Hydrol. Earth Syst. Sci. Discuss., 12, 311–361, 2015
www.hydrol-earth-syst-sci-discuss.net/12/311/2015/
doi:10.5194/hessd-12-311-2015

Hydrology and Earth System Sciences
Open Access
Discussions

Building long-term and high spatio-temporal resolution precipitation and air temperature reanalyses by mixing local observations and global atmospheric reanalyses: the ANATEM method

A. Kuentz[1,2,*], T. Mathevet[1], J. Gailhard[1], and B. Hingray[3]

[1]Électricité de France, DTG, BP 41, 38040 Grenoble CEDEX 09, France
[2]Irstea, Hydrosystems and Bioprocesses Research Unit (HBAN), 1 rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony CEDEX, France
[3]Laboratoire d'étude des Transferts en Hydrologie et Environnement (LTHE), CNRS, BP 53, 38041 Grenoble CEDEX 09, France
[*]now at: Swedish Meteorological and Hydrological Institute, 601 76 Norrköping, Sweden

Received: 2 December 2014 – Accepted: 9 December 2014 – Published: 12 January 2015

Correspondence to: A. Kuentz (anna.kuentz@smhi.se)

Published by Copernicus Publications on behalf of the European Geosciences Union.

HESSD
12, 311–361, 2015

Abstract

Improving the understanding of past climatic or hydrologic variability has received a large attention in different fields of geosciences, such as glaciology, dendrochronology, sedimentology or hydrology. Based on different proxies, each research community produces different kind of climatic or hydrologic reanalyses, at different spatio-temporal scales and resolution. When considering climate or hydrology, numerous studies aim at characterising variability, trends or breaks using observed time-series of different regions or climate of world. However, in hydrology, these studies are usually limited to reduced temporal scale (mainly few decades, seldomly a century) because they are limited to observed time-series, that suffers from a limited spatio-temporal density.

This paper introduces a new model, ANATEM, based on a combination of local observations and large scale climatic informations (such as 20CR Reanalysis). This model allow to build long-term air temperature and precipitation time-series, with a high spatio-temporal resolution (daily time-step, few km^2). ANATEM was tested on the air temperature and precipitation time-series of 22 watersheds situated on the Durance watershed, in the french Alps. Based on a multi-criteria and multi-scale diagnostic, the results show that ANATEM improves the performances of classical statistical models. ANATEM model have been validated on a regional level, improving spatial homogeneity of performances and on independent long-term time-series, being able to capture the regional low-frequency variabilities over more than a century (1883–2010).

1 Introduction

As highlighted by the even larger number of publications in the recent decades, estimating the hydrological impacts of climate change is a key societal requirement for relevant planning and adaptation. It is however difficult because of the numerous sources of uncertainty associated to climate projections. They are related to emission scenarios, models and also to the internal variability of the climate system, intrinsically arising

HESSD
12, 311–361, 2015

from its chaotic and non-linear nature. Internal variability, leading to multi-scale variations – from multi-year to multi-decadal scales, has long been observed for a number of large but also local scale climate features (Madden, 1976).

In a non-stationary climate, multi-decadal variations can remain high above the long-term trend. In climate projections for the coming decades, they often represent a major source of uncertainties (e.g. Hawkins and Sutton, 2009; Deser et al., 2012). For precipitation or hydrometeorological variables such as streamflow, related uncertainty can be as large or even larger than uncertainties due to climate models (e.g. Terray and Boé, 2013; Lafaysse et al., 2014).

Unfortunately, most climate change impact studies still do not account for this uncertainty source. As an illustration, projected climatic and hydrological scenarios for a given future lead time are classically compared to a so-called reference period (around 30 year of data) expected to be representative of the recent climate context. As shown by Hänggi and Weingartner (2011) with a 200 year runoff time-series of the Rhine at Basel, the hydrological reference features are however likely to highly depend on the period used for their estimation. In such a case, the relevance of conclusions and/or adaptation recommendations formulated with the study may be questionable. To our opinion, they at least suffer from a lack of large historical perspective.

Characterizing the multi-scale variability of climate variables appears today to be necessary (if not mandatory) to put into perspective future climate projections. Numerous studies worldwide have investigated past variability of climate and related variables. For hydrology for instance, the following studies could be considered as representative for France (Renard, 2006), Spain (Lorenzo-Lacruz et al., 2012), Germany (Renner and Bernhofer, 2011), Europe (Stahl et al., 2010), Canada (Zhang et al., 2001), west North-America (Rood et al., 2005), or Australia (CSIRO, 2010). They are based on a set of observed time series available for the region of interest. As the density of observations is significantly lower before 1960 (Hannah et al., 2011), most time-series however usually cover a few decades only, which is obviously not sufficient for a relevant analysis of multidecadal variations (Mathevet and Garçon, 2010; Hannaford

HESSD
12, 311–361, 2015

et al., 2013). Long-term historical time-series (covering a period longer than 100 year) are of course the ideal material for this analysis. Such historical series were for instance used for the Loire river in France (Renard, 2006), the Colombia and Missouri rivers in the USA (Rood et al., 2005), the Murray-Darling basin in Australia (CSIRO, 2010) or more recently for a larger panel of French stations by Boé and Habets (2013). Long-term streamflow time-series are obviously rather rare, with a very low spatial density. Some could still be rescued from various national and regional archives sources but the rescue process is long and requires demanding digitizing and quality check. Finally, the temporal homogeneity of data is often questionable (e.g. because of the evolution of measurement practices as shown in Kuentz et al., 2012, 2014, anthropogenic influences, etc.) preventing the use of series for the variability analysis.

Characterizing the long-term variability of climate and related variables from observations is therefore usually not possible. An alternative is to reconstruct the past temporal variations of the variable of interest. A number of reconstruction approaches have been presented for numerous fields of geosciences. They can use environmental markers like for instance tree rings (Frank and Esper, 2005), lake sediments (Wilhelm et al., 2013, 2012), narrative evidences of droughts (Pfister et al., 2006), geochemical tracers in ice core from glaciers (Jouzel et al., 2007). For the reconstruction of past streamflows variations, an efficient way is simulation, where simulated discharges times series are obtained with a hydrological model from past variations of meteorological variables available for the region. In some particular cases, meteorological observations required for such an analysis may cover a much longer time period than the period for which hydrological data are available. They obviously suffer, in most cases, from the same scarcity and length limitations than hydrological data. In such a case, meteorological data can also be reconstructed. A classical reconstruction is obtained using external data (proxy data) from long term series of observations available from one or several neighbouring stations.

Local meteorological data can alternatively be reconstructed from past climate variations. The recent release of two major atmospheric reanalyses for the 20th century

HESSD
12, 311–361, 2015

(from 1871 to present year for the NOAA 20CR, Compo et al., 2011, from 1900 for the ECMWF ERA-20C, Poli et al., 2013) present a great opportunity for such a reconstruction. As their spatial and temporal resolutions do unfortunately rarely fit the resolution standards (typically sub-daily time step, up to 1000 km^2) required for hydrological applications, the reconstruction of local meteorological data is obtained through downscaling.

This study compares 3 different statistical approaches for the reconstruction of high resolution precipitation and temperature data. Reconstructions are respectively obtained from observations available at a neighbouring station, from large scale atmospheric variables extracted from the 20CR reanalyses, and from both data at a time. If the two first approaches have been already applied for similar studies, the last one is original, as it makes use of both local observations and large scale atmospheric information simultaneously. Reconstructions are built at a daily time scale for the 22 subcatchments of the Upper Durance River basin, a mesoscale catchment located in the South-Eastern Alps. They were produced for further hydrological reconstructions for the past 140 year. An exhaustive evaluation of the whole hydrological reconstruction process can be found in Kuentz (2013).

The Upper Durance River basin, meteorological and atmospheric data are presented in Sect. 2. The three reconstruction models are presented in Sect. 3. They are evaluated and compared in Sect. 4. Section 5 shortly discusses the long-term climatic variability reconstructed over the 1870–2010 period with the models. Finally, conclusions and perspectives of this work are given in Sect. 6.

2 Data

2.1 Case study location and spatial climatic inputs

The three methods have been applied for the reconstruction of mean areal temperature and precipitations of 22 sub-basins of the Durance River basin, a mesoscale alpine

HESSD
12, 311–361, 2015

watershed located in South-Eastern France (Fig. 1). The main characteristics of the watersheds are detailed in Table 1.

Limited in the North by the Écrins Alpine massif and in the South by the Mediterranean Sea, the various sub-watersheds highlight very different climate. Upstream hydrological regimes are snow dominated with high snowmelt flows in late spring and early summer. When moving downstream, they become more Mediterranean with additional autumn floods due to large rainfall amounts in this period.

Daily mean areal air temperature and precipitation have been estimated for each watershed over the 1948–2010 period from the SPAZM meteorological analysis produced by Gottardi et al. (2012). In the following the 1948–2010 will be referred to as the "observed period" and the SPAZM series will be referred to as "observations", although it is not direct observations but mean areal air temperature and precipitation series, aggregated at the watershed scale from local observations of temperature and precipitation (Gottardi et al., 2012).

2.2 Local reference long series

To reconstruct the mean areal air temperatures and precipitation of the 22 watersheds, we searched for the longest observed series on or near the Durance watershed. In a technical report published in 1892, Imbeaux (1892) reported four air temperature and forty precipitation measurement stations in the watershed and its neighborhood. Unfortunately, most data have been lost. Very few and incomplete series are still available nowadays. For precipitation, we were able to rebuild a 1883–2010 series for Gap location by merging two sources of data, provided respectively by Electricité de France (EDF) and Météo-France. For air temperature, the nearest daily series, provided by Météo-France, was found in Marseille and covers the period 1868–2010.

For a qualitative assessment of the reconstructed series, we also considered 5 monthly time series from the HISTALP project database (monthly series, Auer et al., 2007). The series also started in the 1870's. For air temperature, the corresponding stations are located around the South-Eastern part of the Alps, in Genova University,

HESSD
12, 311–361, 2015

Milano-Brera, Montpellier, Nice airport and Nîmes airport. For precipitation, they are quite closer to the Durance watershed, located in the cities of Aix-en-Provence, Nice (Cap-Ferrat), Orange, Saint-Paul-les-Durance and Toulon.

2.3 Large scale climatic data

Large scale atmospheric data are extracted from the "20th Century Reanalysis" ("20CR", Compo et al., 2011) from the project of the same name supported by the US Department of Energy and by the Climate Program Office of National Oceanic and Atmospheric Administration (NOAA). This reanalysis has been produced by assimilating only sea level pressure data, which allows it starting as soon as the end of 19th century. The reanalysis covers the 1871–2010 period.

In the present work, large scale variables used for the reconstruction are the fields of geopotential heights at 700 and 1000 hPa in the rectangular spatial domain situated between the points (8° O, 38° N) and (12° E, 50° N).

3 Methodology: mixing two sources of information

In climatology or hydrology, the reconstruction of past climatic data is usually necessary, either to estimate missing values, assess data quality or build long term climatic reanalyses. Different methods are classically used to reconstruct climatic observations. Some of them are only based on the series at reconstruction itself (long-term average or regime, temporal interpolation techniques...), others are based on external data (proxy data) used to calibrate and run a reconstruction model. For climatic reconstructions, proxy data could either be local or regional scale observations and either have the same or a different nature of the reconstructed series.

In the following section, we present the three methods used for the reconstruction. The first one uses local neighbour observation of a similar proxy (respectively, air temperature or precipitation observation). The second is basically a downscaling approach

HESSD
12, 311–361, 2015

using regional large scale information of a different proxy (geopotential fields). The third one uses both proxies at a time.

