%% Version 4.3.2, 25 August 2014

%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Template.tex -- LaTeX-based template for submissions to the

% American Meteorological Society

%

% Template developed by Amy Hendrickson, 2013, TeXnology Inc.,

% amyh@texnology.com, http://www.texnology.com

% following earlier work by Brian Papa, American Meteorological Society

%

% Email questions to latex@ametsoc.org.

%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% PREAMBLE

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% Start with one of the following:

%% DOUBLE-SPACED VERSION FOR SUBMISSION TO THE AMS

\documentclass{ametsoc}

%% TWO-COLUMN JOURNAL PAGE LAYOUT---FOR AUTHOR USE ONLY

%\documentclass[twocol]{ametsoc}

\usepackage{gensymb}

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%% To be entered only if twocol option is used

\journal{mwr}

% Please choose a journal abbreviation to use above from the following list:

%

% jamc (Journal of Applied Meteorology and Climatology)

% jtech (Journal of Atmospheric and Oceanic Technology)

% jhm (Journal of Hydrometeorology)

% jpo (Journal of Physical Oceanography)

% jas (Journal of Atmospheric Sciences)

% jcli (Journal of Climate)

% mwr (Monthly Weather Review)

% wcas (Weather, Climate, and Society)

% waf (Weather and Forecasting)

% bams (Bulletin of the American Meteorological Society)

% ei (Earth Interactions)

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Citations should be of the form ``author year'' not ``author, year''

\bibpunct{(}{)}{;}{a}{}{,}

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%% To be entered by author:

%% May use \\ to break lines in title:

\title{Global Optimization of an Analog Method by Means of Genetic Algorithms}

%%% Enter authors' names, as you see in this example:

%%% Use \correspondingauthor{} and \thanks{Current Affiliation:...}

%%% immediately following the appropriate author.

%%%

%%% Note that the \correspondingauthor{} command is NECESSARY.

%%% The \thanks{} commands are OPTIONAL.

%\authors{Author One\correspondingauthor{Author One,

% American Meteorological Society,

% 45 Beacon St., Boston, MA 02108.}

% and Author Two\thanks{Current affiliation: American Meteorological Society,

% 45 Beacon St., Boston, MA 02108.}}

\authors{Pascal Horton\correspondingauthor{Pascal Horton, University of Bern, Hallerstrasse 12, 3012 Bern, Switzerland.}\thanks{Current affiliation: Institute of Geography \& Oeschger Centre for Climate Change Research, University of Bern, Switzerland}}

%% Follow this form:

% \affiliation{American Meteorological Society,

% Boston, Massachusetts.}

\affiliation{University of Lausanne, Lausanne, Switzerland}

%% Follow this form:

%\email{latex@ametsoc.org}

\email{pascal.horton@alumnil.unil.ch}

%% If appropriate, add additional authors, different affiliations:

%\extraauthor{Extra Author}

%\extraaffil{Affiliation, City, State/Province, Country}

\extraauthor{Michel Jaboyedoff}

\extraaffil{University of Lausanne, Lausanne, Switzerland}

%% May repeat for a additional authors/affiliations:

%\extraauthor{}

%\extraaffil{}

\extraauthor{Charles Obled}

\extraaffil{Universit\'{e} de Grenoble-Alpes, LTHE, Grenoble, France}

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% ABSTRACT

%

% Enter your abstract here

% Abstracts should not exceed 250 words in length!

%

% For BAMS authors only: If your article requires a Capsule Summary, please place the capsule text at the end of your abstract

% and identify it as the capsule. Example: This is the end of the abstract. (Capsule Summary) This is the capsule summary.

\abstract{Analog methods are based on a statistical relationship between synoptic meteorological variables (predictors) and local weather (predictand, to be predicted). This relationship is defined by several parameters, which are often calibrated by means of a semi-automatic sequential procedure. This calibration approach is fast, but has strong limitations. It proceeds through successive steps, and thus cannot handle all parameters dependencies. Furthermore, it cannot automatically optimize some parameters, such as the selection of pressure levels and temporal windows (hours of the day) at which the predictors are compared.

In order to overcome these limitations, the global optimization technique of genetic algorithms is considered, which can jointly optimize all parameters of the method, and get closer to a global optimum, by taking into account the dependencies of the parameters. Moreover, it can objectively calibrate parameters that were previously assessed manually, and can take into account new degrees of freedom.

However, genetic algorithms must be tailored to the problem under consideration. Multiple combinations of algorithms are assessed, and new combinations are developed (e.g., the \textit{chromosome of adaptive search radius}, which is found to be very robust), in order to provide recommendations regarding the use of genetic algorithms for optimizing several variants of analog methods. A global optimization approach provides new perspectives for the improvement of analog methods, and for their application to new regions or new predictands.

}

\begin{document}

%% Necessary!

\maketitle

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% MAIN BODY OF PAPER

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%

%% In all cases, if there is only one entry of this type within

%% the higher level heading, use the star form:

%%

% \section{Section title}

% \subsection\*{subsection}

% text...

% \section{Section title}

%vs

% \section{Section title}

% \subsection{subsection one}

% text...

% \subsection{subsection two}

% \section{Section title}

%%%

% \section{First primary heading}

% \subsection{First secondary heading}

% \subsubsection{First tertiary heading}

% \paragraph{First quaternary heading}

**\section{Introduction}**

**\label{sec:intro}**

Analog methods (AMs) rely on the hypothesis that similar situations, in terms of atmospheric circulation, are likely to result in similar local weather conditions \citep{Lorenz1956, Lorenz1969, Duband1970, Bontron2005}. These methods consist of sampling a certain number of past situations, based on different synoptic-scale meteorological variables (predictors), in order to construct a probabilistic prediction for a local weather variable of interest (predictand). Some common usages of AMs are for operational precipitation forecasting \citep[e.g.,][]{Guilbaud1997, Bontron2005, Hamill2006, Bliefernicht2010, Marty2012, Horton2012, Hamill2015b, BenDaoud2016}, or more recently for precipitation downscaling in a climate perspective \citep[e.g.,][]{Radanovics2013, Chardon2014, Dayon2015}. However, AMs or equivalent methods are also employed to predict temperatures \citep{Radinovic1975, Woodcock1980, Kruizinga1983, DelleMonache2013, Caillouet2016}, wind \citep{Gordon1987, DelleMonache2013, DelleMonache2011, Vanvyve2015, Alessandrini2015, Junk2015, Junk2015c}, solar power \citep{Alessandrini2015a, Bessa2015}, snow avalanches \citep{Obled1980, Bolognesi1993}, insolation \citep{Bois1981}, and the trajectories of tropical cyclones \citep{Keenan1981, Sievers2000, Fraedrich2003}.

Although the method is rather simple, it contains several parameters to be optimized, such as the choice of the predictor variables, pressure levels and temporal windows (hours of the day) at which the predictors are compared along with the spatial domains used for the comparison, the analogy criteria (distance measure), and finally the number of analog situations to retain.

AMs must be adapted to every new considered region because the leading meteorological influences may be location-specific. Even the selection of the pressure levels and the temporal windows should be reconsidered, if not the predictor variable itself. Thus, before being applied in a forecasting or downscaling context, AMs must be calibrated for the given region, which is performed here in the perfect prognosis framework \citep{Klein1963}, in line with the majority of previously cited AM calibration procedures for precipitation predictions \cite[with the exception of][]{Hamill2006,Hamill2015b}.

