Evolution of the Aurelhy method for spatial analysis of climatological fields of precipitation and temperature over France.

Pierre Lassègues[1], Anne-Laure Gibelin, Pascal Simon, Claire Canellas
Météo-France, Direction de la Climatologie et des Services Climatiques, Toulouse, France
[1]Corresponding author: pierre.lassegues@meteo.fr

Introduction

The improvement of the statistical techniques for the production of reference maps of climatological parameters is still a scientific challenge. Météo-France has the responsibility of the production for the french territory of the reference fields of 30 years monthly climate normals and annual means for precipitation, minimum and maximum temperature, number of days of freezing and number of days of precipitation above a threshold. This production is based on Aurelhy, a statistical method using the relation between the meteorological parameter and the parameters of the orography at the climatological scale (Benichou, 1987; Benichou, 1994). Important developments were conducted in the last two years to improve this method.

Data

The data used in this study are the in situ stations monthly climatological normals for precipitation and temperature for the decades 1981-1990, 1991-2000, 2001-2010 over the french territory. The density and the quality of the data are of critical importance. The data are coming from the french climatological data base after QC. The mean density of observations of precipitation is 1station for 160 km² and for temperature 1 station for 320 km². But the lack of stations at high elevation in mountainous zones is an important difficulty.
The second contribution, as input for the spatialization technique, is the digital elevation model (DEM). Several DEMs were tested in the study, coming from the french Institut Géographique National or from the NASA Shuttle Radar Topography Mission (http://www2.jpl.nasa.gov/srtm/), most of them on a 1km grid resolution.

A new software Aurelhy with R language

The Aurelhy method as it was applied by Météo-France until 2015

The Aurelhy method was first implemented by Benichou and Le Breton in 1986. The method is based on a linear model with the principal components of the orography as predictors followed by a kriging of the residuals. The PCA have the interesting advantage of orthogonality, hierarchy and reduced noise. The principal components of the terrain model are produced on a matrix with each line corresponding to the elevations of a selection of points in the vicinity of each point of the target grid. They are called the landscapes in the specific Aurelhy vocabulary. The most important parameters of this selection are range, density and shape. These parameters can be fitted after testing. One usual choice for the landscapes is a square matrix of 11x11 points with a 5 km lag. Elevation itself is also introduced as candidate predictors along with the principal components.
An automatic selection of the candidate predictors is then applied, based on increase of R-squared and Fisher test of significance.
The second step of the Aurelhy method is then a kriging of the residuals and addition of the two contributions. The kriging applied until 2015 was a local kriging with 16 neighbour stations, drift on the coordinates X and Y and a variogram without nugget.
Aurelhy's first implementation was a spatialization technique of the family Regression-Kriging.
The usual grid resolution of the Aurelhy products is 1 km. The model is applied on homogeneous zones (geographicaly and climatologicaly), the french territory is divided in 10 zones.

A new software for the method, based on the R language (https://www.r-project.org/) was implemented in 2015 at Météo-France, offering a large panel of new options.
This new software has improved the possibility to apply the method on very different regions (size of the domain, target grid size, masks for maritime areas) provided that a Numerical Terrain Model is available.
For the PCA of the topography, the spatial scale of the relation between the meteorological variable and the orography can be now carefully scrutinized with the introduction of new possibilities in the building of the Aurelhy landscapes.
The automatic selection of the principal components as best predictors is now available with several new possibilities: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Mallows's CP.
The kriging of the residuals offers extensive possibilities in terms of choosing the kriging parameters. Different values for the range and the nugget of the variogram can be tested as well as the choice of the variogram model fitting. The model was previously of the family Regression-Kriging but Kriging with External Drift is now possible. Elements of comparison between the two methods are presented in Goovaert (1997) and also Hengl et al. (2003).
A possibility of cross-validation is also implemented in the new software.

