Climate Data and Monitoring
WCDMP-No. 86

TENTH SEMINAR FOR HOMOGENIZATION AND QUALITY CONTROL IN CLIMATOLOGICAL DATABASES

AND

FIFTH CONFERENCE ON SPATIAL INTERPOLATION TECHNIQUES IN CLIMATOLOGY AND METEOROLOGY

(Budapest, Hungary, 12–14 October 2020, online)

WEATHER CLIMATE WATER

WORLD
METEOROLOGICAL
ORGANIZATION

Chair, Publications Board
World Meteorological Organization (WMO)
7 bis avenue de la Paix
P.O. Box 2300
CH-1211 Geneva 2, Switzerland

Tel.: +41 (0) 22 730 84 03
Fax: +41 (0) 22 730 81 17
Email: Publications@wmo.int

TENTH SEMINAR

FOR HOMOGENIZATION AND QUALITY CONTROL IN CLIMATOLOGICAL DATABASES

AND

FIFTH CONFERENCE

ON SPATIAL INTERPOLATION TECHNIQUES IN CLIMATOLOGY AND METEOROLOGY

12 – 14 October 2020

Organized by the Hungarian Meteorological Service (OMSZ)

Supported by WMO, OMSZ and EUMETNET

Edited by Mónika Lakatos, Lilla Hoffmann, Andrea Kircsi and Tamás Szentimrey

PREFACE

The Hungarian Meteorological Service organized the jubilee 10th Seminar for Homogenization and Quality Control and 5th Conference on Spatial Interpolation Techniques in Climatology and Meteorology virtually between 12 and 14 October 2020.
Sandor Szalai the head of Climate and Agrometeorological Department of OMSZ initiated the Seminar Series in 1996 with the support of the WMO and under the umbrella of the EUMETNET/ECSN (European Meteorological Network/ European Climate Support Network). Other key person is Tamás Szentimrey from the beginnings who have prepared the scientific programs of the Seminars and have lectured about theoretical background of homogenization and interpolation on the meetings.
The basic questions at starts were the spreading of homogenization methods and the overall use of homogenized time series in climate change studies. The general view has been changed since then. Homogenization became a necessary element of the data management procedures.
The WMO CCl set up team to support quality control and homogenization activities at NMHSs. The main task of the Task Team on Homogenization (TT HOM) was to provide guidance to Members on methodologies, standards and software required for quality control and homogenization of long term climate time-series. WMO Guidelines on Homogenisation (WMO-No. 1245) were published in 2020 and are accessible in English, French and Spanish languages through WMO's electronic library (https://public.wmo.int/en/resources/library). Certainly, Seminars took a major part of the initiation and completion of this guidance, as these occasions created great opportunity for information exchange amongst the homogenization community.
The Homogenization Seminars have been organized together with the Interpolation Conferences since 2004. The reason for joining them was the strong connection between topics, as the homogenization and quality control procedures apply spatial statistics and interpolation techniques for the spatial comparison of the data as well. Besides these the reconstruction of meteorological fields and construction of gridded databases require applying spatial interpolation methods. In addition, the spatial interpolation procedures (e.g. gridding) need homogeneous data series with high quality.
Homogenization of climate data series and spatial interpolation of climate data play an ever-growing role in the meteorology and climatology. The usage of high quality data as an important issue appeared on the European and global level officially.
We hope, that the Seminar series filled a major part in dissemination and development of homogenization methods not only in Europe, but worldwide.

The Organizers

CONTENTS

MATHEMATICAL QUESTIONS OF HOMOGENIZATION AND SUMMARY OF MASH

Tamás Szentimrey

Varimax Limited Partnership, Budapest, Hungary
szentimrey.t@gmail.com

Abstract

There are several methods and software for the homogenization of climate data series but there is not any exact mathematical theory of the homogenization. At the examinations mainly the physical experiences are considered while the mathematical formulation of the problems is neglected in general. Moreover occasionally there are some mathematical statements at the description of the methods in the papers – e. g. capability to adjust the higher order moments – but without any proof and this way is contrary to the mathematical conventions of course. As we see the basic problem of the homogenization is the unreasonable dominance of the practical procedures over the theory and it is the main obstacle of the progress. Therefore we try to formulate some questions of homogenization in accordance with the mathematical conventions. The planned topics to be discussed are as follows.

- The mathematical definition of the inhomogeneity and the aim of homogenization. It is necessary to clarify that the homogenization of climate data series is a distribution problem instead of a regression one.
- Relation of monthly and daily data series homogenization.
- Mathematical overview on the methodology of spatial comparison of series, inhomogeneity detection, adjustment of series in accordance with the publication of WMO Guidelines on Homogenisation (2020).
- Relation of theoretical evaluation and benchmark for methods, validation statistics.

The earlier versions of our method MASH (Multiple Analysis of Series for Homogenization; Szentimrey) were developed formerly at the Hungarian Meteorological Service. These procedures aimed to homogenize the daily and monthly data series in the mean i.e. the first order moment. The new version MASHv4.01 has been developed for joint homogenization of mean and standard deviation using some mathematical results. Theoretically in case of normal distribution the homogenization of mean and standard deviation is sufficient since if the first two moments are homogenous then the higher order moments are also homogeneous. An interactive automatic algorithm also was developed in this new version in order to make the homogenization easier for the users. We will present a summary of the software MASH where our intention was to develop a flexible, interactive automatic, artificial intelligence (AI) system that simulates the human intelligence and mimics the human analysis on the basis of advanced mathematics. We finish the paper with some comments connected to the fact that during the seminar we were obliged to express our skepsis on the credibility of the MULTITEST benchmark results to the authors.

1. INTRODUCTION

I retired from the Hungarian Meteorological Service two years ago but I continue my activity in my VARIMAX Limited Partnership. This activity includes the development of advanced mathematics for meteorology as well as the development of efficient software on the basis of the mathematical results.
Concerning our topic we have the following question. What is the mathematics of homogenization in meteorology? There are several methods and software for the homogenization of climate data series but unfortunately there does not exist any exact well elaborated mathematical theory of this problem. At the climatological examinations mainly the physical experiences are dominated while the mathematical formulation of the problems is neglected in general. We do not argue the importance of the physical aspects but the applied not too advanced mathematics is in contrast with the fact that the methods are declared to be based on the mathematical statistics. Moreover often there are false mathematical statements at the description of the methods in the papers – e.g. capability to adjust the higher order moments – without any proof and this way is contrary to the mathematical conventions of course. We see the basic problem of the homogenization is the unreasonable dominance of the practical procedures over the theory and it is the main obstacle of the progress. As a consequence of this practice the exact evaluation

of the methods is also very problematic or rather it is unrealistic and the progress of the homogenization research activity is doubtful.

In addition also some miracle waiting can be observed namely the Artificial intelligence (AI) and Big Data technology will solve these problems without mathematics. However it is a fake news or bullshit since these procedures need much more mathematics than the traditional ones. “There is no royal road!” (Archimedes) We note that our software MASH was also developed as an interactive automatic, artificial intelligence (AI) system that simulates the human intelligence and mimics the human analysis on the basis of advanced mathematics.

In the paper we try to provide a general approach for the mathematical formulation of homogenization in accordance with the mathematical conventions. We believe the correct mathematical principles can promote understanding and clarifying the questions of homogenization in climatology.

A remarkable good news is the publication of WMO Guidelines on Homogenisation (WMO-No. 1245). The World Meteorological Organization (WMO) established a Task Team on Homogenization (TT-Hom) which operated between 2014 and 2018. Its primary output was to produce a guidance document on monthly climate data homogenization, which was published in 2020. The guidance has five chapters, and the Chapter 5 includes the most important theoretical, mathematical questions in connection of homogenization of mean of monthly series (*Fig. 1*). I was a member of the Task Team therefore allow me to present and discuss these questions in details (Section 4).

CHAPTER 5. THEORETICAL BACKGROUND OF HOMOGENIZATION…………..… 40

5.1 General structure of the additive spatio-temporal models…………………….... 40

5.2 Methodology for comparison of series in the case of an additive model……… 42

5.3 Methodology for breakpoint (changepoint) detection……………………….….... 43

5.3.1 Breakpoint detection based on maximum likelihood estimation.... 43

5.3.2 Breakpoint detection based on hypothesis testing……………….. 43

5.3.3 Attribution of the detected breakpoints for the candidate series… 44

5.4 Methodology for adjustment of series………………………………………… 44

5.5 Possibilities for evaluation and validation of methods………………………... 44

Fig. 1.: WMO Guidelines on Homogenization for monthly data. Content of Chapter 5.

In our conception the meteorological questions and topics cannot be treated separately. Therefore we present a block diagram (*Fig. 2*) to illustrate the possible connection between various important meteorological topics. The software MASH (Multiple Analysis of Series for Homogenization; *Szentimrey*, 1999,2014,2017) and MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; *Szentimrey* and *Bihari*, 2014) were developed by us. These software were applied also in CARPATCLIM project (*Szentimrey et al.*, 2012a,b). We plan to share the new versions MASHv4.01 and MISHv2.01 on the website of VARIMAX this year. Our paper is summary of our conception on homogenization and of the method MASH.

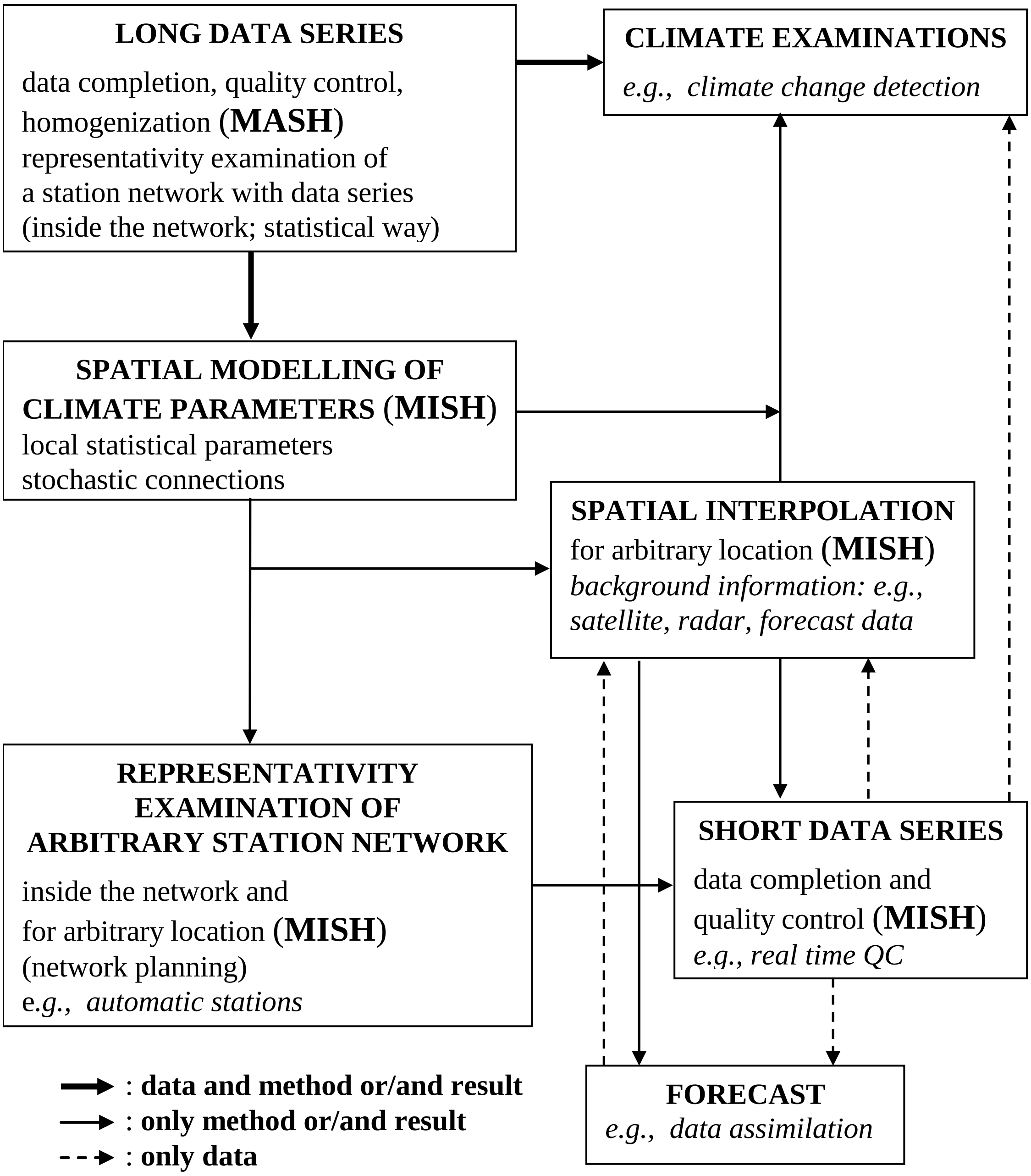


Fig. 2.: Block diagram for the possible connection between various basic meteorological topics and systems.

2. MATHEMATICAL FORMULATION OF CLIMATE DATA HOMOGENIZATION

Unfortunately the exact theoretical, mathematical formulation of the problem of homogenization is neglected at the meteorological studies in general. Therefore we try to formulate this problem in accordance with the mathematical conventions. First of all we emphasize that the homogenization is a distribution problem and not a regression one.

2.1 GENERAL MATHEMATICAL FORMULATION

Notation

Let us assume we have daily or monthly climate data series:

$Y_1(t)\ (t=1,2,..,n)$: candidate time series of the new observing system.

$Y_2(t)\ (t=1,2,..,n)$: candidate time series of the old observing system.

$1 \le T < n$: change-point, series $Y_2(t)\ (t=1,2,..,T)$ can be used before

and series $Y_1(t)\ (t=T+1,..,n)$ can be used after the change-point.

The appropriate theoretical cumulative distribution (CDF) functions are:

$$F_{1,t}(y) = \mathrm{P}(Y_1(t) < y)\ ,\quad F_{2,t}(y) = \mathrm{P}(Y_2(t) < y)\quad y \in (-\infty,\infty)\ ,\ t=1,2,..,n$$

It is very important to remark that as a consequence of some natural changes - e.g. annual cycle, climate change - the series of distribution functions $F_{1,t}(y)$, $F_{2,t}(y)$ $(t=1,2,..,n)$ may change in time! In the statistical climatology the climate change is equivalent with the changing probability of the meteorological events. The inhomogeneity of data series can be defined on the basis of the distribution functions.

Definition 1

The merged series $Y_2(t)\ (t=1,2,..,T)$, $Y_1(t)\ (t=T+1,..,n)$ is inhomogeneous, if the identity of the distribution functions $F_{2,t}(y) \equiv F_{1,t}(y)\ (t=1,2,..,T)$ is not true.

Definition 2

The aim of the homogenization is the adjustment or correction of values $Y_2(t)\ (t=1,2,..,T)$ in order to have the adjusted values $Y_{1,2h}(t)\ (t=1,2,..,T)$ with the same distribution as the elements of series $Y_1(t)\ (t=1,2,..,T)$ have, i.e.:

$$\mathrm{P}(Y_{1,2h}(t) < y) = \mathrm{P}(Y_1(t) < y) = F_{1,t}(y) \qquad y \in (-\infty,\infty)\ ,\ t=1,2,..,T\ . \tag{1}$$

The formula (1) means the equality in distribution: $Y_{1,2h}(t) \overset{d}{=} Y_1(t)\ \ (t=1,2,..,T)$

Remark 1

Within the same climate area, if the variables $Y_1(t), Y_2(t)$ $(t=1,2,..,T)$ have identical distribution, i.e. $Y_2(t) \overset{d}{=} Y_1(t)$ $(t=1,2,..,T)$, then the merged series $Y_2(t)$ $(t=1,2,..,T)$, $Y_1(t)$ $(t=T+1,..,n)$ is homogeneous.

Theorem 1

Let us assume about the random variables Y_1, Y_2 and their distribution functions $F_1(y), F_2(y)$, that $\mathrm{P}(Y_j \in (a_j, b_j)) = 1$ and $F_j(y)$ is a strictly increasing continuous function on the interval (a_j, b_j) $(j=1,2)$. Then applying the transfer function $Y_{1,2h} = F_1^{-1}(F_2(Y_2))$ we obtain that the variable $Y_{1,2h}$ has the same distribution like Y_1 i.e. $\mathrm{P}(Y_{1,2h} < y) = \mathrm{P}(Y_1 < y) = F_1(y)$.

Definition 3

Transfer function: $F_{1,t}^{-1}(F_{2,t}(y))$ and quantile function: $F_{1,t}^{-1}(p)$.

Theoretical formulation of homogenization of $Y_2(t)$ $(t=1,2,..,T)$:

$Y_{1,2h}(t) = F_{1,t}^{-1}(F_{2,t}(Y_2(t)))$, then $\mathrm{P}(Y_{1,2h}(t) < y) = F_{1,t}(y)$.

Remark 2

The basis of the Quantile Matching methods can be integrated into the general theory. However these methods developed in practice mainly for daily data are very weak empiric methods. It is not real mathematics! These methods have good heuristics with poor mathematics.