A common approach to optimizing this method is by means of a semi-automatic sequential calibration procedure, which was developed by \citet{Bontron2004}, and is also described in \citet{BenDaoud2016} and extended by \citet{Radanovics2013}. This determines some parameters of the method sequentially for each consecutive analogy level (e.g., on the atmospheric circulation or on a moisture index). This begins with a manual selection of the meteorological variables (e.g., geopotential height and relative humidity), the pressure levels, the temporal windows, and the initial analog numbers. Then, the spatial window at which the predictors are compared is optimized through an iterative growth of the domain, and the number of analogs is finally reassessed. A succeeding level of analogy can then be introduced, and its spatial window be optimized. The parameters of the preceding levels of analogy are not reassessed, except for the number of analog situations to preserve.

Thus, the sequential calibration procedure allows the optimization of a limited number of parameters (spatial windows and analog numbers), but the selection of predictor variables, pressure levels, and temporal windows must still be made manually. Testing multiple combinations of these presents a very combinatorial problem, which quickly becomes cumbersome, especially when considering multiple predictors within the same level of analogy. Thus, optimizing AMs by means of this sequential technique is laborious if little knowledge is available regarding the predictor-predictand relationship or the leading meteorological influences. Indeed, many combinations of parameters (predictor variables, pressure levels, and temporal windows) must be assessed. Moreover, proceeding to the optimization sequentially ignores potential dependencies between the parameters of the method, whether they are within a single level of analogy or between multiple levels, which could lead to another configuration if the parameters were calibrated together. Thus, there is a high risk of ending in a local optimum, and this cannot be avoided.

When creating this sequential calibration procedure, \citet{Bontron2004} was aware of the problem of dependencies between parameters, and wrote: ''\textit{We perceive here the combinatorial aspect of our problem: variables and spatial windows are not independent. We will present our results by first searching the best variable (e.g., selection of the pressure level and the temporal window for the geopotential height) on a chosen spatial window, and next, the best window for the chosen variable. However, even by repeating the process, are we sure to obtain the optimal combination?}'' Later in his work, he also wrote: ''\textit{Our approach, which is again to vary the parameters one by one -- the others being fixed in a more or less arbitrary manner -- may therefore not exactly lead us to the optimal solution}.'' \citet{Bliefernicht2010} has also confronted the combinatorial issue for the parameters of AMs, and concluded that one must be an expert in order to have a sense of their respective influence, sensitivity, and nonlinear interactions. When calibrating an AM, \citet{BenDaoud2010} also stated that ''\textit{the combinatory aspect related to the calibration was found to be too high for all the parameters to be calibrated simultaneously}.'' The simultaneous calibration of all parameters has not been undertaken so far.

Another optimization strategy, proposed by \citet{Junk2015}, allows for an automatic calibration of weights applied to the different predictors when processing the analogy criteria (distance function). Their strategy consists of a brute-force assessment of all possible combinations. This approach is possible in their implementation, because predictors are considered at a unique point (interpolated to the location of interest), at fixed hours, and at preselected pressure levels, leaving only the weights to be optimized. In the presently employed AM (described in Section \ref{sec:am}), the number of parameters to optimize makes it impossible to proceed with a brute-force strategy.

In order to overcome these limitations, two optimization techniques have been assessed. First, \citet{Horton2012a} assessed the ability of the method of \citet{Nelder1965a}, based on a simplex algorithm. This technique did not provide satisfying results, and failed to converge toward a unique solution. The parameter space of the AMs can be very complex, and is inappropriate for a linear optimization technique. Thus, global optimization techniques are likely to be necessary in order to calibrate most AM variants, as they can optimize all parameters of all analogy levels simultaneously. In addition, they can avoid the systematic manual assessments of all pressure levels and temporal windows. Finally, they allow the testing of new degrees of freedom in AMs. The relevance of genetic algorithms (GAs) is described here, which does not exclude that other global optimization techniques could also work. Although using GAs to optimize AMs may be computationally intensive, once an AM is calibrated its employment in real-time operations or climate downscaling is very fast and lightweight.

This paper specifically describes how GAs should be used in order to successfully optimize several AMs. Indeed, the variants of Gas are numerous, and always need to be tailored to the addressed problem. This requires intensive and systematic comparisons of operators and options, in order to identify the key factors influencing the optimization, and the respective sensitivity of the options. Such analyses are presented here, resulting in recommendations for the use of GAs to optimize several AM implementations. The in-depth analysis of the benefits of such an approach in a specific case study will be the topic of a forthcoming paper.

This document begins by presenting the area of the case study and the relevant data (Section \ref{sec:case\_study\_data}), the considered AM variants (Section \ref{sec:am}), and the assessed GA options (Section \ref{sec:gas}). Comparative analyses of these options are presented in Section \ref{sec:assessment}, leading to the recommendations formulated in Section \ref{sec:use}, and finally the conclusions presented in Section \ref{sec:conclusions}.

**\section{Case study and data}**

**\label{sec:case\_study\_data}**

The area of the study is the alpine upper Rh\^{o}ne catchment in Switzerland (Fig.\ \ref{fig:map}). The altitude ranges from 372 to 4634~m.a.s.l., and the area is 5524~km$^{2}$. Based on various climatological analyses \cite[see][for the details]{Horton2012a}, the gauging stations in the catchment were clustered in 10~subregions (Fig.\ \ref{fig:map}).

This region is the target of the MINERVE (Mod\'{e}lisation des Intemp\'{e}ries de Nature Extr\^{e}me sur les Rivi\`{e}res Valaisannes et de leurs Effets) project, which aims to provide real-time flood management on the upper Rh\^{o}ne catchment \citep{GarciaHernandez2009b}. AMs are used to provide real-time probabilistic precipitation forecasts, based on NWP outputs.

**\subsection{Data}**

**\label{sec:data}**

AMs rely on two types of data: predictors, which are meteorological variables describing the state of the atmosphere at a synoptic scale, and the predictand, which is the local weather time series that is to be predicted.

When working in the perfect prognosis framework \citep{Klein1963}, the meteorological archive from which the predictors are extracted is usually a reanalysis dataset. Conversely, other applications of AMs to wind forecasting are based on a model output statistics method \citep[MOS, see][]{Glahn1972}, and thus employ forecast archives or reforecast products \citep[e.g.,][]{DelleMonache2013, DelleMonache2011, Alessandrini2015, Junk2015, Junk2015c}.

In the present study, the NCEP-NCAR reanalysis I dataset \citep[six-hourly, 17 pressure levels at a resolution of 2.5\degree, see][]{Kalnay1996} is employed, but this could be replaced with any other reanalysis dataset. This dataset is relatively old, although it is still widely used, and better results may be expected with more recent datasets. However, we can safely assume that if an optimization technique works for this reanalysis dataset, it will also work for an improved one.

Here, the predictand (which is to be predicted) is the daily precipitation (0540~UTC to 0540~UTC the following day) measured at the MeteoSwiss network of stations in the catchment of interest. The time series from every available gauging station were averaged over subregions of approximately 500~km$^{2}$ \citep[see][for details]{Horton2012a} in order to smooth local effects \citep{Obled2002, Marty2012}. The time series must be split into calibration and independent validation periods.

**\section{The considered analog method}**

**\label{sec:am}**

AMs are based on the principle that two similar synoptic situations may result in similar local effects \citep{Lorenz1956, Lorenz1969}. Thus, they consist in searching for a certain number of past situations in a meteorological archive that present similar properties for chosen predictors, according to an analogy criteria, in order to extract the observed values of the local weather variable of interest from another archive (predictand time series, see Section \ref{sec:case\_study\_data}\ref{sec:data}). Based on these analog observations, the conditional empirical distribution to be considered as the probabilistic prediction for the target day (the day one wishes to predict) is constructed.