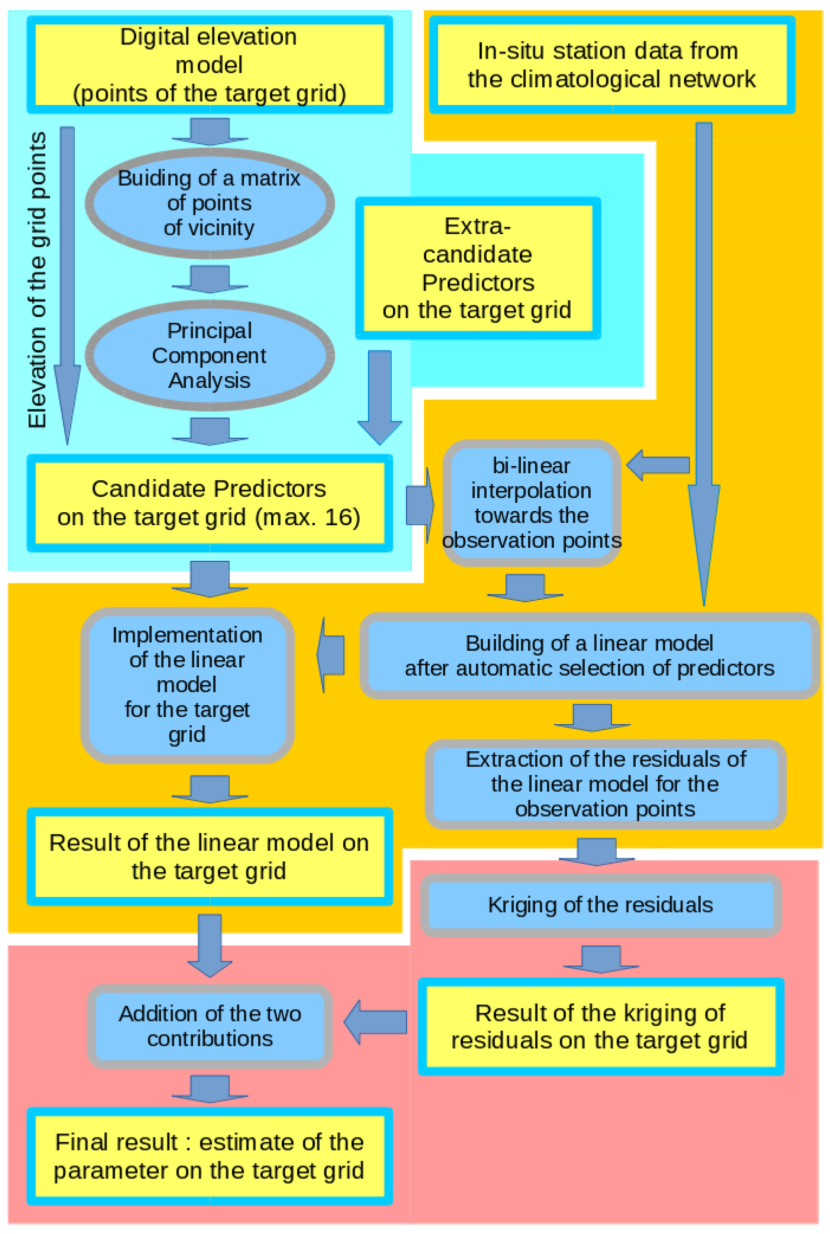

Fig.1 : Flowchart of the Aurelhy method

Tests of new methodological options

A test protocol was designed to produce a comprehensive estimate of the expected advantages of the new methodological options. The different options were tested in several implementations of the Aurelhy analysis for precipitation, minimum and maximum temperature over France. *Table 1* is showing the different steps of this test protocol.
The validation process was based on cross-validation with multiple trials. In this process, each station is put aside one by one and independent estimates with the Aurelhy spatialization are produced. But to assess the uncertainty of the error, multiple trials with a random sample of 20% of the stations were also produced. The most interesting scores analysed were the root mean square error (RMSE), the Peirce skill score (PSS) and the area under ROC curve (AUC). The PSS and AUC are based on a 2x2 contingency table over a threshold. The thresholds analysed in this study are the quantiles 0.1, 0.3, 0.5, 0.7, 0.9 of observations.
Figures 3 is presented to illustrate some results. The results were particularly scrutinized for the mountainous zones. The options showing the best results are the following:
Precipitation: kriging with external drift and free nugget, kriging with variogram fitting maximum likelihood.
Minimum temperature : idem + local kriging with automatic variogram fitting among all the models.
Maximum temperature : idem + all the DEM tested showed better results than the reference DEM, Aurelhy landscape type radial range 25 km and 5 km resolution, association of two different spatial scales for the relation between the meteorological parameter and the topography.
After the individual testing of each new option, their compatibility was also tested in a final new Aurelhy version combining the best ones.

THEME :	Details of implementation :