2.2 ARISING MATHEMATICAL QUESTIONS TO BE SOLVED

Let us suppose the merged series is given that is,

$Y_2(t)$ $(t=1,2,..,T)$, $Y_1(t)$ $(t=T+1,..,n)$

In addition we suppose that the assumptions of the former theorem are fulfilled, consequently the theoretical adjustment or transfer formulas for the series elements are,

$$Y_{1,2h}(t) = F_{1,t}^{-1}(F_{2,t}(Y_2(t))) \quad (t=1,2,..,T) \tag{2}$$

However these transfer formulas are theoretical ones and if we want to apply them in the practice then a number of mathematical statistical estimation problems are arising. The most important problems are as follows.

- Estimation, detection of the change point(s) *T*.
- Estimation of the theoretical distribution functions $F_{1,t}(y)$, $F_{2,t}(y)$ $(t=1,2,..,T)$:
    - i, $F_{1,t}(y)$, $F_{2,t}(y)$ may change in time because of the climate change and the annual cycle, consequently the methodology of the use of the empirical distribution functions is very doubtful.
    - ii, There is no sample for $F_{1,t}(y)$ $(t=1,2,..,T)$ and $F_{2,t}(y)$ $(t=T+1,..,n)$ usually.

These mathematical problems are insolvable generally! Therefore only relative methods can be used with some model assumptions. In addition some simplifications are also necessary. Statistically speaking, some assumptions have to be made!

2.3 MATHEMATICAL FORMULATION FOR NORMAL DISTRIBUTION

The homogenization problem is very complicated in general case however in case of normal distribution a much simpler mathematical formula can be obtained. We emphasize that the normal distribution is a special case but it is basic one in the mathematical statistics as well as in the meteorology. For example the normal distribution model can be accepted for the temperature variables in general.

Theorem 2

Let us assume the data series have normal distribution that is,

$$Y_1(t) \in N\left(E_1(t), D_1(t)\right), \quad Y_2(t) \in N\left(E_2(t), D_2(t)\right) \quad \left(t = 1,2,..,n\right),$$

where $\mathrm{E}\left(Y_1(t)\right) = E_1(t)$, $\mathrm{E}\left(Y_2(t)\right) = E_2(t)$ are the means or expected values and $\mathrm{D}\left(Y_1(t)\right) = D_1(t)$, $\mathrm{D}\left(Y_2(t)\right) = D_2(t)$ are the standard deviations.

Then the transfer formula of homogenization:

$$Y_{1,2h}(t) = F_{1,t}^{-1}\left(F_{2,t}\left(Y_2(t)\right)\right) = E_1(t) + \frac{D_1(t)}{D_2(t)}\left(Y_2(t) - E_2(t)\right) \quad \left(t = 1,2,..,T\right)$$

2.4 MATHEMATICAL QUESTIONS IN CASE OF NORMAL DISTRIBUTION TO BE SOLVED

In case of normal distribution according to the *Theorem 2* we have a much simpler transfer formula for adjustment than the general form (2), that is,

$$Y_{1,2h}(t) = E_1(t) + \frac{D_1(t)}{D_2(t)}\left(Y_2(t) - E_2(t)\right) \quad \left(t = 1,2,..,T\right) \tag{3}$$

This formula is a simple linear one that means if the data series have normal distribution it is sufficient to homogenize the means and standard deviations only that is equivalent with the homogenization of the first two moments. We emphasize that the normal distribution is a basic model in the mathematical statistics as well as in the meteorology and there is no "tail distribution" problem at this important distribution according to the *Theorem 2*! At the normal distribution if the means and standard deviations are homogenous then the higher order moments are also homogeneous and there is not any inhomogeneity in the tails of the distributions. It is in contrast with the popular assumption based on parallel measurements as it is very likely the inhomogeneity in the tails of the distributions at the daily data series. As regards the parallel measurements a mathematical examination for them will be presented at Section 2.5.

Returning to the formula (3) although it is much simpler than (2), there are still a number of mathematical statistical estimation problems to be solved as follows.

- Estimation, detection of the change point(s) *T*.
- Estimation of the statistical parameters $E_1(t), D_1(t), E_2(t), D_2(t)$ $\left(t = 1,2,..,T\right)$:

  i, $E_1(t), D_1(t), E_2(t), D_2(t)$ may change in time because of the climate change and the annual cycle.

  ii, There is no sample for $E_1(t)$, $D_1(t)$ $\left(t = 1,..,T\right)$ and $E_2(t), D_2(t)$ $\left(t = T+1,..,n\right)$ usually.

However these mathematical problems are still very complicated! Therefore only relative methods can be used with some model assumptions. In addition some simplifications are also necessary.

3. RELATION OF DAILY AND MONTHLY HOMOGENIZATION

Practically the theme of homogenization can be divided into two subgroups, such as monthly and daily data series homogenization. These subjects are in strong connection with each other of course, for example the monthly results can be used for the homogenization of daily data.

3.1 THE GENERAL STRUCTURE OF DAILY DATA HOMOGENIZATION

If we have daily data series the general way of homogenization is,

- calculation of monthly series,
- homogenization of monthly series taking advantage of the larger signal to noise ratio,
- homogenization of daily series using the detected monthly inhomogeneities.

So we have the question how can we use the valuable information of detected monthly inhomogeneities for the daily data homogenization?

3.2 A POPULAR PROCEDURE E.G. THE VARIABLE CORRECTION METHODS

The typical steps of the procedure are as follows.

1. Homogenization of monthly series:
   Break points detection, adjustment in the first moment (mean (E)).
   Assumption: homogeneity of higher order moments (e.g. standard deviation (D)).
2. Homogenization of daily series:
   There is a trial to homogenize also in higher order moments.
   (Quantile Matching, Spline methods)
   The used monthly information are only the detected break points.

However the following questions are arising at this procedure:

- Is it adequate model that we have inhomogeneity in higher moments only at daily series but not at monthly ones? Can this model be accepted according to the probability theory? However, this model is false, e.g. it can be proved (*Szentimrey*, 2017b), if there is a common inhomogeneity in the standard deviation (D) of daily data, we may have the same inhomogeneity in monthly data.
- Why are not used the monthly adjustment factors for daily homogenization? It seems to lose some valuable information obtained during the monthly homogenization.

3.3 AN ALTERNATIVE PROCEDURE DEVELOPED IN MASH (SECTION 5 AND (*SZENTIMREY,* 2017B).)

We suggest an alternative procedure to homogenize both the daily and the monthly series.
The steps of the procedure in case of quasi normal distribution (additive model, e.g. temperature) are as follows.

1. Homogenization of monthly series:
   Break points detection, calculation of the adjustment factors for the first two moments and homogenization of mean (E) and standard deviation (D). The adjustment is based on the transfer formula (3).
   Assumption: homogeneity of higher order moments. This assumption is always right in case of normal distribution according to *Theorem 2.*
2. Homogenization of daily series:
   Homogenization of mean (E) and standard deviation (D) on the basis of the monthly results. The used monthly information are the break points and the monthly adjustment factors of the mean and standard deviation. The adjustment is based on the transfer formula (3). If the daily data are normally distributed then there is no inhomogeneity in the higher order moments according to *Theorem2.*

In case of quasi lognormal distribution (multiplicative model, e.g. precipitation) also the above procedure can be applied for the data obtained by certain transformation based on logarithmization.

4. OVERVIEW ON HOMOGENIZATION OF MEAN OF MONTHLY SERIES

This section considers some various theoretical aspects of monthly series homogenization.
In the practice the monthly series are homogenized in the mean mostly. The aim of these homogenization procedures is to detect the inhomogeneities of mean and to adjust the series.
In connection with the such type of homogenization methods we have to give solutions for the following mathematical problems: relative models, statistical spatiotemporal modelling of the series, methodology for comparison of series, break point (changepoint) and outlier detection, methodology for adjustment of series, quality control procedures, missing data completion, usage of metadata, relation of daily and monthly homogenization, manual versus automatic methods, evaluation of methods (theoretical, benchmark).
In practice there are absolute and relative methods applied for homogenization. However the main problem of the application of absolute methods is that the separation of climate change signal and the inhomogeneity is essentially impossible. Relative methods can be applied if there are more station series given, which can be compared mutually. In this case the statistical spatiotemporal modelling of the series is a fundamental question. The adequate comparison, break point detection and adjustment procedures are depending on the chosen statistical model.
The following Sections 4.1-4.4 are related to the Chapter 5 of the WMO Guidelines on Homogenization (*WMO*, 2020).

4.1 GENERAL STRUCTURE OF ADDITIVE SPATIOTEMPORAL MODELS

If the data series are normally distributed (e.g. temperature) then the additive model can be used. In case of relative methods a general form of additive model for more monthly series belonging to the same month in a small climate region can be written as follows,

$$X_j(t) = \mu(t) + E_j + IH_j(t) + \varepsilon_j(t) \qquad \left(j = 1,2,\ldots,N \,;\; t = 1,2,\ldots,n \right) , \tag{4}$$

where $\mu(t)$ is the common and unknown climate change signal, E_j are the spatial expected values, $IH_j(t)$ are the inhomogeneity signals and $\varepsilon_j(t)$ are normal white noise series. The type of inhomogeneity $IH(t)$ is in general a 'step-like function' with unknown break points T and shifts $IH(T) - IH(T+1) \neq 0$, and $IH(n) = 0$ is assumed in general.
The normal distributed vector variables $\boldsymbol{\varepsilon}(t) = [\varepsilon_1(t),..,\varepsilon_N(t)]^{\mathrm{T}} \in N(\mathbf{0},\mathbf{C})$ $(t = 1,...,n)$ are totally independent in time. The spatial covariance matrix $\mathbf{C}$ describes the spatial structure of the series. If the data series are quasi lognormal distributed (e.g. precipitation) then the multiplicative model can be used that can be transformed into the additive one by certain logarithmic procedure.

4.2 METHODOLOGY FOR COMPARISON OF SERIES

The problem of comparison of series is related to the following questions: reference series creation, difference series constitution, multiple comparisons of series etc. This topic is very important for detection as well as for adjustment, because the efficient series comparison can increase both the significance and the power. The development of efficient comparison methods can be based on the examination of the spatial covariance structure of data series. The examined series $X_j(t)$ $(j = 1,...,N)$ have to be taken as candidate and reference series alike, furthermore the reference series are not assumed to be homogeneous at the correct examinations!

The main problem arises from the fact that the shape of climate change signal is unknown. Therefore so-called difference series are examined in order to filter out the climate change signal $\mu(t)$. The simple difference series between pairs are $Z(t) = X_j(t) - X_i(t)$. However the difference series constitution can be formulated in more general way as well. Assuming that $X_j(t)$ is the candidate series and the other ones are the reference series, then the difference series belonging to the candidate series can be constituted as,

$$Z_j(t) = X_j(t) - \sum_{i \neq j} \lambda_{ji} X_i(t) = IH_j(t) - \sum_{i \neq j} \lambda_{ji} IH_i(t) + \varepsilon_{Z_j}(t) \tag{5}$$

with condition of $\sum_{i \neq j} \lambda_{ji} = 1$ for the weighting factors. As a result of the last condition, the unknown climate change signal $\mu(t)$ has been filtered out. Consequently the inhomogeneities can be detected by the examination of the above difference series. In addition if we want to increase the signal to noise ratio in order to increase the power of detection then we have to minimize the variance of noise term $\varepsilon_{Z_j}(t)$.

The covariance matrix $\mathbf{C}$ uniquely determines the optimum weighting factors that minimize the variance, and the optimal difference series created in this manner can be applied efficiently for the detection and adjustment procedures (MASH, *Szentimrey*, 1999, 2014, 2017). We mention that in case of using the generalized-least-squares estimation for the unknown climate change signal $\mu(t)$, also the optimal difference series is obtained with minimal variance. We have to examine more difference series in order to separate the appropriate detected inhomogeneities for the candidate series. More difference series created without common reference series and with minimal variances can be defined as optimal difference series system (MASH).

4.3 METHODOLOGY FOR BREAKPOINT (CHANGEPOINT) DETECTION

One of the basic tasks of the homogenization is the examination of the difference series (5) in order to detect the break points and to attribute the appropriate ones for the candidate series.

The scheme of the breakpoint detection is as follows. Let $Z(t)$ be a difference series according to the formula (5), that is

$$Z(t) = IH_Z(t) + \varepsilon_Z(t) \qquad (t = 1,..,n), \tag{6}$$

where $IH_Z(t)$ is a mixed inhomogeneity of difference series $Z(t)$ with breakpoints. In general the number of breakpoints and their positions with sizes are unknown, furthermore $\varepsilon_Z(t)$ is a normal noise series.

Remark 3 (Outlier detection, QC)

The outlier detection is the quality control (QC) procedure for the monthly data. The outlier detection can be considered as a special part of break points detection, because an outlier is equivalent with two neighboring break points where their sizes are the same in absolute value but with opposite sign.

Returning to the detection procedures the basic types are the stepwise and the multiple break points detection. The stepwise procedure is repeated single-breakpoint detection. The more sophisticated multiple break points detection procedures were developed for joint estimation of the breakpoints. There may be different principles of the detection methods that are classical ways in the mathematical statistics.

4.3.1 Breakpoint detection based on Bayesian Approach

The methods based on Bayesian model selection (segmentation) are the penalized likelihood methods. Then joint maximum likelihood estimation is given for the break points assuming normal distribution of the difference series and using some penalty term. The reason of the penalty term is that the number of breakpoints is unknown. The methods may be different in the penalty terms or criterions e.g. Akaike criterion, Schwarz criterion, Caussinus-Lyazrhi criterion. The penalty terms depend on some 'a priori' probability of break at each time. The PRODIGE procedure (*Caussinus* and *Mestre*, 2004) based on the Caussinus-Lyazrhi criterion is an example for the penalized likelihood methods.

Theoretically this methodology could be also applied for the joint detection of the breakpoints of all the examined series. However the joint likelihood function assuming normal distribution depends on the inverse of the spatial covariance matrix what can cause complicated technical problems in case of larger networks.

4.3.2 Breakpoint detection based on Test of Hypothesis

Another way is to use test of hypothesis methods for the detection of breakpoints of the difference series. In accordance with the statistical conventions then the null hypothesis is the homogeneity. These methods also assume the normal distribution therefore the test statistics are derived from the t-type statistics in general. The significance and the power of the procedures can be defined according to the probabilities of type of errors. Type one error means detection of false breakpoint while type two error means neglecting some real breakpoint. The test statistics can be compared to the critical value that depends on the given significance level. In case of multiple test statistics the critical values can be calculated by Monte Carlo method.

Most of these methods are stepwise ones that is repeated single-breakpoint detection. However multiple breakpoint detection procedures also can be developed. The essence of this procedure is that between the neighboring detected break points the homogeneity can be accepted and between not neighboring detected break points the homogeneity cannot be accepted. In addition confidence intervals also can be given for the breakpoints that make possible automatic use of metadata. The method MASH (*Szentimrey*, 1999, 2014, 2017) is an example for this type of hypothesis procedure.

There is a favorable property of these methods that the result can be evaluated, validated by comparison of the test statistics before and after homogenization.

4.3.3 Attribution of the detected breakpoints for the candidate series

During the breakpoints detection procedures the difference series are examined with mixed inhomogeneity (5) potentially. If more difference series are examined for a candidate series then the mixed inhomogeneity is less problem but the attribution of the breakpoints for the candidate series is not a trivial task. Then a synthesis is necessary and the key question of the homogenization software is to develop automatic procedures for this attribution problem.

4.4 METHODOLOGY FOR ADJUSTMENT OF SERIES

Beside the detection another basic task of the homogenization is the adjustment of series. Calculation of the adjustment factors can be based on the examination of difference series for estimation of shifts at the detected break points. In general the methods use point estimation for the shifts at the detected break points.

There are methods that use the standard least squares technique after breakpoint detection procedure for joint estimation of the shifts of all the examined series. For example the method PRODIGE (*Caussinus* and *Mestre*, 2004). Probably the generalized least squares estimation technique based on spatial covariance matrix would be more efficient and it would be equivalent with the maximum likelihood estimation for the shifts in the case of normal distribution. Another way is that the calculation of the adjustment factors is based on some confidence intervals given for the shifts at the detected break points. For example the method MASH (*Szentimrey*, 1999, 2014, 2017). The confidence intervals given for the break points and shifts make possible also the automatic use of metadata.

5. SUMMARY OF SOFTWARE MASH

(Multiple Analysis of Series for Homogenization; *Szentimrey* 1999, 2008, 2014, 2017)

5.1 GENERAL COMMENTS

The new version MASHv4.01 has been developed for homogenization of daily and monthly series. The most important novelty of this version is the homogenization in standard deviation (*D*) beside the mean (*E*), see Section 3.3 and (*Szentimrey*, 2017b). The basic conception of the MASH system is the homogenization of monthly series derived from daily series. The procedures depend on the distribution of climate elements, and additive or multiplicative model can be used. Homogenization of daily series is based on the detected monthly inhomogeneities.

Quasi normal distribution (e.g. temperature)

Beside the monthly mean series another type of monthly series are also derived to estimate the inhomogeneity of standard deviation (*D*). These series can be homogenized in standard deviation (*D*) by multiplicative model and the monthly mean series can be adjusted with the estimated inhomogeneity of standard deviation (*D*). The adjusted monthly mean series can be homogenized in mean (*E*) by additive model.

Quasi lognormal distribution (e.g. precipitation)

Monthly mean or sum series can be homogenized by multiplicative model.

