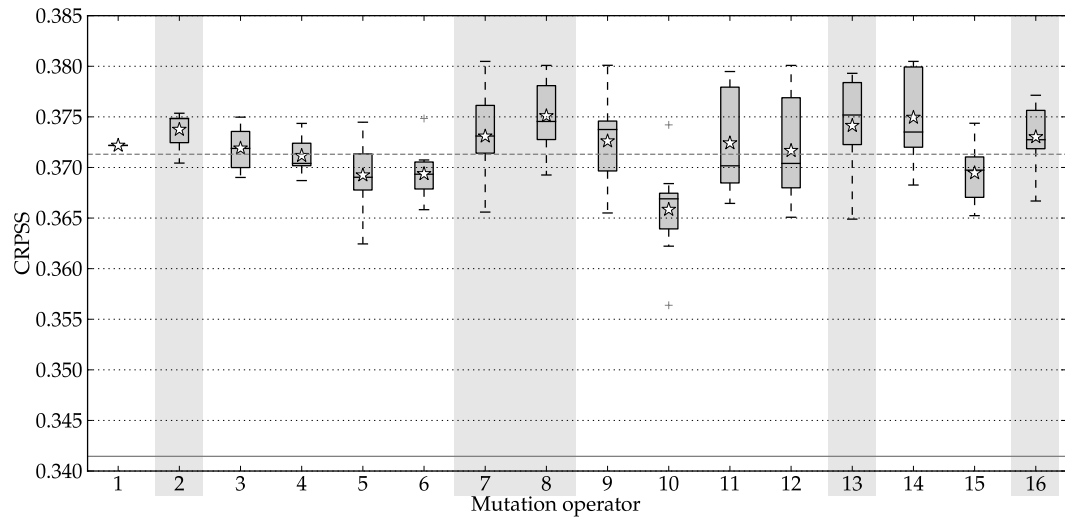


FIG. 5. Influence of the chromosome crossover operators (Table 2) on the optimization performance (improvement of the score). Same conventions as Fig. 4.



881 FIG. 6. Influence of the mutation operators (Table 3) on the optimization performance. Same conventions as  
882 Fig. 4.



883 FIG. 7. Influence of the mutation operators (Table 4) on the optimization performance, leaving the optimizer  
 884 choose the pressure level of the atmospheric circulation analogy (single level of analogy). The predictand  
 885 is precipitation over a subcatchment in the Swiss Alps (Binn-Simplon region). The continuous bottom line  
 886 represents the score of the sequential calibration and the dashed superior line, the score of the optimization  
 887 without automatic selection of the pressure levels. Same conventions as Fig. 4.



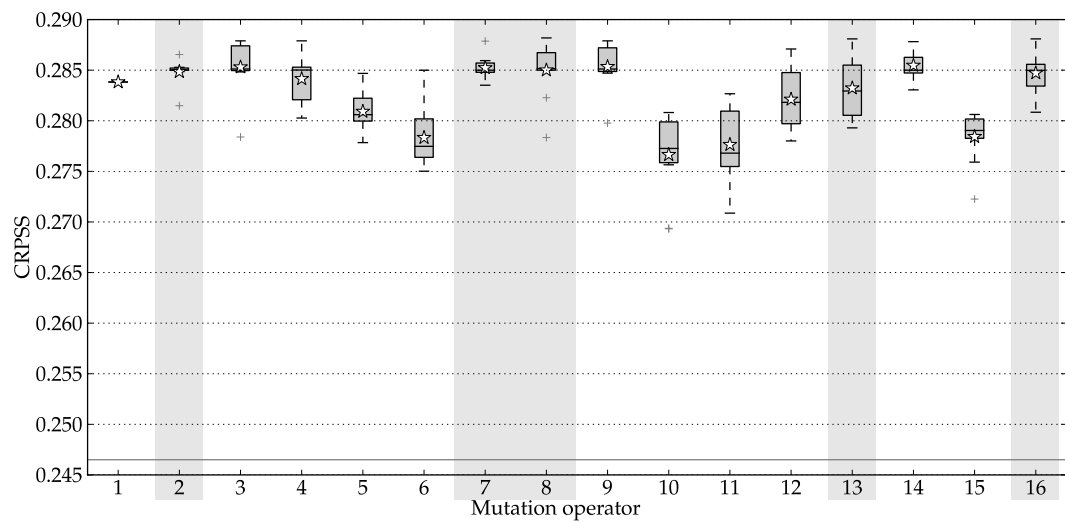


FIG. 8. Same as Fig. 7, but for another region in the Swiss Alps, with different atmospheric influences.