Predictors for precipitation predictions can be varied. For example, the geopotential height at different pressure levels and temporal windows, or humidity variables \citep[see][for a more detailed list of predictors]{BenDaoud2016}. This method often contains several levels of analogy, which constitute of successive subsamplings on predictors of different natures (e.g., atmospheric circulation, moisture variables, vertical motion, and air temperature).

The basis of the AM implementation considered here is the following \citep[the same approach as][]{Guilbaud1997, Bontron2005, Marty2012, Horton2012, Radanovics2013, Chardon2014, Dayon2015, BenDaoud2016}:

\begin{enumerate}

\item Preselection: To cope with seasonal effects, $n\_{0}$ candidate dates are extracted from the archive within a period of four months centered around the target date, for every year of the archive. Alternatively, the candidate dates can be selected based on similar air temperatures \citep[see][]{BenDaoud2016}.

\item First level of analogy: $n\_{1}$ dates are selected out of the preselected $n\_{0}$, by means of an analogy ranking. The first level of analogy for precipitation prediction is often based on the atmospheric circulation. The similarity between the atmospheric circulation of the target date and the candidate situations is assessed in terms of the geopotential height (at specific pressure levels, such as 500 and 1000~hPa, and over a certain spatial window, e.g., 12~h and 24~h) by means of the S1 criteria \citep[Eq. (\ref{eq:S1}), ][]{Teweles1954, Drosdowsky2003}, which is a comparison of gradients over a defined spatial window. Various studies have found the S1 criteria to be more relevant than Euclidean distances for pressure fields \citep{Wilson1980, Woodcock1980, Guilbaud1998, Bontron2004}.

\begin{equation}

\**label{eq:S1}**

S1=100 \frac {\displaystyle \sum\_{i} \vert \Delta\hat{z}\_{i} - \Delta z\_{i} \vert}

{\displaystyle \sum\_{i} max( \vert \Delta\hat{z}\_{i} \vert , \vert \Delta z\_{i} \vert ) }

\end{equation}

where $\Delta \hat{z}\_{i}$ is the forecast geopotential height difference between the \textit{i}th pair of adjacent points from the grid of the target situation, and $\Delta z\_{i}$ is the corresponding observed geopotential height difference in the candidate situation. The differences are processed separately in both directions. The smaller the S1 values, the more similar the pressure fields are.

The $n\_{1}$ dates with the lowest S1 values are considered as analog situations, in terms of the atmospheric circulation, for the target day. The analog number $n\_{1}$ is a parameter to be calibrated.

\item Subsequent level(s) of analogy: Beyond the similarity of airflows, one may look for analogies in other variables of interest, such as moisture variables. Therefore, the $n\_{1}$ analogs are subsampled once more on the basis of another variable, in order to obtain a lower number of analog dates, $n\_{2}$. When the predictor is not a pressure field, the analogy criteria usually represents absolute distances to the mean absolute error (MAE) or root mean squared error (RMSE), with the latter being most often employed.

A second level of analogy, based on thermodynamic variables, was introduced by \citet{Vallee1986} and \citet{Gibergans-Baguena2007}. After a systematic assessment of variables, \citet{Bontron2004} pointed out that a moisture index consisting of the product of relative humidity at 850~hPa and total precipitable water achieves the best performance. This index does not represent an actual physical quantity, but expresses the water content of the air column and its proximity to saturation.

This process can be repeated by subsampling a decreasing number of analog dates, $n\_{i}$, according to various meteorological variables.

\item Probabilistic prediction: Then, the daily observed amounts of precipitation for the $n\_{i}$ resulting dates provide the empirical conditional distribution, considered as the probabilistic forecast for the target day.

\end{enumerate}

The parameters to be defined (manually or automatically), for every level of analogy, are as follows:

\begin{itemize}

\item The selection of meteorological variables (predictors), containing synoptic scale information, and having a direct or indirect dependency with the target predictand.

\item The pressure level, or height, at which the predictor is selected.

\item The temporal window is the hour(s) of the day at which the predictors are considered.

\item The spatial window is the domain on which predictors are compared. The ideal size of this area is that which maximizes the useful information and minimizes noise.

\item The analogy criterion, needed to compare the variables on the chosen spatial and temporal windows, is a distance measure, used to rank observed situations according to their degree of similarity with the target situation.

\item Eventual weights between the predictors \cite[e.g.,][]{Horton2012a, Junk2015}.

\item The ideal number of analog situations, which is the best compromise in order to take into account local variability and maximize useful synoptic information \citep{Bontron2004}.

\end{itemize}

In order to calibrate the method, the continuous ranked probability score \citep[CRPS,][]{Brown1974, Matheson1976, Hersbach2000} is often employed to assess the performances of AMs \citep[see, e.g.,][]{Bontron2004, Bontron2005, BenDaoud2008, Horton2012, Marty2012, Radanovics2013, Chardon2014, Junk2015, BenDaoud2016, Caillouet2016}. This allows the evaluation of predicted cumulative distribution functions $F(y)$. For example, of the precipitation values $y$ from analog 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 as follows:

\begin{equation}

\**label{eq:CRPS}**

CRPS = \frac{1}{l} \sum\_{i=1}^{l} \left\{ \int\_{-\infty}^{+\infty} \left[ F\_{i}(y)-H\_{i}(y-y\_{i}^{0})\right]^{2} dy \right\}

\end{equation}

where $H(y-y\_{i}^{0})$ is the Heaviside function, which is zero when $y-y\_{i}^{0}<0$ and has a value of one otherwise. The mean CRPS is averaged over the calibrations for the respective validation periods.

In order to compare the value of the score with regard to a reference, one often considers its skill score expression, and employs the climatological distribution of daily precipitation as the reference. The CRPSS (\textit{continuous ranked probability skill score}) is thus defined as follows:

\begin{equation}

\**label{eq:CRPSS}**

CRPSS = 1-\frac{CRPS}{CRPS\_{r}}

\end{equation}

where $CRPS\_{r}$ is the score value for the reference. An increase in the CRPSS indicates a better prediction.

**\section{Assessed genetic algorithm variants}**

**\label{sec:gas}**

Genetic algorithms \citep[GAs,][]{Holland1992b, Goldberg1989} are part of the family of evolutionary algorithms \citep{Back1993b, Schwefel1993}, which are inspired by some mechanisms of biological evolution, such as reproduction, genetic mutations, chromosomal crossovers, and natural selection. Unlike linear or local optimizations, GAs seek the global optimum on complex surfaces, theoretically without restriction, but with no guarantee of reaching it \citep{Haupt2004}. The objective function to be optimized (often named the fitness function in this context) can be of various types \citep{Joines1996a}, but GAs must be adapted in order to perform optimally.

A key element of the configuration of GAs is to determine the correct balance between exploration and exploitation \citep{Back1992a, Smith1997a}. Exploration is characterized by a relatively high probability of assessing 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 to be located. Exploitation is characterized by a local search in an area of interest, and generally makes small movements. The latter aims at finding small improvements at the end of the optimization.

**\subsection{Structure and operators}**

**\label{sec:gas:operators}**

GAs optimize a population of $N$ individuals. Each individual contains a chromosome, which here is a vector of the AM parameters. Genes are the individual parameters constituting the chromosome. These can be categorical (e.g., choice of the meteorological variable), discrete (e.g., number of analog dates to select), or continuous.