As in most reconstruction works, the methods rely on a period on which both proxy data and data at the reconstruction point are available. This concomitant period will be referred as the observation period. The reconstruction period is the period on which the reconstruction model is applied: it corresponds to the period where proxy data are available while data is missing at the reconstruction point.

3.1 Local information

A classical method used for climatic reconstructions is based on regression like models, where predictors are local observed data well correlated with the data at reconstruction. This model is calibrated against observations on their common period of availability (observation period) and the series at reconstruction is either filled or extended by the modelled series. In the following, the series to be reconstructed will be called target series (Tg), and the local series used as reference will be called neighbour series (Ne), although the neighbourhood of these two series is variable in space or time and concretely depends on the availability of other local observed series.

In this paper, to reconstruct the target series, a local observed neighbour series will be used through a simple linear regression model, called local model (LM). The estimate $\widehat{X}_{\mathrm{LM},d}$ of the target variable X obtained with LM for a given day d has the classical following expression:

$$\mathrm{LM}: X_{\mathrm{LM},d} = \alpha \cdot X_{\mathrm{Ne},d} + \beta + \varepsilon_d \tag{1}$$

where $X_{\mathrm{Ne},d}$ is the value of the neighbour series for the day d, α is a multiplicative correction factor, β is an additive correction factor and ε_d is a residual assumed to have zero mean. Depending on the nature of the reconstructed series (air temperature or precipitation), the correction factor is either only multiplicative (i.e. $\beta = 0$) or additive (i.e. $\alpha = 1$).

HESSD
12, 311–361, 2015

3.1.1 Air temperature reconstruction

For air temperature reconstruction, LM classically uses an additive correction, assumed to be constant over time and mainly influenced by the difference of altitude between the target and neighbour series. However, even when target and neighbour series are very well correlated, residuals of such models usually exhibits a strong seasonal pattern. Then, LM can be slightly improved assuming that the additive correction varies over time. In the present case, it is represented by a daily harmonic function, calibrated on the interannual mean monthly residuals of the differences between the target series and the neighbour series.

The Local Model for air temperature reconstruction can thus be written as:

$$\mathrm{LM} : T_{\mathrm{LM},d} = T_{\mathrm{Ne},d} + \beta_d + \varepsilon_d \quad (2)$$

where $T_{\mathrm{LM},d}$ is the estimate of the target air temperature for the day d, $T_{\mathrm{Ne},d}$ is the value of the neighbour series temperature for this same day, β_d is the correction, function of the calendar day of the year, and ε_d is a residual assumed to have zero mean.

In this paper, we chose to use this model in a deterministic way, that is without considering the residual term. Uncertainty is accounted for in the mixed model as explained in Sect. 3.3.

3.1.2 Precipitation reconstruction

For precipitation reconstruction, LM classically uses a multiplicative correction, assumed to be constant over time. This multiplicative correction is more adequate for precipitation and compatible with the asymmetrical distribution of precipitation values (no negative values). The correction factor is taken constant all over the year. The improvement brought by a variable correction has nevertheless been assessed and shown as negligible (Kuentz, 2013). The constant multiplicative correction factor is calibrated over the common data availability period as:

HESSD
12, 311–361, 2015

$$\alpha = \frac{\overline{P_{\mathrm{Tg}}}}{\overline{P_{\mathrm{Ne}}}} \tag{3}$$

where $\overline{P_{\mathrm{Ne}}}$ is the mean value of the neighbour series and $\overline{P_{\mathrm{Tg}}}$ is the mean value of the target series, both calculated on the common data availability period.

The Local Model used for precipitation reconstruction reads:

$$\mathrm{LM}_d : P_{\mathrm{LM},d} = \alpha \cdot P_{\mathrm{Ne},d} + \varepsilon_d \tag{4}$$

where $P_{\mathrm{LM},d}$ is the estimate of the target precipitation for the day d, $P_{\mathrm{Ne},d}$ is the value of the neighbour series precipitation for this same day and ε_d is a residual with zero mean. As explained previously for the air temperature reconstruction, a simple version of this model with a residual term considered equal to zero is used in this paper, as uncertainty will be taken into account in another way.

3.2 Large scale climatic information: the Analog method

The second reconstruction method is the analog method introduced by Lorenz (1969). The method is currently largely used to produce meteorological scenarios in the context of weather forecasting (Van Den Dool, 1989; Horton et al., 2012) or climate projections (Teng et al., 2012; Bourqui et al., 2011; Hingray et al., 2013). The method is seldom applied to the reconstruction of climatic series over the past as made by Timbal et al. (2006). Nevertheless, the release of the long atmospheric reanalyses for the 20th century opens doors for more development of such uses, allowing for the reconstruction of long climatic series covering the entire 20th century.

The analog method is based on the fact that local meteorological variables are strongly influenced by the state of the atmosphere and its circulation at the synoptic scale. As long as a long enough archive with concomitant local and large scale observations is available, it is then possible to produce local meteorological scenarios for

HESSD
12, 311–361, 2015

any other day for which the required large scale atmospheric predictors are available. For this, the k days that are the most similar to the target day in terms of atmospheric circulation are first identified in the archive. The surface meteorological variables observed for one of those analog days are next used as weather scenario for the target day.

In our case, the archive is the 1948–2010 observation period. As large scale atmospheric predictors are available for each day of the 1883–2010 period covered by the 20CR atmospheric reanalysis, the method allows for the reconstruction of 127 year time series of daily and local meteorological variables.

The Analog Model has some parameters to be set such as the type or the level of predictors, the number of analog days selected for the prediction, the spatial domain used to compute the similarity criterion or the similarity criterion itself. Numerous variations of the analog method have been developed. In the present work, we use the Analog Model (ANA) presented by Obled et al. (2002) and further explored by Bontron and Obled (2005), Ben Daoud et al. (2010) and Horton et al. (2012). Its main features are presented below.

- *The predictors* are the geopotential fields at 700 and 1000 hPa for the times 0 and 24 h.
- *The similarity criterion* is Teweles and Wobus (1954). This score is based on the shape of the geopotential fields and have been shown to perform better than a classical Euclidean distance for this type of use (e.g. Wetterhall et al., 2005).
- *The spatial domain* used to estimate the similarity includes all grid points between the longitudes 8° W and 12° E and the latitudes 38 and 50° N, with a step of 2°.
- *A moving seasonal filter* is used for the determination of candidate analog days. For each target day, candidate analog days are days which calendar day is included in a 60 days interval around the calendar day of the target.

HESSD
12, 311–361, 2015

The reconstruction is deterministic when only one analog is used (classically the nearest analog). The analog day can be also selected among the n nearest analogs. As a result, an ensemble of reconstructions can be produced. This allows characterizing the uncertainty in the reconstruction. The ensemble of reconstructions obtained with ANA model for the variable X and day d will be written in the following $[X_{\mathrm{ANA}_d^k}]_{k=1\ldots n}$ where $k = 1\ldots n$ refers to the n nearest analogs selected for the day d. In the present case, the selection is done among the 50 nearest analogs ($n = 50$).

3.3 Mixing formulation: the ANATEM method

Both local and large scale predictors are available for the 1870–2010 period. The Local Model LM and the Analog Model ANA can be therefore used to produce two different reconstructions of precipitation or air temperature for this period, one based on local observed data (another station with available data), the other from large scale atmospheric information (synoptic variables). The originality and the strength of the ANATEM method introduced here is to combine the two previous models and to therefore take advantage of both local and synoptic information.

The principle of ANATEM is the following: the local variable reconstructed for the target day d is the local variable estimate obtained by the local model, corrected by the errors of the Local Model identified when it is applied for the prediction of the local variable on the n analogs days. In other words, for any target day, the Analog Model ANA allows the identification of n analog days in terms of atmospheric circulation (see Sect. 3.2). The n prediction errors respectively obtained when the Local Model LM is used for predicting the local observed value for each of these n days are used to define the error distribution associated to the prediction obtained with the Local Model for the target day d. The prediction obtained with ANATEM for the target day is therefore probabilistic.

HESSD
12, 311–361, 2015

3.3.1 Air temperature reconstruction

Let ANA_d^k be the kth analog day selected for the target day d. Let's write $T_{\mathrm{ANA}_d^k}$ the observed air temperature for this day, and $T_{\mathrm{LM,ANA}_d^k}$ the air temperature estimate that would have been obtained with the local model LM when applied for the prediction of temperature for this same kth analog day.

The ANATEM method assumes that the error made by the local model LM for the target day d could be the same as the error made by this same model for the day ANA_d^k. In such a case, the estimate of the local temperature $\widehat{T}_d^k$ obtained for day d through ANATEM using the kth analog day would read:

$$\widehat{T}_d^k = T_{\mathrm{LM},d} + T_{\mathrm{ANA}_d^k} - T_{\mathrm{LM,ANA}_d^k} \quad (5)$$

where $T_{\mathrm{LM},d}$ is the air temperature estimate that would have been obtained with LM for the target day.

According to Eq. (5), the reconstructed temperature $\widehat{T}_d^k$ obtained with the kth analog day can be written as a linear function of the value of $T_{\mathrm{LM},d}$ as follows:

$$\widehat{T}_d^k = f(x_d) = x_d + a_d^k \quad (6)$$

with $x_d = T_{\mathrm{LM},d}$ and $a_d^k = [T_{\mathrm{ANA}_d^k} - T_{\mathrm{LM,ANA}_d^k}]$.

Thus, the value of $\widehat{T}_d^k$ can be read in the (T_{LM},T) space as the ordinate of the point with abscissa $T_{\mathrm{LM},d}$ in the line which slope is 1 and intercept is $T_{\mathrm{ANA}_d^k} - T_{\mathrm{LM,ANA}_d^k}$. Such representation is shown in Fig. 2 (left): each analog day selected for the day d is represented by a black point in the (T_{LM},T) space, and the line described above is plotted in red. The vertical distance between this line and the first bisector plotted in green represents the error a_d^k made by the local model for the analog day ANA_d^k, which is transposed to the target day to obtain the ANATEM estimation.

By repeating this operation for each analog day, a set of parallel lines is obtained as shown in Fig. 2 (right). The ensemble of reconstructed temperatures for day d is

HESSD
12, 311–361, 2015

the local model estimate $T_{\mathrm{LM},d}$ corrected by the ensemble of errors made by LM over the n analog days. It can be represented by the ensemble of points with abscissa $T_{\mathrm{LM},d}$ in each of the n parallel lines plotted in Fig. 2 (right). The distribution of these reconstructed values $[\widehat{T}_d^k]_{k=1\ldots n}$ is represented by a boxplot (10, 25, 50, 75 and 90 % quantiles) on the right part of the figure, and the mean value is shown by a blue point.

Finally, the probabilistic prediction from the ANATEM method for day d has the following expression:

$$\left[\widehat{T}_d^k\right]_{k=1\ldots n} = T_{\mathrm{LM},d} + \left[T_{\mathrm{ANA}_d^k} - T_{\mathrm{LM,ANA}_d^k}\right]_{k=1\ldots n}. \quad (7)$$

For the example shown in Fig. 2, over the n analog days having a similar synoptic situation than day d, the local model estimate T_{LM} was in average higher than the observed temperature at the target point. This lead to negatively correct the model for the reconstructed day d. While the value of the local model was −9.0 °C, the fifty air temperature values produced by the ANATEM method have a mean of −11.2 °C and their 10 and 90 % quantiles are respectively −13.1 and −9.3 °C.

3.3.2 Precipitation reconstruction

The ANATEM method uses the same principle for precipitation reconstruction. Another formulation was however proposed, due to the specific features of precipitations (asymmetric distribution, lot of zero values).

The additive correction formulation used for the probabilistic reconstruction of temperature is not suitable for precipitation. It can actually produce negative values as illustrated in Fig. 3 (left), elaborated on the same principle than Fig. 2 (right).