Reference experiment (Aurelhy operationnal production for France climate normals 1981-2010)	Aurelhy landscapes 11x11 points resolution 5 km, PCA based on eigenvectors pre-calculated for the french orography, linear model on target domain without margin, predictors selection on significant increase of R square, kriging without nugget, drift on the coordinates X and Y, local kriging with 16 neighbor stations, variogram range 80 km.
Principal Component Analysis parameterization (2 experiments)	- dynamical PCA realised on the target domain with the variance-covariance matrix. - idem with correlation matrix.
Digital Elevation Model (5 experiments)	Five DEM are tested from French Institut Géographique National (IGN) or NASA Shuttle Radar Topography Mission (SRTM) with grid size 1km or 250m.
Automatic selection of predictors (3 experiments)	Akaike Information Criterium, Bayesian Information Criterium, Mallows's CP.
Kriging parameters (5 experiments)	Free nugget, Kriging with External Drift with free nugget, idem with variogram fitting maximum likelihood, automatic variogram fitting among all the models possible, drift on the coordinates X-Y put aside.
Aurelhy landscapes parameters (5 experiments)	11x11 points 10 km resolution, 11x11 points 3 km resolution, radial range 50 km 10 km resolution, radial range 25 km 5 km resolution, radial range 15 km 3 km resolution.
New candidate predictors (5 experiments)	Monthly climatology of solar irradiation from Météo-France Ajonc or from SAF CLIMAT, land surface cover classification Corine Land Cover (8 classes or 15 classes), distance to the sea.
Other methodological options (4 experiments)	True elevation of the stations (instead of interpolated from the DEM), linear model built on the target domain + margins, better precision of horizontal localization of the stations, association of two different spatial scales for the relation between the meteorological parameter and the topography.

Table 1 : Test protocol with 30 experiments for the analysis of performance of the new Aurelhy methodological options