There are numerous implementation variants of GAs, which are often optimal for a given problem \citep{Hart1991a, Schraudolph1992a}. The differences are found in the implementation of the operators, through significantly different algorithms, which has an important effect on the results \citep{Gaffney2010a}. Here, operators are defined as the mechanisms that modify the values of the genes, to try to bring individuals (or chromosomes) closer to an optimum of the fitness function. The structure of the method (Fig. \ref{fig:structure\_gas}) that results from the work of \citet{Holland1992b} is common to most applications \citep{Back1993b}, and consists of the following steps. (1) A population of $N$ individuals (parameter sets of the AM to be optimized) is randomly generated, which constitutes the initial population. The fitness (performance score or objective function) of every individual is assessed. (2) A natural selection algorithm is applied, after which only the best individuals remain. This constitutes the intermediate generation (IG), from which (3) couples are formed according to given rules. Then, (4) these couples proceed to the reproduction or chromosome crossover stage, to mix their genes according to the selected operator version. New children are generated, in order to refill the IG back to $N$ individuals. (5) Parents and children are then subjected to mutation, where some genes are randomly changed. Finally, (6) the new generation is then re-assessed, with the best individual restored if degraded. Then, (7) according to the ending criteria, the optimization either ends or begins again for another iteration.

All of the considered operators and their options are described in the following sections (only briefly for operators of less importance). Many other operators exist, but only those that are evaluated are described here.

**\subsubsection{Genesis of the population}**

The most current version of the initial population generation is a random initialization based on a uniform sampling. The size $N$ of the population is often a compromise between the computation time and the quality of the solution. The chosen $N$ must allow sufficient sampling of the solutions field \citep{Beasley1996a}, and thus should vary as a function of the chromosome size (i.e., the number of genes or parameters to be optimized).

**\subsubsection{Natural selection}**

**\label{sec:gas:nat\_selection}**

Natural selection is performed on the basis of the objective function values. The selection allows for only a certain portion of the population to be kept, usually half ($N/2$), which can access the IG (with $N\_{IG}$ members). Several techniques exist, such as:

\begin{itemize}

\item $N\_{IG}$-elitism \citep{Michalewicz1996}: Only the better half is preserved.

\item Tournament selection \citep{Michalewicz1996, Zitzler2004a}: Two individuals are randomly selected, and the best of these is chosen, but with a certain probability.

\end{itemize}

**\subsubsection{Selection of the couples}**

**\label{sec:gas:selection\_couples}**

Individuals of the IG can then reproduce, which begins with the selection of pairs (the parents). The techniques implemented in this study are the following:

\begin{itemize}

\item Rank pairing: Individuals are gathered into pairs according to their ranks.

\item Random pairing: Individuals are randomly selected, according to a uniform law.

\item Roulette wheel selection \citep{Goldberg1989}: The selection probability assigned to each individual is proportional to their fitness, so that the most adapted individuals have a greater probability of reproduction. There are two weighting techniques. The first proceeds according to the rank, and the second according to the fitness value.

\item Tournament selection: A number of individuals (two or three) are randomly picked, and the best is kept, with a certain probability. This operation is performed twice, once for each partner.

\end{itemize}

**\subsubsection{Chromosome crossover}**

**\label{sec:gas:crossover}**

Once the two parents have been selected for breeding, they combine their chromosomes and produce two children, bringing the number of individuals in the population back to $N$. The combination of chromosomes is carried out using a crossover operator, thereby generating two offspring that have characteristics derived from both parents. This allows the mixing of genes, and the potential accumulation of positive mutations. The evaluated crossover operators are the following:

\begin{itemize}

\item Single-point crossover \citep{Goldberg1989}: The genes located after a randomly chosen point within the chromosome are exchanged between the two parents.

\item Two-point crossover: Similar to the single-point crossover, but with two intersections defining the segments to be exchanged. This is considered to be more efficient than the previous method \citep{Beasley1993a}.

\item Multiple-point crossover \citep{DeJong1975a}: A generalization of the previous method, with any number of crossover points up to the number of genes.

\item Uniform crossover \citep{Syswerda1989}: For each gene of the chromosome, a random decision is made to exchange the values between the parents or not.

\item Binary-like crossover \citep{Haupt2004}: In order to reproduce the behavior present in the canonical crossover algorithm, which is applied to a binary representation of the genes \citep{Goldberg1989, Goldberg1990a, Herrera1998a}, \citet{Haupt2004} propose an operator that combines the standard crossover operation with an interpolation approach. The genes located after a crossover point are exchanged, but the gene located at the intersection is modified according to the following:

\begin{equation}

\left\lbrace \begin{array}{l}

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{array} \right.

\**label{eq:binary\_like\_crossover}**

\end{equation}

where $g\_{o1,n}$ and $g\_{o2,n}$ are the $n$-th genes of the two new offspring, and $g\_{p1,n}$ and $g\_{p2,n}$ are those of the two parents. Furthermore, $\beta$ is a random value between zero and one, which can either be unique for the whole chromosome or can change for every gene.

\item Blending method \citep{Radcliffe1991a}: In this approach, instead of exchanging the genes between the chromosomes after some crossover points, they are combined through a linear combination, also using a random value $\beta$.

\item Linear crossover \citep{Wright1991a}: This widens the range of gene values, and produces three children from two parents.

\item Heuristic crossover \citep{Michalewicz1996}: This is a variation of the blending method.

\item Linear interpolation: Unlike the previous techniques, this one does not rely on crossover points, but rather on a linear interpolation on every gene of the couple. This also uses a random value $\beta$, which is the same for every gene.

\item Free interpolation: This technique performs interpolation on each gene, like the previous one. However, in this case the weighting factor $\beta$ changes for each gene.

\end{itemize}

**\subsubsection{Mutation}**

**\label{sec:gas:mutation}**

The combination of strong genes through the chromosomes crossover operation is theoretically the most important operating mechanism in conventional GAs \citep{Holland1992b,Back1993b}. However, many studies identify the mutation process as the key operator, and crossovers as secondary \citep[see][]{Back1992a, Back1996a, Back1996b, Smith1997a, Deb1999, Costa2005a, Costa2007a}.

The mutation operator modifies some gene values. It adds diversity to the population, and prevents a freeze of the evolution, or a genetic drift to a local optimum. Thus, it makes the convergence to the global optimum theoretically possible \citep{Beasley1993a}, as it allows exploration beyond the current region of the parameters space by introducing new characteristics that were not present in the original population \citep{Haupt2004}.

The evaluated and developed mutation operators are listed in the following (equations are only provided for the operator versions that were found to be relevant, the reader may refer to the corresponding literature for details regarding the other options). These apply to genes formed of continuous or discrete variables, but not those that are categorical (e.g., the choice of the meteorological variable or analogy criterion). In the latter case, the random choice of a new value is always based on a uniform distribution, without a notion of distance in the parameters space.

\begin{itemize}

\item Uniform mutation: The mutation rate ($p\_{mut}$) is constant and equal for every gene of each individual. They each have the same probability of mutating. When a gene is selected for mutation, a new random value is assigned, according to a uniform law.

\item Variable uniform mutation \citep{Fogarty1989}: A variable mutation rate over the generations was first suggested by \citet{Holland1992b}, and evaluated by \citet{Fogarty1989}. In most applications, the mutation rate decreases with each generation, in a deterministic and global (for all individuals) manner \citep{Back1992b}.

\item Constant normal mutation: This method employs normal distributions to generate new values of the gene, based on an estimated standard deviation.

\item Variable normal mutation \citep{Horton2012a}: With the same logic as the variable uniform mutation, a mutation operator is tested using a normal distribution with a variable mutation rate and standard deviation, which decrease linearly over generations. This operator has six parameters.