5.2 THE MOST IMPORTANT FEATURES OF MASH SYSTEM

Homogenization of monthly series:

- Relative homogeneity test procedure.
- Step by step iteration procedure: the role of series (candidate, reference) changes step by step in the course of the procedure.
- Interactive automatic, artificial intelligence (AI) system.
- Additive or multiplicative model can be used depending on the distribution.
- Including automatic Quality Control and missing data completion.
- Providing the homogeneity of the seasonal and annual series as well.
- Metadata (probable dates of break points) can be used automatically.
- The homogenization results and the metadata can be verified.

Homogenization of daily series:

- Based on the detected monthly inhomogeneities.
- Including automatic Quality Control and missing data completion for daily data.

5.3 REMARK ON AUTOMATION OF METHODS AND SOFTWARE

One of the fundamental questions of homogenization procedures is the relation of the manual, interactive and automatic methods. Because of the large station networks the development of automatic procedures is essential. However the aim of MASH is not the full automation and we also are skeptical in such an aspect. Our intention was to develop a flexible, interactive automatic, artificial intelligence (AI) system that simulates the human intelligence and mimics the human analysis on the basis of advanced mathematics. The mechanic, labor-intensive procedures are fully automated, moreover the operating process can be controlled simply and the accidental mistakes can be corrected interactively. The basic idea of this conception is controlling the results via the verification, hypothesis test tables generated automatically during the automatic procedures. As regards the automation the specialty of MASH software is the automatic usage of metadata furthermore automatic data quality control and missing data completion. The automatic quality control results also can be evaluated via the verification tables and can be modified interactively.

The elder version of MISH-MASH software can be downloaded from:
http://www.met.hu/en/omsz/rendezvenyek/homogenization_and_interpolation/software/

We plan to share the new version MASHv4.01 this year 2021.
Some results for joint homogenization of mean (E) and standard deviation (D) can be found in the paper (*Szentimrey*, 2017b).

6. POSSIBILITIES FOR EVALUATION, VALIDATION OF THE METHODS

6.1 THEORETICAL EVALUATION

If want to obtain a real image of the homogenization methods, then the theoretical evaluation of their mathematical basis is indispensable. It is a very serious scientific deficiency in the meteorology that no such overall study has been conducted to date.

6.2 BENCHMARK

Another possibility is a blind comparison and validation study for the homogenization methods. This practical way is much more popular in the meteorology than the theoretical one. Then the methods are tested on a realistic benchmark dataset. The benchmark contains simulated data with inserted inhomogeneity. Testing the methods on a generated benchmark dataset seems to be an objective validation procedure however we have to know also the limits of such type of examinations.
The interpretation of benchmark results is not a trivial problem, since these are depending on different factors, such as:

- tested methods (quality, manual, interactive, automatic),
- testing benchmark dataset (quality, adequacy),
- testers (skilled or unskilled),
- methodology of evaluation (validation statistics).

The creation of adequate benchmark dataset and the development of appropriate validation statistics are critical points and they need also strong theoretical mathematical background. “There is no royal road!” (Archimedes) In the following items some important, fundamental questions will be discussed about the relations of the above factors.

6.2.1 Relation of tested methods and testers

Because of the large station networks the development of automatic procedures is essential.
However the interactive automatic methods may be more efficient than the full automatic ones as a consequence of the human intuition and control. But how can we compare the manual,

interactive and automatic methods? The results of the manual and interactive methods depend also on the testers. The method or the user is tested if we evaluate the test results? The comparison of manual methods to automatic ones seems similar to the comparison of handmade and factory products. Moreover a manual time consuming method with a skilled tester or a full automatic method may have better results than an interactive automatic method with an unskilled tester. Unfortunately we had noticed such a problem at the MULTITEST benchmark project (*Domonkos at al.*, 2020; *Guijarro et al.*, 2017) where our method MASH was tested in the simple automatic mode of operation without using the interactive capability of this method. We indicated this problem to the authors many times however it was without any effect. Therefore during the seminar we expressed our skepsis on the credibility of the MULTITEST benchmark results to the authors. We remark that simple mechanic benchmark examinations may lead to false conclusions.

6.2.2 Relation of tested methods and testing benchmark datasets

The quality and adequacy of the benchmark datasets is a key question of the benchmark testing. The generated benchmark data should be similar to the real climate data. Advanced mathematics is needed for appropriate spatiotemporal modelling of the data. The inadequate, oversimplified models may lead false testing results. These simple models favour the similarly simple methods! There may be a further confusing problem namely the interrelationship between methods and benchmark data. Sometimes the development of methods are affected by the development of benchmark datasets or the methods are optimized on some benchmark data. There may be a problem if the methods and the benchmark datasets are based on the same hypothetic, unjustified assumptions. Then such methods can have unfair advantage over the others during the test procedure. For example the MULTITEST benchmark data (*Domonkos et al.,* 2020; *Guijarro et al.,* 2017) and the ACMANT method (*Domonkos,* 2017) are based on the same assumption that the annual cycle or the seasonal change of the inhomogeneity is a sinusoid. Thus the other methods developed without this hypothetic assumption were handicapped during the test. Moreover in my opinion, if a method developer is both the developer of the benchmark dataset and the tester of the methods then probably his method will be also the winner of the competition. “I only believe in statistics that I doctored myself.” (Churchill)

6.2.3 Methodology of evaluation, validation statistics

The so called “Hit rate, False alarm rate” statistics are often applied at the benchmark studies.
However we do not think that the statistics hit rate and false alarm rate would be adequate validation statistics for evaluation of the homogenization methods. These simple statistics are very trendy in the meteorology, but the validation is much more complex mathematical task, since the aim of the homogenization is not the precise break point detection. The key question is the similarity of the homogenized and the original clean series. For this comparison several validation statistics can be used, e.g. RMSE, CRMSE, trend statistics etc., but it would be recommended to develop other special test statistics too.

References

Caussinus, H, Mestre, O. 2004: Detection and correction of artificial shifts in climate series, Appl. Statist., 53, Part 3, pp. 405-425.

Domonkos, P., 2017: Time series homogenisation with optimal segmentation and ANOVA correction: past, present and future, Proceedings of the 9th Seminar for Homogenization and Quality Control in Climatological Databases and 4th Conference on Spatial Interpolation Techniques in Climatology and Meteorology (Ed. Szentimrey T, Lakatos M, Hoffmann L), Budapest, Hungary, 2017, WCDMP-No. 85, pp. 29–45.

Domonkos, P., Guijarro, Venema, V., Sigro, J., Brunet., M., 2020: Efficiency of time series homogenization: method comparison with 12 monthly temperature test datasets, Journal of Climate, Published-online: 29 Dec 2020, DOI 10.1175/JCLI-D-20-0611.1.

Guijarro, J.A., Lopez, J.A, Aguilar, E., Domonkos, P., Venema, V., Sigro, J., Brunet., M., 2017: Comparison of homogenization packages applied to monthly series of temperature and precipitation: the MULTITEST project, Proceedings of the 9th Seminar for Homogenization and Quality Control in Climatological Databases and 4th Conference on Spatial Interpolation Techniques in Climatology and Meteorology (Ed. Szentimrey T, Lakatos M, Hoffmann L), Budapest, Hungary, 2017, WCDMP-No. 85, pp. 46-62.

Szentimrey, T., 1999: Multiple Analysis of Series for Homogenization (MASH), Proceedings of the Second Seminar for Homogenization of Surface Climatological Data, Budapest, Hungary; WMO, WCDMP-No. 41, pp. 27-46.

Szentimrey, T., 2008: Development of MASH homogenization procedure for daily data. Proceedings of the Fifth Seminar for Homogenization and Quality Control in Climatological Databases, Budapest, 2006, WCDMP-No. 71, WMO/TD-NO. 1493, 123–130.

Szentimrey T, Lakatos M, Bihari Z, Kovacs T, Nemeth A, Szalai S, Auer I, Hiebl J, Milkovic J, Štěpánek P, Zahradnicek P, Tolasz R, Kilar P, Pyrc R, Limanowka D, Cheval S, Matei M, Kajaba P, Ivanakova G, Bochnicek O, Nejedlik P, Statsny P, Mihic D, Petrovic P, Savic T, Skrynyk O, Nabyvanets Y, Gnatiuk N, 2012a: Final report on quality control and data homogenization measures applied per country, including QC protocols and measures to determine the achieved increase in data quality. Carpatclim Project, Deliverable D1.12, http://www.carpatclim-eu.org/docs/deliverables/D1_12.pdf.

Szentimrey T, Lakatos M, Bihari Z, Kovacs T, Nemeth A, Szalai S, Auer I, Hiebl J, Milkovic J, Štěpánek P, Zahradnicek P, Tolasz R, Kilar P, Pyrc R, Limanowka D, Cheval S, Matei M, Kajaba P, Ivanakova G, Bochnicek O, Nejedlik P, Statsny P, Mihic D, Petrovic P, Savic T, Skrynyk O, Nabyvanets Y, Gnatiuk N, 2012b: Final report on the creation of national gridded datasets, per country. Carpatclim Project, Deliverable D2.9, http://www.carpatclim-eu.org/docs/deliverables/D2_9.pdf.

Szentimrey, T. 2013: Theoretical questions of daily data homogenization, Időjárás Vol. 117. No. 1, January-March 2013. pp. 113-122.

Szentimrey,T., Bihari, Z., 2014: Manual of interpolation software MISHv1.03, Hungarian Meteorological Service, p. 60.

Szentimrey, T., 2017a: Manual of homogenization software MASHv3.03, Hungarian Meteorological Service,p.71.

Szentimrey, T., 2017b: Some theoretical questions and development of MASH for homogenization of standard deviation, Proceedings of the 9th Seminar for Homogenization and Quality Control in Climatological Databases and 4th Conference on Spatial Interpolation Techniques in Climatology and Meteorology (Ed. Szentimrey T, Lakatos M, Hoffmann L), Budapest, Hungary, 2017, WCDMP-No. 85, pp. 63-73

Venema, V. K. C. et al., 2012: Benchmarking homogenization algorithms for monthly data, Climate of the Past, 8, 89-115

World Meteorological Organization (WMO), 2020: Guidelines on Homogenization, WMO-No. 1245, p. 63.

CAPACITY OF ACMANTv4 FOR HOMOGENIZING CLIMATIC DATASETS OF NATIONAL METEOROLOGICAL SERVICES

Peter Domonkos

Tortosa, Spain, dpeterfree@gmail.com

1. INTRODUCTION

National meteorological services usually possess large climatic databases, which are expected to be of high quality in all aspects including the temporal homogeneity of the data. The homogeneity of national and regional mean temperature data is particularly important, as the related climatic trends show the intensity of global warming signal. Nevertheless, the homogeneity is important for all climatic variables, in all spatial scales, and for the data of all kinds of temporal resolution. ACMANT homogenization method has been developed in the recent decade to solve the fast and accurate homogenization of large climatic datasets, and was used in World Meteorological Organization trainings in 2014 and 2015. The newest version ACMANTv4 can be applied for the homogenization of surface air temperature, precipitation total, relative humidity, sunshine duration, radiation, wind speed and atmospheric pressure time series of either daily or monthly resolution. ACMANTv4 is a fully automatic, relative homogenization method. In the homogenization of large datasets, the program selects the most appropriate set of partner series for each candidate series. The method can be applied for the joint homogenization of time series of different lengths, and it is characterised by high missing data tolerance.
The most important characteristic of a homogenization method is its accuracy. ACMANT has repeatedly produced excellent results in international method comparison tests. However, the role of efficiency test results seems to decline in the method selection, and ACMANT has rarely been used in the recent years. Its reasons will be discussed in the study.

2. PROPERTIES OF ACMANTv4

2.1. GENERAL PROPERTIES OF ACMANT

The development of ACMANT started during the European Action COST ES0601, Advances in Homogenisation Methods of Climate Series: an Integrated Approach (“HOME”, 2007-2011). It is a fully automatic method for the homogenization of climatic datasets in daily or monthly resolution. Its development was continuous in the last decade, and the most recent version ACMANTv4 is freely available from https://github.com/dpeterfree/ACMANT, together with the manual and scientific description of the method.
The ACMANT development is based on three main sources: i) earlier knowledge; ii) own ideas; iii) tests with benchmark datasets.
i) Earlier knowledge. It includes experts’ suggestions and experiences with other homogenization methods published in the past. The most important earlier work is the PRODIGE homogenization method (Caussinus and Mestre, 2004). Both its break detection method and correction method are inbuilt to ACMANT (ACMANT = Adapted Caussinus – Mestre Algorithm for homogenizing Networks of climatic Time series). Further examples are the adaptation of the optimal weighting of neighbor series in time series comparisons (Szentimrey, 2010) and the Vincent’s method to downscale correction terms from monthly to daily resolution (Vincent et al. 2002).

ii) Own ideas. The structure of ACMANT is rather complex. When it arrived to its fourth version, the length of the scientific description is 71 pages (Domonkos, 2020). Naturally, it includes several own ideas. One important idea is the bivariate detection, which is applied in ACMANT from its first version. The ACMANT bivariate detection searches breaks jointly in the series of the annual means and series of summer – winter differences. Its application is recommended for the homogenization of climatic elements which are observed in mid- or high latitudes and whose variations are related to radiation effects. Another important idea is the ensemble homogenization, which is applied in varied form in ACMANT. The risk of the occurrences of large errors for random effects can be notable reduced by ensemble homogenization, as large error generating random effects usually affect only a small part of the ensemble members.
iii) Tests with benchmark datasets. ACMANT is a complex homogenization procedure including empirically set parameters. The performance of such complex procedures cannot be concluded by clean mathematical deductions, but they need experimental control.
All these aspects of ACMANT provide its high performance, i.e. ACMANT often gives more accurate results than other homogenization methods.
Further favourable characteristics of ACMANT are: iv) Input time series may cover varied periods; v) Tolerance of high missing data ratio; vi) Several output options.
iv) Input time series may cover varied periods, and the automatic composite reference series creation in ACMANT always selects the one comprising the highest number of neighbor series possible (up to 30 neighbour series). If a section of the candidate series is not covered with that, further composite reference series are created comprising longer neighbor series to cover a larger part of the candidate series (and these steps are repeated until the break detection has been performed for all parts of the candidate series).
v) Tolerance of high missing data ratio. ACMANT needs an amount of monthly or daily data that equals approx. 10 years of continuous record. The first and last 5 years of the data must be in blocks of at least 25% compactness (i.e. 75% missing data ratio is allowed for these blocks), while the missing data ratio is unlimited between the first and last data blocks.
vi) Output options: User may choose the homogenized output completed with interpolated data to a pre-defined period, or completed only within the homogenized period, or data gaps left the same as in the input dataset. It is also possible to choose vertical output form, in which each homogenized value is written to a separate line.
ACMANTv4 cannot use metadata. I plan to include the automatic use of metadata in the following ACMANT version. Although the role of metadata in automatic homogenization procedures is usually not decisive, they may be important when the number of comparable time series or their spatial correlations are relatively low.

2.2. NOVELTIES OF ACMANTV4

The methodological differences between ACMANT and HOMER (Mestre et al. 2013) and the novelties in ACMANTv3 are presented by Domonkos and Coll (2017). More recently, ACMANTv4 has been published, and we summarize its novelties here.

- More kinds of climatic elements can be homogenized with ACMANTv4, they are: temperature, precipitation total, air humidity, wind speed, sunshine duration, radiation and atmospheric pressure. All these climatic elements can be homogenized either in daily or monthly resolution.
- The elimination of possible physical outlier values before the homogenization procedure is automatic. The homogenization procedure treats any value falling out of the physically acceptable range of values as it would be a missing data.
- ACMANTv4 can treat large datasets in one only run of its main program, up to 5000 time series. The software divides large datasets to smaller networks of 30 – 60 time series

belonging to zones of the same or nearly the same climate. In case of relatively low spatial correlations indicating higher spatial variability of climate, the automatically formed networks can be smaller.

- The accuracy of ACMANT has been improved more with some modifications in the algorithm. The two most important changes are a) Ensemble homogenization in the third (last) round of the homogenization procedure; and b) Inclusion of the weighted ANOVA model for the calculation of correction terms.
  a) In the second round of the procedure the ensemble homogenization is performed in the same way as in ACMANTv3, except that the significance threshold is set lighter by the modification of a parameter in the Caussinus – Lyazrhi criterion (see more details about the parameterization of the Caussinus – Lyazrhi criterion in ACMANT in Domonkos, 2020). The ANOVA correction (this time the simple, unweighted version) is applied to the annual series in each ensemble member. Then the ensemble minimum and ensemble average are calculated for each annual value of each time series. Note that the difference between the ensemble minimum and ensemble average tends to be larger for periods of time series where the break detection results show instability, therefore better estimations can be achieved using these two characteristics together than using only one of them. In the final homogenization round, 9 linear combinations of the ensemble minimum and ensemble average are taken to form 9 sets of pre-homogenized reference series to the ensemble homogenization of each candidate series. The parameterization of the linear combinations is empirical (Domonkos, 2020).
  b) The weighted ANOVA model differs from the common ANOVA correction model (Lindau and Venema, 2018) in a way that the spatial variation of climate is taken into consideration by weighting the neighbor series. Szentimrey (2010) proposes the use of ordinary kriging for the optimal weighting of neighbor series in the estimation of correction terms. However, the weighted ANOVA is the most time consuming routine in ACMANT, therefore the weighting of time series is simplified to the use of squared correlations, i.e. the same weighting algorithm is applied here as the traditional weighting in composite reference series (Peterson and Easterling, 1994), and applied only in the last round of the homogenization procedure (see also Domonkos, 2017,2020).
- Input data preparation has been made easier, as the justifying of time series to the same time period is no longer needed (see more details in the Manual of ACMANTv4).