An alternative formulation is to use a multiplicative correction for each analog date. The probabilistic reconstruction is here defined by the following expression:

$$\left[\widehat{P}_d^k\right]_{k=1\ldots n} = P_{\mathrm{LM},d} \cdot \left[\frac{P_{\mathrm{ANA}_d^k}}{P_{\mathrm{LM,ANA}_d^k}}\right]_{k=1\ldots n}. \quad (8)$$

HESSD
12, 311–361, 2015

The multiplicative formulation obviously avoids the estimation of negative precipitation values. A graphical representation of this reconstruction strategy is given in Fig. 3 (right). As illustrated, the reconstructed values seem to be reasonable for common values, but the reconstruction can produce unreasonable high values of precipitation.

In the following, we have therefore chosen to build the probabilistic reconstruction of precipitation with a correction model that has a multiplicative behaviour for low values of LM_d and an additive behaviour for high values of LM_d. Its analytical formulation and its asymptotic behaviours when x_d tends to zero or infinity have the following expressions:

$$\widehat{P}_d^k = f(x_d) = \frac{x_d^2 + a_d^k \cdot x_d}{x_d + b_d^k} \tag{9}$$

where $x_d = \mathrm{LM}_d$ and where a_d^k and b_d^k are parameters to be expressed in function of $P_{\mathrm{ANA}_d^k}$ and $\mathrm{LM}_{\mathrm{ANA}_d^k}$.

$$\widehat{P}_d^k = \frac{x_d^2 + a_d^k \cdot x_d}{x_d + b_d^k} = x_d \cdot \frac{x_d + a_d^k}{x_d + b_d^k} \underset{x_d \to 0}{\sim} x_d \cdot \frac{a_d^k}{b_d^k} \tag{10}$$

$$\widehat{P}_d^k = \frac{x_d^2 + a_d^k \cdot x_d}{x_d + b_d^k} = x_d \cdot \left(1 + \frac{a_d^k}{x_d}\right) \cdot \left(1 + \frac{b_d^k}{x_d}\right)^{-1} \underset{x_d \to +\infty}{\sim} x_d + (a_d^k - b_d^k) \tag{11}$$

We define the two model parameters a_d^k and b_d^k following the work of Dufour and Garçon (1997) for the assimilation of streamflow data in a hydrological model. The parameters are expressed as a function of $P_{\mathrm{ANA}_d^k}$ and $P_{\mathrm{LM,ANA}_d^k}$ in order to reach a compromise between a good multiplicative behaviour for low values and a good additive behaviour for high values:

HESSD
12, 311–361, 2015

$$a_d^k = P_{\mathrm{ANA}_d^k} \text{ and } b_d^k = \frac{\left(P_{\mathrm{LM,ANA}_d^k}\right)^2}{P_{\mathrm{ANA}_d^k}}. \quad (12)$$

The probabilistic reconstruction obtained with ANATEM for precipitation finally reads:

$$\left[\widehat{P}_d^k\right]_{k=1\ldots n} = \left[\frac{{P_{\mathrm{LM},d}}^2 + P_{\mathrm{ANA}_d^k} \cdot P_{\mathrm{LM},d}}{P_{\mathrm{LM},d} + \frac{{P_{\mathrm{LM,ANA}_d^k}}^2}{P_{\mathrm{ANA}_d^k}}}\right]_{k=1\ldots n}. \quad (13)$$

The graphical representation of this formulation is shown in Fig. 4. The left side graphic shows the curve corresponding to Eq. (13) applied for one analog day ANA_d^k, and the right side graphic shows the ensemble of curves associated respectively to the n analog days. The distribution of the reconstructed values $[\widehat{P}_d^k]_{k=1\ldots n}$ is represented by the boxplot.

In the case of very different values of $P_{\mathrm{ANA}_d^k}$ and $P_{\mathrm{LM,ANA}_d^k}$, Eq. (13) can potentially produce unreasonably high values of corrected precipitation $\widehat{P}_d^k$. In order to avoid such values we applied the following filters:

- if $P_{\mathrm{ANA}_d^k} > 10 \cdot P_{\mathrm{LM,ANA}_d^k}$ then the value of $P_{\mathrm{ANA}_d^k} = 10 \cdot P_{\mathrm{LM,ANA}_d^k}$, and
- if $P_{\mathrm{ANA}_d^k} < \frac{1}{10} \cdot P_{\mathrm{LM,ANA}_d^k}$ then the value of $P_{\mathrm{ANA}_d^k} = \frac{1}{10} \cdot P_{\mathrm{LM,ANA}_d^k}$.

The filtering threshold (10) has been chosen arbitrary. A sensitivity analysis with different values from 2 to 100 showed that this threshold has fairly no impact on the reconstruction performances because very few analog days are generally affected by this filtering operation. The filters are represented by blue zones on Fig. 4.

For the example day shown in Fig. 4, the local model LM gives a reconstructed value of 15.0 mm. The mean, the 10 and 90 % percentiles of the probabilistic reconstruction obtained with ANATEM are respectively 14.8, 7.8 and 21.0 mm.

HESSD
12, 311–361, 2015

4 Analysis of ANA, LM and ANATEM performances

4.1 Evaluation process

The data presented in Sect. 2 allow reconstructing the daily air temperature and precipitation series for the 22 selected watersheds over the period 1883–2010. The reconstruction is deterministic for the LM model. For ANATEM and ANA, 50 reconstruction time series are generated from the probabilistic reconstructions obtained each day of the period. In the present section, the three reconstruction models are evaluated based on their reconstruction skill for the 1948–2010 observed period

The evaluation is based on three criteria. The ratio β between the mean estimated value and the mean observed value of the variable evaluates the bias of the reconstruction. The ratio between the SD of the reconstructed and of the observed values (α) evaluates the ability of the reconstruction to reproduce the observed variability of the variable. The coefficient of correlation r between observed and reconstructed series additionally measures the ability of the reconstruction to reproduce the observed temporal variations (e.g. alternating dry/wet or warm/cold periods). The overall performance obtained for these three criteria is additionally summarized within the Kling–Gupta Efficiency criterion (KGE; Gupta et al., 2009) defined as following:

$$\mathrm{KGE} = \sqrt{(1-r)^2 + (1-\alpha)^2 + (1-\beta)^2}. \quad (14)$$

The ability of the reconstruction to reproduce the variability and variations of observations is carried out for multiple temporal resolutions: daily (high frequency variability), monthly (accounting thus for the infra-annual variability) and annual (low-frequency variability) resolutions. For annual resolution, series are aggregated by hydrological year, i.e. from 1 October to 30 September.

In the following sections, we first present the performance of the three models for an illustrative watershed (L'Ubaye at Barcelonette). The evaluation relies (1) on the graphical comparison of the observed and reconstructed annual series for the 1948–

HESSD
12, 311–361, 2015

2010 period and (2) on the distributions obtained for r, α, β and KGE when estimated for the daily, monthly and annual time-step from the 50 ensembles. We next present results obtained for the 22 watersheds of the Durance Basin.

4.2 Performance for L'Ubaye at Barcelonnette

4.2.1 Air temperature reconstruction

Figure 5 presents the mean annual time-series of mean air temperature of the watershed. These figures firstly show that the observed temperature has increased during last 60 year, with a mean value around 3 °C in the 50's and a mean value around 4 °C nowadays. The ANA model do not capture the temporal evolution and the variability of air temperature, contrary to LM and ANATEM. We also notice that ANA ensembles are much wider than ANATEM ensembles. This is highlighted with the distributions obtained for the different criteria at the annual time-step (Fig. 6, right part):

- ANA has a limited mean bias (β close to 1), but has rather bad temporal correlation and significant bias of variability, which is exhibited by rather low mean values of r, α, KGE (between 0.2 and 0.6, not visible on the figure);
- LM and ANATEM present very good temporal correlations (mean r greater than 0.9) and limited mean and variability bias. The two models have slightly different skills: LM has no mean bias (by construction) but a significant variability bias (mean α less than 0.9) whereas ANATEM has a limited mean and variability bias (mean β and mean α greater than 0.95).

Figure 6 presents also the distributions of the criteria for daily and monthly time-steps. The hierarchy between the three model is comparable at these time-steps, with KGE values ranging from 0.77 to 0.87 for ANA, ranging from 0.95 to 0.98 for LM and ranging from 0.93 to 0.97 for ANATEM. Moreover, the mean criteria are higher at a monthly time-step and then a daily time-step, compared to the annual time-step finally. This means that models have more difficulties to reproduce annual or daily variability than

HESSD
12, 311–361, 2015

intra-annual variability (this is partly due to the seasonality of air temperature). LM model is performing slightly better than ANATEM at daily and monthly time-steps. Conversely, ANATEM is performing better than LM at annual time-steps.

4.2.2 Precipitation reconstruction

Figure 7 presents the observed and reconstructed annual time-series of mean precipitation of the watershed. As highlighted, observed annual precipitation present a strong variability, ranging from 1000 mm yr^{-1} for given periods to 1500 to 2000 mm yr^{-1} for some exceptional years around. The three models capture rather well this variability and are able to reproduce wet or dry periods. ANA ensembles are much wider than ANATEM ensembles.

The distributions of criteria at the annual time-step (Fig. 8, right part) confirms these statements:

- ANA has a moderate correlation (mean r close to 0.5), LM and ANATEM have a rather good correlation (mean r greater than 0.8);
- ANA and LM have a limited mean bias (less than 0.02), ANATEM has a moderate mean bias (less than 0.05);
- ANA and ANATEM have a limited variability bias (less than 0.02), LM has a moderate variability bias (less than 0.05).

The hierarchy between the three models is comparable at daily and monthly time-steps, with KGE values ranging from 0.4 to 0.7 for ANA, ranging from 0.8 to 0.9 for LM and ranging from 0.78 to 0.88 for ANATEM (Fig. 8). ANA is clearly poor at a daily time-step, with a very limited correlation (r less than 0.4). The mean criteria are higher at a monthly time-step and similar at daily and annual time-steps. As for air temperature, this highlights the difficulty of the models to reproduce the low and high frequency variability while the infra-annual variability is well-captured.

HESSD
12, 311–361, 2015

4.3 Performance for all 22 watersheds

For the sake of conciseness, we consider here for the evaluation only one reference time-series for each model. For the Local Model, this is simply the reconstruction obtained with the model. For the probabilistic reconstruction models ANA and ANATEM, this is the mean time-series derived from the ensemble of 50 time-series reconstructions (the daily value for a given day is the mean of the probabilistic reconstruction for this day). As it will be noticed later, these mean time series obviously present a much lower variability than each time series of the reconstruction ensemble. For the sake of simplicity, these mean time series will be also referred to as reconstructed time series. In the following, the performance of a given model will be presented with the distributions of r, α, β and KGE criteria obtained for the 22 watersheds at the daily, monthly and annual time-steps.

4.3.1 Air temperature reconstruction

The main results obtained for air temperature reconstruction are (Fig. 9):

- at a daily and a monthly time-steps, ANA suffers from a limited positive mean bias (mean β around 1.03) and a significant negative variability bias (mean α from 0.85 to 0.88). Correlation with observations is very good (mean r greater than 0.93). At the annual time-step, ANA is not able to capture the long-term variability and trend, with a very low correlation (mean r close to 0.53, not shown on the figure) and a very strong negative variability bias (mean α close to 0.42, not shown on the figure). At different time-steps and for different criteria, ANA also exhibits a rather good spatial robustness of performances (spread of the distribution: distance between quantile 0.1 and 0.9), compared to LM and ANATEM models.

- at the different time-steps, LM model presents very satisfactory results. It has no mean bias (by construction) and a moderate to limited variability bias (mean α

HESSD
12, 311–361, 2015

between 0.91 and 0.99). The high to low frequency variability is very well captured (mean r between 0.92 and 0.99). LM has moderate spatial robustness for correlation and variability bias, for daily and annual time-steps.

