Centre de recherche sur l'environnement alpin
Zentrum für alpine Umweltforschung
Research center on alpine environment

TeREsA

A Toolbox in R for Environmental Analysis

Technical Manual v1.0

July 2016

For software version 1.0.1.0

Authors:

Javier FLUIXÁ SANMARTÍN	CREALP
Alain FOEHN	CREALP
Eric TRAVAGLINI	CREALP
Samuel ALESINA	CREALP
Javier GARCÍA HERNÁNDEZ	CREALP

Please refer to this technical manual as:

Fluixá-Sanmartín, J., Foehn, A., Travaglini, E., Alesina, S., Montero-Rubert, R. and García Hernández, J. (2016). *TeREsA – Technical Manual v1.1*. CREALP Group, Switzerland.

Index

Chapter 1. Introduction

The *TeREsA* software is an R-based tool that was developed by the CREALP to optimize **environmental data analysis**. Originally conceived for hydro-meteorological data, some of the modules may be also used to process other types of dataset, such as air pollutant distribution or mineral concentration. TeREsA is a free software running on Windows (Windows Vista or later versions required).

This Technical Manual is a theoretical basis for the different 13 analysis modules available in TeREsA. It should be seen as a reference for the different methodologies and calculations used in the software, commonly applied in the technical literature.

Document structure

This manual is composed of eight main chapters:

1. Introduction
2. Basic statistics
3. Frequency analysis for extreme values
4. Trend analysis
5. Drought Analysis
6. Interpolation methods
7. Elevation sampling
8. Aggregation

For the TeREsA software utilisation, the reader can also use the TeREsA User Manual (Travaglini et al., 2016).

Chapter 2. Basic statistics

Basic statistics are useful to characterize time series. The following statistics are calculated with this module in TeREsA:

- **Minimum, maximum**: the minimum and maximum values of the series.
- **Quartiles**: the quartiles of a ranked set of data values are the three points that divide the data set into four equal groups, each group comprising a quarter of the data. The first quartile (Q1) is defined as the middle number between the smallest number and the median of the data set. The second quartile (Q2) is the median of the data. The third quartile (Q3) is the middle value between the median and the highest value of the data set.
- **Median**: it is the number separating the higher half of a data series from the lower half. The median can be used as a measure of location when a distribution is skewed, when end-values are not known, or when one requires reduced importance to be attached to outliers, e.g., because they may be measurement errors.
- **Mean**: it is a central value of a discrete set of numbers: specifically, the sum of the values divided by the number of values, typically denoted by $\bar{x}$.
- **Mode**: it is the value that appears most often in a set of data. The mode of a continuous probability distribution is the value x at which its probability density function has its maximum value, so the mode is at the peak.
- **Standard Deviation**: used to quantify the amount of variation or dispersion of the time series. The standard deviation σ of a random variable is the square root of its variance and, for a discrete variable $(x_1, x_2, \dots, x_N)$, it is expressed as:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

  where $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$

- **Coefficient of skewness**: it is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. Negative skew, for instance, indicates that the tail on the left side of the probability density function is longer or fatter than the right side. For a sample of N values, a natural method of moments estimator of the population skewness is:

$$b_1 = \frac{m_3}{s^3} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^3}{\left[\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2\right]^{3/2}}$$

  where $\bar{x}$ is the sample mean, s is the sample standard deviation, and the numerator m_3 is the sample third central moment.

- **Kurtosis**: it is a measure of the "tailedness" of the probability distribution of a real-valued random variable. In a similar way to the concept of skewness, kurtosis is a descriptor of the shape of a probability distribution. The kurtosis is the fourth standardized moment defined as:

$$Kurtosis = \frac{\mu_4}{\sigma^4} = \frac{E[(X-\mu)^4]}{(E[(X-\mu)^2])^2}$$

where μ_4 is the fourth moment about the mean and σ is the standard deviation.

Chapter 3. Frequency analysis for extreme values

3.1 Characterization of extreme values

When analyzing environmental data, one of the main issues is the characterization of extreme values in time series using information from past events to update statistical representations of possible future events. Usually, we are interested in a specific extreme event (flood, extreme rainfall…), typically the largest of the year. When we consult only the maximum values within a single year, the dataset is called an annual series. Within an annual series, only the largest value per year is allowed, even if an additional significant peak occurred.

The goal of these analyses, widely used in risk analysis, structure design, or flood protection territory planning, is to use hydro-meteorological data in order to characterize extreme events as follows:

1. Assessing the frequency distribution of extreme events appearance, that is finding the relation between one characteristic variable (generally the peak flow or peak rainfall intensity) and its annual exceedance probability (or AEP). This may be used to predict the possible magnitude of a variable over a certain time period and/or to estimate the frequency with which an event of a certain magnitude may occur.
2. Calculating the main dimensional aspects of the hydrograph/hyetograph corresponding with a given event: rising limb, recession (or falling) limb, peak value, lag time, time to peak.

3.2 Frequency distribution

Frequency analyses involve using observed annual extreme data to calculate statistical information such as mean values, standard deviations, skewness, and recurrence intervals. These statistical data are then used to construct frequency distributions, which are graphs and tables that tell the likelihood of maximum variables as a function of recurrence interval or exceedance probability.

Frequency distributions can take on many forms according to the equations used to carry out the statistical analysis. Let X denote a random variable, and x a possible value of X. For a random variable X, its cumulative distribution function (cdf), denoted $F_X(x)$, is the probability the random variable X is less than or equal to x:

$$F_X(x) = P(X \leq x)$$

where $F_X(x)$ is the non-exceedance probability for the value x.

The probability density function (pdf) describes the relative likelihood that a continuous random variable X takes on different values, and is the derivative of the cdf:

$$f_X(x) = \frac{dF_X(x)}{dx}$$

In hydrology the *percentiles* or *quantiles* of a distribution are often used as design events. The pth quantile x_p is the value with cumulative probability p:

$$F_X(x_p) = p$$

The annual exceedance probability (or AEP) is then calculated as the probability with which an extreme value x_p will be exceeded:

$$AEP = 1 - p$$

The return period (denoted T) is often specified rather than the exceedance probability. In general, x_p is the T-year flood (or rainfall) for:

$$T = \frac{1}{1-p} = \frac{1}{AEP}$$

If extreme events are independent from year to year, the probability that the first exceedance of level x_p occurs in year k is the probability of $(k-1)$ years without an exceedance followed by a year in which the value of X exceeds x_p:

$$P\big(exactly\ k\ years\ unitl\ X \geq x_p\big) = p^{k-1}.(1-p)$$

The probability that x_p is exceeded in a L-year period is:

$$p_L = 1 - \left(1 - \frac{1}{T}\right)^L$$

In hydrology we often speak of the 20-year flood or the 1000-year rainfall, rather than events exceeded with probabilities of 5% or 0.1% in any year. Return period may be incorrectly understood to mean that one and only one T-year event should occur every T years. Actually the probability of the T-year event being exceeded is $1/T$ in every year: on average one event greater than the T-year level occurs in a T-year period. Usually, return periods include the 2–, 10–, 25–, 50–, 100–, and even 500–year events.

The common frequency distributions are:

- **Normal**: this distribution is widely used in hydrology, as well as in other civil engineering applications. It is symmetrical about the mean and is therefore only suitable for data where the skewness coefficient g is equal to or close to zero.
- **Lognormal**: while hydrological data are usually strongly skewed, the logarithms of the data have a near-symmetrical distribution. The log-normal distribution is a normal distribution using the logarithms of the data.
- **Gumbel (EV1)**: the extreme value Type 1 distribution has a constant positive skewness and is commonly used for hydrological analyses. The maxima from any distribution that converges on an exponential function at the positive tail (normal, Chi-square, lognormal, etc.) will have a Gumbel distribution if the basic assumptions are satisfied.
- **Weibull (EV3)**: the Weibull distribution is negatively skewed.

- **Generalized Extreme Value (GEV)**: the general extreme value distribution is the generalized form of the extreme value distributions. It is a family of three sub-types of distribution, which are classified according to the value of the skewness coefficient g.
- **Exponential**: This is the simplest of the one-tailed distributions and is based on the equation $y = e^x$ which is equal to 1 when $x = 0$ and decays rapidly to 0.006 when $x = 5$. It is seldom used directly in hydrological analyses, but like the gamma distribution, it is incorporated in the more complex models derived from it.
- **Pearson type III**: this is essentially a gamma distribution (distribution of the sum of a number of independent, exponentially distributed random variables), but with the mean displaced by a constant x_0 from the origin. It includes the normal distribution as a special case when the skew is zero.
- **log-Pearson type III**: as with the lognormal distribution, this is the distribution of the logarithms of the values, and is the form in which the Pearson type III distribution is most commonly used in hydrological analyses.