2.3. HOMOGENIZATION OF VARIOUS CLIMATIC ELEMENTS WITH ACMANTV4

Relative humidity, sunshine duration, radiation, wind speed and atmospheric pressure are homogenized largely with the same algorithm of additional inhomogeneity model as which is applied to temperature homogenization. However, some small differences occur:

- For every climatic elements (also for temperature and precipitation), specific physical threshold values are set (e.g., 1 and 100 for relative humidity), which are not allowed to crossed. User can define own threshold values if he/she opts to use own defined thresholds (see Manual). If an input data falls out of the range of the physically accepted values, that value will be substituted with the missing data code. If a homogenized value falls out of this range, then that will be substituted with the closest acceptable value.
- In daily homogenization of sunshine duration and radiation (and precipitation) the monthly values are sums of the daily values, while for the other climatic elements the monthly values are averages of the daily values.

- In the homogenization of relative humidity, sunshine duration, radiation (and mean temperature and maximum temperature) the model seasonal change of inhomogeneity size is sinusoid with modes at the solstices when the data of mid- or high-latitude observing stations are homogenized. For wind speed and sea level pressure (and minimum temperature) the seasonal cycle is irregular, and this latter model is applied also for the other climatic elements when the data are collected from tropical observing sites. For atmospheric pressure at the station level, the seasonal change of inhomogeneity is considered zero.

All the differences relative to temperature homogenization are relatively small, and the unification of treating in the same way or almost the same way the inhomogeneities of different climatic elements may need reasoning. The homogenization of temperature time series is performed according to the additive inhomogeneity model, it is typical not only for ACMANT, but for most contemporary homogenization methods, first of all for monthly homogenization algorithms. The additive inhomogeneity model means that the deviation caused by the inhomogeneity is considered constant between two adjacent breaks, hence the optimal correction term is presumed to be independent from weather and climate anomalies. This model performs well for long-term temperature means, still adequately for seasonal and monthly means, while it is less accurate for the correction of daily temperature data.

Earlier I did not agree to use the additive inhomogeneity model for other climatic elements, for their differing probability distributions from that of temperature data. For instance, while temperature data have near Gaussian distribution, the distributions of relative humidity and sunshine duration are limited to relatively narrow ranges. The application of additive model easily can produce physically impossible daily values. However, the correction terms calculated for longer sections of time series still give good approaches for section means, as the change of weather and climate anomalies much less affect the long-term means than the individual daily values. As a consequence, after the post-correction of possible physical outliers, the additive inhomogeneity model included in ACMANT algorithm can be applied to the homogenization of several other climatic elements with similar efficiency to that of the temperature homogenization.

The use of additive inhomogeneity model in wind speed homogenization needs further discussion. The probability distribution of wind speed presents higher similarity to the probability distribution of precipitation total than to the probability distribution of temperature. In spite of this, the selection of the model inhomogeneity type is not straightforward, as true wind speed inhomogeneities are often not fully additive, either multiplicative. I made experiments with 15-20 year long wind speed time series from 44 stations of the Catalan meteorological service (Servei Meteorològic de Catalunya) homogenizing them both with additive inhomogeneity model and multiple inhomogeneity model. Then I calculated the spatial correlations of the homogenized series both for their natural form and for the logarithm series, and I compared the results belonging to the two different model applications. I did not find significant differences between the compared correlations, and the relation between them was even dependent on the observation period considered. This topic needs further study with larger datasets of more varied geographical regions, but I decided to include the additive model for wind speed homogenization in ACMANTv4.

3. ACMANT PERFORMANCE IN EFFICIENCY TESTS

The efficiency of ACMANT was tested in five method comparison projects, with five benchmark datasets: i) HOME benchmark; ii) USA daily temperature benchmark; iii) Austrian daily relative humidity benchmark; iv) MULTITEST benchmark; v) INDECIS benchmark.

i) The HOME monthly benchmark dataset comprises of 15 surrogate European temperature networks, 15 surrogate European precipitation networks, and also synthetic networks which were examined less than the surrogate networks. The homogenization experiments were performed in

2010-2011. At that time, ACMANT homogenized only temperature. A very early version of ACMANT gave excellent results in root mean square error (RMSE) reduction, but more modest results in the reduction of trend bias. However, that version did not include yet the ANOVA correction model. The official ACMANTv1 method (Domonkos, 2011) including the ANOVA correction was completed a bit later, and the closing publication of HOME (Venema et al. 2012) refers to ACMANTv1 as "late ACMANT", as the tests with ACMANTv1 were performed later than the blind tests of most methods and method versions. ACMANTv1 produced the most accurate temperature homogenization results among all the methods tested against the whole surrogate temperature benchmark dataset. Note, however, that a slight overfitting effect might be present in those results. Its explication is that even if a method is fully automatic, an overfitting might occur when the developer specifies the method parameters using previous knowledge about the general properties of the benchmark (general properties of benchmark for instance the number of time series in network, break frequency, signal to noise ratio, etc.). This kind of overfitting is unintentional.

ii) A daily temperature benchmark dataset was developed for four geographical regions of the USA by Willett et al. (2014) and Killick (2016). Its homogeneous section is likely the most realistic homogeneous benchmark dataset that has ever been created, as all climatic characteristics, even low frequency changes such the ones generated by ENSO variability are realistic in that. Four inhomogeneous datasets were created to each homogeneous set, and Killick (2016) evaluated the homogenization results. The residual RMSE was the lowest with Climatol (Guijarro, 2018) and ACMANTv2, so that these two methods were tied for the first place, while the accuracy of MASH (Szentimrey, 1999) was slightly lower. Network mean errors cannot be evaluated for this benchmark, as the number of statistically independent networks is only 4. Its explanation is that the four versions of the inhomogeneous sets for any given geographical region always yielded similar rank orders between methods (while the rank order was more varied according to geographical regions), which indicates that the network properties were too similar to consider the homogenization experiments with them statistically independent.

iii) A surrogate daily relative humidity benchmark dataset was developed for six Austrian regions (Chimani et al. 2018), and the homogenization accuracy was evaluated for three methods, ACMANTv3, MASH and HOMOP. HOMOP (Nemec et al. 2013) is a combination of the homogenization method PRODIGE and the daily data adjustment method SPLIDHOM (Mestre et al. 2011). The median of the residual RMSE was the lowest with ACMANT, although the overall efficiency was found similar for HOMOP and ACMANT. Note: The manual of ACMANTv3 suggested using the method only for the homogenization of temperature and precipitation.

iv) Between 2015 and 2017 a Spanish project was dedicated to the multiple testing of automatic homogenization methods. Only monthly temperature and monthly precipitation total test datasets were developed, but the size and variety of the created datasets were much larger than in any other benchmarking experiment. A large part of the homogeneous temperature benchmark was a synthetic set composed of the climate of one Spanish site and white noise. In six additional temperature datasets the homogeneous part was a monthly adaptation of the USA dataset of Killick (2016). We could test only freely accessible, automatic homogenization programs: ACMANTv3, two versions of Climatol_v3.0 (Guijarro, 2018), MASH monthly (identical with MASHv3.03, Szentimrey, 2014), MASH annual (monthly detection is omitted from the original MASHv3.03), Pairwise Homogenization Algorithm version 52d (PHA, Menne and Williams, 2009), Penalised t-test "simple" of RHtests_v4.0, and Penalised t-test with quantile matching of RHtests_v4.0 (Wang et al. 2007; Chan and Feng, 2019). Beyond them, three automated versions of HOMERv2.6 were also tested. José Guijarro ran most of the tests, I contributed only with testing MASH monthly: that program ran extremely slowly in the adaptation to Linux environment of José's computer, while I use the same Windows environment and FORTRAN

language as MASH, so running MASH was easier to me. In the 9th homogenization seminar we presented the tests evaluated until early 2017 (Guijarro et al. 2017). In those tests mostly ACMANT showed the highest performance both in temperature and precipitation homogenization and both in RMSE reduction and in trend bias reduction. In the method performance Climatol closely followed ACMANT, while the rank order was more mixed for the other methods. For datasets of low signal-to-noise ratio the results were different, there all homogenization methods gave rather similar results without clear "best method".
The early evaluations did not include results about the reduction of network mean errors, in spite of their importance in understanding the homogenization impact on the estimations of regional and global mean trends and variability. José has an automatic efficiency evaluation program inbuilt to the main program of running homogenizations with various homogenization methods. In that program once the evaluation of RMSE and trend bias reduction has been output, the time series are discarded, thus network mean errors cannot be calculated with that program. In early 2017 the research group agreed on generating a benchmark for archiving experimental results, and also for providing a wider efficiency evaluation including the residual RMSE and trend bias calculations for network means. These are the "late experiments" of MULTITEST.
For the late experiments, twelve kinds of homogenization problems of temperature series were selected with at least 100 networks for each homogenization problem. This benchmark comprises of 1900 networks of overall 24500 monthly temperature series. All the data and homogenization results are accessible from: https://zenodo.org/record/3934835#.XwTjF-dS_IU. For this benchmark, I calculated the method efficiencies. As the monthly temperature homogenization part of ACMANTv4 was ready at that time, ACMANTv4 was also tested. I ignored the HOMER results for the revealed problem of its Joint Detection routine (Gubler et al. 2017; Domonkos, 2017). The RMSE reduction and trend bias reduction results for individual time series were similar to those of the early MULTITEST experiments: The best results were provided by ACMANTv4 followed closely by ACMANTv3 and Climatol. However, the reduction of network mean biases shows different picture (Fig. 1).

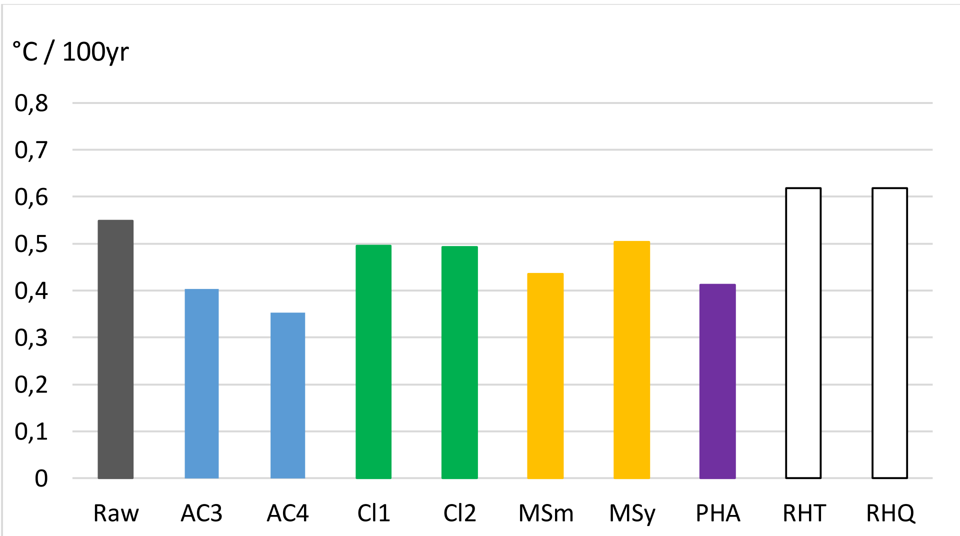


Fig. 1.: Network mean trend bias for the 1900 monthly temperature networks of MULTITEST benchmark. Raw – inhomogeneous data; AC3 – ACMANTv3; AC4 – ACMANTv4; Cl1 and Cl2 – Climatol versions; MSm – MASH monthly; MSy – MASH annual; RHT and RHQ – Rhtests versions.

The highest error reduction was achieved with ACMANTv4, but even with this method 64% of the raw data bias was remained. ACMANTv3 and PHA removed 25-27% of the raw data bias, MASH monthly removed 20%, while the other methods could not improve significantly the accuracy of network mean trends, moreover RHtests increased the raw data bias. These results

are generally weaker than the HOME results for 15 networks (Venema et al. 2012), likely because the test datasets here are more difficult to homogenize.
The results and the rank orders between homogenization methods vary between the 12 datasets of the benchmark. The most important exception is a test dataset with concerted breaks in its networks. There the PHA method removed significantly larger ratio of the raw data bias than any other method.
When the results for individual time series and those for network means are synthesized, we find that only ACMANT produces nearly even efficiency. Beyond the fact that ACMANT gives the highest efficiency in most experiments, all the other tested methods show weaknesses either in the accuracy of station series homogenization or in the reduction of network mean bias. Climatol gives good results for individual time series, but does not for network means. The opposite is true for PHA. MASH monthly is the third best method in reducing network mean bias, but for individual time series the MASH annual gives better results than MASH monthly.
v) In the INDECIS project, a daily benchmark dataset of various climatic elements has been created (http://www.indecis.eu/benchmarking.php). The benchmark dataset has temperature maximum, temperature minimum, precipitation total, relative humidity, sunshine duration, wind speed, atmospheric pressure, cloud coverage and snow depth sections. In the homogeneous section of the benchmark, one big network of 100 time series with Swedish data and another network of 30 time series with Slovenian data are allocated to the set of each climatic element. Regarding the inhomogeneous data, four inhomogeneous sets belong to each homogeneous set: one with breaks only, one with breaks and data quality issues, another with data gaps and data quality issues, and finally one more in which all the problems of breaks, data quality issues and data gaps occur. Although the overall size of this benchmark is large, the number of statistically independent networks for a given climatic variable is only two.
The INDECIS benchmark was homogenized only with ACMANTv4 and Climatol, in fact the cloud coverage homogenization and snow depth homogenization were skipped even with ACMANT. It seems that the other homogenization methods available at present are not fully prepared to the homogenization of such large daily datasets including diverse climatic elements and data quality issues.
For all the tested climatic elements the homogenization results are good both with ACMANTv4 and Climatol. The efficiency was evaluated for the RMSE of the mean values, RMSE of the 0.05 and 0.95 percentiles, and trend bias of individual time series. The rank order between ACMANT and Climatol varies according to climatic elements and networks, and the overall experience is that the efficiency is very similar for these two methods, at least for this benchmark.

4. ACMANT IN PRACTICE

The ACMANT homogenization method is used in the Catalan meteorological service, but it is rarely used in other countries. I know about occasional use of ACMANT in Greece, Italy, Ireland, Israel and in the cited study of Austria. I think that ACMANT should be used more frequently, as far as test results show that it is the only method which provides high error reduction both in station time series and regional average series. ACMANT would be particularly useful for national meteorological services, as the data of national meteorological services are used in various research and service tasks, hence they should be of high quality in all aspects, as far as it is possible.
The relatively rare use of ACMANT has various objective and subjective reasons. The most important objective reason is that ACMANT, and especially ACMANTv4, is a new method, and it needs more time for having its positive characteristics learnt and widely accepted. On the other hand, ACMANTv4 still have some shortcomings which may somewhat restrict its use.

- ACMANT does not allow human intervention. – ACMANT is not recommended to homogenize small networks (i.e. up to 7-8 time series) when metadata are available. When metadata are not available, human intervention is not needed, and ACMANT can be used from 4 time series per network. Note here that experts' views are diverse about the benefits of automatic vs. interactive methods. Nothing shows better the need of using objective methods whenever they are applicable, than the problem of Joint Detection of HOMER: It had been passed 5-6 years from the creation of HOMER when the error was first discovered, as being an interactive method it was not subjected to objective tests. I note also that the idea that homogenization methods should be fully objective is much older than the history of efficiency tests with benchmark datasets: "An ideal test procedure should be able to detect all non-homogeneities, characterize them properly, estimate their correct size, and date them accurately. It should not assign non-homogeneities where these do not exist. Furthermore, it should be objective, reproducible, and automated." (Moberg and Alexandersson, 1997). Nevertheless, the metadata use with small networks is exception, as metadata are non-quantitative pieces of information, thus they can be used more efficiently in interactive mode than within any automated procedure.
- ACMANT does not use metadata. – I admit that the use of metadata could improve the method accuracy, although the expected improvement is small within an automated procedure, and very small or absent for large networks (a scientific problem that never has been tested yet). I plan to provide a metadata use option in the next ACMANT version.
- The accuracy of ACMANT is suboptimal when concerted breaks occur within a short time period in many time series of a network. – The MULTITEST results show that for this homogenization problem the PHA method gives more accurate results. I plan to improve the accuracy of ACMANT by modifying its algorithm. The algorithm of the next ACMANT version will start with a pairwise comparison of the time series. Note, however, that ACMANTv4 mostly gives more accurate results than PHA, so the problem with concerted breaks should not be a reason to exclude ACMANTv4 from practical homogenization.
- ACMANT tends to detect too many breaks. – Various examinations show that the false alarm rate of ACMANT is higher than with some other homogenization methods (Killick, 2016; Coll et al. 2020). In addition, the output break list of ACMANT does not contain all pieces of information which are used to the final adjustments of the data. It has two explanations: a) In the final homogenization round, 9 sets of break positions are generated by the 9 ensemble members of this phase. For every set the ANOVA correction terms are calculated, then the averages of the 9 sets are applied. However, there would be no sense to average break positions, among other reasons because the number of breaks often differs between the detection results for individual ensemble members. Therefore, the break list of the most probable scenario is output by ACMANT; b) The output break list does not inform about the seasonal variation of adjustment terms when the model of irregular seasonality of inhomogeneities is applied. As a whole, the break detection in ACMANT cannot be considered one of the aims of the homogenization, but a statistical tool to find the optimal correction terms.
- An ACMANT software package does not contain programs of complete quality control procedures, hence it is expected from users to subject their data to basic quality control (manually or by the use of another software) before running ACMANT. – I do not plan to change this, since there exist several excellent quality control programs which can be applied before using ACMANT. Note also that while ACMANT is fully automatic, complete quality control procedures are usually not (and should not) so.