– at the different time-steps, ANATEM model has very satisfactory results. It has a moderate mean negative bias (mean β close to 0.97) and a limited to moderate variability bias (mean α between 0.95 and 0.98). The short-term to long-term variability is very well captured (mean β between 0.95 and 0.99). ANATEM has moderate robustness concerning mean bias, but also correlation and variability bias, for daily and annual time-steps.

LM and ANATEM models thus clearly outperforms ANA model. LM is characterised by a very good correlation and no mean bias, but a moderate variability bias. ANATEM is characterised by a very good correlation, limited mean and variability bias. Model performances are better and more robust at a monthly time-step, compared to daily and annual time-step. ANATEM has a slightly better spatial robustness of performances than LM. This is also expressed by mean KGE values, ranging from 0.52 to 0.97 for ANA, from 0.92 to 0.99 for LM and from 0.95 to 0.99 for ANATEM respectively.

4.3.2 Precipitation reconstruction

The three model presents slightly different results for precipitation reconstruction (Fig. 11):

– at a daily time-step ANA suffers from a very moderate mean negative mean bias (mean β close to 0.95) and a strong variability bias (mean α around 0.55). It also has a limited correlation (mean r close to 0.65). At monthly and annual time-step, ANA has a moderate to limited mean bias (mean β close to 0.95), a significant variability bias (mean α around 0.8) and an acceptable level of correlation (mean r between between 0.7 to 0.8).

HESSD
12, 311–361, 2015

– at the different time-steps, LM model has very satisfactory results. It has a no mean bias (by construction), a limited variability bias (mean α from 0.97 to 1.05). The short-term to long-term variability is well captured (mean r between 0.77 and 0.84).

– at the different time-steps, ANATEM model has very satisfactory results. It has a limited negative mean bias (mean β around 0.96), a limited variability bias (mean α from 0.94 to 1.02). The short-term to long-term variability is well captured (mean r between 0.75 and 0.87).

LM and ANATEM models perform thus better than ANA model, particularly concerning the correlation. LM is characterised by a good correlation, no mean bias and a limited variability bias. ANATEM is also characterised by a good correlation, limited mean and variability bias. Model performances are better and more robust at a monthly time-step, compared to daily and annual time-step. The spatial robustness of performances is slightly lower for the variability criteria than for the others criterion. LM has the lowest robustness, then ANATEM and finally ANA. This is again illustrated by the mean KGE values ranging from 0.43 to 0.71 for ANA, from 0.75 to 0.83 for LM and from 0.76 to 0.84 for ANATEM.

4.4 Spatial patterns of models performance

In the present section, we discuss the spatial pattern of performances (in terms of correlation, at a daily time-step) of the three models. We also present the spatial pattern of the gain in performance obtained with ANATEM reconstructions when either compared to ANA or LM reconstructions.

4.4.1 Air temperature reconstruction

For temperature reconstructions, the spatial patterns of model performance are presented in Fig. 12. For ANA, the performance of the reconstruction is fairly not influ-

HESSD
12, 311–361, 2015

enced by the location of the watershed, with a mean correlation ranging from to 0.92 to 0.94 (Fig. 12a). For LM (Fig. 12b), the location of the watershed has a slightly higher influence on the performance, with a mean correlation ranging from 0.95 to 0.98. This spatial pattern has a clear south-west to north-east structure, with a decrease of model performances driven by the distance to the local reference time series (located in Marseille, south-west of the watersheds). Finally, for ANATEM model (Fig. 12c), the location of the watershed (i.e., distance to Marseille) also influences the performance of the reconstruction, with a mean correlation ranging from 0.97 to 0.99. However, ANATEM has slightly better performances than LM and then ANA and the range of correlation values is slightly thinner than the range observed for LM.

The contribution of LM model (Fig. 12d) to the performance of ANATEM, decreases from south-west to north-east, ranging from 0.06 to 0.04. Conversely, the contribution of ANA model (Fig. 12e) to the performance of ANATEM, slightly increases from south-west to north-east, ranging from 0.0 to 0.02. The contribution of large scale information (through ANA model) is stronger when LM model (local information) is less efficient, that is, when the location at reconstruction is far from the reference temperature station.

4.4.2 Precipitation reconstruction

The spatial patterns of model performance obtained for precipitation are slightly different than those obtained for temperature (Fig. 13). For ANA, the location of the watershed does not really influences the performance, with a mean correlation ranging from to 0.62 to 0.68 (Fig. 13a). For LM (Fig. 13b) and ANATEM (Fig. 13c), watersheds close to the local reference station highlight conversely better performances than the others (the correlation ranges from 0.62 to 0.88 for LM and from 69 to 0.89 for ANATEM). However, the similarity in terms of large scale forcing influences probably influences the performance. Hence, two watersheds at the same distance to the Gap station have rather different performances (i.e. Buech watershed have a very good correlation of 0.88 and the Durance at Briançon a moderate correlation of 0.77).

HESSD
12, 311–361, 2015

ANATEM increases the global reconstruction performance but it also notably smooths local contrasts. The contribution of LM to the performance of ANATEM decreases as the distance to Gap increases, ranging from 0.22 to 0.02 (Fig. 13d). Conversely, the contribution of ANA to the performance of ANATEM slightly increases from south-west to north-east, ranging from 0.0 to 0.07 (Fig. 13e). As observed for air temperature reconstruction and here in a more pronounced way, the contribution of large scale information (through ANA model) is stronger when LM model (local information) is less efficient, as a result of an increasing distance to the reference station.

5 Climatic variability assessment

5.1 1883–2010 reconstructions of air temperature

Figure 14 presents the 1883–2010 annual anomaly time-series of air temperature reconstructed by the ANATEM method (mean model) for the 22 watersheds of the Durance river. Anomalies have been computed as the differences to the 1883–2010 mean. This figure exhibits a pseudo-stationary period from 1880 to 1940, then a slight temperature increase between 1940 and 1980 and a stronger increase from 1980 until nowadays. In order to better characterize low-frequency variability, a smoothed mean of the 22 series reconstructed for the 22 watersheds respectively has been computed by LOESS (Cleveland, 1979). This series (red curve in Fig. 14) highlights a relatively strong variability: mean air temperature can vary of nearly 1 °C in less than 10 year (e.g.: 1890–1900, 1940–1950).

The ANATEM reconstructions have been qualitatively compared to five series of air temperature anomalies obtained from homogeneised series of the HISTALP project (Genova University, Milano-Brera, Montpellier, Nice airport and Nîmes airport).

The ANATEM model reproduces fairly well the annual and low-frequency variability of air temperature anomalies from HISTALP stations (the mean correlation of ANATEM and HISTALP annual series is close to 0.8). However, the warming trend

HESSD
12, 311–361, 2015

in the HISTALP series is larger than in the ANATEM reconstructions, HISTALP temperatures being significantly lower than ANATEM temperatures before 1900 and significantly higher after 1980. ANATEM reconstructions and HISTALP time-series are obviously sensitive to the reference time-series (i.e. Marseille for ANATEM) and the homogenisation process applied to the observations (for both Marseille and HISTALP stations). Further research is required to explore the sensitivity of the ANATEM reconstructions to these key features (partly tested in Kuentz, 2013).

5.2 1883–2010 reconstructions of precipitation

Figure 15 presents the 1883–2010 precipitation reconstructions obtained with ANATEM for the 22 watershed along with five precipitation HISTALP series (Aix-en-Provence, Nice (Cap-Ferrat), Orange, Saint-Paul-les-Durance and Toulon). For both the reconstructions and the HISTALP series, the mean smoothed series is also given.

ANATEM series present a very homogeneous temporal behaviour when compared to the high dispersion observed between the five HISTALP series. This may be partly explained by the fact that ANATEM series are reconstructed for all watersheds based on a same reference series (Gap). The main reason is however probably the high spatial variability of precipitation and the fact that HISTALP series cover a much wider spatial domain than ANATEM series. The low dispersion between the reconstructed series is otherwise coherent with the limited dispersion obtained between time-series observed for the same 22 watersheds on the observation period (not shown here).

Besides, the smoothed time series from ANATEM reconstructions is highly correlated to the smoothed time series from HISTALP data. ANATEM reconstruction is therefore able to reproduce the low frequency variability of precipitation resulting from climate variability. Some differences can be observed for example between 1920 and 1930 or between 1970 and 1980. They may be due again to the large spatial variability of precipitation which would also translate to different precipitation indexes, as long as they are estimated from different stations. As noticed for air temperature reconstructions, these differences could also be due to the reference series used in ANATEM and to

HESSD
12, 311–361, 2015

the homogenisation process for HISTALP series. Additional work would be worth to explore the importance of these issues.

6 Conclusions

Reconstructing local scale meteorological variables over long periods is challenging and necessary for better understanding the low frequency variability of regional climate and climate driven variables. Three models are compared in the present work, using different kind of data for the reconstruction: the regression based Local Model (LM) uses local observations of the variable from neighbouring stations as predictor, the Analog Model (ANA), a so-called downscaling model, uses large scale information of atmospheric circulation, the ANATEM model uses a mix of both local and large scale atmospheric information combining therefore both the Local Model and the Analog Model.

The three models have been developed and applied for the reconstruction of mean air temperature and precipitation time-series of a sample of 22 watersheds situated in the Durance watershed, south-east of France. This sample of watersheds represents a wide range of climatic conditions, from highly mountainous to Mediterranean. The local observation used for the reconstruction are respectively Marseille air temperature, Gap precipitation historical time-series and geopotential fields from the 20CR reanalysis.

The multicriteria and multiscale performance assessment highlights that the best reconstructions are obtained when local information is used. ANA is clearly less efficient than the two others methods, particularly concerning long-term (annual) air temperature variability or short-term (daily) precipitation variability. The regression based model and the ANATEM model have conversely very satisfactory results for all criteria. ANATEM has a slight advantage and the spatial patterns of the reconstruction skills show that it takes benefit from the qualities of both underlying models. Hence, the ANATEM method allow to reconstruct very satisfactory air temperature and precipita-

HESSD
12, 311–361, 2015

tion reanalyses at a high temporal resolution (daily) and different spatial scales (from 4 to 3500 km^2), while improving the spatial robustness of performances.

Time series of air temperatures reconstructed for the 1883–2010 period exhibit the well-known warming experienced since the middle of last century, with a higher rate since the 1980's. Reconstructed precipitation time-series highlight the large inter-annual variability of annual precipitation for the Durance region. Long-term climatic reanalyses exhibits some particular periods with rather strong rainfall anomalies, such as wet periods at the beginning of 1910's and mid 1930's (known for floods events), or rather dry periods such as 1940's and 1970's (known for droughts events).

This work has a number of interesting perspectives and raises a variety of challenging questions.

The potential for improving the method is not negligible. The ANA method we used here has been firstly developed for precipitation forecasts (e.g. Obled et al., 2002). The poor reconstruction skill obtained for temperature was therefore not a surprise and other large scale predictors could potentially allow for a better reconstruction of air temperature variations. This also applies for the precipitation reconstruction. The predictors used here do only inform on the atmosphere dynamics. The inclusion of thermodynamic predictors and of humidity predictors for the identification of analog days has been proved to improve for the studied region the performance of the method (Marty et al., 2011; Chardon et al., 2014).

Another possibility of progress concerns the combined formulation used for the ANATEM method. The formulation presented in this paper has been applied straightforwardly and has not been modified on the basis of these results. However, we are convinced that fancy statistical developments on the way both models are combined (e.g. forecast combination methods as in Winkler and Makridakis, 1983 or Hoeting et al., 1999) would allow to improve temporal correlation, or reduce mean and variability bias and consider probabilistic calibration (not shown in this paper).