Table 1 provides a summary of the pdf or cdf of these distributions, and their ranges of applicability. For more information on these distributions please refer to WMO (World Meteorological Organization, 2008) or to the vast related bibliography.

Distribution	pdf and/or cdf	Range
Normal	$f_X(x) = \frac{1}{\sqrt{2\pi{\sigma_X}^2}} * exp\left[-\frac{(x-\mu_X)^2}{2{\sigma_X}^2}\right]$	$-\infty < x < \infty$
Lognormal	$f_X(x) = \frac{1}{x\sqrt{2\pi{\sigma_Y}^2}} * exp\left[-\frac{(x-\mu_Y)^2}{2{\sigma_Y}^2}\right]$ where $Y = \ln(X)$	$x > 0$
Gumbel	$f_X(x) = \frac{1}{\alpha} exp\left[-\frac{x-\xi}{\alpha} - exp\left(-\frac{x-\xi}{\alpha}\right)\right]$ $F_X(x) = exp\left[-exp\left(-\frac{x-\xi}{\alpha}\right)\right]$	$-\infty < x < \infty$
Weibull	$f_X(x) = \left(\frac{\kappa}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\kappa-1} exp\left[-\left(\frac{x}{\alpha}\right)^{\kappa}\right]$ $F_X(x) = 1 - exp[-(x/\alpha)^{\kappa}]$	$x > 0$ $\alpha, \kappa > 0$
GEV	$F_X(x) = exp\left\{-\left[1 - \frac{\kappa(x-\kappa)}{\alpha}\right]^{1/\kappa}\right\}$	(${\sigma_X}^2$ exists for $\kappa > -0.5$) if $\kappa > 0, x < \left(\xi + \frac{\alpha}{\kappa}\right)$ if $\kappa < 0, x > \left(\xi + \frac{\alpha}{\kappa}\right)$
Exponential	$f_X(x) = \beta . exp[-\beta(x-\xi)]$ $F_X(x) = 1 - \exp\left(-\beta(x-\xi)\right)$	$x > \xi$ for $\beta > 0$

Distribution	pdf and/or cdf	Range
Pearson III	$f_X(x) = \lvert\beta\rvert.[\beta.(x-\xi)]^{\alpha-1}.\frac{exp[-\beta(x-\xi)]}{\Gamma(\alpha)}$ $\Gamma(\alpha)$ is the gamma function (for $\beta > 0$ and $\xi = 0$: $\gamma_x = 2CV_x$)	$\alpha > 0$ for $\beta > 0$: $x > \xi$ for $\beta < 0$: $x < \xi$
log-Pearson III	$f_X(x) = \lvert\beta\rvert.\{\beta.[ln(x) - \xi]\}^{\alpha-1}.\frac{exp\{-\beta[ln(x)-\xi]\}}{x.\Gamma(\alpha)}$	for $\beta < 0$: $0 < x < \exp(\xi)$ for $\beta > 0$: $exp(\xi) < x < \infty$

***Table 1**. Commonly used Frequency Distributions in hydrology*

In *TeREsA*, the Weibull distribution has been used for the determination of the empirical distribution based exclusively on data.

3.3 Parameters estimation

Once the frequency distribution (or distributions) has been selected, the next task is to estimate their parameters by fitting this distribution to the datasets of the annual maximum values. It is then possible to evaluate the required quantiles using the fitted model equations.

Several general approaches are available for estimating the parameters of a distribution. Traditionally, the method of ordinary moments (MOM) has been popular in hydrology even though it has been recognized as statistically inefficient in comparison to the method of maximum likelihood (ML). The method of probability-weighted moments (PWMs), introduced by Greenwood, et al. (1979) is, in many cases, convenient to apply, and it has been found by Hosking, et al. (1985) to be comparable with ML in its statistical properties for sample sizes which are normally encountered in hydrology.

A more recent methodology employing L-moment statistics (Hosking, 1990) shows considerable improvement over the more conventional maximum likelihood or method at moments techniques.

According to their availability and their convenience in the functions used in *TeREsA*, the corresponding parameters estimation methods chosen for each frequency distribution have been selected and are shown in **Table 2**.

Distribution	Method
Normal	Maximum Likelihood
Lognormal	Maximum Likelihood
Gumbel	Maximum Likelihood

Distribution	Method
GEV	Probability weighted moments
Exponential	Method of ordinary moments
Pearson III	Maximum Likelihood
log-Pearson III	Maximum Likelihood

***Table 2**. Parameter estimation methods chosen for each frequency distribution in TeREsA*

Once the distribution is chosen and its parameters are defined, it is possible to estimate the quantile –or annual extreme value– for any given return period using the equation shown in **Table 1** or graphically.

3.4 Selecting a distribution – Goodness of fit tests

After the dataset of empirical information has been adjusted by one (or several) frequency distribution, it is needed to describe how well it fits the observations by assessing the discrepancy between observed values and the values expected under the model in question.

An initial evaluation of the adequacy of a fitted probability distribution is best done by generating a probability plot of the observations and the distributions. Such a plot serves both as an informative visual display of the data and a check to determine whether the fitted distribution is consistent with the data.

A more objective procedure is the use of analytical goodness-of-fit criteria, useful to determine whether it is reasonable to conclude that a given set of observations was drawn from a particular family of distributions, or whether a particular departure of the data from a model is statistically significant (Stedinger et al., 1993). Several rigorous statistical tests are available, some of which are described below.

Kolmogorov-Smirnov goodness-of-fit test

The non-parametric Kolmogorov–Smirnov test provides bounds within which every observation on a probability plot should lie if the sample is actually drawn from the assumed distribution (Kottegoda and Rosso, 1997). It is one of the most useful and general methods for comparing two samples, as it is sensitive to differences in both location and shape of the empirical cumulative distribution functions (ECDF) of the two samples.

The test statistic, in a two-sided test, is the maximum absolute difference (that is, usually the vertical distance) between the empirical and hypothetical CDFs (**Figure 1**).

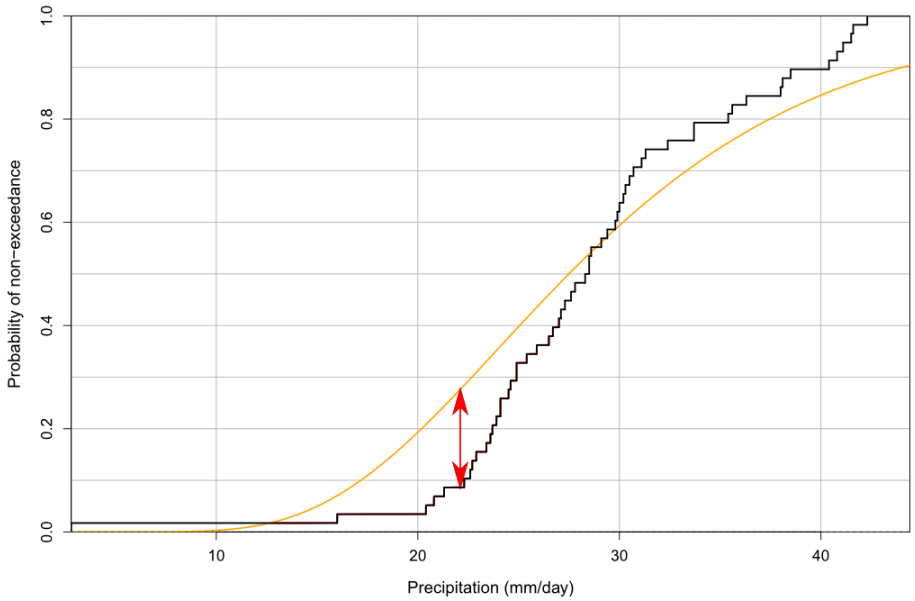

***Figure 1**. Example of the Kolmogorov-Smirnov statistic. Orange line is CDF, black line is the ECDF, and the red arrow is the Kolmogorov-Smirnov statistic.*

For a continuous variate X let $x_{(1)}, x_{(2)}, \ldots x_{(n)}$ represent the order statistics of a sample of size n, that is, the values arranged in increasing order. The empirical or sample distribution function $F_n(x)$ is a step function. This gives the proportion of values not exceeding x and is defined as:

$$\begin{aligned} F_n(x) &= 0, \quad for\ x < x_{(1)}, \\ &= \frac{k}{n}, \quad for\ x_{(k)} \leq x \leq x_{(k+1)}; k = 1, 2, \ldots, n-2, \\ &= 1, \quad for\ x \geq x_{(n)} \end{aligned}$$

Let $F_0(x)$ denote a completely specified theoretical continuous CDF. The null hypothesis H_0 is that the true CDF of X is the same as $F_0(x)$. That is, under the null hypothesis

$$\lim_{n\to\infty} Pr[F_n(x) = F_0(x)\ ] = 1$$

The test criterion is the maximum absolute difference between $F_n(x)$ and $F_0(x)$, formally defined as

$$D_n = \sup_x |F_n(x) - F_0(x)\ |$$

where sup_x is the supremum of the set of distances.