\item Non-uniform mutation \citep{Michalewicz1996}: Two random numbers are chosen 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:

\begin{equation}

g\_{n}^{'} =

\left\lbrace \begin{array}{l l}

g\_{n} + \left(b\_{n}-g\_{n}\right) r\_{2} \varphi^{2} & if \; r\_{1} < 0.5 \\

g\_{n} - \left(g\_{n}-a\_{n}\right) r\_{2} \varphi^{2} & if \; r\_{1} \geq 0.5 \\

\end{array} \right.

\**label{eq:nonuniform\_mutation}**

\end{equation}

with

\begin{equation}

\varphi = 1 - \dfrac{G}{G\_{m}}

\end{equation}

where $a\_{n}$ and $b\_{n}$ are the lower and upper bounds, respectively, of the $n$-th gene, $G$ is the present generation, and $G\_{m}$ is the maximum number of generations.

This operator is adapted for this application, which is not based on a predefined number of generations, by swapping $\varphi$ with $\varphi'$:

\begin{equation}

\varphi' = 1 - \min \left( \dfrac{G}{G\_{m,r}}, 1 \right) \left(1-\omega\right)

\end{equation}

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

\item Individual adaptive mutation rate \citep{Back1992a}: Based on the ideas of evolution strategies \citep[see][]{Rechenberg1973, Schwefel1981}, \citet{Back1992a} introduced the concept of self-adaptive GAs. The idea is to distribute the control parameters within the individuals themselves, which partially decentralizes the control of the evolution. This allows for a reduction in the manual tuning of GAs, and introduces a notion of self-management. The first approach is the introduction of a mutation rate for each individual, which mutates itself under its own probability \citep{Back1992a}. Then, the resulting new rate is employed to mutate the genes of the individual. Thus, as this rate decreases it has a lower probability of being mutated itself. Mutations are performed according to a constant uniform distribution. The initial mutation rates are randomly chosen \citep{Back1992a}, and the method has no parameters. Other approaches exist for introducing a self-adaptation \citep[see][]{Smith1997a, Deb1999, Deb2001a}.

\item Individual adaptive search radius: Based on the ideas of non-uniform mutation, we introduce a search radius to the approach of individual adaptive mutation rates. This search radius $r\_{a}$, bounded between zero and one (relative to the parameter ranges), is also adaptive, and behaves similarly to the adaptive mutation rates. In order to separate its evolution from that of the mutation rate, its own value is also considered as a self-mutation rate for eventually mutating 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:

\begin{equation}

g\_{n}^{'} =

\left\lbrace \begin{array}{l l}

g\_{n} + \left(b\_{n}-g\_{n}\right) r\_{2} r\_{a} & if \; r\_{1} < 0.5 \\

g\_{n} - \left(g\_{n}-a\_{n}\right) r\_{2} r\_{a} & if \; r\_{1} \geq 0.5 \\

\end{array} \right.

\**label{eq:adaptive\_search\_radius\_mutation}**

\end{equation}

where $r\_{1}$ and $r\_{2}$ are randomly selected, such as in non-uniform mutation. Therefore, no external parameter is necessary.

\item Chromosome of adaptive mutation rate \citep[or \textit{n adaptative mutation rate},][]{Back1992a}: Analogously to the individual adaptive mutation rate, this approach leaves the control of the evolution rate to individuals. The difference here is that each gene has a specific mutation rate. The main advantage is that the automatic tuning of the mutation can be made much more precise \citep{Smith1997a}. Thus, a second chromosome containing the mutation rate for each gene of the first chromosome is considered. The operations of mutation and self-mutation are similar to the case of the individual adaptive mutation rate, but proceed in a distributed manner within the chromosome. Moreover, the same crossover operations are applied to this new chromosome as to the main chromosome, and for the same crossing points. Thus, during an exchange of genes, children also inherit the mutation rates specific to each of these genes.

\item Chromosome of adaptive search radius: In this operator, we combine the operations of the chromosome of adaptive mutation rate with the adaptive search radius approach. Similarly, an individual has three chromosomes. The first contains the values to be optimized, the second the distributed mutation rate, and the third the distributed search radius. Again, no external parameter is required.

\item Multi-scale mutation: Finally, we develop another approach that is also based on the search radius concept. However, in this one the rate does not decrease with time. Methods based on a reduction of the mutation rate or radius simulate a transition from the exploration phase to the exploitation phase. Here, the idea is to test an approach that combines both exploration and exploitation during the whole optimization process. Thus, the search radius $r\_{a}$ of Eq. (\ref{eq:adaptive\_search\_radius\_mutation}) is considered as a random value for each individual, but restricted to four equiprobable values, 1, 0.5, 0.1, and 0.02, representing a range from full exploration to fine exploitation. The only external parameter is the mutation rate, which is fixed.

\end{itemize}

**\subsubsection{Elitism}**

The process of elitism is introduced after both natural selection and mutations. This ensures the survival of the best individual, so that a superior solution is never lost. After the natural selection operation, if the previous best individual has not been selected, then it is copied to the IG in place of a randomly selected individual. After mutation, if the previous best individual has mutated and its new version has a lower performance score than the original, then the original is also reinserted in the IG in place of a randomly selected individual.

**\subsubsection{Ending the optimization}**

The convergence check determines whether the solution is acceptable, and if the algorithm may stop. Here, the optimization is stopped if the best individual does not change for $x$ generations. This value should not be too low, in order to allow the algorithm to escape from local optima, and because the rate of improvement decreases with the progression of the optimization. Thus, it is common that the best individual does not evolve over several generations when getting closer to the global solution. A value of $x=20$ generations is chosen.

**\subsection{Implementation and constraints}**

GAs are computationally very intensive because they require many evaluations of the objective function. These assessments require some time (a couple of seconds each) in this application, as they require the calculation and assessment of a prediction for every day of the calibration period, which spans several decades. In order to reduce the computation time, the performance score of an individual, who has previously been evaluated and has not changed, is preserved. Thus, the score of each individual living in the selection is preserved until it mutates.

Because the calculation of the objective function for each member of the population of a generation is completely independent, they are processed in parallel on different CPUs. In order to perform optimizations for multiple time series, the use of a cluster is necessary.

Some constraints must be taken into account. For example, when a crossover or mutation operation results in a parameter value that is outside of the authorized bounds, this must be brought back within the limits. Moreover, the parameters are of different types: continuous, discrete, or categorical (e.g., the selection of the meteorological variable). New values resulting from the optimizer need to respect the type of data that they represent. Other constraints exist between the parameters, e.g., the temporal windows of the relative humidity and the precipitable water predictors must match when processing the moisture index.

**\section{Assessment process and results}**

**\label{sec:assessment}**

Choosing GA options, such as the mutation rate, population size, natural and selection, appears to be difficult given the high number of existing variants, each developed for a specific problem \citep{Haupt2004, Costa2007a}. Thus, different studies suggest significantly different configurations \citep{DeJong1975a, Grefenstette1986, Back1996a, Back1996b}.

In the present study, the choice of meteorological variables (predictors) was still performed manually. Then, GAs were required to jointly optimize, for all levels of analogy, (i) the spatial windows (position and size, which differ between the pressure levels), (ii) the temporal windows (hours of observation of the predictors), (iii) the number of analog situations, and (iv) the selection of the pressure levels (two in this case). The first experiments considered a single level of analogy (on the geopotential height, see Section \ref{sec:am}), and a second level (on moisture) was added later in Section \ref{sec:assessment}\ref{sec:assessment:results}\ref{sec:assessment:mutation}.