The objective arguments do not fully explain why ACMANT is rarely used nowadays. I am worried about the future of ACMANT not because I am the creator of ACMANT, but because ACMANT is likely the most accurate homogenization method available at present. I am worried about the future of ACMANT due to the appearance of some strange ideas related to the solution of time series homogenization in high impact scientific publications. Now I mention two of them:

- A modern idea is that reanalysis data should substitute the role of neighbor series in relative homogenization. – Reanalysis data are not homogeneous, and the inhomogeneities in them are spatially dependent, therefore the information value of reanalysis time series equals the information value of 1 neighbour series. So that reanalysis time series can really be useful for homogenization, however not instead of neighbor series, but together with them. Once we had an excellent expert who explained that no time series can be presumed fully homogeneous, therefore he advised "use as many neighbour series as possible" (Alexandersson, 1986).
- Sceptic views about the efficiency tests with benchmark datasets (e.g., Szentimrey, 2017). – When we test automatic methods the data are saved, the evaluation can be controlled, new evaluations can be performed, etc. The only true problem is the similarity of the synthetic or surrogate data to the real world data. In this respect, the multiplication of scenarios helps, i.e. an ensemble of appropriately developed scenarios will surely contain the real world homogenization problems with adequate approximation. But this strategy would need more tests and not less tests. At present, the small variety and often oversimplified spatio-temporal structures of homogeneous benchmarks limit most the reliability of method comparison test results, at this point more concerted community effort is needed. The role of tests is important, as mathematics, although give us several good tools, do not offer perfect solutions, "there is no royal road" (Szentimrey, 2004). Of course, tests are also needed to recognize and correct possible programming errors.

5. CONCLUSIONS

The ACMANTv4 homogenization method should be used more frequently in practice, since according to present knowledge this method provides the most accurate results in most kinds of homogenization tasks. ACMANT would be particularly useful for national meteorological services, as their data often need automatic homogenization for the dataset size. The data of national meteorological services are typically used for various purposes, hence it is not enough if the data accuracy is sufficient only for station level data or only for regional averages.
Finally, I copy here an extract from the executive summary of the HOME Progress Report (https://www.cost.eu/actions/ES0601/#tabs|Name:parties):
"At present only a limited number of publications intercompare some common methods and their impact on the climate record. The large number of different methods could be seen as a weakness in the science and is a challenge for the climatological community to address. There is therefore a need for a coordinated European initiative in order to produce standard methods designed to facilitate such comparisons and promote the most efficient methods of homogenisation."
We will not forget what the predecessors taught us, will we?

Acknowledgements
I thank the Servei Meteorològic de Catalunya for the access to the observed wind speed data.

References

Alexandersson H (1986) A homogeneity test applied to precipitation data. *J Climatol* 6:661-675.

Caussinus H, Mestre O (2004) Detection and correction of artificial shifts in climate series. *J Roy Stat Soc* C 53: 405-425. https://doi.org/10.1111/j.1467-9876.2004.05155.x

Chan R, Feng Y (2019) https://github.com/ECCC-CDAS/RHtests

Chimani B, Venema V, Lexer A, Andre K, Auer I, Nemec J (2018) Inter-comparison of methods to homogenize daily relative humidity. *Int J Climatol* 38:3106–3122. https:/doi.org/10.1002/joc.5488.

Coll J, Domonkos P, Guijarro J, Curley M, Rustemeier E, Aguilar E, Walsh S, Sweeney J (2020) Application of homogenization methods for Ireland's monthly precipitation records: comparison of break detection results. *Int J Climatol* 40:6169-6188. https://doi.org/10.1002/joc.6575

Domonkos P (2011) Adapted Caussinus-Mestre Algorithm for Networks of Temperature series (ACMANT). *Int J Geosci* 2:293-309. https://doi.org/10.4236/ijg.2011.23032

Domonkos P (2017) Time series homogenisation with optimal segmentation and ANOVA correction: past, present and future. Ninth Seminar for Homogenization and Quality Control in Climatological Databases (Ed. T. Szentimrey, M. Lakatos, and L. Hoffmann), WMO WCDMP-85, Geneva, Switzerland, 29–45.

Domonkos P (2020) ACMANTv4: Scientific content and operation of the software. 71pp, https://github.com/dpeterfree/ACMANT

Domonkos P, Coll J (2017) Homogenisation of temperature and precipitation time series with ACMANT3: Method description and efficiency tests. *Int J Climatol* 37:1910–1921. https://doi.org/10.1002/joc.4822

Gubler S, Hunziker S, Begert M, Croci-Maspoli M, Konzelmann T, Brönnimann S, Schwierz C, Oria C, Rosas G (2017) The influence of station density on climate data homogenization. *Int J Climatol* 37:4670–4683. https://doi.org/10.1002/joc.5114

Guijarro JA (2018) Homogenization of climatic series with Climatol, https://www.climatol.eu/homog_climatol-en.pdf

Guijarro JA, López JA, Aguilar E, Domonkos P, Venema V, Sigró J, Brunet M (2017) Comparison of homogenization packages applied to monthly series of temperature and precipitation: the MULTITEST project. Ninth Seminar for Homogenization and Quality Control in Climatological Databases (Ed Szentimrey T, Lakatos M, Hoffmann L) WMO WCDMP-85:46-62.

Killick RE (2016) Benchmarking the Performance of Homogenisation Algorithms on Daily Temperature Data. PhD thesis, University of Exeter, UK. https://ore.exeter.ac.uk/repository/handle/10871/23095

Lindau R, Venema V (2018) On the reduction of trend errors by the ANOVA joint correction scheme used in homogenization of climate station records. *Int J Climatol* 38:5255-5271. https://doi.org/10.1002/joc.5728

Menne MJ, Williams CN (2009) Homogenization of temperature series via pairwise comparisons. *J Climate* 22:1700–1717. https://doi.org/10.1175/2008JCLI2263.1

Mestre O, Domonkos P, Picard F, Auer I, Robin S, Lebarbier E, Böhm R, Aguilar E, Guijarro JA, Vertacnik G, Klancar M, Dubuisson B, Štěpánek P (2013) HOMER: homogenization software in R – methods and applications. *Idöjárás* Q J Hung Meteorol Serv 117:47–67.

Mestre O, Gruber C, Prieur C, Caussinus H, Jourdain S (2011) SPLIDHOM: a method for homogenization of daily temperature observations. *J Appl Meteor Climatol* 50:2343–2358. https://doi.org/10.1175/2011JAMC2641.1

Moberg A, Alexandersson H (1997) Homogenization of Swedish temperature data. Part II: homogenized gridded air temperature compared with a subset of global gridded air temperature since 1861. *Int J Climatol* 17:35–54.

Nemec, J, Gruber C, Chimani B, Auer I (2013) Trends in extreme temperature indices in Austria based on a new homogenized dataset. *Int J Climatol* 33:1538–1550. https://doi.org/10.1002/joc.3532

Peterson TC, Easterling DR (1994) Creation of homogeneous composite climatological reference series. *Int J Climatol* 14:671-679.

Szentimrey T (1999) Multiple Analysis of Series for Homogenization (MASH). Second Seminar for Homogenization of Surface Climatological Data (Ed Szalai S, Szentimrey T, Szinell Cs). WMO WCDMP-41:27-46.

Szentimrey T (2004) Something like an introduction. Fourth Seminar for Homogenization and Quality Control in Climatological Databases (Ed Szalai S). WMO WCDMP-56:5-16.

Szentimrey T (2010) Methodological questions of series comparison. Sixth Seminar for Homogenization and Quality Control in Climatological Databases (Ed Lakatos M, Szentimrey T, Bihari Z, Szalai S). WMO WCDMP-76:1-7.

Szentimrey T (2014) Manual of homogenization software MASHv3.03, Hungarian Meteorological Service, pp. 69.

Szentimrey T (2017) Introduction on homogenization, quality control, spatial interpolation, gridding. Ninth Seminar for Homogenization and Quality Control in Climatological Databases (Ed Szentimrey T, Lakatos M, Hoffmann L), WMO WCDMP-85:5-19.

Venema V, Mestre O, Aguilar E, Auer I, Guijarro JA, Domonkos P, Vertacnik G, Szentimrey T, Štěpánek P, Zahradnicek P, Viarre J, Müller-Westermeier G, Lakatos M, Williams CN, Menne M, Lindau R, Rasol D, Rustemeier E, Kolokythas K, Marinova T, Andresen L, Acquaotta F, Fratianni S, Cheval S, Klancar M, Brunetti M, Gruber C, Duran MP, Likso T, Esteban P, Brandsma T (2012) Benchmarking monthly homogenization algorithms. *Clim Past* 8:89–115. https://doi.org/10.5194/cp-8-89-2012

Vincent LA, Zhang X, Bonsal BR, Hogg WD (2002) Homogenization of daily temperatures over Canada. *J Climate* 15:1322–1334. https://doi.org/10.1175/1520-0442(2002)015<1322:HODTOC>2.0.CO;2

Wang XL, Wen QH, Wu Y (2007) Penalized maximal t test for detecting undocumented mean change in climate data series. *J Appl Meteor Climatol* 46/6: 916-931. https://doi.org/10.1175/JAM2504.1

Willett KM, Williams CN, Jolliffe I, Lund R, Alexander L, Brönniman S, Vincent LA, Easterbrook S, Venema V, Berry D, Warren R, Lopardo G, Auchmann R, Aguilar E, Menne M, Gallagher C, Hausfather Z, Thorarinsdottir T, Thorne PW (2014) A framework for benchmarking of homogenisation algorithm performance on the global scale. *Geosci Instrum Meth Data Sys* 3:187-200. https://doi.org/10.5194/gi-3-187-2014

ANALYSIS OF PARALLEL MEASUREMENTS IN GREECE

Athanassios Argiriou[1], Panagiotis Ioannidis[1], Anna Mamara[2]

[1]Laboratory of Atmospheric Physics, University of Patras, Greece, [2]Hellenic National Meteorological Service, Athens, Greece

Abstract

This work presents the preliminary results of the comparison between parallel measurements of daily maximum and minimum temperatures in Greece. The data comes from the WMO compatible network of the Hellenic National Meteorological Service. It was found that transition to AWS introduces break points to the original time series.

1. INTRODUCTION

Climate change studies rely on long term homogenized time series of meteorological (WMO, 2011). There are several sources of inhomogeneities. Some of them are related to the instrumentation used, i.e. replacement of an instrument (Aguilar et al., 2003), degradation of instrument performance, etc. The replacement of conventional meteorological stations by automatic ones is also a major inhomogeneity issue needed to be considered carefully by establishing a common measurement period between the two types of stations and maintaining them the longest possible (Aguilar et al., 2003, WMO, 2017). This study aims in contributing to the effort of the scientific community (e.g. http://www.surfacetemperatures.org/databank/parallel_measurements; Kaspar et al. (2016); Acquaotta et al. (2016); Baumgartner et al. 2018) of collecting and analysing the available parallel measurement time series. We present the preliminary analysis of all the available parallel measurement time series of daily maximum (TX) and minimum (TN) values, from the weather station network of the Hellenic National Meteorological Service (HNMS).

2. DATA

The HNMS started replacing the conventional (Stevenson screen) weather stations of its network with automatic ones (AWS) in the begining of the previous decade. However and despite the fact that this replacement was rather extensive and for reasons that are beyond the scope of this paper, extensive parallel measurement records have been kept for a limited number of stations and meteorological parameters. In this work we limit our analysis to TX and TN daily time series, for the stations and periods listed in Table 1.

The last column of Table 1 gives the climate zone of the station location, as defined by Mamara et al. (2013). Figure 1a shows this classification; Figure 1b shows the location of the stations on the map of Greece. From Fig. 1, it can be deduced that the parallel measurement stations cover only four out of the eight climate zones of Greece.

3. RESULT

The daily values compared here are the 06:00 for TN and at 18:00 UTC for TX readings for both the conventional and the automatic weather stations. Their differences (ΔTN and ΔTX) are defined as the readings of the automatic weather station minus that of the conventional

(Stevenson screen) station; a positive temperature difference indicates that the value recorded by the AWS is higher than that of the conventional station.

Table 1.: HNMS stations with parallel measurements. Climate zone classification as in Mamara et al. (2013).

WMO ID	Station name – ICAO code	Parallel measurements period	Latitude (°)	Longitude (°)	Altitude (m)	Climatic zone
16632	Kozani - LGKZ	1/2009 – 12/2011	40º 17’ 22.20’’	21º 50’ 29.40’’	621	B
16642	Ioannina - LGIO	11/2010 – 12/2012	39º 41’ 41.70’’	20º 49’ 09.50’’	483.36	B
16643	Aktio - LGPZ	1/2013 - 12/2014	38º 55’ 19.17’’	20º 46’07.79’’	1.47	C
16648	Larissa - LGLR	1/2009 – 3/2012	39º 38’ 45.80’’	22º 27’ 36.55’’	71.15	D
16699	Tanagra - LGTG	1/2009 – 2/2012	38º 20’ 07.44’’	23º 33’ 46.44’’	138.05	D
16710	Tripoli - LGTP	1/2012 – 12/2014	37º 31’ 28.92’’	22º 23’ 49.92’’	650.57	B
16726	Kalamata - LGKL	11/2012 – 12/2014	37º 04’ 09.12’’	22º 01’ 21.36’’	6.20	C
16766	Paros - LGPA	08/2010 – 12/2013	37º 00’ 41.24’’	25º 07’ 44.85’’	33.30	E

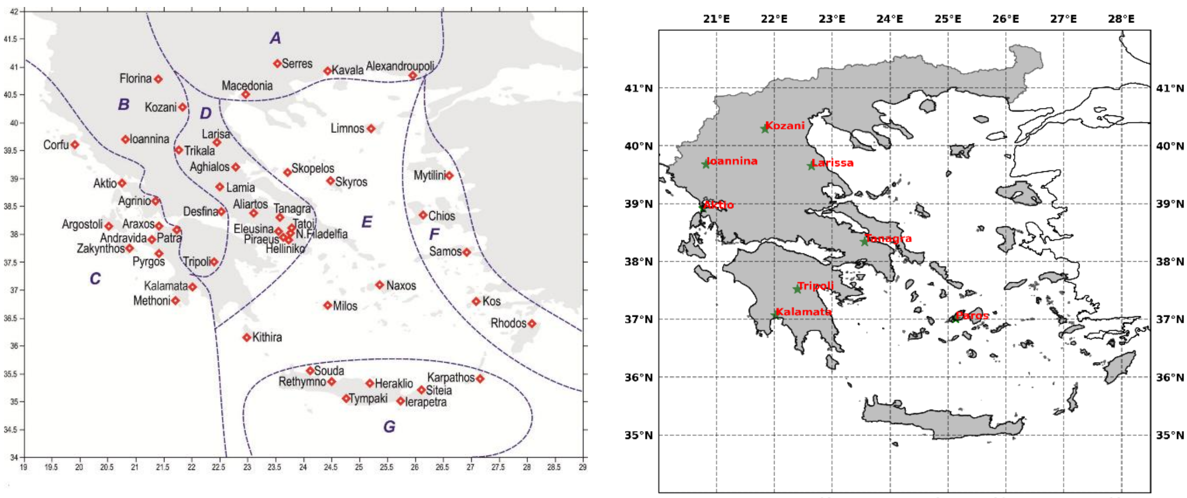


Fig. 1.: Climatic zones (a); Stations with parallel measurements (b).

Tables 2 and 3 provide summary statistics for ΔTN and ΔTX respectively.

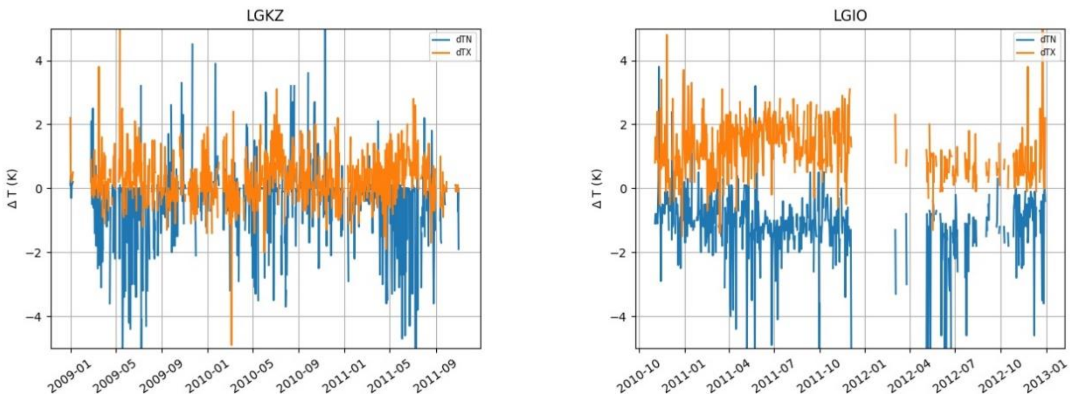
Table 2.: ΔTN Summary statistics.