The foregoing measure of deviation is for a two-sided test which is commonly used. If for some reason a one-sided test is required to test whether, for instance, $F_n(x) > F_0(x)$, then the statistic is modified as

$$D_n^+ = \sup_x [F_n(x) - F_0(x)\ ]$$

Likewise one can define the statistic D_n^- . One of the advantages of the test is that the test statistic is distribution-free. For large values of n, Smirnov (1948) gives the limiting distribution of $\sqrt{n}\ D_n$ as

$$\lim_{n\to\infty}\left[Pr\left(\sqrt{n}\,D_n \le z\right)\right] = \left(\frac{\sqrt{2\pi}}{z}\right)\sum_{k=1}^{\infty} exp\left[-(2k-1)^2\frac{\pi^2}{(8z^2)}\right]$$

Thus, one can compute that the critical values $D_{n,\alpha}$ for large samples, say $n > 35$, are $1.3581/\sqrt{n}$ and $1.6276/\sqrt{n}$ for $\alpha = 0.05$ and 0.01, that is, for probabilities of 0.95 and 0.99 respectively.

The test is applied on the assumption that $F_0(x)$ denotes a completely specified theoretical continuous CDF, that is with known parameters.

However, various studies have found that, even in this corrected form, the test is less powerful for testing normality than the Anderson–Darling test, for instance.

Anderson-Darling goodness-of-fit test

The Anderson-Darling test is a modification of the Kolmogorov-Smirnov test, giving more weight to the tails. This test is devised to give heavier weighting to the tails of a distribution where unexpectedly high or low values, called outliers, are sometimes located. This is made possible if one divides the difference between the empirical CDF $F_n(x)$ and theoretical CDF $F_0(x)$ to be tested (that is, $(F_n(x) - F_0(x))$,which approaches zero in each tail) by $\sqrt{F_0(x)\,[1 - F_0(x)]}$ (Anderson and Darling, 1954). After squaring the test statistic becomes

$$A^2 = \int_{-\infty}^{\infty} [F_n(x) - F_0(x)]^2 \cdot \frac{1}{F_0(x)[1 - F_0(x)]} \cdot F_0(x) \cdot dx$$

It is shown that this is equivalent to

$$A^2 = -n - \sum_{i=1}^{n} \frac{[2i-1]\left[ln\{F_0(x_{(i)})\} + ln\{1 - F_0(x_{(n-i+1)})\}\right]}{n}$$

where $x_{(1)}, x_{(2)}, \ldots x_{(n)}$ are the observations ordered in ascending order. Because the CDFs are in the range 0–1, their logarithms are negative and hence the summation on the right-hand side of the previous equation is negative. The absolute value of the summation is also greater than n, thus resulting in a positive value of A^2.

For large values of the test statistic A^2, the null hypothesis that $F_n(x)$ and $F_{n0}(x)$ have the same distribution is rejected.

Chi-squared goodness-of-fit test

The chi-squared test (also referred to as χ^2 test or chi-square test) is a test of significance based on the chi-squared statistic with CDF given by:

$$F(\chi^2) = \frac{1}{2}\int_0^{\chi^2} \frac{(\frac{t}{2})^{(\nu/2)-1} \cdot e^{-t/2}}{\Gamma(\nu/2)}\,dt$$

The statistic is derived by the sum of squares of independent standard normal variates. The main steps are the ranking of a sample of data, division into a number of classes depending on the magnitudes and the range, and the fitting of a probability distribution. The statistic

comes from the weighted sum of squared differences between the observed and theoretical frequencies.

The observed frequencies O_i and expected frequencies E_i are found by multiplying the relative frequencies, for each class i from a total of l classes, by the sample size n. To test whether the differences between the observed and expected frequencies are significant, we use the statistic

$$X^2 = \sum_{i=1}^{l} \frac{(O_i - E_i)^2}{E_i}$$

A large value of this statistic indicates a poor fit; so we need to know what values are acceptable. The sampling distribution of X^2 tends, as n approaches infinity, to a χ_v^2 v distribution, where $v = l - k - 1$ represents the degrees of freedom and k is the number of parameters estimated from the same data used for the test.

The test gives satisfactory results when there is no significant dependence between the variables, if $n \geq 50$ and for each class i, $n_i \geq 5$. It is versatile and does not require one to know the values of the parameters before the test, as in the classical form of the Kolmogorov-Smirnov goodness-of-fit test,

For the application of the chi-squared test to a continuous variable, the expected frequencies E_i are the products of the total sample size n and the areas under the pdf, as specified by the null hypothesis, between the bounds of each class i.

Although this is an attractive test, the choice of classes will affect the power of the test. Furthermore, it is not the best approach to have equal class intervals for the purpose. An equitable allocation of the frequencies is obtained if we divide the total area under the pdf into equal areas and hence find the class boundaries. This is the equal-probabilities method of constructing classes as proposed by Mann and Wald (1942) and clarified by Williams (1950) who also suggest for a level of significance $\alpha = 0.05$, values of classes $l =$ 39, 35, 30, 23, 15, 12, and 9 for total sample sizes $n =$2000, 1500, 1000, 500, 200, 100, and 50, respectively. For other values of n and α, we use the formula

$$l = 2\left[\frac{2(n-1)^2}{z_\alpha^2}\right]^2$$

where z_α is the value which a standard normal variate exceeds with probability α. Also $n/l \geq 5$; although this requirement is relaxed somewhat by recent authors, too many classes tend to reduce the power of the test.

The Mean Square Prediction Error

The Mean Square Prediction Error is a simple but effective indicator of the goodness-of-fit of a distribution. It is calculated as the expected value of the squared difference between the fitted values and the (unobservable) distribution:

$$MSPE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{x}_i)^2$$

where n is the size of the sample, x_i are the observations and $\hat{x}_i$ the predictions made with the probability distribution.

For the validation of the fitted parameters, TeREsA presents a report for each station including:

- The probability plot of the observations *vs* the distributions.
- A table containing the values of the parameters, the Mean Square Prediction Error, and the statistic and the p-value of the Kolmogorov-Smirnov goodness-of-fit tests.

Chapter 4. Trend analysis

Among the methodologies to analyze trends in times series, the following methods are the most common: Linear regression, Mann-Kendall test, Spearman's rho test, Cox Stuart test, Pettitt test and Sen's non parametric estimator of slope.

For this module, the trend analyses are applied to the annual means of each time series (or seasonal means), so the white noise of the time series can be reduced and their autoregressive component avoided without using any pre-whitening method (von Storch, 1999).

4.1 Linear regression

Linear regression attempts to model the relationship between two variables (usually data versus time) by fitting a linear equation to observed data.

If plots of data versus time suggest a simple linear increase or decrease over time, a linear regression of the variable against time may be fit to the data.

$$X = \beta_0 + \beta_1 * t + \varepsilon$$

The most common method for fitting a regression line is the method of least-squares. This method calculates the best-fitting line for the observed data by minimizing the sum of the squares of the vertical deviations from each data point to the line (if a point lies on the fitted line exactly, then its vertical deviation is 0).

The null hypothesis is that the slope coefficient $\beta_1 = 0$. Regression makes stronger assumptions about the distribution of X over time than does Mann-Kendall. It must be checked for normality of residuals, constant variance and linearity of the relationship (best done with residuals plots). If X is not linear over time, a transformation will likely be necessary. If all is ok, the t-statistic on β_1 is tested to determine if it is significantly different from 0. If the slope is nonzero, the null hypothesis of zero slope over time is rejected, and we conclude that there is a linear trend in Y over time.

4.2 Mann-Kendall test

The purpose of the Mann–Kendall test (Gilbert, 1987; Kendall, 1975; Mann, 1945) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. This test can be used in place of a parametric linear regression analysis, which requires that the residuals from the fitted regression line be normally distributed; an assumption not required by the Mann-Kendall test as it is a non-parametric (distribution-free) test.

Commonly used when analyzing various types of environmental data (Hipel and McLeod, 1994; McLeod et al., 1991), the Mann-Kendall test is best viewed as an exploratory analysis and is most appropriately used to identify stations where changes are significant or of large magnitude and to quantify these findings (Hirsch et al., 1982).