**\subsection{Comparison process}**

With the main goal of the present paper being to provide recommendations regarding the use of GAs to optimize different AM variants, a systematic procedure was adopted. The results are summarized in the following \citep[see][for the details]{Horton2012a}. In order to compare different configurations of GAs, a factorial design approach was applied, in a similar manner as in \citet{Costa2005a,Costa2007a} and \citet{Mariano2010a}. This allows the isolation of the effects of a parameter under different combinations of the other options.

In order to evaluate a combination of operators/options, 10 optimizations were processed per configuration of GAs. For all combinations of GA options, 21~630 optimizations were performed on an Intel Xeon based HPC cluster with 96 CPU, running Linux RedHat. Such an assessment was not possible on the whole archive length, and had to be performed on a reduced calibration period of five years from 1998 to 2002, while looking for analog dates over a 48 year period (1961-2008). The total resources required for this comparison amounted to 229~539 h CPU (26.2 years CPU). One should remember that these intensive resources were required in order to assess multiple combinations of GA options, in order to provide recommendations. Based on these, applying GAs to optimize AMs requires some computational power, but not as much (the required time is a function of the archive length, the population number, GA options, and the complexity of the AM to be optimized).

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, (iii) number of generations: characterization of the convergence speed, and (iv) number of evaluations of the objective function: characterization of the required calculation time.

**\subsection{Convergence}**

After a first quick assessment, GAs were found to be successful in optimizing the considered AM implementations. Figure \ref{fig:gas\_evolution\_good} illustrates the evolution of the score of the best individual over progressive generations for the 10 optimizations processed for a given configuration of GAs (with the same operator options). Seven out of 10 optimizations converged to the global optimum (dashed line at the top of the plot in Fig. \ref{fig:gas\_evolution\_good}) after between 40 and 55 generations, which is not the best performance we will observe. The performance score of the sequential approach (bottom line) was quickly exceeded, without introducing any new parameters to the method (Fig. \ref{fig:gas\_evolution\_good}). The improvement was around 8.8\% in this case, which is substantial.

Even GA configurations that were later rejected, because they did not converge (Fig. \ref{fig:gas\_evolution\_bad}), performed significantly better than the sequential approach. GAs displayed a promising potential for optimizing AMs automatically, globally, and objectively.

**\subsection{Results of the comparison}**

**\label{sec:assessment:results}**

The results of the factorial design procedure illustrate the effect of an operator when its contribution is isolated from the other operators. This means that we analyzed the effects of a given operator for equivalent conditions (with the same settings for other operators), while assessing multiple combinations of these. This contribution was then summarized as a percentage of gain/loss with respect to the mean of all variants, for equivalent external conditions. For example, to evaluate the performance of the uniform crossover operator, its performance was compared to the average of all crossover operators, while retaining the same population size, same mutation, natural selection, and couples selection operators.

From the start of the assessments, the importance of the mutation operator was obvious \cite[see][for details]{Horton2012a}, and its leading influence on the optimization performance was evident. Its role is analyzed later, in Section \ref{sec:assessment}\ref{sec:assessment:results}\ref{sec:assessment:mutation}.

**\subsubsection{Breeding operators}**

Each combination of the six options for the couples selection (Fig. \ref{fig:operator\_selectcoupl\_score}) and 21 options for the chromosome crossover operators (Fig. \ref{fig:operator\_crossover\_score}) was evaluated, along with variants of the other operators. This resulted in 1008 combinations, requiring 10080 optimizations.

The performances of the couples selection operators were relatively similar (Fig. \ref{fig:operator\_selectcoupl\_score}). Overall, the tournament selection with three candidates performed slightly better than the others, along with the roulette wheel weighting. However, the latter was a little less effective in terms of convergence and the number of evaluations (not shown). The couples selection operator played no significant role in this application.

As for the crossover operators (Fig. \ref{fig:operator\_crossover\_score}), binary-like crossover (especially with two points of intersection, whether $\beta$ is shared or not) performed slightly better than the others, especially in terms of convergence (not shown).

**\subsubsection{Mutation operator}**

**\label{sec:assessment:mutation}**

Having identified the leading role of the mutation operator, the following sensitivity analysis focuses on it. The 10 different implementations (see Section \ref{sec:gas}\ref{sec:gas:operators}\ref{sec:gas:mutation}) were tested with different options (Fig. \ref{fig:operator\_mutation\_score}), bringing the number of variations to 109. Some optimizations with no mutation were also performed as a reference. Along with variants of the other operators \citep[see][for details]{Horton2012a}, this resulted in 660 combinations (and so 6600 optimizations).

Figure \ref{fig:operator\_mutation\_score} presents the results of this analysis, and illustrates the important role of mutation on the optimization performance. Configurations without mutation (the last box in Fig. \ref{fig:operator\_mutation\_score}) achieved an inferior performance compared to most mutation operators, and the magnitude of this operator influence was significantly higher than those of the other options. The details of the analysis \citep[see][]{Horton2012a} indicate that the other reproduction operators seem to be of secondary importance. This observation is in line with the conclusions reached by several other authors (see Section \ref{sec:gas}\ref{sec:gas:operators}\ref{sec:gas:mutation}).

The mutation operators based on variable normal or variable uniform laws performed particularly poorly here, and were difficult to configure. Many operators presented more or less similar performance scores, and required a variable amount of assessments. The convergence analysis \citep[see][]{Horton2012a} highlighted the three best operators: non-uniform mutation, chromosome of adaptive search radius, and multi-scale mutation. Thus, further optimizations (an additional 4950) were performed, using variants of these three operators (Fig. \ref{fig:operator\_mutation\_score\_atmlevel}-\ref{fig:operator\_mutation\_score\_r4}).

The first analysis was the optimization of the precipitation prediction over a subcatchment (the Binn-Simplon region) in the Swiss Alps (Fig. \ref{fig:operator\_mutation\_score\_atmlevel}). The optimizer also chose the two pressure levels of the atmospheric circulation analogy (this method has a single level of analogy). The resulting CRPSS performance score (see Section \ref{sec:am}) was obviously superior to that obtained using the sequential calibration (bottom line in Fig. \ref{fig:operator\_mutation\_score\_atmlevel}). For most options, it also achieved slightly better results than the optimization without the selection of the pressure levels (dashed line). A clear breakthrough in the performance was not expected, as the former selection of pressure levels is the result of intensive comparative work \citep{Bontron2004}. However, this application demonstrates that when correctly configured, GAs can automatically and successfully choose the pressure levels. However, the automatic selection of the pressure levels significantly increased the difficulty for the GAs to converge to a unique solution, ideally the global optimum. A likely explanation for this is that the pressure levels were considered as categorical values within the optimization (sampled with a uniform law), and thus the approaches relying on a distance in the parameters space, such as the search radius, could not fully exploit the properties that made them efficient. However, even though the results exhibit a certain variability, most of them present very good performance scores, despite different selections of the AM parameters.

The same experiment was performed for another region (Swiss Chablais, see Fig. \ref{fig:map}), which is sensitive to other meteorological influences (Fig. \ref{fig:operator\_mutation\_score\_rhoneamont}), in order to assess the eventual dependencies of the operators with the predictand. Even though differences can be observed with Fig. \ref{fig:operator\_mutation\_score\_atmlevel}, the same options perform better globally.

Next, a second level of analogy was introduced (Fig. \ref{fig:operator\_mutation\_score\_r2}), based on moisture variables (see Section \ref{sec:am}). The GAs had to optimize both levels of analogy (geopotential height and moisture index) simultaneously. Once again, the results were better than those for the sequential calibration (bottom line on Fig. \ref{fig:operator\_mutation\_score\_r2}). Finally, a preselection of air temperature was added, instead of the fixed calendar window, as proposed by \cite{BenDaoud2016}. The results showed generally higher scores (Fig. \ref{fig:operator\_mutation\_score\_r4}), demonstrating the success of the optimizer in taking advantage of this new degree of freedom, and its ability to handle the optimization of three analogy levels simultaneously. Again, the most relevant options were the same globally.