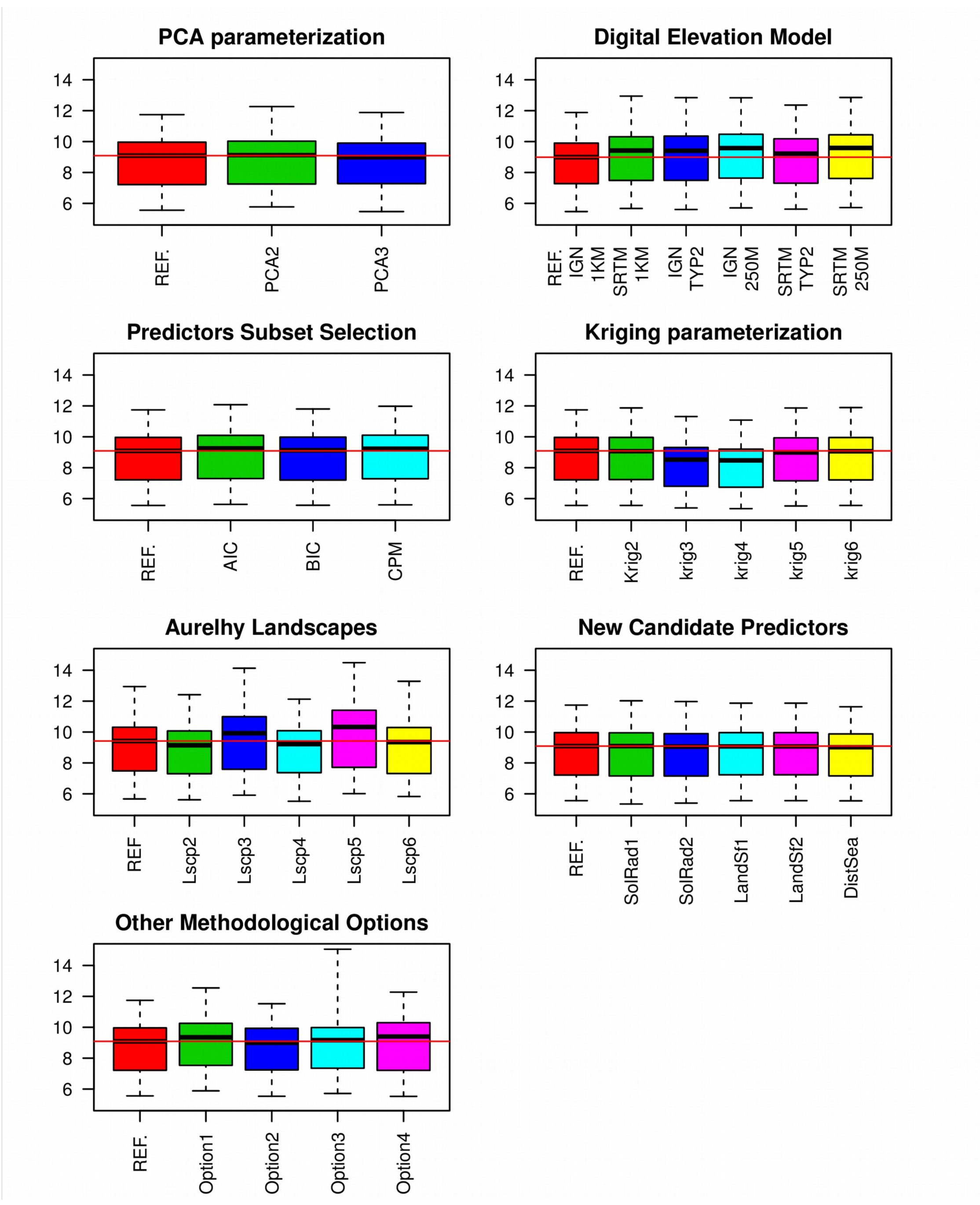

Fig. 3 : Boxplots of RMSE for precipitation for 30 Aurelhy methodological options (see table1) after cross-validation with multiple trial of random sample of 20% of stations. The stations are covering continental France, month of january, april, july and october, three decades 1981-1990, 1991-2000, 2001-2010. Unit: mm/month.

Search of the best spatial scale of the parameters of the orography:

The parameters of the orography resulting of the principal component analysis are very different according to the range and resolution of the Aurelhy landscapes. The goal is to find the best spatial scale for the relation between the orography parameters and the meteorological parameter (and best predictors for the linear model). This point is discussed in Gyalistras (2003) and also in Masson and Frei (2014). To illustrate this point, maps of different fields of principal components produced with different Aurelhy landscapes are presented *Figure 2*, showing very different structures.