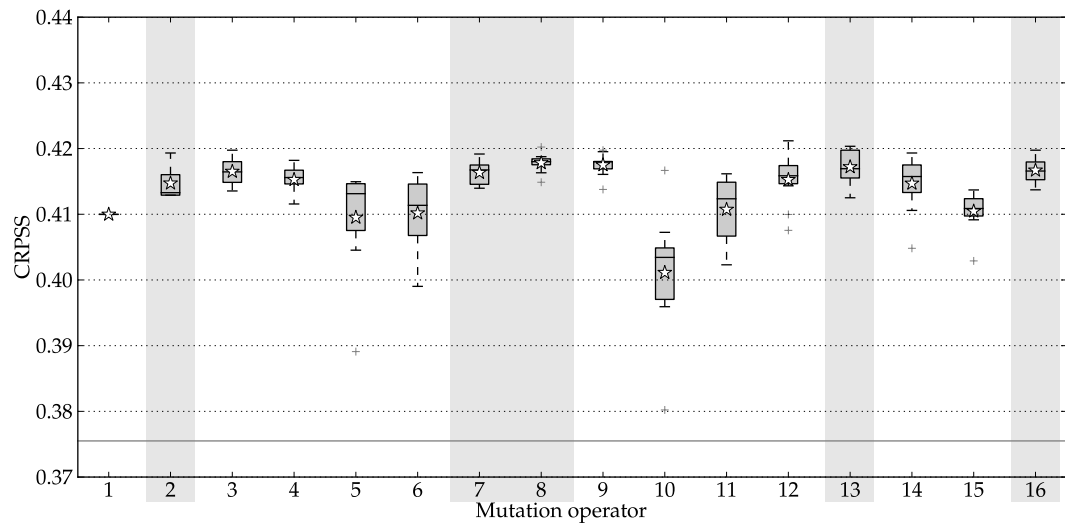


FIG. 9. Same as Fig. 7, but with a second level of analogy on moisture variables.

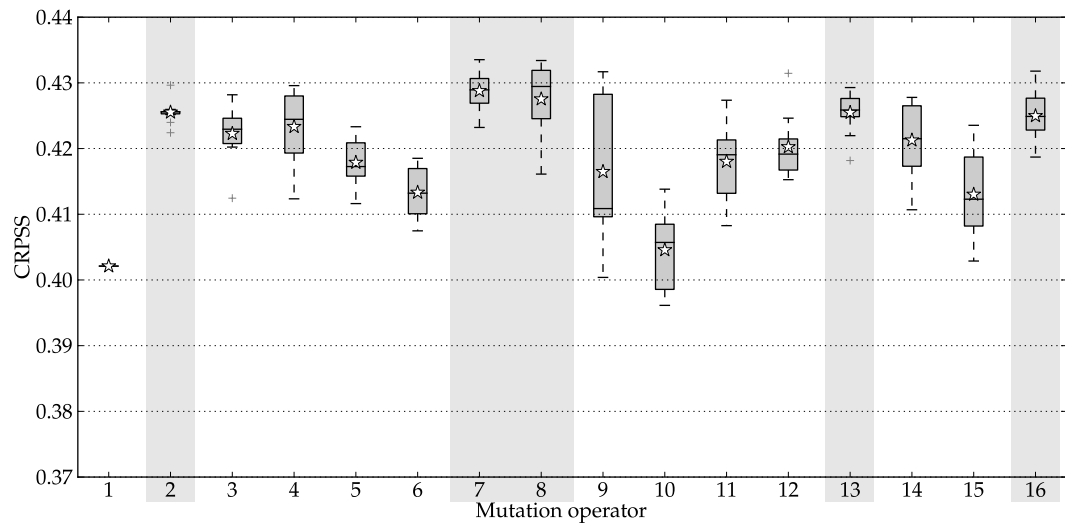
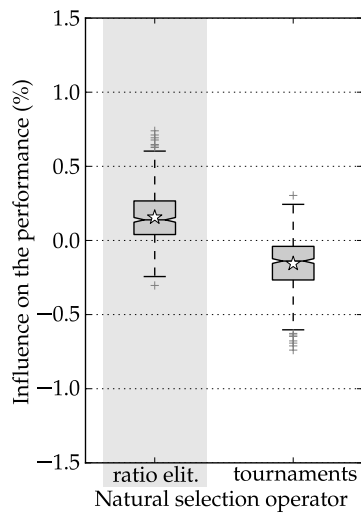


FIG. 10. Same as Fig. 9, but with a preselection on air temperature rather than a fixed calendar window.



888 FIG. 11. Influence of the natural selection operators on the optimization performance. Same conventions as  
 889 Fig. 4.

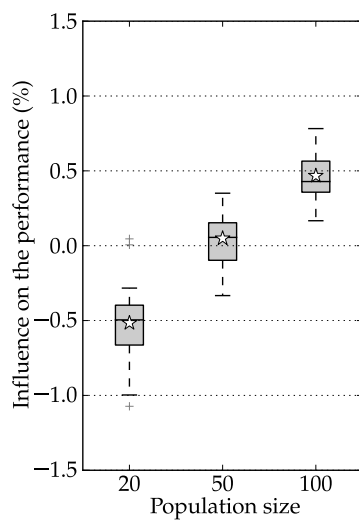
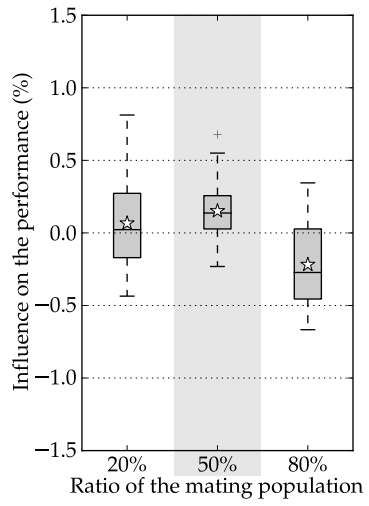


FIG. 12. Influence of the population size on the optimization performance. Same conventions as Fig. 4.



890 FIG. 13. Influence of the intermediate population (IG) ratio on the optimization performance. Same conven-  
891 tions as Fig. 4.