Following these various tests of the relevance of the mutation operators, the following advice can be presented (detailed options are provided in Section \ref{sec:use}\ref{sec:recommendations}):

\begin{itemize}

\item \textit{Non-uniform mutation}: This operator performed well in terms of convergence, mainly when the number of parameters to optimize was rather low. However, the number of required evaluations can be fairly substantial. The main disadvantage of the non-uniform mutation is the number of parameters it requires, which are difficult to estimate a priori. The mutation rate was found to be more important than for the others. The difficulty is that its optimal value may be case-related.

\item \textit{Chromosome of adaptive search radius}: Unlike in the above case, this proposed operator is very robust, as it requires no options and is auto-adapting. It is interesting to note that the insertion of an extra chromosome representing the search radius resulted in a better performance than other self-adaptive operators (such as the chromosome of adaptive mutation rate). If one had to choose a single option for the mutation operator, we would recommend this one, as it was proven effective and requires no parameter.

\item \textit{Multi-scale mutation}: Finally, the multi-scale mutation, which also performed fairly well, requires one parameter, the mutation rate. However, it can also be difficult to estimate a correct value a priori.

\end{itemize}

For this application, the mutation operator has a leading effect, and should be chosen with care. It may be wise to perform multiple optimizations, and to consider these three operators in parallel, in order to obtain results from options that are sometimes either more efficient or more robust. It is interesting to note that the three best techniques incorporate a notion of search distance. It is likely that this notion is the key to these algorithms in this application, and allows them to initially explore the parameter space and then converge. In fact, the search radius directly represents the notion of transition between exploration and exploitation, in our opinion more than a possible evolution of mutation rates.

**\subsubsection{Other options}**

The analysis of the natural selection operator (Fig. \ref{fig:operator\_selectnat\_score}) revealed a slight preference for ratio-elitism compared to the tournament selection, but this was not very significant. This operator, or at least the two assessed versions, did not appear to significantly influence the optimization performance.

The size of the population ($N$) had an effect on the optimization performance (Fig. \ref{fig:option\_taillepop\_score}). A bigger population led to better results, but also to longer optimizations. The required number of evaluations, and thus the required time, was approximately proportional to the population size. Thus, if the population size doubled, then the time required for the optimization also nearly doubled (e.g., for a five year calibration period, approximately 4~h CPU were required when $N = 50$, and 14~h CPU when $N = 200$). The optimal size seems to depend on the complexity of the AM to be optimized. A more complex AM (i.e., with more degrees of freedom) requires a bigger population size. A rule of thumb based on a limited number of case studies (not presented here) is provided as follows:

\begin{itemize}

\item $N\approx100$ for very simple AM implementations (one level of analogy with two pressure levels),

\item $N\approx200$ for a slightly more complex AM (one level of analogy with four pressure levels, or two levels of analogy with fewer pressure levels),

\item $N\approx500$ for significantly more complex AMs (two or three levels of analogy with four pressure levels for the atmospheric circulation, and two to four levels for the moisture analogy).

\end{itemize}

The influence of the size of the IG (proportion of the total population) selected for mating was also assessed (Fig. \ref{fig:option\_popratio\_score}). This option had little influence on the performance of the optimizations. A value of 50~\% seems to be a wise choice.

**\section{Use of GAs to optimize AMs}**

**\label{sec:use}**

**\subsection{Recommended configuration of GAs}**

**\label{sec:recommendations}**

Optimizations using GAs of AMs of varying complexities were performed with a large number of combinations of operators, in order to make recommendations for optimizing AMs (see Section \ref{sec:assessment}). The conclusions are as follows:

\begin{itemize}

\item The population size should be in accordance with the complexity of the AM to be optimized, ranging from 100 for simple cases up to 500 for the most complex AMs.

\item The value of the IG ratio is not significantly important. A value of 50\% seems appropriate.

\item Ratio-elitism performs slightly better than tournaments for the natural selection operator, but this is not decisive.

\item The performance of the operators for the couples selection are relatively similar. The roulette wheel weighting and the tournament selection are more efficient in terms of the convergence and required number of evaluations.

\item Most of the crossover operators yield a relatively similar performance, but the binary-like crossover with two points of intersection is slightly better, especially in terms of convergence.

\item Mutation clearly yields a dominant influence. Three mutation operators stand out: the non-uniform mutation, multi-scale mutation, and chromosome of adaptive search radius. The latter is the most robust, as it has no controlling parameter.

\end{itemize}

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

\begin{itemize}

\setlength\itemsep{-4px}

\item the non-uniform mutation once, with $p\_{mut}=0.05$, $G\_{m,r}=50$, $\omega=0.1$

\item the non-uniform mutation once, with $p\_{mut}=0.05$, $G\_{m,r}=100$, $\omega=0.1$

\item the non-uniform mutation once, with $p\_{mut}=0.1$, $G\_{m,r}=100$, $\omega=0.1$

\item the multi-scale mutation once, with $p\_{mut}=0.1$

\item the chromosome of adaptive search radius twice

\end{itemize}

where $p\_{mut}$ is the mutation rate (or mutation probability), $G\_{m,r}$ is the maximum number of generations during which the magnitude of the research varies, and $\omega$ is a threshold chosen by the user in order to maintain a minimum search radius when the number of generations $G>G\_{m,r}$.

**\subsection{Illustration of application}**

In order to illustrate the achievable gain by using GAs on an AM, precipitation prediction for the southeast ridges region of the alpine Rh\^{o}ne catchment (Fig. \ref{fig:map}) was optimized for the whole calibration period (40 years in the period 1961-2008, with eight years omitted for validation), instead of the smaller five year period.

The optimizer could select the geopotential height at four pressure levels, at any time of day, on unconstrained spatial windows. Moreover, a weighting was introduced to the combination of the criteria processed on each pressure level \citep[such as in][]{Horton2012a, Junk2015}. In this case, no new meteorological variable was added, and the method still contained a unique level of analogy.

The performance (CRPSS) of the reference method, calibrated by means of the sequential procedure, amounts to 32.00\%, with a 0.95 confidence interval (assessed by bootstrapping on 10000 samples) of (29.21\%; 34.94\%). The method optimized by means of GAs achieved a performance of 37.62\%, with a 0.95 confidence interval of (35.21\%; 40.11\%). Thus, the resulting gain is statistically significant. The results on all subregions and the physical meaning of the optimized parameters will be discussed in a forthcoming paper.

**\section{Conclusions}**

**\label{sec:conclusions}**

In order to automatically optimize several AM variants, and to avoid the limitations of the usual sequential calibration, GAs were evaluated. Given the large number of existing operators and options, multiple variants were assessed systematically in order to identify which operators are important, and which variants perform best for the considered AM implementations. The mutation operator was identified as a key element for this application, and new variants were developed that proved efficient, such as the chromosome of adaptive search radius, which is considerably robust (with no control parameter). Recommendations were established for the relevant employment of GAs for optimizing AMs. Although using GAs to optimize AMs may be computationally intensive, once an AM is calibrated, its use in real-time operations it very fast and lightweight.

The possibility that that a different global optimization method or other operators of GAs may perform even better cannot be excluded. Still, the relevance of global optimization techniques for AMs has now been proven, as they provide relevant AM parameters that are automatically, globally, and objectively established. A global optimization is the only way to take into account all of the dependencies between parameters and levels of analogy.