The null hypothesis H_0 is that a sample of data ordered chronologically is independent and identically distributed:

$$H_0: \quad Prob\left[Y_j > Y_i\right] = 0.5, \ \ where\, T_j > T_i$$

$$H_1: \quad Prob\left[Y_j > Y_i\right] \neq 0.5 \ \ (2 - sided\ test)$$

The Mann-Kendall test statistic S is defined as follows (Yue et al., 2002b):

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sgn(x_j - x_j)$$

where:

$$sign(x) = \begin{cases} +1 & if\ x > 0 \\ 0 & if\ x = 0 \\ -1 & if\ x < 0 \end{cases}$$

The Mann-Kendall test statistic Z is estimated by the following formula as:

$$Z = \begin{cases} \dfrac{S-1}{\sigma} & S > 0 \\ 0 & S = 0 \\ \dfrac{S+1}{\sigma} & S < 0 \end{cases}$$

where the variance σ^2 is defined as:

$$\sigma^2 = \left[n(n-1)(2n+5) - \sum_{j=1}^{m} t_j.(t_j-1)(2t_j-5)\right] \Big/ 18$$

Here, n is the length of the time series $x_1, .., x_n$;

x_i and x_k are the data values in years i and k;

m is the number of tied groups;

t_j is the number of data values in the jth group.

A positive (negative) value of Z indicates an upward (downward) monotone trend for the test time series. The test statistic Z is used a measure of significance of trend. This test statistic is used to test the null hypothesis, H_0. If $|Z| > Z_{\alpha/2}$, where α represents the chosen significance level[1], then the null hypothesis is invalid implying that the trend is significant.

As a matter of good scientific practice, a significance level is chosen before data collection and is often set to 0.05 (5%) (Craparo, 2007), although other significance levels (e.g., 0.01) may be used, depending on the field of study (Sproull, 2002).

[1] The significance level (denoted α, alpha) is the probability of rejecting the null hypothesis given that it is true. $Z_{\alpha/2}$ is obtained from the standard normal cumulative distribution tables (eg: 5% with $Z_{0.025} = 1.96$)

This test is particularly useful since missing values are allowed and the data need not conform to any particular distribution. Also, data reported as trace or less than the detection limit can be used by assigning them a common value that is smaller than the smallest measured value in the data set.

4.3 Spearman's rho test

Spearman's rho (Lehmann and D'Abrera, 2006; Sneyers, 1990) test is another rank-based nonparametric method used for trend analysis and is similar to the Mann-Kendall method. The Spearman's rho test is a simple method with uniform power for linear and non-linear trends and is commonly used to verify the absence of trends.
In this test, the null hypothesis (H_0) is that all the data in the time series are independent and identically distributed, while the alternative hypothesis (H_1) is that increasing or decreasing trends exist (Yue et al., 2002a). The SR test statistic D and the standardized test statistic *ZSR* are expressed as follows:

$$D = 1 - \frac{6\sum_{i=1}^{n}(R_i - i)^2}{n(n^2-1)}$$

$$Z = D * \sqrt{\frac{n-2}{1-D^2}}$$

where R_i is the rank of observation X_i in the time series and n is the length of the time series. Positive values of Z indicate upward trends, while negative Z indicate downward trends in the time series. When $|Z| > t_{(n-2,1-\alpha/2)}$, the null hypothesis is rejected and a significant trend exists in the time series. $t_{(n-2,1-\alpha/2)}$ is the critical value of t from the t-student table, for α significant level.

4.4 Cox Stuart test

The Cox-Stuart test (Cox and Stuart, 1955) is a modified sign test. It is defined as a little powerful test (power equal to 0.78), but very robust for the trend analysis. It is therefore applicable to a wide variety of situations, to get an idea of the evolution of values obtained. The proposed method is based on the binomial distribution.

Given a set of ordered observations $X_1, X_2, \ldots, X_n$. We group the variables into pairs $(X_1, X_{1+c}), (X_2, X_{2+c}), \ldots, (X_{n-c}, X_n)$, where:

$$c = \begin{cases} \dfrac{n}{2} & if\ n\ is\ even \\ \dfrac{(n+1)}{2} & if\ n\ is\ odd \end{cases}$$

If there is an upward trend, then:

$$P(X_i < X_{i+c}) > P(X_i > X_{i+c}) \quad \text{for all } i$$

Define

$$sign(X_i, X_{i+c}) = \begin{cases} +1 & if\ X_i < X_{i+c} \\ 0 & if\ X_i = X_{i+c} \\ -1 & if\ X_i > X_{i+c} \end{cases}$$

We can now apply the sign test to test if $P(Xi < Xi + c) = P(Xi > Xi + c)$, or equivalently if $P(+) = P(-)$.

4.5 Pettitt test

The approach after Pettitt (1979) is commonly applied to detect a single change-point in hydrological series or climate series with continuous data.

It tests the null hypothesis H_0: The T variables follow one or more distributions that have the same location parameter (no change), against the alternative: a change point exists. The non-parametric statistic is defined as:

$$K_T = max|U_{t,T}|$$

where:

$$U_{t,T} = \sum_{i=1}^{t} \sum_{j=t+1}^{T} sgn\ (X_i - X_j)$$

The change-point of the series is located at K_T, provided that the statistic is significant. The significance probability of K_T is approximated for $p \leq \alpha$ with:

$$p = 2 * exp\left(\frac{-6\ {K_T}^2}{T^3 + T^2}\right)$$

α being the significance level of the test.

4.6 Sen's non parametric estimator of slope

If a linear trend is present, the true slope (change per unit time) may be estimated by computing the Sen's slope (Sen, 1968; Theil, 1992). This method chooses the median slope among all lines through pairs of two-dimensional sample points.

Closely related to the Mann-Kendall test, it can be computed efficiently, and is insensitive to outliers; it can be significantly more accurate than non-robust simple linear regression for skewed and heteroskedastic data, and competes well against non-robust least squares even for normally distributed data in terms of statistical power (Wilcox, 2001).

The Theil–Sen estimator of a set of two-dimensional points (x_i, y_i) is the median m of the slopes $(y_j - y_i)/(x_j - x_i)$ determined by all pairs of sample points. Once the slope m has been determined, one may determine a line from the sample points by setting the y-intercept b to be the median of the values $y_i = m * x_i$.

A confidence interval for the slope estimate may be determined as the interval containing the middle $(1 - \alpha)$ of the slopes of lines determined by pairs of points, and may be

estimated quickly by sampling pairs of points and determining the $(1-\alpha)*100$ interval (in %) of the sampled slopes.

Chapter 5. Drought Analysis

Drought is a natural phenomenon that results from persistent lower levels of precipitations than what is considered normal, generally affecting larger areas than other hazards. When this phenomenon extends over a season or a longer period of time, precipitation is insufficient to meet the demands of human activities and the environment (World Meteorological Organization, 2012) and shortages of water availability may turn into disaster events.

5.1 Drought classification

Droughts are commonly classified in three main categories (Barua et al., 2011; Zargar et al., 2011):

- **Meteorological drought**: it is expressed solely as the level of dryness measured in rainfall deficiency (Keyantash and Dracup, 2002) as the precipitation's departure from normal over some period of time, reflecting one of the primary causes of a drought.
- **Agricultural drought**: it is associated with a shortage of available water for plant growth, and is assessed as insufficient soil moisture to replace evapotranspirative losses (World Meteorological Organization, 1975).
- **Hydrological drought**: it is associated with the effects of periods of precipitation (including snowfall) shortfalls on surface or subsurface water supply (i.e., streamflow, reservoir and lake levels, groundwater). Hydrological droughts are usually out of phase with or lag the occurrence of meteorological and agricultural droughts: it takes longer for precipitation deficiencies to show up in components of the hydrological system.

A supplemental division may be adopted: the **socio-economical droughts** (Heim, 2002; Wilhite and Glantz, 1985), as a consequence of the other types. This category occurs when physical water shortages start to affect the health, well-being, and quality of life of people and/or when the drought starts to affect the supply and demand of economic products such as water, fish production, and hydroelectric power generation (Barua et al., 2011).

The relationship between the different drought categories can be illustrated as in **Figure 2**. A meteorological drought in terms of lack of precipitation is the primary cause of a drought. It usually first leads to an agricultural drought due to lack of soil water. If precipitation deficiencies continue a hydrological drought in terms of surface water deficits develops. The groundwater is usually the last to be affected, but also the last to return to normal water levels.