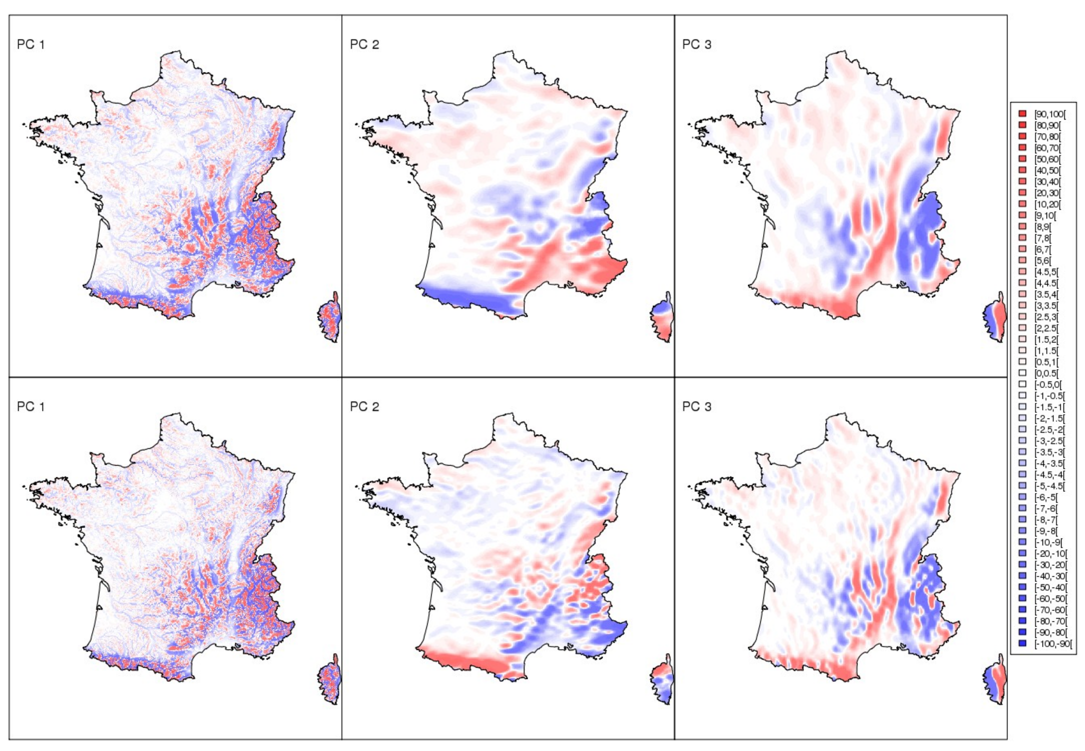

Fig.2 : Fields of the first 3 Principal Component Analysis Scores of the DEM over France with two different Aurelhy landscapes. Top panels: Aurelhy landscapes 11x11 points 5 km resolution. Lower panels : Aurelhy landscapes 11x11 points 3 km resolution.

Conclusion

A new software for the Aurelhy method with R language is now available, offering the possibility to implement very different model of spatialization, not only for the climatological time scales.
The new parameterization of the Aurelhy method, after multiple testing of new options, allowed a decrease of the root mean square error after cross-validation between 10 and 15% for the french reference climate normal fields of precipitation and temperature.
Future improvements with the gap filling of missing values of the time-series of in-situ stations are under consideration.
However, the Aurelhy method is not the only one for the production of climate normals reference maps. Météo-France is also planning new developments for spatial analysis at the climatological, daily and hourly time scale with very different methods.

REFERENCES

- Benichou P. and Le Breton O. (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *La Météorologie* **7**, 23-34.
- Benichou P. (1994) Cartography of statistical pluviometric fields with automatic allowance for topography. *Global Precipitations and Climate Change NATO ASI Series* Volume 26, 1994, pp 187-199.
- Goovaert P. (1997) Geostatistics for natural resources evaluation. *Oxford University Press,* New York.
- Gyalistras D. (2003) Development and validation of a high-resolution monthly gridded temperature and precipitation data set for Switzerland (1951-2000) *Climate research* 25, 55-83, doi:10.3354/cr025055
- Hengl T, Geuvelink G.B.M., Stein A. (2003) Comparison of kriging with external drift and regression kriging. *Technical note, ITC*, http://www.itc.nl/library/Academic_output/ Accessed 26 june 2015
- Masson D. and Frei C. (2014) Spatial analysis of precipitation in a high-mountain region: exploring methods with multi-scale topographic predictors and circulation types. *Hydrol. Earth Syst. Sci. Discussions.* 11, 4639-4694, doi:10.5194/hessd-11-4639-2014

Poster EMS2016-215 september 2016 Trieste