Station name – ICAO code	Number of Points (-)	Mean value ± Standard deviation (°C)	Minimum value (°C)	Maximum value (°C)
Kozani - LGKZ	776	-0,4 ± 1,3	-5,9	9,4
Ioannina - LGIO	531	-1,2 ± 1,2	-7,4	3,8
Aktio - LGPZ	517	-0,3 ± 0,4	-2,4	2,7
Larissa - LGLR	1079	-0,5 ± 0,5	-3,0	4,0
Tanagra - LGTG	1032	-0,5 ± 0,5	-3,6	1,9
Tripoli - LGTP	584	-0,3 ± 0,3	-1,9	3,5
Kalamata - LGKL	608	-0,2 ± 0,3	-1.7	1,0
Paros - LGPA	796	-0,3 ± 0,7	-5,6	6,1

Table 3.: ΔTX Summary statistics.

Station name – ICAO code	Number of points (-)	Mean value ± Standard deviation (°C)	Minimum value (°C)	Maximum value (°C)
Kozani - LGKZ	776	0,3 ± 0,8	-4,9	5,4
Ioannina - LGIO	531	1,2 ± 0,9	-1,5	5,3
Aktio - LGPZ	517	0,3 ± 0,4	-2,8	2,2
Larissa - LGLR	1079	0,6 ± 0,4	-1,1	4,2
Tanagra - LGTG	1032	0,6 ± 0,4	-1,4	4,3
Tripoli - LGTP	584	0,0 ± 0,2	-2,7	1,9
Kalamata - LGKL	608	0,4 ± 0,4	-0.1	2,1
Paros - LGPA	796	0,3 ± 0,5	-6,1	3,1

The difference time series are visualized by plotting them as time series (Fig. 2), histograms (Fig. 3) and boxplots (Fig. 4).



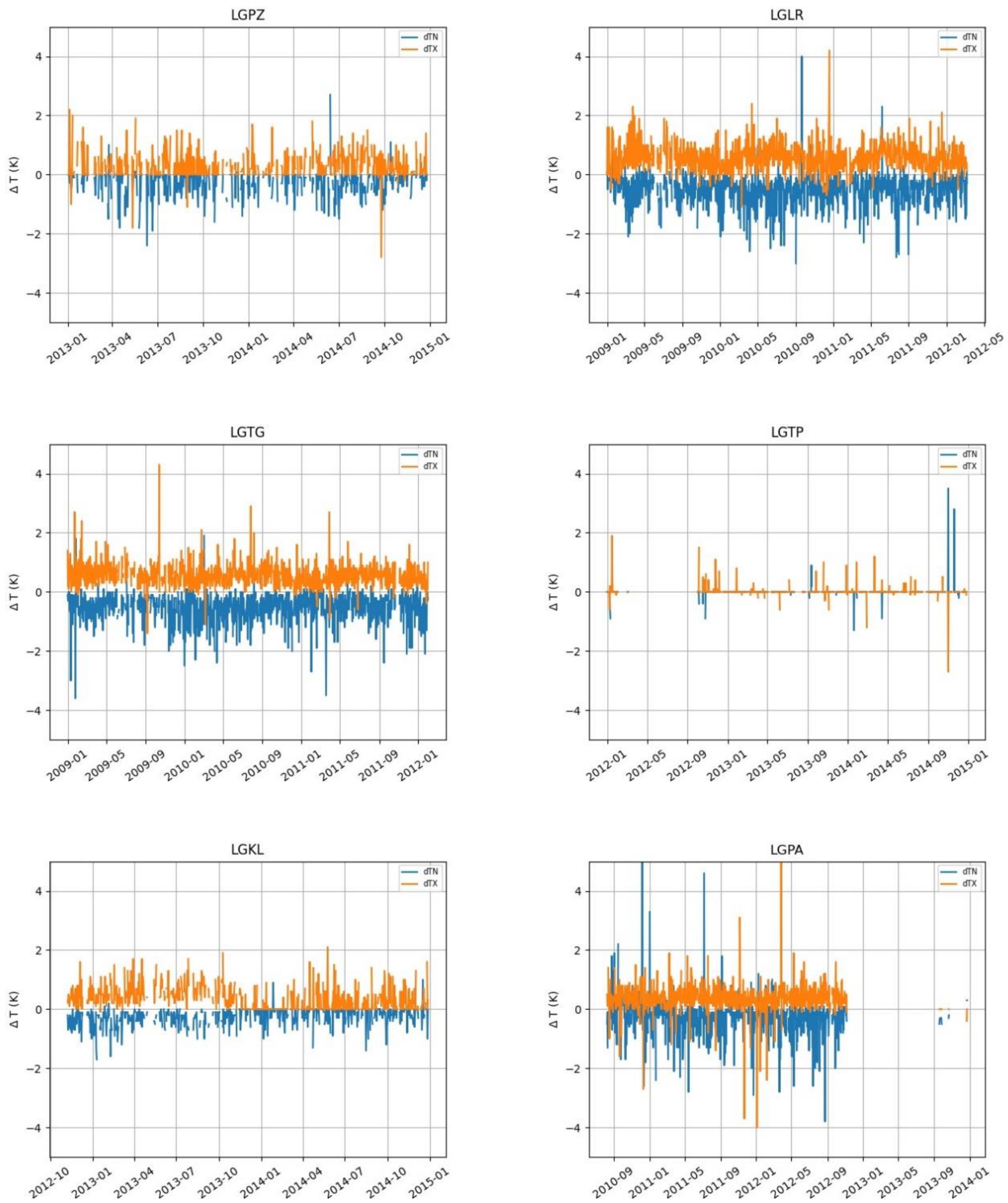
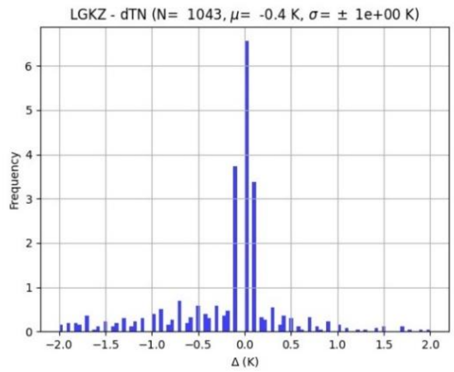
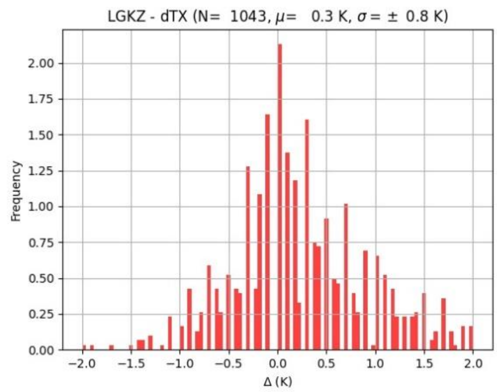
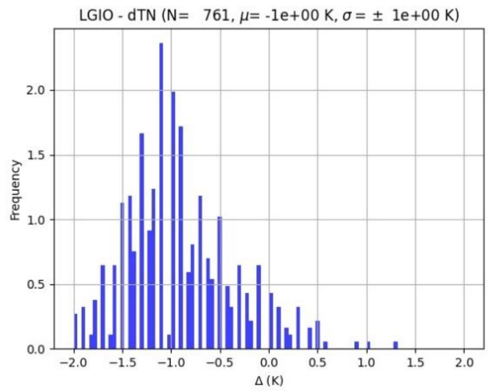
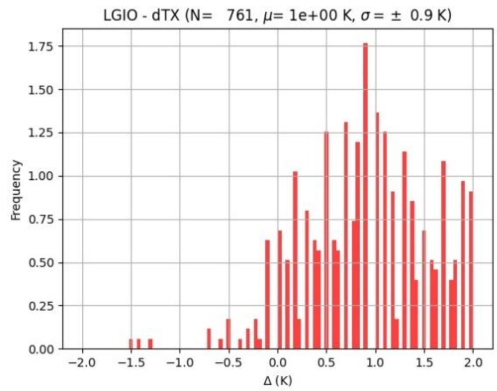


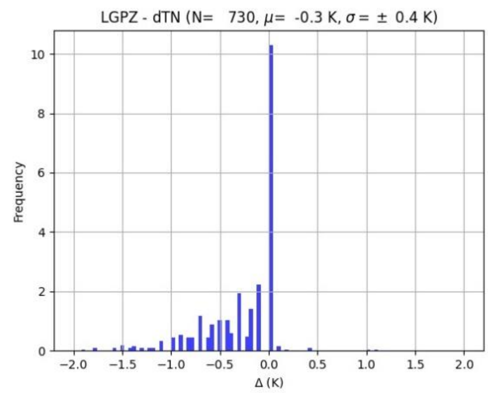
Fig. 2.: Temperature difference time series plots (ΔTN blue line, ΔTX orange line).

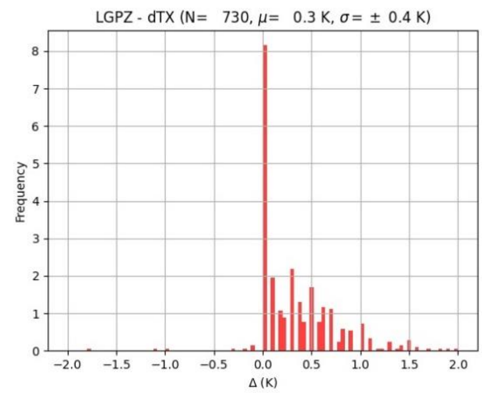

LGKZ - dTN (N= 1043, μ= -0.4 K, σ= ± 1e+00 K)
Frequency
Δ (K)

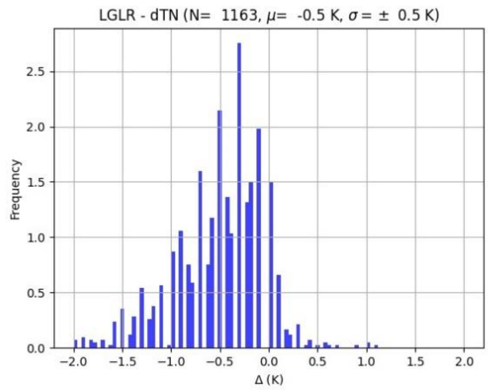

LGKZ - dTX (N= 1043, μ= 0.3 K, σ= ± 0.8 K)
Frequency
Δ (K)

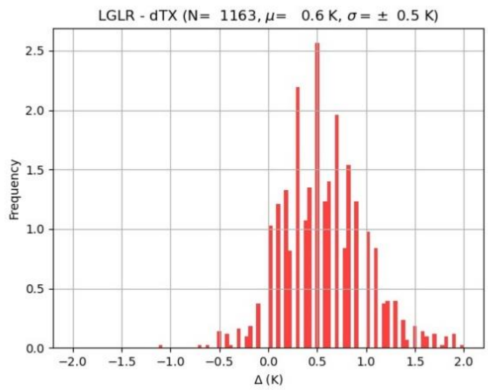

LGIO - dTN (N= 761, μ= -1e+00 K, σ= ± 1e+00 K)
Frequency
Δ (K)


LGIO - dTX (N= 761, μ= 1e+00 K, σ= ± 0.9 K)
Frequency
Δ (K)


LGPZ - dTN (N= 730, μ= -0.3 K, σ= ± 0.4 K)
Frequency
Δ (K)


LGPZ - dTX (N= 730, μ= 0.3 K, σ= ± 0.4 K)
Frequency
Δ (K)


LGLR - dTN (N= 1163, μ= -0.5 K, σ= ± 0.5 K)
Frequency
Δ (K)


LGLR - dTX (N= 1163, μ= 0.6 K, σ= ± 0.5 K)
Frequency
Δ (K)

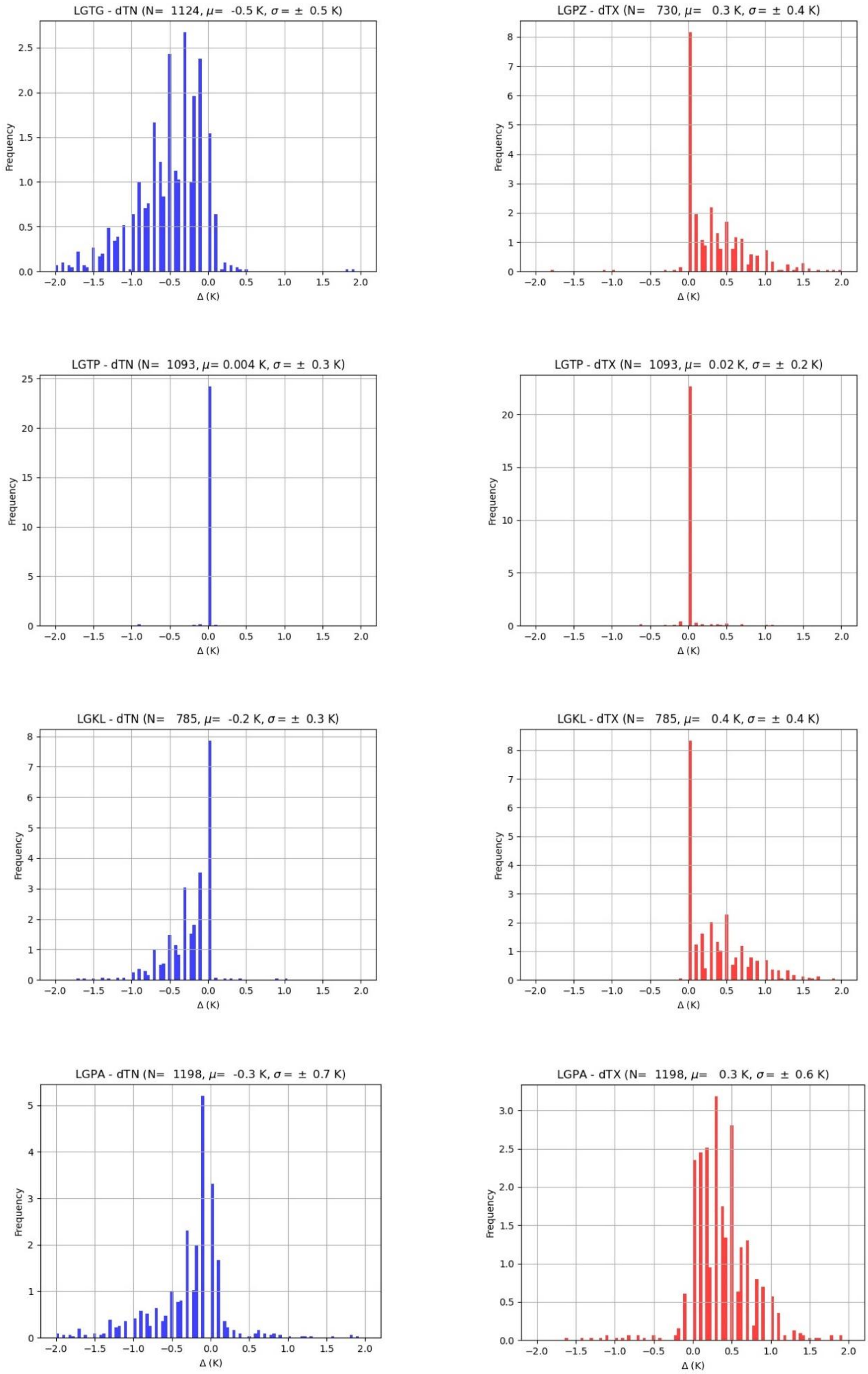
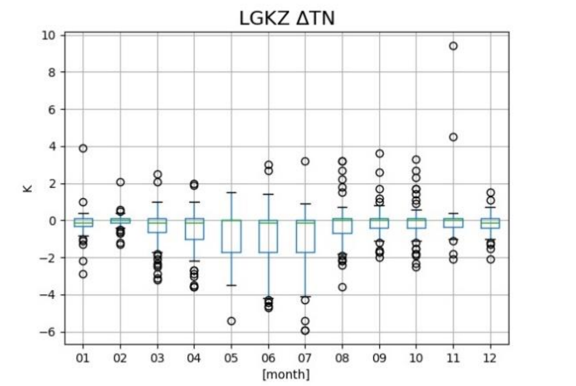
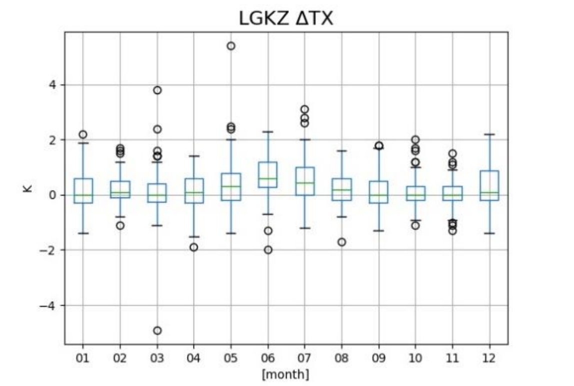
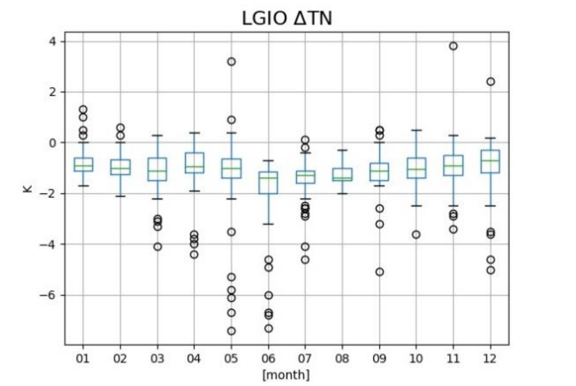
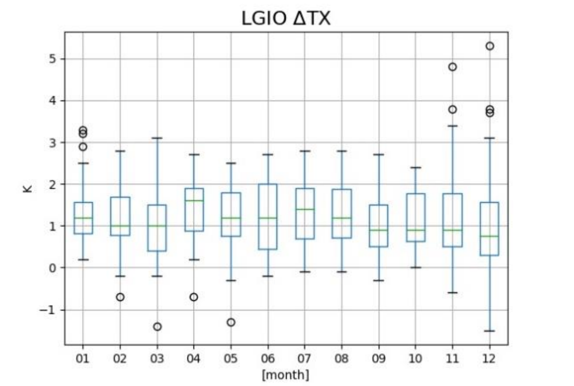


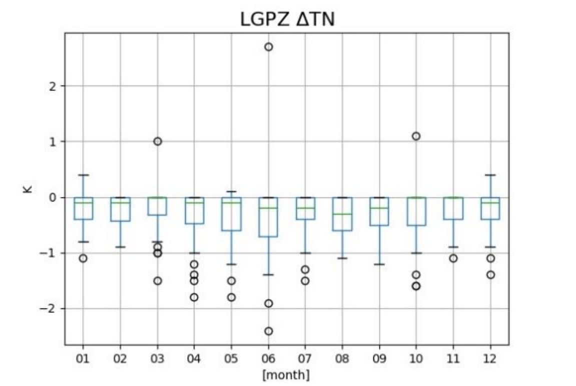
Fig. 3.: Temperature difference histograms for TN (blue bars) and TX (red bars); N denotes the number of points, μ the mean value and σ the standard deviation.