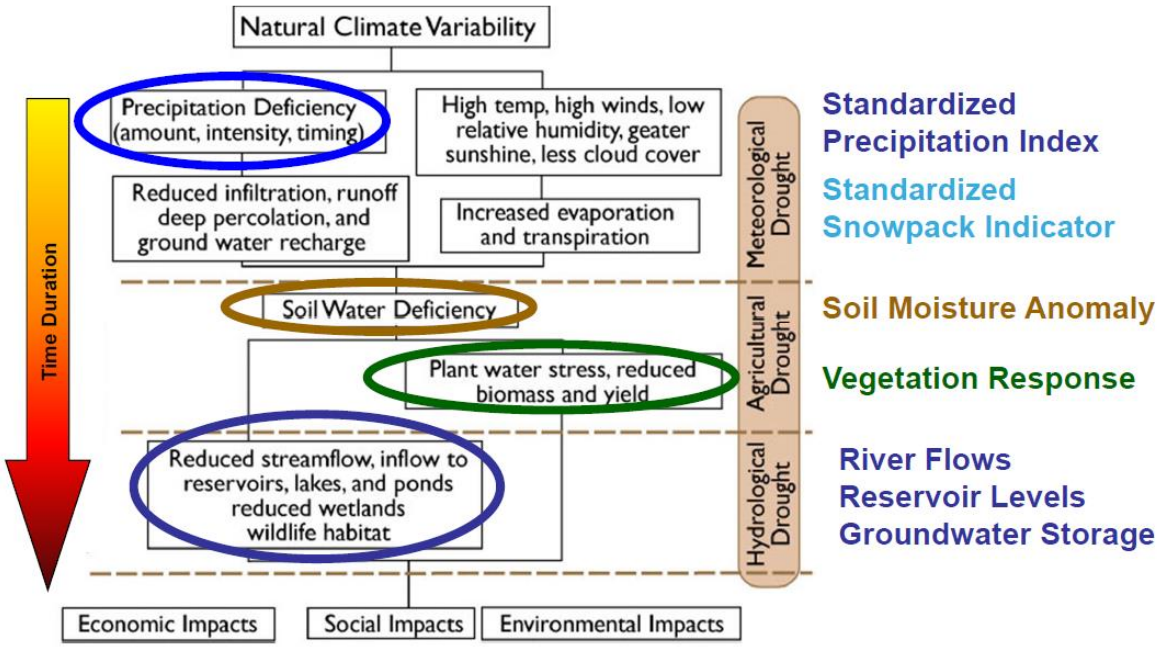

***Figure 2**. Flow chart illustrating the progression of drought, and the relationship between Meteorological, Agricultural, and Hydrological Drought (source: National Drought Mitigation Centre, USA).*

The drought module of *TeREsA* is focused on the calculation and exploitation of meteorological drought indices, described hereafter: Percentage of Normal precipitation (PN), Standardized Precipitation Index (SPI), Rainfall Anomaly Index (RAI) and Deciles.

All indices require a reference period used to define what is considered the "normal precipitations". *TeREsA* is able to set this period by setting its initial and end years or by defining a "moving" period of *n* years previous to the date we are calculating the index.

5.2 Drought indices

The definition of meteorological drought indices rests basically upon the comparison of a given weather variable (in this case the precipitation) to its normal value (the definition of "normality" may vary from index to index), resulting in a single number. This allows assimilating large data into a comprehensible picture for drought analysis and thus facilitating decision making.

Drought indices are calculated for specific stations and results are in point data format. These results serve their purposes, but it is often in map form that the data best communicate a message based on a geographic context to the decision-maker trying to understand drought severity and spatial extent (World Meteorological Organization, 2012).

A variety of techniques can be used to generate a continuous map of meteorological drought with information between stations. One such technique generates an interpolated surface of estimated values at locations between sites based on mathematical relationships of the indicator or index between the original point data. Often this produces a map that appears

"natural", but is still based on the data from specific points and is only as accurate as the original data and the interpolation technique. No single interpolation method can be applied to all situations, and the most commonly used interpolation techniques include Kriging and Inverse Distance Weighting (IDW).

The Kriging method, which has its origins in geological applications and the mining industry, assumes that there is a relationship between points that is non-random and changes over space. Inverse Distance Weighting (IDW) is used when the data points are scattered but dense enough to represent local variations. The data, as the name implies, are weighted to favour data closer in proximity to the point being processed.

Percent of Normal Precipitation (PN)

The percent of normal precipitation is one of the simplest measurements of rainfall for a location. Analyses using percent of normal are very effective when used for a single region or a single season.

It is calculated by dividing actual precipitation by normal precipitation (typically considered to be a 30-year mean) and multiplying by 100%. This can be calculated for a variety of time scales, including monthly, seasonal or annual. For PN values over 100%, the precipitation is higher than the average precipitation (and viceversa): the higher PN value, the wetter the considered period is.

One of the disadvantages of using the percent of normal precipitation is that the mean, or average, precipitation is often not the same as the median precipitation, which is the value exceeded by 50% of the precipitation occurrences in a long-term climate record. The reason for this is that precipitation on monthly or seasonal scales does not have a normal distribution. Use of the percent of normal comparison implies a normal distribution where the mean and median are considered being the same.

Standardized Precipitation Index (SPI)

The Standardized Precipitation Index (SPI) was formulated by Mckee, Doesken and Kleist in 1993 to quantify the precipitation deficit for multiple time scales.

It is computed by considering the precipitation anomaly with respect to the mean value for a given time scale, divided by its standard deviation. The precipitation is not a normal distribution, at least for time-scales less than one year. Therefore, the variable is adjusted so that the SPI is a Gaussian distribution with zero mean and unit variance.

The index calculation is based on the following expressions:

$$SPI = +\left(t - \frac{c_0 + c_1 * t + c_2 * t^2}{1 + d_1 * t + d_2 * t^2 + d_3 * t^3}\right),$$

$$t = \sqrt{ln\left(\frac{1}{H(P)^2}\right)} \quad for\ 0 < H(P) < 0.5$$

$$SPI = -\left(t - \frac{c_0 + c_1 * t + c_2 * t^2}{1 + d_1 * t + d_2 * t^2 + d_3 * t^3}\right),$$

$$t = \sqrt{ln\left(\frac{1}{\left(1 - H(P)\right)^2}\right)} \quad for\ 0.5 < H(P) < 1$$

where P is the cumulated precipitation for the given time-scale, $H(P)$ is the cumulative probability of the observed precipitation and $c_0, c_1, c_2, d_1, d_2, cd_3$ are mathematical constants. **Table 3** presents the interpretation of the possible values of the SPI.

SPI	Description
≥ 2.0	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
−0.99 to 0.99	Near normal
−1.49 to −1.0	Moderately dry
-1.99 to -1.5	Severely dry
≤ -2.0	Extremely dry

***Table 3**. Classification of the period according to the values of the SPI.*

As mentioned earlier, the SPI was designed to quantify the precipitation deficit for multiple timescales, or moving averaging windows (World Meteorological Organization, 2012). These time scales reflect the impact of a drought on different water resources needed by various decision-makers: meteorological and soil moisture conditions (agriculture) respond to precipitation anomalies on relatively short timescales (1-6 months), whereas streamflow, reservoirs, and groundwater respond to longer-term precipitation anomalies (6-24 months or longer):

- 1-month SPI: a 1-month SPI map is very similar to a map displaying the percentage of normal precipitation for a 30-day period. Its application can be related closely to meteorological types of drought along with short-term soil moisture and crop stress.
- 3-month SPI: the 3-month SPI provides a comparison of the precipitation over a specific 3-month period with the precipitation totals from the same 3-month period for all the years included in the historical record. A 3-month SPI reflects short- and medium-term moisture conditions and provides a seasonal estimation of precipitation. Looking at longer timescales can prevent misinterpretation believing that a drought might be over when in fact it is just a temporary wet period.
- 6-month SPI: the 6-month SPI indicates seasonal to medium-term trends in precipitation. A 6-month SPI can be very effective in showing the precipitation over distinct seasons. Information from a 6-month SPI may also begin to be associated with anomalous streamflows and reservoir levels, depending on the region and time of year.

- 12-month up to 24-month SPI: the SPI at these timescales reflects long-term precipitation patterns. SPIs of these timescales are usually tied to streamflows, reservoir levels, and even groundwater levels at longer timescales.