The issue of the reference series used for the Local Model is also challenging. We have shown that the good performance of the methods is achieved thanks to this local

HESSD
12, 311–361, 2015

information. A thorough sensitivity analysis to the selection of the reference time-series should be carried out. Considering the importance of local information, an extension of the method should also consider the possibility to make use of all historical stations available in the close or farer neighbourhood of the region under construction. Of course, historical local scale data covering long historical periods are very scarce and sparse. Our results also highlight that the reconstruction skill decreases when the distance to the reference station increases. The region we have considered covers a rather narrow domain. However, we can expect that the interest of the reference station is much lower if we would do reconstructions for much more distant locations. We can expect conversely that the relative interest of the large scale information would be much larger for distant sites. Additional work is definitively required to assess the relative interest of both components of the ANATEM model in this context.

Because of the numerous scientific and operational stakes associated with the characterization of long term variability, we are confident that all of these questions will be tackled by the scientific community in the next years. A major application of such reconstructions will be obviously the possibility to reconstruct long term variations for a number of climate driven variables. As an illustration, we used the long-term climatic time-series produced in the present work for reconstructing long-term hydrological time-series at multiple hydrometric stations of the Durance basin (Kuentz, 2013; Mathevet et al., 2013). An outstanding result of this reconstruction is that the time series obtained for the whole 20th century present a very high correlation level with historical discharges time series obtained from rescued hydrometric archives for the catchment. In our case, the availability of historical streamflow time series allowed us to demonstrate the overall quality of the meteorological reconstruction. This independent hydrological validation is not expected to be feasible everywhere but it gives high confidence in this hydrometeorological reconstruction approach. Even when such an independent validation cannot be carried out, the reconstructed time series definitively produce a high-value information for researchers or water resources managers. Fur-

HESSD
12, 311–361, 2015

ther works for other hydroclimatic contexts are therefore also worth to better identify the potential of the method and the possibility for improving it.

Acknowledgements. The authors are grateful to Meteo-France and the HISTALP project (Historical Instrumental Climatological Surface Time Series Of The Greater Alpine Region) for providing long-term climatic data, as well as to the NOAA which provided the Twentieth Century Reanalysis Project dataset. The authors also wish to thank Vazken Andréassian, Anne-Catherine Favre and Cristian Perret for their support and their helpful comments during different phases of this work.

References

Auer, I., Böhm, R., Jurkovic, A., Lipa, W., Orlik, A., Potzmann, R., Schöner, W., Ungersböck, M., Matulla, C., Briffa, K., Jones, P., Efthymiadis, D., Brunetti, M., Nanni, T., Maugeri, M., Mercalli, L., Mestre, O., Moisselin, J.-M., Begert, M., Müller-Westermeier, G., Kveton, V., Bochnicek, O., Stastny, P., Lapin, M., Szalai, S., Szentimrey, T., Cegnar, T., Dolinar, M., Gajic-Capka, M., Zaninovic, K., Majstorovic, Z., and Nieplova, E.: HISTALP historical instrumental climatological surface time series of the Greater Alpine Region, Int. J. Climatol., 27, 17–46, doi:10.1002/joc.1377, 2007. 317

Ben Daoud, A., Sauquet, E., Lang, M., Obled, C., and Bontron, G.: La prévision des précipitations par recherche d'analogues: état de l'art et perspectives, Houille Blanche, 2009-6, 60–65, doi:10.1051/lhb/2009079, 2010. 322

Boé, J. and Habets, F.: Multi-decadal river flow variations in France, Hydrol. Earth Syst. Sci., 18, 691–708, doi:10.5194/hess-18-691-2014, 2014. 315

Bontron, G. and Obled, C.: A probabilistic adaptation of meteorological model outputs to hydrological forecasting, Houille Blanche, 2005-1, 23–28, 2005. 322

Bourqui, M., Mathevet, T., Gailhard, J., and Hendrickx, F.: Hydrological validation of statistical downscalling methods applied to model projections, IAHS-AISH Publ., 344, 32–38, 2011. 321

Chardon, J., Hingray, B., Favre, A.-C., Autin, P., Gailhard, J., Zin, I., and Obled, C.: Spatial similarity and transferability of analog dates for precipitation downscaling over France, J. Climate, 27, 5056–5074, doi:10.1175/JCLI-D-13-00464.1, 2014. 338

HESSD
12, 311–361, 2015

Cleveland, W. S.: Robust locally weighted regression and smoothing scatterplots, J. Am. Stat. Assoc., 74, 829–836, doi:10.1080/01621459.1979.10481038, 1979. 335

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, Q. J. Roy. Meteorol. Soc., 137, 1–28, doi:10.1002/qj.776, 2011. 316, 318

CSIRO: Climate variability and change in south-eastern Australia: a synthesis of findings from Phase 1 of the South Eastern Australian Climate Initiative (SEACI), Tech. rep., South Eastern Australian Climate initiative, Australia, 2010. 314, 315

Deser, C., Phillips, A., Bourdette, V., and Teng, H.: Uncertainty in climate change projections: the role of internal variability, Clim. Dynam., 38, 527–546, doi:10.1007/s00382-010-0977-x, 2012. 314

Dufour, C., and Garçon, R.: Méthode statistique de recalage du modèle de prévision au pas journalier MORDOR dans le cadre du projet Vienne, Rapport technique, Electricité de France, Grenoble, France, 1997. 326

Frank, D. and Esper, J.: Temperature reconstructions and comparisons with instrumental data from a tree-ring network for the European Alps, Int. J. Climatol., 25, 1437–1454, doi:10.1002/joc.1210, 2005. 315

Gottardi, F., Obled, C., Gailhard, J., and Paquet, E.: Statistical reanalysis of precipitation fields based on ground network data and weather patterns: Application over French mountains, J. Hydrol., 432–433, 154–167, doi:10.1016/j.jhydrol.2012.02.014, 2012. 317

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: implications for improving hydrological modelling, J. Hydrol., 377, 80–91, doi:10.1016/j.jhydrol.2009.08.003, 2009. 328

Hänggi, P., and Weingartner, R.: Inter-annual variability of runoff and climate within the Upper Rhine River basin, 1808–2007, Hydrolog. Sci. J., 56, 34–50, doi:10.1080/02626667.2010.536549, 2011. 314

Hannaford, J., Buys, G., Stahl, K., and Tallaksen, L. M.: The influence of decadal-scale variability on trends in long European streamflow records, Hydrol. Earth Syst. Sci., 17, 2717–2733, doi:10.5194/hess-17-2717-2013, 2013. 314

HESSD
12, 311–361, 2015

Hannah, D. M., Demuth, S., van Lanen, H. A. J., Looser, U., Prudhomme, C., Rees, G., Stahl, K., and Tallaksen, L. M.: Large-scale river flow archives: importance, current status and future needs, Hydrol. Process., 25, 1191–1200, doi:10.1002/hyp.7794, 2011. 314

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, B. Am. Meteorol. Soc., 90, 1095–1107, doi:10.1175/2009BAMS2607.1, 2009. 314

Hingray, B., Hendrickx, F., Bourqui, M., Creutin, J.-D., François, B., Gailhard, J., Lafaysse, M., Lemoine, N., Mathevet, T., Mezghani, A., and Monteil, C.: RIWER 2030 – climats régionaux et incertitudes, ressource en eau et gestion associée de 1960 à 2030, Rapport final projet ANR VMCS 2009-2012, ANR, CNRS/LTHE, EDF/LNHE, EDF/DTG, Grenoble, France, 2013. 321

Hoeting, J. A., Madigan, D., Raftery, A. E., and Volinsky, C. T.: Bayesian model averaging: a tutorial, Stat. Sci., 14, 382–401, 1999. 338

Horton, P., Jaboyedoff, M., Metzger, R., Obled, C., and Marty, R.: Spatial relationship between the atmospheric circulation and the precipitation measured in the western Swiss Alps by means of the analogue method, Nat. Hazards Earth Syst. Sci., 12, 777–784, doi:10.5194/nhess-12-777-2012, 2012. 321, 322

Imbeaux, E.: La Durance: Régime, Crues et Inondations, Vve Ch. Dunod, Paris, France, 1892. 317

Jouzel, J., Masson-Delmotte, V., Cattani, O., Dreyfus, G., Falourd, S., Hoffmann, G., Minster, B., Nouet, J., Barnola, J. M., Chappellaz, J., Fischer, H., Gallet, J. C., Johnsen, S., Leuenberger, M., Loulergue, L., Luethi, D., Oerter, H., Parrenin, F., Raisbeck, G., Raynaud, D., Schilt, A., Schwander, J., Selmo, E., Souchez, R., Spahni, R., Stauffer, B., Steffensen, J. P., Stenni, B., Stocker, T. F., Tison, J. L., Werner, M., and Wolff, E. W.: Orbital and millennial antarctic climate variability over the past 800,000 years, Science, 317, 793–796, doi:10.1126/science.1141038, 2007. 315

Kuentz, A.: Un siècle de variabilité hydro-climatique sur le bassin de la Durance: recherches historiques et reconstitutions, Thèse de doctorat, AgroParisTech, Paris, France, 2013. 316, 320, 336, 339

Kuentz, A., Mathevet, T., Perret, C., Gailhard, J., and Andréassian, V.: Uncertainty estimation and reconstruction of historical streamflow records, Geophys. Res. Abstr., 14, EGU2012-5632-1, 2012. 315

HESSD
12, 311–361, 2015

Kuentz, A., Mathevet, T., Coeur, D., Perret, C., Gailhard, J., Guérin, L., Gash, Y., and Andréassian, V.: Hydrométrie et hydrologie historiques du bassin de la Durance, Houille Blanche, 2014-4, 57–63, doi:10.1051/lhb/2014039, 2014. 315

Lafaysse, M., Hingray, B., Mezghani, A., Gailhard, J., and Terray, L.: Internal variability and model uncertainty components in future hydrometeorological projections: The Alpine Durance basin, Water Resour. Res., 50, 3317–3341, doi:10.1002/2013WR014897, 2014. 314

Lorenz, E. N.: Atmospheric predictability as revealed by naturally occurring analogues, J. Atmos. Sci., 26, 636–646, doi:10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2, 1969. 321

Lorenzo-Lacruz, J., Vicente-Serrano, S., López-Moreno, J., Morán-Tejeda, E., and Zabalza, J.: Recent trends in Iberian streamflows (1945–2005), J. Hydrol., 414–415, 463–475, doi:10.1016/j.jhydrol.2011.11.023, 2012. 314

Madden, R. A.: Estimates of the natural variability of time-averaged sea-level pressure, Mon. Weather Rev., 104, 942–952, doi:10.1175/1520-0493(1976)104<0942:EOTNVO>2.0.CO;2, 1976. 314

Marty, R., Zin, I., Obled, C., Bontron, G., and Djerboua, A.: Toward real-time daily PQPF by an analog sorting approach: application to flash-flood catchments, J. Appl. Meteorol. Clim., 51, 505–520, doi:10.1175/JAMC-D-11-011.1, 2011. 338

Mathevet, T., and Garçon, R.: Tall tales from the hydrological crypt: are models monsters?, Hydrolog. Sci. J., 55, 857–871, doi:10.1080/02626667.2010.503934, 2010. 314

Mathevet, T., Kuentz, A., Gailhard, J., and Andreassian, V.: Characterizing a century of climate and hydrological variability of a mediterranean and mountainous watersheds: the Durance River Case-Study, in: AGU Fall Meeting Abstracts, #H31B-1156, San Francisco, California, USA, 2013. 339