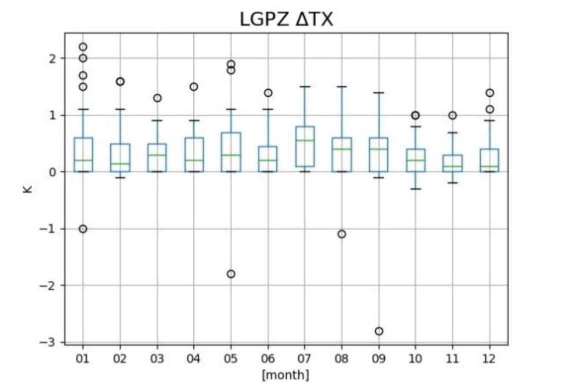

LGKZ ΔTN
K
[month]

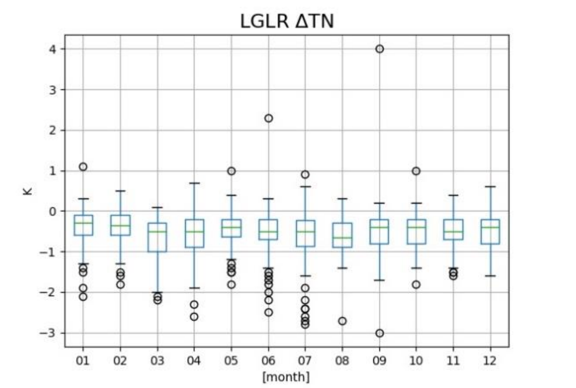

LGKZ ΔTX
K
[month]

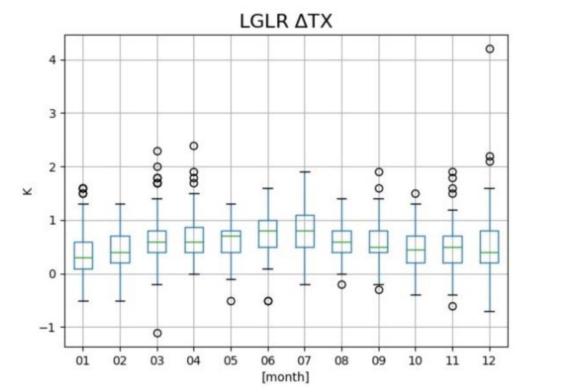

LGIO ΔTN
K
[month]


LGIO ΔTX
K
[month]


LGPZ ΔTN
K
[month]


LGPZ ΔTX
K
[month]


LGLR ΔTN
K
[month]


LGLR ΔTX
K
[month]

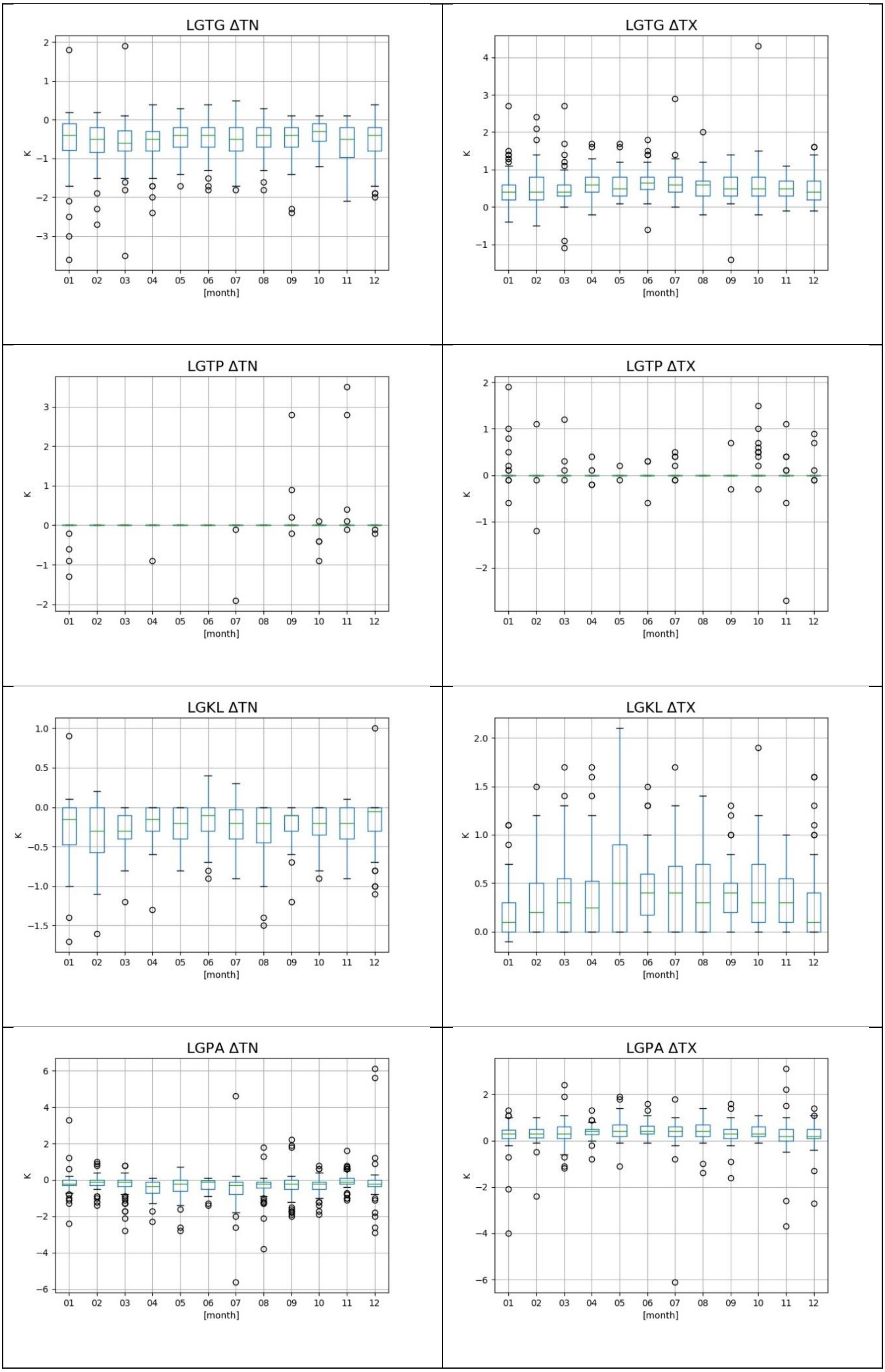


Fig. 4.: ΔTN and ΔTX boxplots.

4. CONCLUSIONS

The main findings of this preliminary analysis are summarized as follows:

- The differences between TN and TX reported by conventional and AWS are of about 0.5 K in average for all stations.
- AWS tend to record lower TN values and higher TX values compared to conventional station readings throughout the year (no seasonality).
- This behavior results to higher DTR values via the AWS readings.
- The recorded differences are expected to introduce breaks in long term time series of daily measurements.

References

Acquaotta F, Fratianni S and Venema V (2016) Assessment of parallel precipitation measurements networksin Piedmont, Italy. Int. J. Climatol. 36: 3963–3974 (doi: 10.1002/joc.4606)

Aguilar E, Auer I, Brunet M, Peterson TC and Jon Wieringa (2003) Guidelines on climate, metadata and homogenization. WMO/TD No. 1186, (https://library.wmo.int/index.php?lvl=notice_display&id=11635#.XlkJ5agzY2w), p. 10, 21

Baumgartner D J, Pötzi W, Freislich H, Strutzmann W. Veronig A M, Foelsche U, Rieder H E (2018) A comparison of long-term parallel measurements of sunshine duration obtained with a Campbell-Stokes sunshine recorder and two automated sunshine sensors. Theor Appl Climatol (2018) 133:263–275 (doi: 10.1007/s00704-017-2159-9)

Kaspar F, Hannak L, Schreiber K-J (2016) Climate reference stations in Germany: Status, parallel measurements and homogeneity of temperature time series. Adv. Sci. Res., 13, 163–171 (doi:10.5194/asr-13-163-2016)

Mamara A, Argiriou A, Anadranistakis M (2013) Homogenization of mean monthly temperature series of Greece. Int. J. Climatol. 33: 2649–2666 (doi: 10.1002/joc.3614)

World Meteorological Organization – WMO (2011) Guide to Meteorological Practices. Report WMO-No. 100, Geneva, p. 5.1.

World Meteorological Organization – WMO (2017) Challenges in the Transition from Conventional to Automatic Meteorological Observing Networks for Long-term Climate Records. Report WMO-No. 1202, Geneva, p. 1.4.

HOMOGENIZATION OF MONTHLY TEMPERATURE SERIES IN NORWAY

Elinah Khasandi Kuya, Herdis Motrøen Gjelten and Ole Einar Tveito

Norwegian Meteorological Institute, Henrik Mohns plass 1, 0371 Oslo, Norway, elinahkk@met.no

Abstract

Climatological standard normals are applied to describe expected weather and climate conditions at given locations and also acts as a reference for current conditions. Homogenization of Norway's monthly temperature time series for the period 1961-2018 was therefore undertaken for the purpose of calculating new climatological standard normals for the period 1991-2020. The Norwegian observation network has changed considerably during the last 20-30 years, introducing non-climatic changes such as automation and relocation. Homogenization was therefore necessary to provide a consistent basis for the new normals. HOMER was applied to detect and adjust inhomogeneities in 145 monthly temperature series (including 30 series from Sweden and 7 from Finland). The results of the homogeneity testing indicate that approximately 92% of the temperature series were inhomogeneous. The annual adjustment factor ranged from −0.94°C to 1.01 °C. 99% of the breaks were confirmed by metadata. Relocation of the station was the most common reason for inhomogeneity, explaining more than 40 % of the inhomogeneities found by HOMER. Results further demonstrated the benefits of including Swedish and Finnish series as reference series in the homogeneity testing of Norway's temperature series. Results also showed a wider range of anomalies in the raw series than in homogenized series confirming that homogenization contributes to better spatial coherence of the temperature series. This clearly provides a strong guidance on the reliability of the adjusted dataset.

1. INTRODUCTION

Climate normals play an important role in weather and climate studies and have two major purposes: (a) as an indicator of the conditions most likely to be experienced at a given place under the current climate and (b) as a reference against which climate conditions at a given region in a given time can be compared (WMO 2017). These purposes require good basic data that are consistent and homogenous. Climate normals are calculated for 30-year periods (WMO 2011) and up to now, WMO has operated with standard normal periods that are subsequent 30-year periods (1901-1930, 1931-1960, 1961-1990 and the coming 1991-2020). However, the WMO Congress in 2015 decided that the standard normal period should now be the last 30-year period ending with a year ending with 0 (WMO 2017).

One of the main prerequisites to establish climate normals is to ensure that the data used is as homogenous as possible, i.e. that it represents only natural variations in weather and climate. However, climatic observations are often influenced by inhomogeneities because of a variety of external influences. The main causes of inhomogeneity are relocation of stations and the change of instrumentation. Other causes include changes of observer and/or observing practices, and changes in the surrounding environment (Aguilar et al. 2003). Most of such changes introduce artificial shifts (change points) in the time series that are abrupt, while some, particularly changing surroundings, lead to gradual increasing biases. Time series with such changes no longer describe natural variability or trends and therefore require homogenization before making any climate assessment.

Several efforts have been made to homogenize temperature and precipitation series at MET Norway in later years (Andresen 2011, Lundstad 2016, Lundstad and Tveito 2016). These studies showed promising results, however, they either did not cover the entire period 1961-2018 and/or were done for a small number of stations. Considerable work therefore remained to establish robust and homogenous data for calculating new climate normals.

This study presents homogenization of monthly mean temperature time series for Norway for the period 1961-2018. The primary objective of the study is to establish a high quality temperature reference dataset for calculating new standard climate normal for the period 1991-2020 and for climate monitoring services.

2. DATA AND METHODS

Temperature dataset

This study used a temperature data set obtained from ClimNorm - a network activity under the Nordic Framework for Climate services (Tveito et al. 2020). The objective of ClimNorm includes exchange of data, methods and experiences to provide the best possible data basis for calculating the new 1991-2020 standard normal, and covers the area Denmark, Estonia, Finland, Latvia, Norway and Sweden.

The objective of the analysis was to produce a homogenous data set as a basis for calculating new standard climate normals. The analysis covered the period 1961 - 2018 since the study was conducted in 2019 and complete series were available up to 2018. Data back to 1961 was used because it is better to homogenize with a longer time series. Many of the temperature series did not cover the entire study period, mostly because the observation networks had been modernized and thus a large number of stations relocated in the process. It was therefore necessary to merge incomplete series where necessary and possible. The criteria for merging time series was a maximum horizontal distance of 10 km and maximum vertical distance of 100 m between the stations. In addition, a criterion of data coverage of 80 % (no more than ten years missing in the period 1961-2018) had to be fulfilled for a series to be included in the analysis. This ensured that the resulting dataset would be a good reference for further interpolation and homogenization of series with poorer data coverage. The resulting dataset consisted of 108 series from Norway, 30 from Sweden and 7 from Finland. 46% of all the series used were merged. More details on the data set can be found in Kuya et al. (2020) and Tveito et al. (2020).

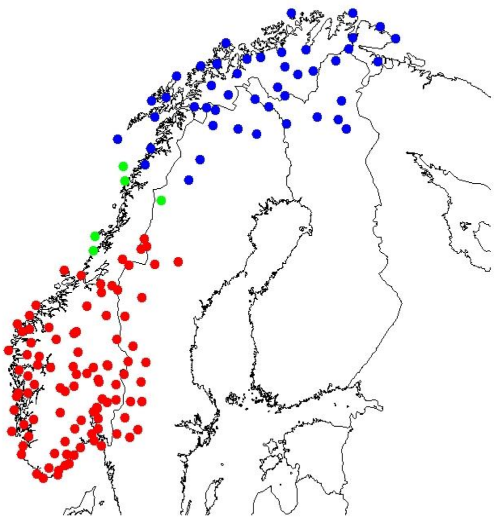


Fig. 1.: Location of stations in network 1 (red), network 2 (blue) and stations used in both networks (green).

Table 1.: Number of series in each of the two analysed networks. Number in parenthesis shows the number of stations when including the stations used in both networks.

	Norway	Sweden	Finland
Network 1 - south	76 (78)	20 (21)	0
Network 2 - north	32 (34)	10	7

The Homogenization analysis was conducted in two networks - one covering the northern part of mainland Norway and one covering the southern part. A check was conducted to assess if any of the border stations in network 1 should be included in network 2 and vice versa, and as a result five stations were included in both networks. Neighbouring stations from Sweden and Finland were also included. Fig. 1 and Table 1 shows the data series used in this analysis.

Homogenization procedure

The homogenization software HOMER (Mestre et al. 2013) was used in the analysis. HOMER is a relative homogenization method, meaning that the analysis relies on neighbouring stations (reference stations). It is an iterative semi-automatic method that takes advantage of metadata when accepting and rejecting detected breaks and it is well described in e.g. Kuya et al. (2020) and Lundstad and Tveito (2016).