Rainfall anomaly Index (RAI)

The Rainfall anomaly Index (RAI) was developed by Van Rooy (1965). The positive and negative RAI indices are computed by using the mean of ten extremes. Let $\bar{M}$ be the mean of the ten highest precipitation records for the period under study, $\bar{P}$ the mean precipitation of all the records for the period, and P the precipitation for the specific year. Then the positive RAI (for positive anomalies) for that year is:

$$RAI = 3 * \frac{P - \bar{P}}{\bar{M} - \bar{P}}$$

Let $\bar{m}$ be the mean of the ten lowest precipitation records for the period under study. Then the negative RAI (for negative anomalies) for that year is:

$$RAI = -3 * \frac{P - \bar{P}}{\bar{m} - \bar{P}}$$

The classification of the RAI is as follows:

RAI	Period
≥ 3.00	Extremely wet
2.00 to 2.99	Very wet
1.00 to 1.99	Moderately wet
0.50 to 0.99	Slightly wet
-0.49 to 0.49	Near normal
-0.99 to -0.50	Slightly dry
-1.99 to -1.00	Moderately dry
-2.99 to -2.00	Very dry
≤ -3.00	Extremely dry

***Table 4**. Classification of the period according to the values of the RAI.*

Deciles

Arranging monthly precipitation data into deciles is another drought-monitoring technique. It was developed by Gibbs and Maher (1967) to avoid some of the weaknesses within the "percent of normal" approach.

The technique divides the distribution of occurrences over a long-term precipitation record into tenths of the distribution. Each of these categories is called a decile. The first decile is the rainfall amount not exceeded by the lowest 10% of the precipitation occurrences. The second decile is the precipitation amount not exceeded by the lowest 20% of occurrences. These deciles continue until the rainfall amount identified by the tenth decile is the largest

precipitation amount within the long-term record. By definition, the fifth decile is the median, and it is the precipitation amount not exceeded by 50% of the occurrences over the period of record.

Its classification is shown in **Table 5**.

Class	Percent	Period
Deciles 1-2	lowest 20%	Much below normal
Deciles 3-4	next lowest 20%	Below normal
Deciles 5-6	middle 20%	Near normal
Deciles 7-8	next highest 20%	Above normal
Deciles 9-10	highest 20%	Much above normal

***Table 5**. Classification of the period according to the values of the deciles.*

5.3 Characterization of a drought event

We can define drought occurrence based on the values of the drought indices.

According to McKee et al. (1993) a drought event is defined as a period in which the SPI is continuously negative and the SPI reaches a value of -1.0 or less. The drought begins when the SPI first falls below zero and ends with the positive value of SPI following a value of -1.0 or less (**Figure 3**). The event ends when the SPI becomes positive. This definition has been adopted and is used by the Joint Research Centre of the European Commission or by the National Drought Mitigation Center of the USA.

Each drought event, therefore, has a duration defined by its beginning and end, and a drought magnitude which is the positive sum of the SPI for each month during the drought event. The intensity of a drought event is then defined as the magnitude of this event divided by its duration.

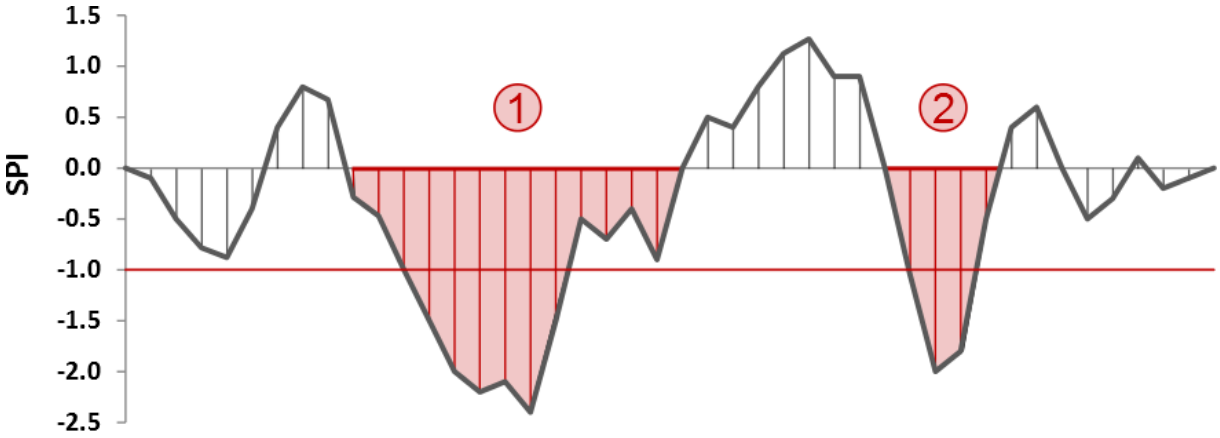

***Figure 3**. Example of 3 drought events according to the SPI (source: Vogt et al.).*

The criterion for defining an event has been extrapolated to the rest of the drought indices as follows:

Index	Lower limit (start of the drought event)	Upper limit (end of the drought event)
SPI	*-1*	*0*
PN	*75*	*100*
RAI	*-1*	*0*
Deciles	*40*	*60*

***Table 6**. Definition of the lower and upper limits of the indices defining the start and the end (respectively) of a drought event.*

Chapter 6. Interpolation methods

6.1 Thiessen interpolation

This method is also known under the names of Voronoi diagram.

This is the simplest interpolation method for meteorological data, and uses the nearest available elevation data (ED) point as a reference, without performing any further adjustment on the data. The interpolated data at each time step thus has polygonial structure (Figure 4).

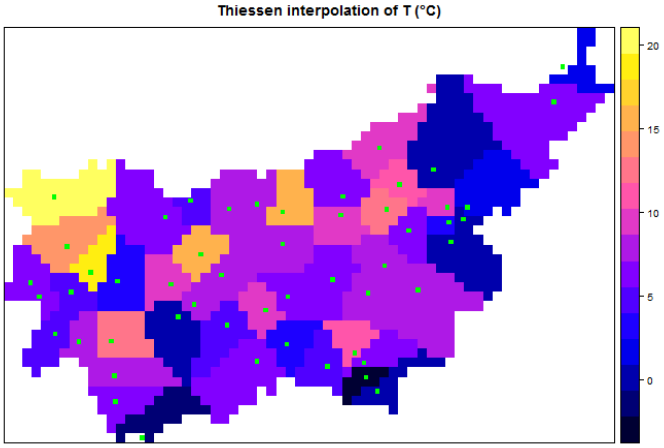

***Figure 4**. Thiessen interpolation of measurement points (green) on interpolation points (other colors)*

Method

The code is optimised for the situation in which the interpolation needs to be repeated numerous times with almost identical[2] data configurations.

1. First, a lookup table is generated that lists for each interpolation point the nearest measurement stations available, for the sensor of interest (P, T and ETP). The distance between points is estimated as the horizontal distance, without taking elevation changes into account:

$$d_{ij} = \sqrt{\Delta X^2 + \Delta Y^2} \qquad \text{DS.1}$$

   Thanks to this lookup table, the distances only need to be calculated once.
2. The program then loops through each time step and fetches for each interpolation point the data value from the reference measurement station, indicated by the first element of the lookup table for that station.

[2] Alterations can occur if measurement points are missing.

3. In the situation where one of the data points is missing, data from the next nearest measurement stations is used.

It can be expected that this method runs faster than other interpolation methods.

6.2 Inverse distance weighting interpolation

Note: Shephard interpolation is a special case of inverse distance weighting.

In this method, the weights are proportional to the inverse power of the distance between measurement stations and interpolation points. The user can choose between a power of 1 and a power of 2.

Method

For this method, the idw function from the gstat package (Pebesma, 2004) for R (R Development Core Team, 2014) was used. The weight w_{ij} of the j[th] measurement station for the i[th] interpolation point is given by:

$$w_{ij} = \frac{1}{d_{ij}^{p}} \qquad \text{DS.2}$$

Where d_{ij} is the 2D distance as defined in DS.1. The data estimation at the i[th] interpolation point x_i is:

$$x_i = \frac{1}{\sum_j w_{ij}} \sum_j w_{ij} \times y_j \qquad \text{DS.3}$$

Where y_j is the data measurement at the j[th] measurement station.

When, for a given interpolation point, all measurement stations are relatively far away, the estimated value at the interpolation point tends towards the global average.

6.3 Kriging

For this method, the autokrige function from the automap package (Hiemstra et al., 2009) for R was used. This function automatically estimates the variogram and then calls the krige function from the gstat package (Pebesma, 2004) to do the actual kriging.

Ordinary Kriging method

Ordinary Kriging assumes that the random variable Z(s) to be estimated respects the model

$$Z(s) = \mu(s) + \varepsilon(s) \qquad \text{DS.4}$$

Where $\mu(s)$ is an unknown constant. Documentation on kriging can be found in Cressie (1993).

Universal Kriging method

When the option "remove trend with elevation data" is activated, Universal Kriging is used instead of Ordinary Kriging. The conceptual difference is that $\mu(s)$ is assumed to be a deterministic function (of the elevation) and not an unknown constant. See Matheron (1993) for more details.