Obled, C., Bontron, G., and Garçon, R.: Quantitative precipitation forecasts: a statistical adaptation of model outputs through an analogues sorting approach, Atmos. Res., 63, 303–324, doi:10.1016/S0169-8095(02)00038-8, 2002. 322, 338

Pfister, C., Weingartner, R., and Luterbacher, J.: Hydrological winter droughts over the last 450 years in the Upper Rhine basin: a methodological approach, Hydrolog. Sci. J., 51, 966–985, doi:10.1623/hysj.51.5.966, 2006. 315

Poli, P., Hersbach, H., Tan, D., Dee, D., Thépaut, J. N., Simmons, A., Peubey, C., Laloyaux, P., Komori, T., and Berrisford, P.: The data assimilation system and initial performance evaluation

HESSD
12, 311–361, 2015

of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C), ERA Report Series 14, ECMWF, UK, 2013. 316
Renard, B.: Détection et prise en compte d'éventuels impacts du changement climatique sur les extrêmes hydrologiques en France, Thèse de doctorat, Institut National Polytechnique de Grenoble, Grenoble, 2006. 314, 315
Renner, M. and Bernhofer, C.: Long term variability of the annual hydrological regime and sensitivity to temperature phase shifts in Saxony/Germany, Hydrol. Earth Syst. Sci., 15, 1819–1833, doi:10.5194/hess-15-1819-2011, 2011. 314
Rood, S. B., Samuelson, G., Weber, J., and Wywrot, K.: Twentieth-century decline in streamflows from the hydrographic apex of North America, J. Hydrol., 306, 215–233, doi:10.1016/j.jhydrol.2004.09.010, 2005. 314, 315
Stahl, K., Hisdal, H., Hannaford, J., Tallaksen, L. M., van Lanen, H. A. J., Sauquet, E., Demuth, S., Fendekova, M., and Jódar, J.: Streamflow trends in Europe: evidence from a dataset of near-natural catchments, Hydrol. Earth Syst. Sci., 14, 2367–2382, doi:10.5194/hess-14-2367-2010, 2010. 314
Teng, J., Chiew, F., Timbal, B., Wang, Y., Vaze, J., and Wang, B.: Assessment of an analogue downscaling method for modelling climate change impacts on runoff, J. Hydrol., 472–473, 111–125, doi:10.1016/j.jhydrol.2012.09.024, 2012. 321
Terray, L., and Boé, J.: Quantifying 21st-century France climate change and related uncertainties, CR Geosci., 345, 136–149, doi:10.1016/j.crte.2013.02.003, 2013. 314
Teweles, S. and Wobus, H.: Verification of prognostic charts, B. Am. Meteorol. Soc., 35, 455–63, 1954. 322
Timbal, B., Arblaster, J. M., and Power, S.: Attribution of the late-twentieth-century rainfall decline in Southwest Australia, J. Climate, 19, 2046–2062, doi:10.1175/JCLI3817.1, 2006. 321
Van Den Dool, H. M.: A new look at weather forecasting through analogues, Mon. Weather Rev., 117, 2230–2247, doi:10.1175/1520-0493(1989)117<2230:ANLAWF>2.0.CO;2, 1989. 321
Wetterhall, F., Halldin, S., and Xu, C.-Y.: Statistical precipitation downscaling in central Sweden with the analogue method, J. Hydrol., 306, 174–190, doi:10.1016/j.jhydrol.2004.09.008, 2005. 322
Wilhelm, B., Arnaud, F., Sabatier, P., Crouzet, C., Brisset, E., Chaumillon, E., Disnar, J.-R., Guiter, F., Malet, E., Reyss, J.-L., Tachikawa, K., Bard, E., and Delannoy, J.-J.: 1400 years

HESSD
12, 311–361, 2015

of extreme precipitation patterns over the Mediterranean French Alps and possible forcing mechanisms, Quaternary Res., 78, 1–12, doi:10.1016/j.yqres.2012.03.003, 2012. 315
Wilhelm, B., Arnaud, F., Sabatier, P., Magand, O., Chapron, E., Courp, T., Tachikawa, K., Fanget, B., Malet, E., Pignol, C., Bard, E., and Delannoy, J. J.: Palaeoflood activity and climate change over the last 1400 years recorded by lake sediments in the north-west European Alps, J. Quaternary Sci., 28, 189–199, doi:10.1002/jqs.2609, 2013. 315
Winkler, R. L. and Makridakis, S.: The combination of forecasts, J. Roy. Stat. Soc. A-G., 146, 150–157, doi:10.2307/2982011, 1983. 338
Zhang, X., Harvey, K. D., Hogg, W. D., and Yuzyk, T. R.: Trends in Canadian streamflow, Water Resour. Res., 37, 987–998, doi:10.1029/2000WR900357, 2001. 314

HESSD
12, 311–361, 2015

Table 1. Main characteristics of the 22 selected watersheds. # correspondance with Fig. 1, P_A annual mean precipitations (on period 1948–2010), T mean air temperature (on period 1948–2010)

#	Name	Altitude m	Area km^2	$\bar{P}_A$ $mm\,yr^{-1}$	$\bar{T}$ °C
1	The Durance river at Val-des-Près	1360	203	1322	2.5
2	The Guisane river at Monêtier-les-Bains	1510	78	1627	2.3
3	The Durance river at Briançon	1187	548	1381	2.9
4	The Guil river at Montdauphin	895	725	1087	3.3
5	The Durance river at La Clapière	787	2170	1352	3.5
6	The Riou de Crachet river at Saint-Paul	2020	4	1532	1.6
7	The Ubaye river at Roche-Rousse	790	946	1235	4.1
8	The Ubaye river at Barcelonnette	1132	549	1201	3.6
9	The Durance river at Serre-Ponçon	652	3582	1301	4.0
10	The Buech river at Les Chambons	662	723	1259	7.4
11	The Méouge river at Méouge	545	221	1094	8.9
12	The Jabron river at Piedguichard	593	89	1206	9.1
13	The Bes river at La Javie	805	165	1085	6.6
14	The Lauzon river at Villeneuve	341	124	1097	10.4
15	The Asse river at La Clue de Chabrières	605	375	1077	8.6
16	The Verdon river at Allos	1780	10	1592	2.7
17	The Verdon river at Colmars	1230	158	1453	4.3
18	The Issole river at Saint-André-les-Alpes	931	137	1229	6.8
19	The Verdon river at Castillon	790	657	1319	6.2
20	The Artuby river at La Bastide	1008	91	1272	8.4
21	The Jabron river at Comps-sur-Artuby	782	66	1116	9.0
22	The Verdon river at Sainte-Croix	400	1625	1176	8.2

HESSD
12, 311–361, 2015

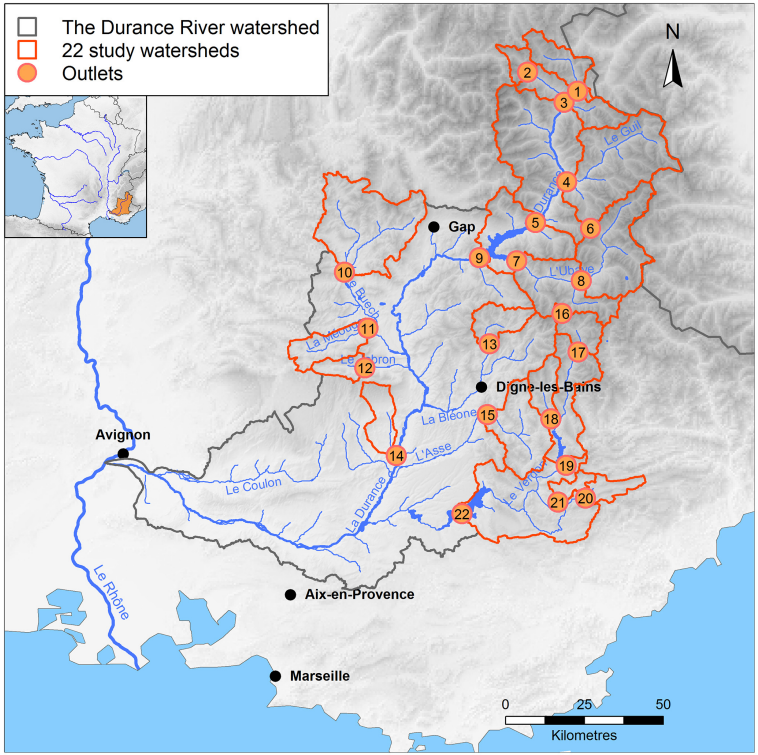


Figure 1. Map of the study area with the 22 watersheds selected.

HESSD
12, 311–361, 2015

Building long term precipitation and air temperature reanalyses: the ANATEM method

A. Kuentz et al.

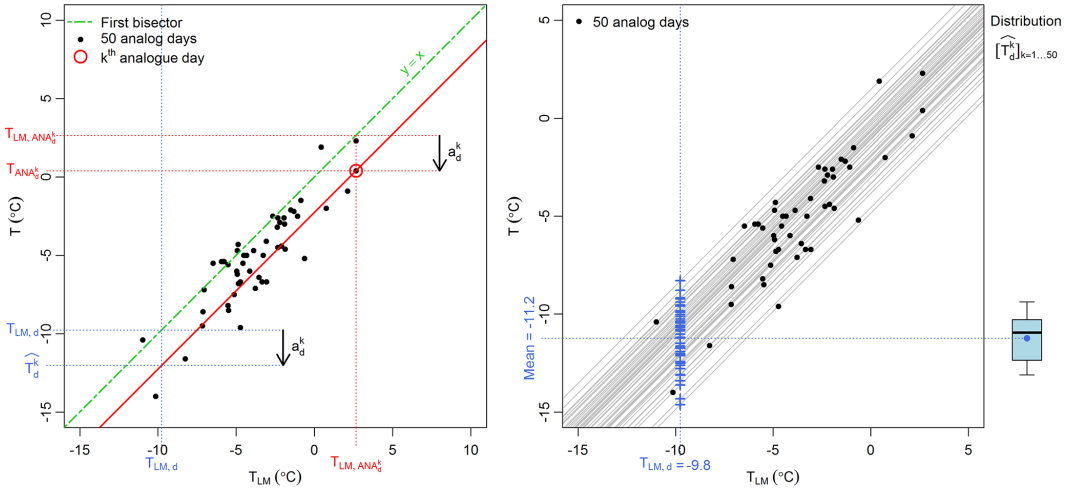
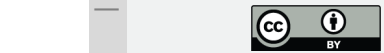


Figure 2. Representation of the ANATEM formulation for air temperature reconstruction of a given day d. Left panel: observed temperature for the target as a function of the temperature estimate from LM. For a given analog k, the corrected estimate for day d is $\widehat{T}_d^k$ which is obtained as $\widehat{T}_{\mathrm{LM},d} + a_d^k$ where $a_d^k = T_{\mathrm{ANA}_d^k} - T_{\mathrm{LM,ANA}_d^k}$. Right panel: probabilistic prediction obtained for d from the 50 analogs. The corresponding boxplot (10, 25, 50, 75 and 90 % quantiles) is given on the right of the figure.