The global optimization approach allows the easy adaptation of AMs to new regions by potentially taking into account local meteorological influences, and thus has a significant potential for application. Moreover, it can be employed to automatically explore datasets, in order to extract the most relevant variables. Thus, this method can make it easier to assess other predictands, such as the temperature, limit of snowfall, or occurrence of hail, while allowing the algorithms to select the best variables and associated parameters.

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% ACKNOWLEDGMENTS

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%

\acknowledgments

Thanks to 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 Dominique B\'{e}rod for his support, and to Michel Bierlaire for his advice regarding optimization methods.

Thanks to the Roads and Water Courses Service, 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 the MINERVE (Mod\'{e}lisation des Intemp\'{e}ries de Nature Extr\^{e}me des Rivi\`{e}res Valaisannes et de leurs Effets) project, which started this research. The fruitful collaboration with the Laboratoire d'Etude des Transferts en Hydrologie et Environnement of the Grenoble Institute of Technology (G-INP) was made possible thanks to the Herbette Foundation. NCEP reanalysis data was provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/. Precipitation time series were provided by MeteoSwiss.

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% APPENDIXES

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%

% Use \appendix if there is only one appendix.

%\appendix

% Use \appendix[A], \appendix}[B], if you have multiple appendixes.

%\appendix[A]

%% Appendix title is necessary! For appendix title:

%\appendixtitle{}

%%% Appendix section numbering (note, skip \section and begin with \subsection)

% \subsection{First primary heading}

% \subsubsection{First secondary heading}

% \paragraph{First tertiary heading}

%% Important!

%\appendcaption{<appendix letter and number>}{<caption>}

%must be used for figures and tables in appendixes, e.g.,

%

%\begin{figure}

%\noindent\includegraphics[width=19pc,angle=0]{figure01.pdf}\\

%\appendcaption{A1}{Caption here.}

%\end{figure}

%

% All appendix figures/tables should be placed in order AFTER the main figures/tables, i.e., tables, appendix tables, figures, appendix figures.

%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% REFERENCES

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Make your BibTeX bibliography by using these commands:

\bibliographystyle{ametsoc2014}

\**bibliography{references}**

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% TABLES

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% Enter tables at the end of the document, before figures.

%%

%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% FIGURES

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%% Enter figures at the end of the document, after tables.

%%

%

\begin{figure}[t]

\centerline{**\includegraphics**[width=19pc,angle=0]{fig01.pdf}}

\caption{Location of the alpine Rh\^{o}ne catchment in Switzerland, and its discretization into 10 subregions: (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. (source: Swisstopo)}

**\label{fig:map}**

\end{figure}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=19pc,angle=0]{fig02.pdf}\\

\end{center}

\caption{Genetic algorithms operational flowchart.}

**\label{fig:structure\_gas}**

\end{figure}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=19pc,angle=0]{fig03.pdf}\\

\end{center}

\caption{Evolution of the score of the best individual over each generation for the 10 optimizations processed for a given configuration. The continuous bottom line represents the score of the sequential approach, and the dashed line (top) shows the supposed global optimum. The circles represent the end of the optimization (when the best individual did not progress during 20 generations).}

**\label{fig:gas\_evolution\_good}**

\end{figure}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=19pc,angle=0]{fig04.pdf}\\

\end{center}

\caption{Same as Fig. \ref{fig:gas\_evolution\_good}, but for a GA configuration considered to be less relevant.}

**\label{fig:gas\_evolution\_bad}**

\end{figure}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=19pc,angle=0]{fig05.pdf}\\

\end{center}

\caption{Influence of the couples selection operators (Section \ref{sec:gas}\ref{sec:gas:operators}\ref{sec:gas:nat\_selection}) 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.}

**\label{fig:operator\_selectcoupl\_score}**

\end{figure}

\begin{figure\*}[t]

\begin{center}

\noindent\includegraphics[width=39pc,angle=0]{fig06.pdf}\\

\end{center}

\caption{Influence of the chromosome crossover operators (Section \ref{sec:gas}\ref{sec:gas:operators}\ref{sec:gas:crossover}) on the optimization performance (improvement of the score). $\beta$s represent a shared $\beta$ parameter, and $\beta$u the unshared version. The same conventions apply as in Fig. \ref{fig:operator\_selectcoupl\_score}.}

**\label{fig:operator\_crossover\_score}**

\end{figure\*}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=19pc,angle=0]{fig07.pdf}\\

\end{center}

\caption{Influence of the mutation operators (Section \ref{sec:gas}\ref{sec:gas:operators}\ref{sec:gas:mutation}) on the optimization performance. In parentheses is the number of variants considered (combination of options). The same conventions apply as in Fig. \ref{fig:operator\_selectcoupl\_score}.}

**\label{fig:operator\_mutation\_score}**

\end{figure}

\begin{figure\*}[t]

\begin{center}

\noindent\includegraphics[width=33pc,angle=0]{fig08.pdf}\\

\end{center}

\caption{Influence of the mutation options (Section \ref{sec:gas}\ref{sec:gas:operators}\ref{sec:gas:mutation}) on the optimization performance, letting the optimizer choose the pressure level of the atmospheric circulation analogy (single level of analogy). For the non-uniform mutation, $\omega=0.1$ in every case. The continuous bottom line represents the score of the sequential calibration, and the dashed superior line is the score of the optimization without automatic selection of the pressure levels. The same conventions apply as in Fig. \ref{fig:operator\_selectcoupl\_score}.}

**\label{fig:operator\_mutation\_score\_atmlevel}**

\end{figure\*}

\begin{figure\*}[t]

\begin{center}

\noindent\includegraphics[width=33pc,angle=0]{fig09.pdf}\\

\end{center}

\caption{Same as Fig. \ref{fig:operator\_mutation\_score\_atmlevel}, but for another region in the Swiss Alps (Swiss Chablais), with different atmospheric influences.}

**\label{fig:operator\_mutation\_score\_rhoneamont}**

\end{figure\*}

\begin{figure\*}[t]

\begin{center}

\noindent\includegraphics[width=33pc,angle=0]{fig10.pdf}\\

\end{center}

\caption{Same as Fig. \ref{fig:operator\_mutation\_score\_atmlevel}, but with a second level of analogy on moisture variables.}

**\label{fig:operator\_mutation\_score\_r2}**

\end{figure\*}

\begin{figure\*}[t]

\begin{center}

\noindent\includegraphics[width=33pc,angle=0]{fig11.pdf}\\

\end{center}

\caption{Same as Fig. \ref{fig:operator\_mutation\_score\_r2}, but with a preselection on air temperature rather than a fixed calendar window.}

**\label{fig:operator\_mutation\_score\_r4}**

\end{figure\*}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=11pc,angle=0]{fig12.pdf}\\

\end{center}

\caption{Influence of the natural selection operators on the optimization performance. The same conventions apply as in Fig. \ref{fig:operator\_selectcoupl\_score}.}

**\label{fig:operator\_selectnat\_score}**

\end{figure}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=11pc,angle=0]{fig13.pdf}\\

\end{center}

\caption{Influence of the population size on the optimization performance. The same conventions apply as in Fig. \ref{fig:operator\_selectcoupl\_score}.}

**\label{fig:option\_taillepop\_score}**

\end{figure}

\begin{figure}[t]

\begin{center}

\noindent\includegraphics[width=11pc,angle=0]{fig14.pdf}\\

\end{center}

\caption{Influence of the intermediate generation (IG) ratio on the optimization performance. The same conventions apply as in Fig. \ref{fig:operator\_selectcoupl\_score}.}

**\label{fig:option\_popratio\_score}**

\end{figure}

\end{document}