The option of kriging with trend removal for precipitation was removed because in preliminary tests, the results proved unreliable (total volume about 3 times higher than for other methods).

Chapter 7. Elevation sampling

The user can provide elevation data that is necessary for certain interpolation methods (currently only the Universal kriging method). The elevation data (ED) is used to estimate the elevation at each interpolation point before the interpolation takes place. Because both the elevation data and the interpolation points are spatial points, the program uses a Gaussian sampling method to estimate the local elevation. The method aims to account for the fact that each interpolation point actually represents an area.

Method

This method relies on a Gaussian filter to estimate the local elevation of a point. The computation steps at each interpolation point are:

1. Calculate distance of ED points to the current interpolation point.
2. Calculate a weight for each ED point based on its distance. For this, a normal distribution function with standard deviation equal to twice the interpolation grid cell size is used. The weighting function used is:

$$w_{ij} = \frac{1}{2\Delta g\sqrt{2\pi}} e^{-\frac{{d_{ij}}^2}{2(2\Delta g)^2}} \qquad \text{DS.5}$$

   Where $\boldsymbol{w_{ij}}$ is the weight of the j[th]ED point for the i[th] interpolation point, $\boldsymbol{d_{ij}}$ is the Euclidean distance in 2D space between those two points, and $\Delta \boldsymbol{g}$ is the minimum distance between interpolation points.

3. Do a weigted sum of the elevations $\boldsymbol{h_j}$ of the ED to find an estimation of the elevation $\boldsymbol{h_i}$ at the current interpolation point:

$$h_i = \frac{1}{\sum_j w_{ij}} \sum_j w_{ij} \times h_j \qquad \text{DS.6}$$

As shown in the figures below, the use of a Gaussian filter smoothens the elevation estimation in comparison to block sampling, where the elevation is estimated by averaging DE data within a rectangular block around each interpolation point, thus introducing a strong sampling effect.

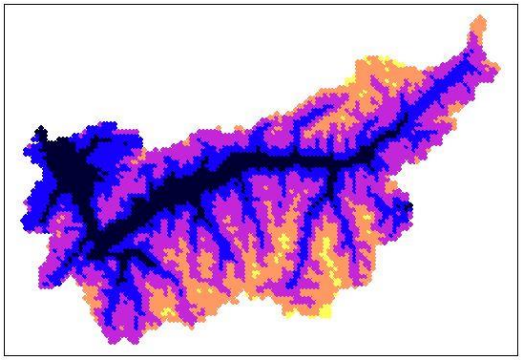

***Figure 5**. Elevation estimation by block sampling*

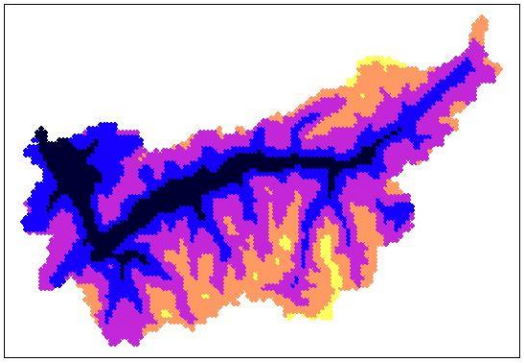

***Figure 6**. Elevation estimation with a gaussian filter*

Because the weighting function uses a standard deviation that is independent from the resolution of the DE data, the method will not work if DE points are much sparser than interpolation points: indeed, the weights rapidly fall to zero for all DE points further than $\mathbf{2}\Delta\boldsymbol{g}$ from the interpolation point.

Chapter 8. Aggregation

Once the measured time series are interpolated at each interpolation point, the data is aggregated: one time series is computed for each catchment. Each aggregated time series is computed by averaging the data of all interpolation points situated within the boundaries of each catchment (Figure 7). Each aggregated time series is associated to an "interpolated station" located at the center of gravity of the corresponding polygon.

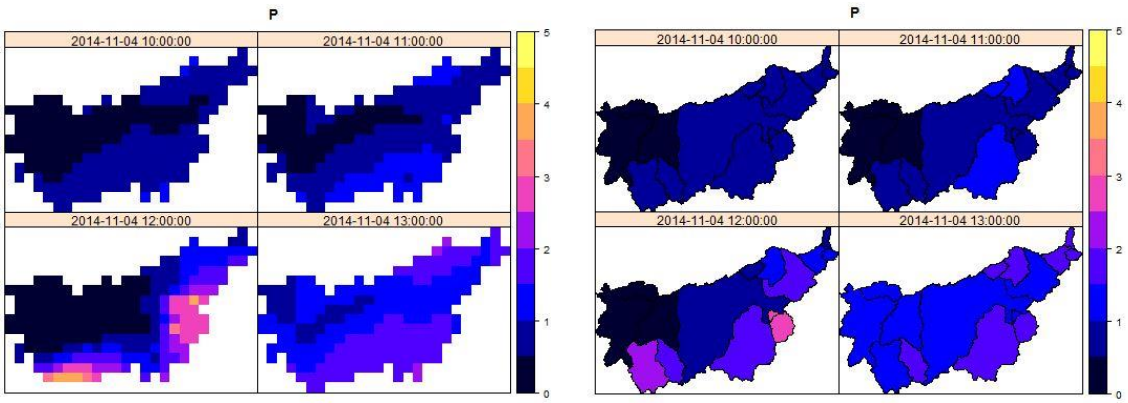

***Figure 7**. Mean aggregation of the interpolated temperature, by polygon. (left: interpolated data, right: aggregation of interpolated data)*

Bibliography

Anderson, T.W., Darling, D.A., 1954. A Test of Goodness of Fit. J. Am. Stat. Assoc. 49, 765–769. doi:10.1080/01621459.1954.10501232

Barua, S., Ng, A.W.M., Perera, B.J.C., 2011. Comparative Evaluation of Drought Indexes: Case Study on the Yarra River Catchment in Australia. J. Water Resour. Plan. Manag. 137, 215–226. doi:10.1061/(ASCE)WR.1943-5452.0000105

Best, D.J., Roberts, D.E., 1975. Algorithm AS 89: The Upper Tail Probabilities of Spearman's Rho. Appl. Stat. 24, 377. doi:10.2307/2347111

Conover, W.J., 1999. Practical nonparametric statistics, 3. ed. ed, Wiley series in probability and statistics Applied probability and statistics section. Wiley, New York, NY.

Cox, D.R., Stuart, A., 1955. Some quick sign test for trend in location and dispersion. Biometrika 42, 80–95. doi:10.1093/biomet/42.1-2.80

Craparo, R.M., 2007. Significance level, in: Encyclopedia of Measurement and Statistics 3, A Sage Reference Publication. Salkind, Neil J., pp. 889–891.

Foehn, A., García Hernández, J., Roquier, B., Paredes Arquiola, J., 2016. RS MINERVE – User's manual v2.4.

García Hernández, J., Paredes Arquiola, J., Foehn, A., Roquier, B., 2016. RS MINERVE - Technical Manual v2.5.

Gilbert, R.O., 1987. Statistical methods for environmental pollution monitoring. Wiley, New York.

Greenwood, J.A., Landwehr, J.M., Matalas, N.C., Wallis, J.R., 1979. Probability weighted moments: Definition and relation to parameters of several distributions expressable in inverse form. Water Resour. Res. 15, 1049–1054. doi:10.1029/WR015i005p01049

Hamed, K.H., Ramachandra Rao, A., 1998. A modified Mann-Kendall trend test for autocorrelated data. J. Hydrol. 204, 182–196. doi:10.1016/S0022-1694(97)00125-X

Heim, R.R., 2002. A Review of Twentieth-Century Drought Indices Used in the United States. Bull. Am. Meteorol. Soc. 83, 1149–1165. doi:10.1175/1520-0477(2002)083<1149:AROTDI>2.3.CO;2

Hiemstra, P.H., Pebesma, E.J., Twenhöfel, C.J.W., Heuvelink, G.B.M., 2009. Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network. Comput. Geosci. 35, 1711–1721. doi:10.1016/j.cageo.2008.10.011

Hipel, K.W., McLeod, A.I., 1994. Time series modelling of water resources and environmental systems. Elsevier, Amsterdam; New York.

Hirsch, R.M., Slack, J.R., Smith, R.A., 1982. Techniques of trend analysis for monthly water quality data. Water Resour. Res. 18, 107–121. doi:10.1029/WR018i001p00107

Hollander, M., Wolfe, D.A., Chicken, E., 2014. Nonparametric statistical methods, 3. ed. ed, Wiley series in probability and statistics. Wiley, Hoboken, NJ.