HESSD
12, 311–361, 2015



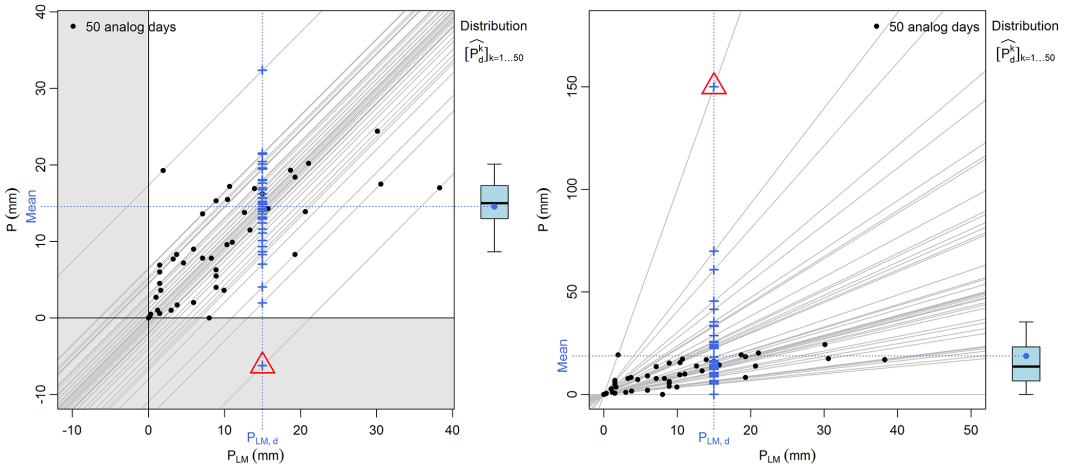


Figure 3. Representation of additive and multiplicative formulation for precipitation reconstruction from a local model and 50 analog days of a given day d. Left panel: case of an additive formulation for the correction. Right panel: case of a multiplicative correction. The triangles highlight aberrant or potentially aberrant corrected predictions.

HESSD
12, 311–361, 2015

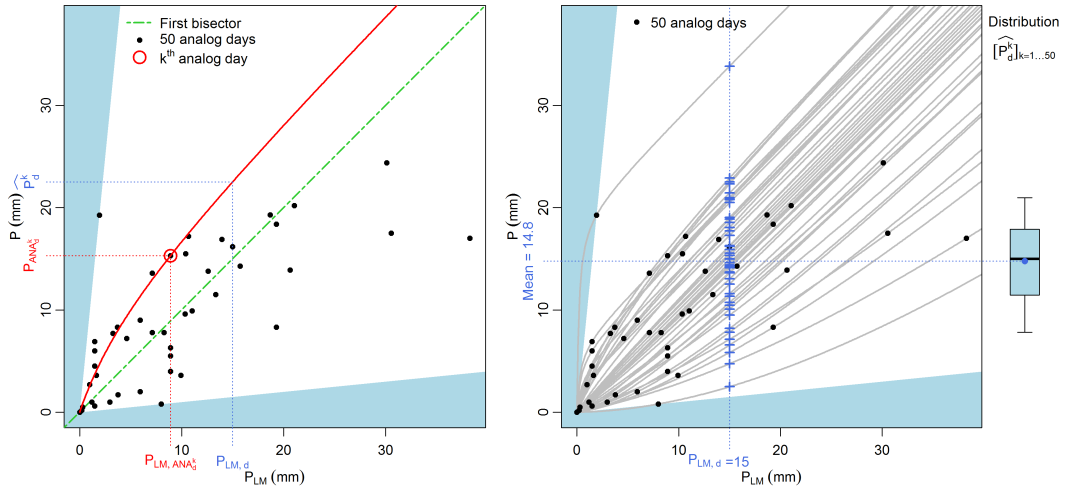


Figure 4. Representation of the ANATEM formulation for precipitations reconstruction of a given day d. Left panel: observed precipitation for the target as a function of the precipitation estimate from LM. For a given analog k, the corrected estimate for day d can be read as the ordinate of the point of abscissa $P_{LM,d}$ on the red curve which equation is Eq. (13). Right panel: probabilistic prediction obtained for d from the 50 analogs. The corresponding boxplot (10, 25, 50, 75 and 90 % quantiles) is given on the right of the figure.

HESSD
12, 311–361, 2015

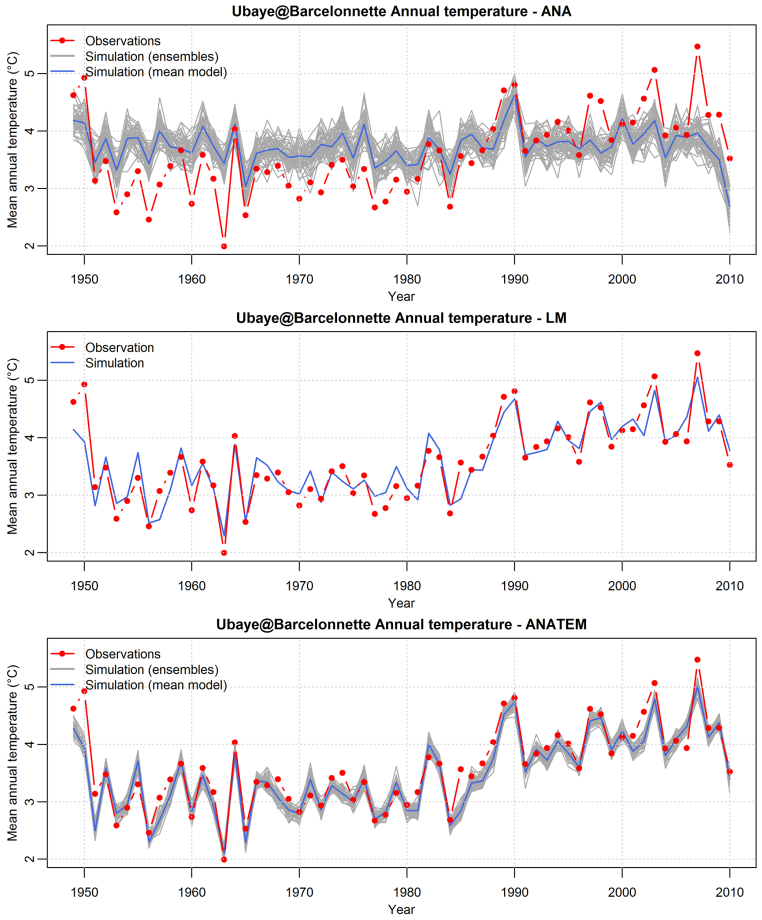


Figure 5. Annual time-series of air temperature reconstructions for the Ubaye River at Barcelonnette watershed by analog method (ANA), local model (LM) and ANATEM method.

HESSD
12, 311–361, 2015

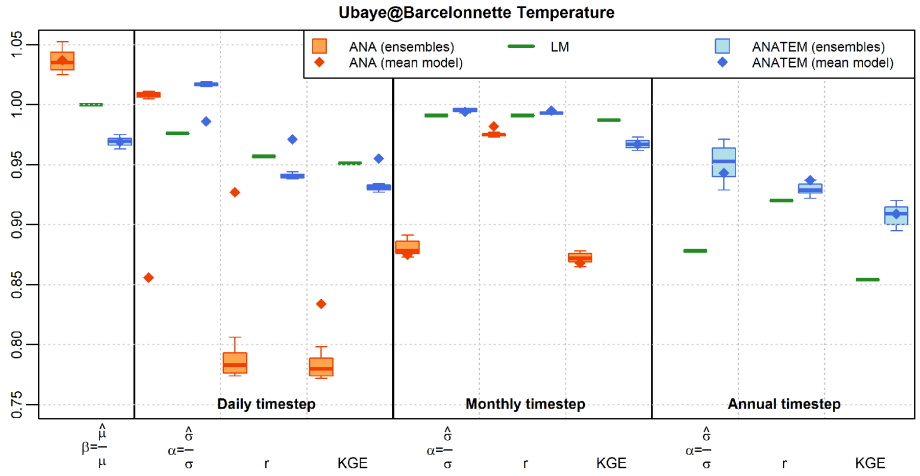


Figure 6. Daily, monthly and annual performance criteria of air temperature reconstructions for the Ubaye River at Barcelonnette watershed by analog method (ANA), local model (LM) and ANATEM method.

HESSD
12, 311–361, 2015

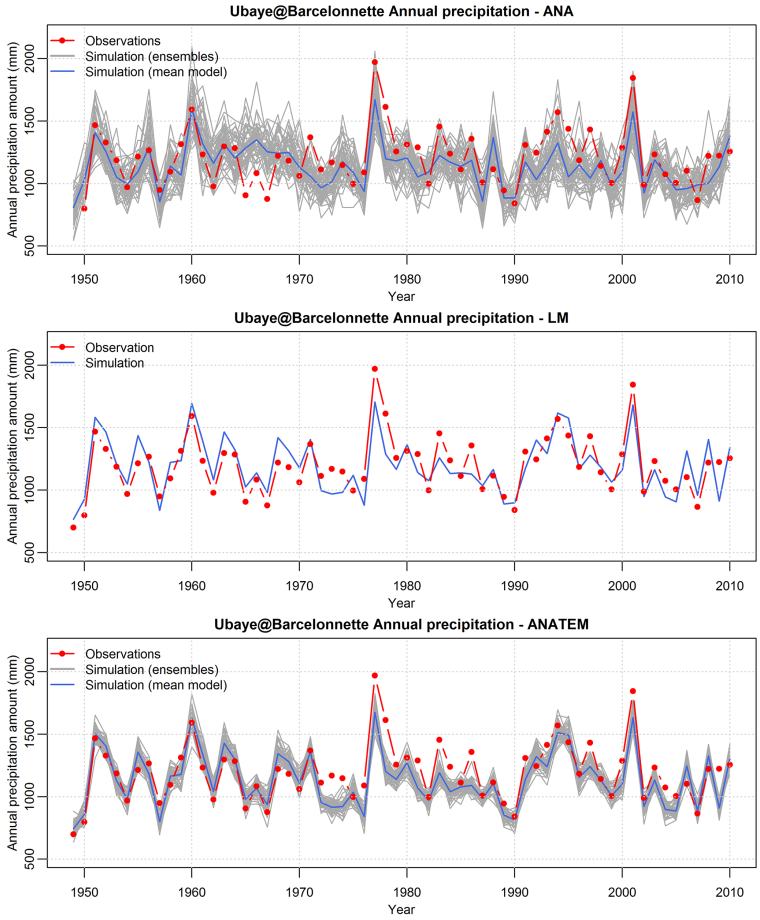


Figure 7. Annual time-series of precipitation reconstructions for the Ubaye River at Barcelonnette watershed by analog method (ANA), local model (LM) and ANATEM method.

HESSD
12, 311–361, 2015

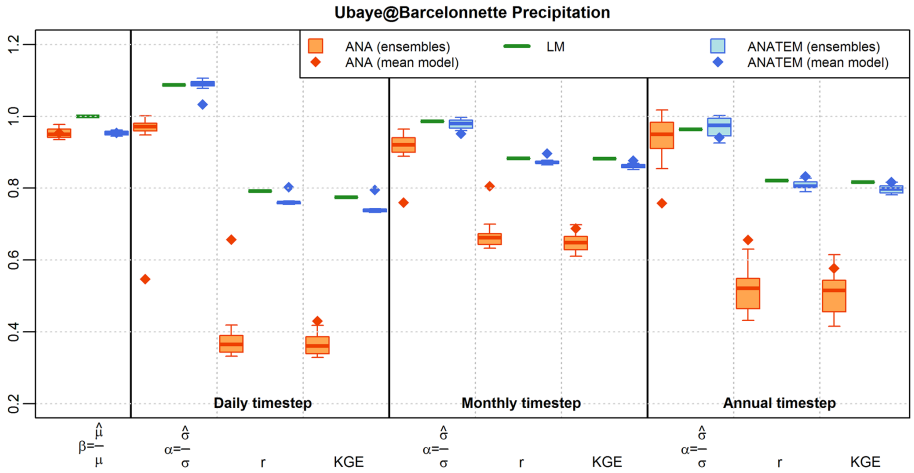


Figure 8. Daily, monthly and annual performance criteria of precipitation reconstructions for the Ubaye River at Barcelonnette watershed by analog method (ANA), local model (LM) and ANATEM method.