Hosking, J.R.M., 1990. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. J. R. Stat. Soc. Ser. B Methodol. 52, 105–124.

Hosking, J.R.M., Wallis, J.R., Wood, E.F., 1985. Estimation of the Generalized Extreme-Value Distribution by the Method of Probability-Weighted Moments. Technometrics 27, 251. doi:10.2307/1269706

Kendall, M.G., 1975. Rank correlation methods, 4. ed., 2. impr. ed. Griffin, London.

Keyantash, J.A., Dracup, J.A., 2002. The quantification of drought: an evaluation of drought indices. Bull. Am. Meteorol. Soc. 1167–1180.

Kottegoda, N.T., Rosso, R., 1997. Statistics, probability, and reliability for civil and environmental engineers. McGraw-Hill, New York, NY.

Lehmann, E.L., D'Abrera, H.J.M., 2006. Nonparametrics: statistical methods based on ranks, Rev. 1. ed. ed. Springer, New York, NY.

Libiseller, C., Grimvall, A., 2002. Performance of partial Mann-Kendall tests for trend detection in the presence of covariates. Environmetrics 13, 71–84. doi:10.1002/env.507

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometrica 13, 245. doi:10.2307/1907187

Mann, H.B., Wald, A., 1942. On the Choice of the Number of Class Intervals in the Application of the Chi Square Test. Ann. Math. Stat. 13, 306–317. doi:10.1214/aoms/1177731569

Matheron, G., 1971. The Theory of Regionalized Variables and Its Applications, École Nationale Supérieure des Mines de Paris. ed, Les cahiers du Centre de morphologie mathématique de Fontainebleau.

McCuen, R.H., 2003. Modeling hydrologic change: statistical methods. Lewis Publ, Boca Raton, Fla.

McKee, T.B., Doesken, N.J., Kliest, J., 1993. The relationship of drought frequency and duration to time scales. Presented at the Proceedings of the 8th Conference of Applied Climatology, American Meterological Society, Anaheim, CA, pp. 179–184.

McLeod, A.I., Hipel, R.W., Bodo, B. A, 1991. Trend analysis methodology for water quality time series. Environmetrics 2(2), 169–200.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30, 683–691. doi:10.1016/j.cageo.2004.03.012

Pettitt, A.N., 1979. A Non-Parametric Approach to the Change-Point Problem. Appl. Stat. 28, 126. doi:10.2307/2346729

R Development Core Team, 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Sen, P.K., 1968. Estimates of the Regression Coefficient Based on Kendall's Tau. J. Am. Stat. Assoc. 63, 1379–1389. doi:10.1080/01621459.1968.10480934

Smirnov, N., 1948. Table for Estimating the Goodness of Fit of Empirical Distributions. Ann. Math. Stat. 19, 279–281. doi:10.1214/aoms/1177730256

Sneyers, R., 1990. On the statistical analysis of series of observations, Technical note / World Meteorological Organization. Secretariat of the World Meteorological Organization, Geneva.

Sproull, N.L., 2002. Handbook of research methods: a guide for practitioners and students in the social sciences, 2. ed.; 1. Paperback ed. ed. Scarecrow, Lanham, Md.

Stedinger, J.R., Vogel, R.M., Foufoula-Georgiou, E., 1993. Frequency analysis of extreme events, in: Handbook of Hydrology. David R. Maidment, New York, pp. 18.1–18.66.

Theil, H., 1992. A Rank-Invariant Method of Linear and Polynomial Regression Analysis, in: Raj, B., Koerts, J. (Eds.), Henri Theil's Contributions to Economics and Econometrics. Springer Netherlands, Dordrecht, pp. 345–381.

Travaglini, E., Fluixá-Sanmartín, J., Alesina, S., Foehn, A., García Hernández, J., 2016. TeREsA – User Manual v1.1.

Van Rooy, M.P., 1965. A rainfall anomaly index independent of time and space. Notos 14, 43–48.

Vogt, J., Carrao, H., Spinoni, J., Drought Team, n.d. How to define Drought Hazard? Which global datasets could be used?

von Storch, H., 1999. Misuses of Statistical Analysis in Climate Research, in: von Storch, H., Navarra, A. (Eds.), Analysis of Climate Variability. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 11–26.

Wilcox, R.R., 2001. Fundamentals of modern statistical methods: substantially improving power and accuracy. Springer, New York, NY.

Wilhite, D.A., Glantz, M.H., 1985. Understanding: the Drought Phenomenon: The Role of Definitions. Water Int. 10, 111–120. doi:10.1080/02508068508686328

Williams, C.A., 1950. On the Choice of the Number and Width of Classes for the Chi-Square Test of Goodness of Fit. J. Am. Stat. Assoc. 45, 77. doi:10.2307/2280429

World Meteorological Organization, 2012. Standardized Precipitation Index - User Guide. World Meteorological Organization, Geneva, Switzerland.

World Meteorological Organization, 2008. Guide to hydrological practices. World Meteorological Organization, Geneva, Switzerland.

World Meteorological Organization, 1975. Drought and agriculture: report of the CAgM working group on the assessment of drought, Technical Note. Secretariat of the World Meteorological Organization, Geneva.

Yue, S., Pilon, P., Cavadias, G., 2002a. Power of the Mann–Kendall and Spearman's rho tests for detecting monotonic trends in hydrological series. J. Hydrol. 259, 254–271. doi:10.1016/S0022-1694(01)00594-7

Yue, S., Pilon, P., Phinney, B., Cavadias, G., 2002b. The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrol. Process. 16, 1807–1829. doi:10.1002/hyp.1095

Zargar, A., Sadiq, R., Naser, B., Khan, F.I., 2011. A review of drought indices. Environ. Rev. 19, 333–349. doi:10.1139/a11-013

Acknowledgments

TeREsA is developed by the research center **CREALP** with the collaboration of the **Universitat Politècnica de València** and the **Hydro10** Association. We would like to thank sincerely these organizations for their support. In addition, we would like to express our gratitude to the people who have personally contributed to the improvement of the program and its documentation, in particular: Stéphane Micheloud[3], Nestor Lerma[4], Edgar Belda Ibáñez[4], Alex Dionisio-Calado[3] and Matthew Moy-de-Vitry[3].

Our gratitude also goes to the developers of all the complementary R packages used in TeREsA, listed hereafter.

- ✓ automap
- ✓ bitops
- ✓ chron
- ✓ class
- ✓ colorspace
- ✓ data.table
- ✓ devtools
- ✓ dichromat
- ✓ digest
- ✓ evaluate
- ✓ e1071
- ✓ fBasics
- ✓ fExtremes
- ✓ fGarch
- ✓ fitdistrplus
- ✓ fTrading
- ✓ FNN
- ✓ formatR
- ✓ geosphere
- ✓ ggmap
- ✓ ggplot2
- ✓ gridBase
- ✓ gridExtra
- ✓ gstat
- ✓ gtable
- ✓ highr
- ✓ hydroTSM
- ✓ intervals
- ✓ jpeg
- ✓ knitr
- ✓ labeling
- ✓ latticeExtra
- ✓ lmomco
- ✓ magrittr
- ✓ manipulate
- ✓ mapproj
- ✓ maps
- ✓ maptools
- ✓ markdown
- ✓ mime
- ✓ munsell
- ✓ ncdf4
- ✓ PearsonDS
- ✓ plyr
- ✓ png
- ✓ pracma
- ✓ proto
- ✓ randtests
- ✓ raster
- ✓ RColorBrewer
- ✓ Rcpp
- ✓ Rcurl
- ✓ reshape
- ✓ reshape2
- ✓ rgdal
- ✓ rgeos
- ✓ rgl
- ✓ RgoogleMaps
- ✓ rjson
- ✓ RJSONIO
- ✓ rkt
- ✓ RSAGA
- ✓ scales
- ✓ shapefiles
- ✓ signal
- ✓ sm
- ✓ sp
- ✓ spacetime
- ✓ SPEI
- ✓ stringi
- ✓ stringr
- ✓ svDialogs
- ✓ svGUI
- ✓ tkrplot
- ✓ trend
- ✓ timeDate
- ✓ timeSeries
- ✓ wq
- ✓ XML
- ✓ xts
- ✓ zoo

---

[3] Centre de recherche sur l'environnement alpin - CREALP

[4] Universitat Politècnica de València – UPV

TeREsA

CREALP - Centre de recherche sur l'environnement alpin

Rue de l'Industrie 45
CH-1951 Sion

Tél. +41 (0) 27 607 11 80
Fax +41 (0) 27 607 11 94
crealp@crealp.vs.ch