HESSD
12, 311–361, 2015

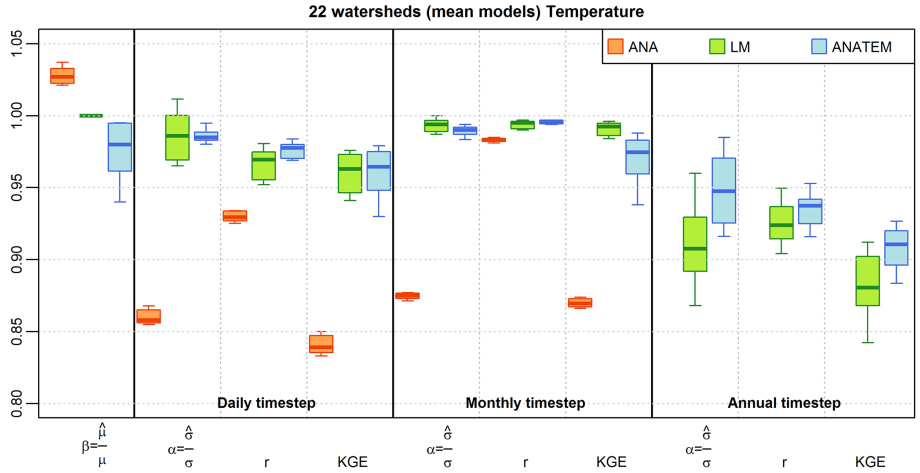


Figure 9. Daily, monthly and annual performance criteria of air temperature mean reconstructions for 22 watersheds by analog method (ANA), local model (LM) and ANATEM method.

HESSD
12, 311–361, 2015

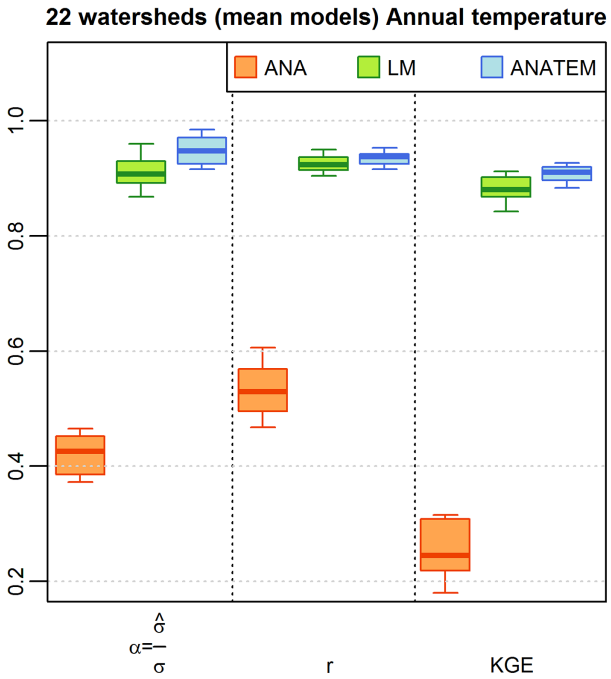


Figure 10. Annual performance criteria of air temperature mean reconstructions for 22 watersheds by analog method (ANA), local model (LM) and ANATEM method (larger scale).

HESSD
12, 311–361, 2015

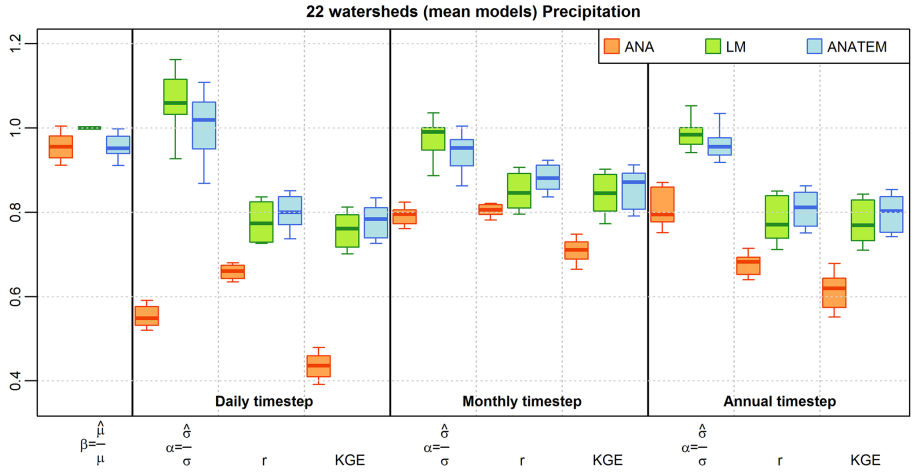


Figure 11. Daily, monthly and annual performance criteria of precipitation mean reconstructions for 22 watersheds by analog method (ANA), local model (LM) and ANATEM method.

HESSD
12, 311–361, 2015

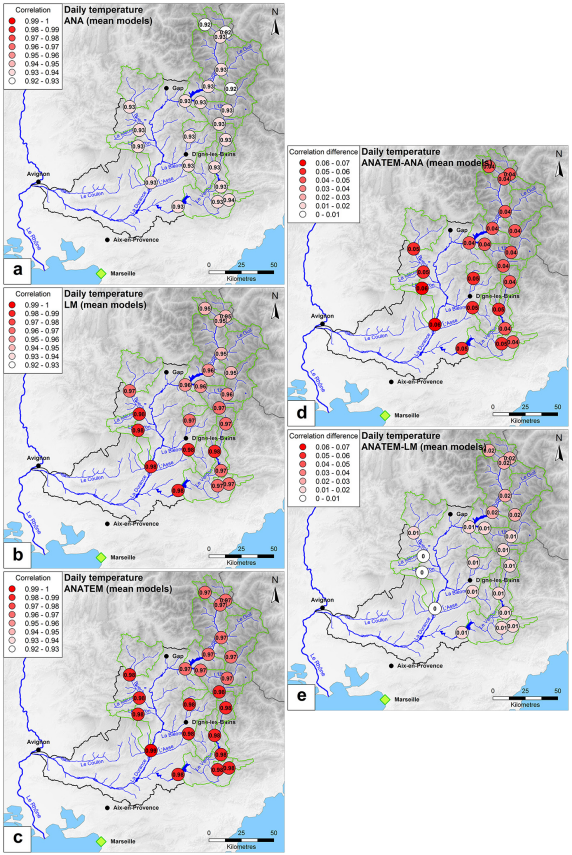


Figure 12. Regional correlation patterns of air temperature mean reconstructions by **(a)** analog method (ANA), **(b)** local model (LM) and **(c)** ANATEM method. Spatial pattern of the gain in correlation obtained with ANATEM reconstructions compared to ANA **(d)** or LM **(e)** reconstructions.

HESSD
12, 311–361, 2015

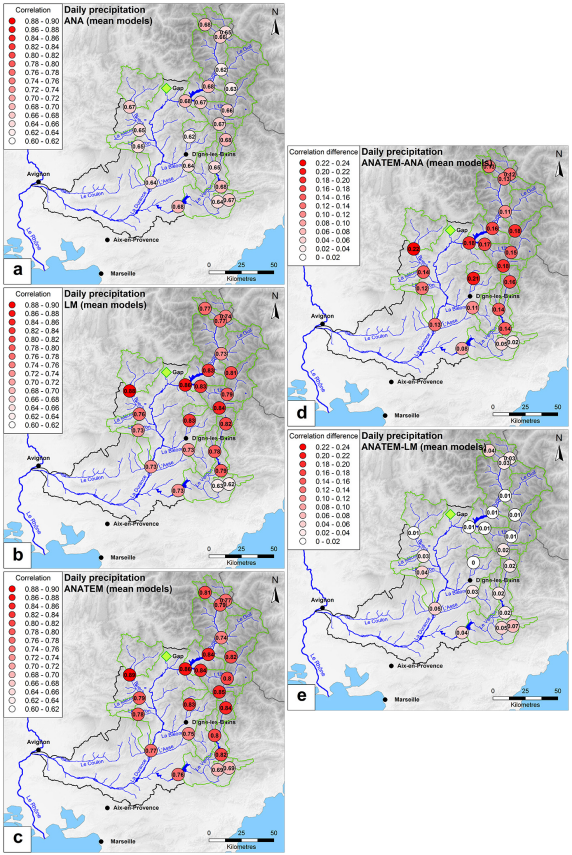


Figure 13. Regional correlation patterns of precipitation mean reconstructions by **(a)** analog method (ANA), **(b)** local model (LM) and **(c)** ANATEM method. Spatial pattern of the gain in correlation obtained with ANATEM reconstructions compared to ANA **(d)** or LM **(e)** reconstructions.

HESSD
12, 311–361, 2015

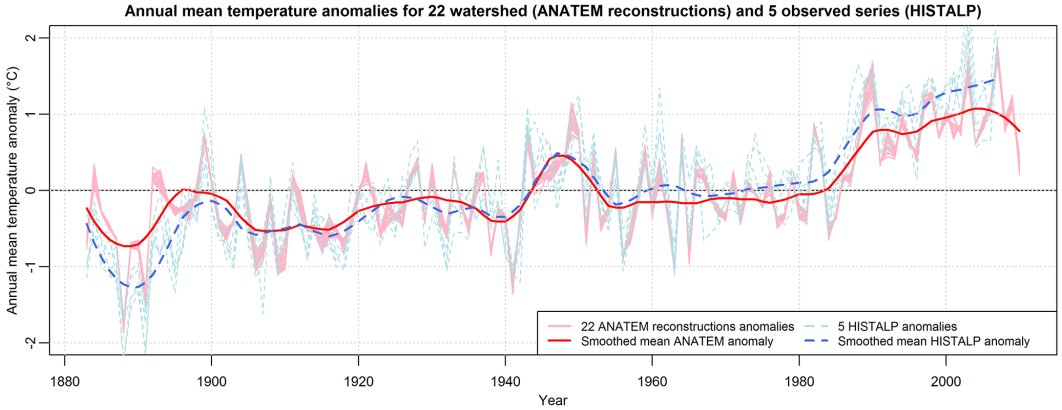


Figure 14. Mean annual air temperature anomaly for the 22 watersheds (ANATEM) and 5 stations (HISTALP).

HESSD
12, 311–361, 2015

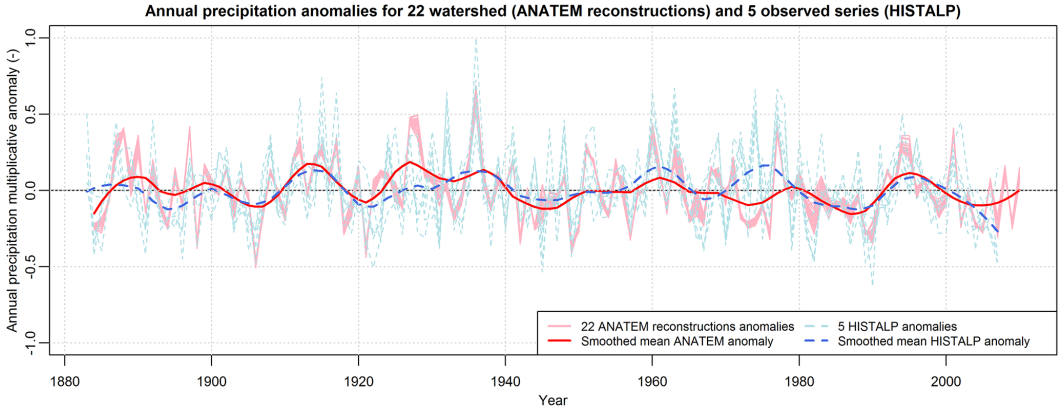


Figure 15. Mean annual precipitation anomaly for the 22 watersheds (ANATEM) and 5 stations (HISTALP).

HESSD
12, 311–361, 2015