The settings for selecting reference series for homogeneity testing was based on their cross correlation coefficient with certain distance restriction to ensure that the candidate and reference series represent the same climatological features. In this study, a minimum correlation of 0.95 was set, but also a minimum of eight reference series for each candidate series, meaning that all candidate series would have at least eight reference stations regardless of the correlation coefficient. The lowest correlation coefficient recorded was 0.9, which shows that the data series were quite highly correlated.

The built in quality control function in HOMER was used to identify anomalous series and detect outliers. This consisted of visual analysis of the graphical outputs provided in HOMER. Candidate series with very high/low values above/below a reference range were further inspected to determine if they were legitimate outliers. Some of these suspected outliers were a result of data entry or processing errors and were corrected. Others were modelled values or thought to be observations of very poor quality and were consequently removed. The remaining suspected outliers were kept because they were assumed to be correct observations.

The break points detected by HOMER were checked against the documented metadata to validate if they were actually inhomogeneities and determine the optimum position of the break. This is a time consuming and subjective process that requires careful considerations.

In order to avoid over-adjusting the dataset, a set of quite strict guidelines was followed:

- Breaks within the first and last five years of a series were rejected.
- With no metadata explaining a break, the change-point should be seen in > 80% of the pairwise difference series.
- With supporting evidence from metadata, the change-point should be seen in at least three of the pairwise difference series.
- Joint and ACMANT detected breaks were accepted if they corresponded to metadata or if they were present in at least three pairwise difference series.
- Close breaks were assessed carefully. Weight was placed on the break point with metadata (or the most obvious or dominant reason for a break). Very close breaks, especially those within two years of each other and with adjustment factors with opposite signs of each other, were rejected.

Both annual and seasonal series were analyzed. This was important because not all breaks would be detected in the annual series, some may have only been apparent in one or two seasons. When all detected break points had been either rejected or accepted, monthly adjustment factors would be calculated and applied in the final steps of the homogenization procedure. More details on the procedure can be found in Kuya et al. (2020).

3. RESULTS

The HOMER Homogenization method was performed on 145 monthly series. This includes 108 temperature series from Norway, 30 from Sweden and 7 from Finland that were organized into two networks representing the northern and southern part of Norway (Fig. 1). The data quality control revealed 90 detected outlier values; however, the great majority of the outliers (61 values) were confirmed to be legitimate observations with higher/lower daily averages with respect to their neighbouring stations. Eight values were found to be outliers because of data entry or processing error and thus adjusted and corrected. Only 21 values were characterized as true outliers and therefore removed.

The results for cluster analysis in HOMERs quality control function showed the classification of Norway into six temperature regions. This was similar to those identified by Hanssen-Bauer and Nordli (1998) where a combination of principal component analysis and cluster analysis was used to divide Norway into six temperature regions.

The results of the homogeneity testing indicate that approximately 92 % of the series analysed were inhomogeneous. Only nine of the Norwegian temperature series were found to be homogenous. All of these nine stations had at some point either relocated, painted the radiation screen or changed instruments. This clearly shows that in some instances external changes in the station may not necessarily cause inhomogeneities in climate series. The inhomogeneous series had breakpoints ranging from one to four, with most series having two breakpoints. Over 200 detected breaks in the Norwegian series and approximately 70 detected breaks in Swedish and Finnish series were homogenized.

Almost all of the accepted breaks could be explained by metadata. Relocation of the station was the most common reason for inhomogeneity, explaining more than 40 % of the inhomogeneities, see Fig. 2. This was no surprise as about half of the series in the study were merged series. New observer, which is listed as a reason for inhomogeneities, should in theory not affect the measurements. However, sometimes a change in observer leads to a change in routines, and that might cause inhomogeneities. Only about 2 % of the breaks could not be explained by the documented station history. The annual adjustment factors in the analysed series ranged from -0.94 to 1.01 °C. The highest adjustments in the series were connected to relocations of the station.

It should be noted that the Swedish and Finnish series were adjusted without any metadata except for those that had been merged (implying that the station had been relocated).

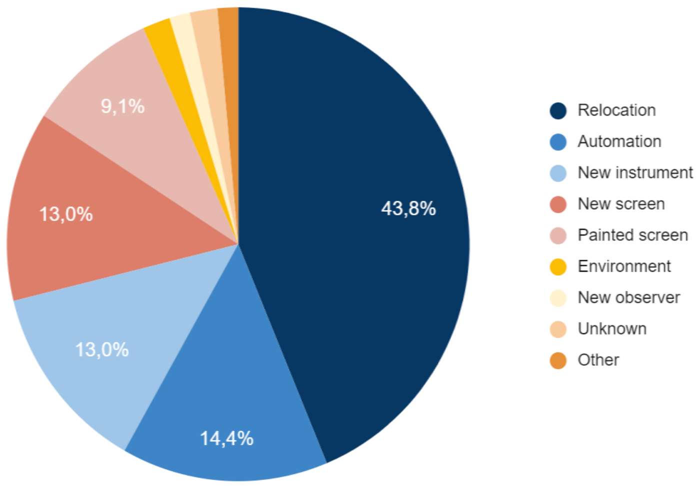


Fig. 2.: Main reasons for inhomogeneities in percent of the total amount of breaks. The reasons are relocation, automation, change of instruments, change or painting of radiation screen, change of observer, changes in the surrounding environment, in addition to unknown and other causes.

THE LATE 1980S CLIMATE SHIFT

HOMER detected a break around 1986-89 in about 70 % of the analysed series, and in most cases these breaks were not supported by metadata or pairwise comparison. This period corresponds to a “climate shift” in the late 1980s, when there was a strong increase in temperature, especially in winter. A study by Lundstad and Tveito (2016) demonstrated that the joint detection algorithm in HOMER is sensitive to shifts in climate such that whenever there is an abrupt regional change in temperature, e.g. a strong increase as in this case, the detection algorithm identifies this as a break since the response at the stations in the network might differ. This abrupt change in the climatic system has been observed in many places in Europe between 1985 and 1991, and Skelton et al. (2020) suggests a possible cause to be Arctic sea ice loss, potentially linked to reduced sulfate aerosol emissions and coupled to temperature by an albedo feedback mechanism. Therefore, the breaks around 1986-89 that did not have metadata supporting them were rejected to avoid masking possible climate changes in the series.

The raw and corrected Gardermoen annual series is provided in Fig. 3. The figure shows an annual increase in the average temperature for Gardermoen throughout the study period for both the homogenized and non-homogenized series, but the increase in the homogenized series is smaller. The late 1980s climate shift is clearly observed in the figure, where several cold years before the late 80s are followed by unusually warm years after the early 90s.

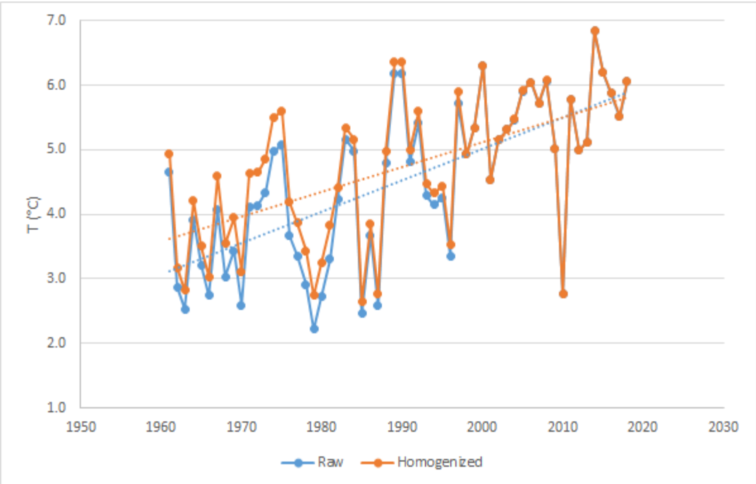


Fig. 3.: Raw and homogenized annual temperature series for Gardermoen. Linear trends for both series are included.

IMPACTS OF ADDING SWEDISH AND FINNISH REFERENCE STATIONS

A case study was conducted to determine the impacts the Swedish and Finnish reference series would have on the homogenization analysis. This was done by comparing analyses with and without Swedish and Finish stations to ascertain the differences in the dependence on the reference series number based on their geographical location and its impact on the detection of breaks in the candidate series. It was apparent that adding bordering stations from Sweden and Finland had a positive impact on the break detection, especially (as expected) in the Norwegian series along the border. They would get more and better-correlated reference stations and would therefore be able to easier and more accurately detect breaks in the series. Using a larger reference dataset that includes several reference series from adjacent similar climatic areas avoids to greater extent single incidents in the reference series that cause false inhomogeneities (Lindau and Venema 2016).

IMPACTS OF HOMOGENIZATION

To assess the impact of homogenization on the dataset, anomaly series of the homogenized and non-homogenized data were compared. 1961-1990 was used as a reference period and the series were filtered using a Gaussian density function to show temperature variations on a five-year scale. The filtered anomaly series from six stations are presented in Fig. 4, one for each temperature region in Norway. The comparison shows a smaller amplitude in the anomalies for the homogenized series. This confirms that homogenization of a data set gives better spatial coherence and contributes to the understanding of climate variation both in space and in time. This clearly provides a strong guidance on the reliability of the adjusted dataset. Nevertheless, the homogenous series preserves the general statistical distribution of the raw series.

Another comparison was done by Tveito et al. (2020), who compared maps of the change of monthly normal values between 1961-1990 and 1991-2020 based on non-homogenized series with maps based on the homogenized series for January, see Fig. 5. It was seen that the homogenized temperature data set gave a smoother spatial pattern, meaning that the regional climate trends were well represented, rather than local variations. This is especially clear in the southern part of Norway where local patterns are seen in the non-homogeneous map. Tveito et al. (2020) concludes that the maps based on homogenized data are clearly more trustworthy in explaining the large-scale climate variations that should explain the change of "normal" climatologies.

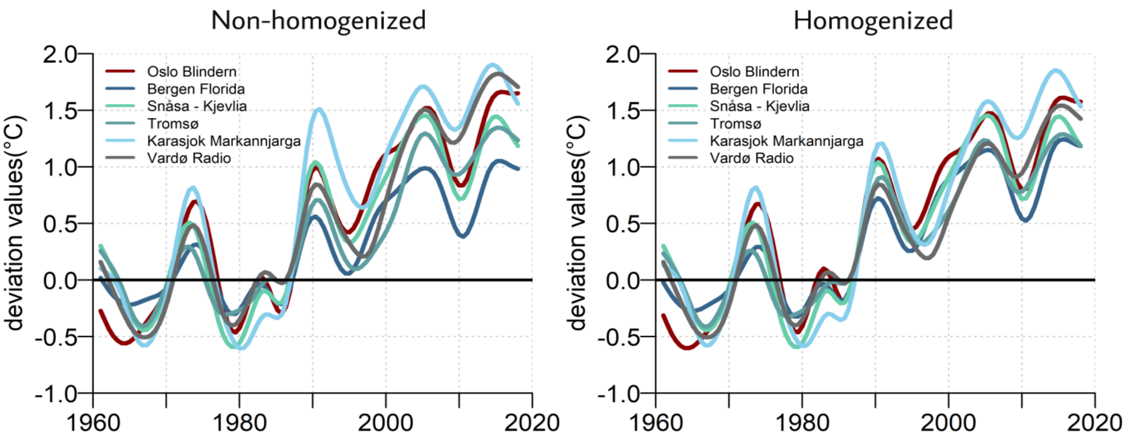


Fig. 4.: Deviation of the annual temperature with respect to the 1961-1990 mean before (left) and after homogenization (right). The anomaly series have been filtered using a 5-year Gaussian density function.

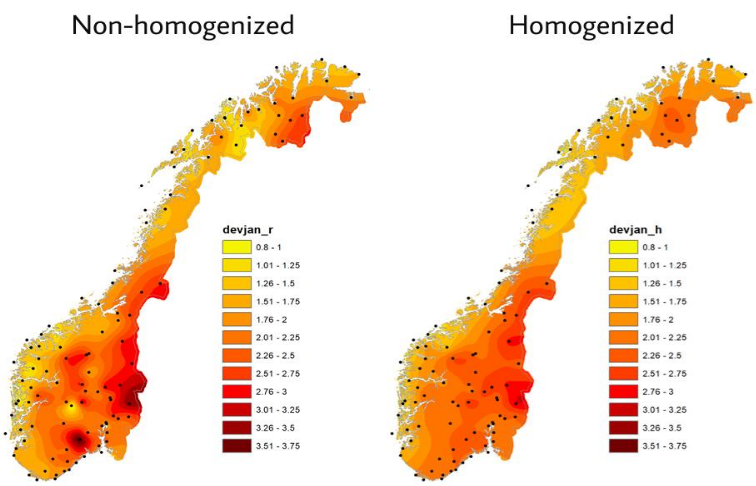


Fig. 5.: Differences between the 1961-1990 and 1991-2020 mean monthly January temperature. Non-homogenized data (left) and homogenized data (right) (Tveito et al. 2020).

4. CONCLUSIONS

A 58-year long homogenous dataset for 108 Norwegian monthly mean temperature series was produced. This dataset will serve as a robust basis for further homogenization, interpolation and calculation of new normals for series with poorer data cover. Homogenization proved to be beneficial for spatial coherence and describing regional climate characteristics and development. The homogenized series are more reliable in explaining the large-scale climate variations that should explain the change of "normal" climatologies.

References

Aguilar E, Auer I, Brunet M, Peterson TC, Wieringa J (2003) Guidelines on climate metadata and homogenization. WMO-TD No. 1186, WCDMP No. 53. World Meteorological Organization, Geneva.

Andresen, L.: 2011 Homogenization of monthly long-term temperature series of mainland Norway. Met.no Note May 2011.

Hanssen-Bauer, I. and Nordli, P.Ø. 1998. Annual and seasonal temperature variations in Norway 1876-1997. DNMI Report No. 25/1998 Klima.

Kuya E.K., Gjelten H M., Tveito O.E. 2020. Homogenization of Norway's mean monthly temperature series. MET Report 3/2020.

Lindau R., Venema V. 2016. The uncertainty of break positions detected by homogenization algorithms in climate records. Int J Climatol 36: 576−589

Lundstad, E. 2016. Homogenization of daily precipitation in Norway, MET Report 12/2016 Climate

Lundstad, E., Tveito O.E. 2016: Homogenization of daily mean temperature in Norway. DNMI Report No. 6/2016 Klima

Mestre O., Domonkos P., Picard F., Auer I., Robin S., Lebarbier E., Böhm R., Aguilar E., Guijarro J., Vertachnik G., Klancar M., Dubuisson B., Stepanek P. 2013. HOMER: a homogenization software - methods and applications. Idojaras-Quarterly Journal of the Hungarian Meteorological Service. 117, 47-67.

Skelton A., Kirchner N., Kockum I. 2020. Skewness of temperature data implies an abrupt change in the climate system between 1985 and 1991. Geophysical Research Letters 47 (23). https://doi.org/10.1029/2020GL089794

Tveito O.E., Aņiskeviča S., Cappelen J., Engström E., Gjelten H.M., Jensen C.D., Jokinen P., Kuya E.K., Lussana C., Mäkelä A., Mändla K., Vint K., Wern L., Zandersons V. 2020. ClimNorm - temperature data set, gap filling methods and regional analysis to prepare new climate normals. MET Report 4/2020.

WMO 2011. Guide to Climatological Practices, Vol.100. World Meteorological Organization: Geneva, Switzerland

WMO 2017. WMO Guidelines on the Calculation of Climate Normals. WMO-No. 1203 Geneva, Switzerland

JOINT HOMOGENIZATION OF TIME SERIES WITH UNEQUAL LENGTH BY APPLYING THE MASH PROCEDURE

Beatrix Izsák [1*, 2], Mónika Lakatos [1], Rita Pongrácz [4], Tamás Szentimrey [3] and Olivér Szentes [1]

[1]Hungarian Meteorological Service, Kitabel P. s 1, H-1024, Budapest, Hungary
[2] ELTE Faculty of Science, Doctoral School of Earth Sciences Budapest, Hungary
[3] Varimax Limited Partnership Budapest, Hungary
[4] ELTE Institute of Geography and Earth Sciences, Department of Meteorology Budapest, Hungary
izsak.b@met.hu

Abstract

The Hungarian Meteorological Service (HMS) is celebrating its 150th anniversary this year. Thanks to the continuous recording of archive data, new data was added to the database. These should be checked and homogenized for the period 1871-1900 before being subjected to climatic analyses.
Homogenization of the data series raises the problem that how to homogenize together the long and short data series, since the meteorological observation system was upgraded substantially in the last decades. It is possible to solve these problems with method MASH (Multiple Analysis of Series for Homogenization, Szentimrey) due to its adequate mathematical principles for such purposes. When the station network is upgraded and we have short data series besides the long series, the common section must be homogeneous together with the long as well as with the short data series in addition these two or more systems have to be homogeneous themselves too. MASH is able to fulfill these criteria, as it is based on hypothesis testing and it involves an iteration procedure. The solution is that we synchronize the common part's inhomogeneities within two or more different MASH processing for the two or more datasets with different length.

1. INTRODUCTION

In recent years, as the importance of climate change research has grown, more and more databases for climatological purposes have emerged. These are mainly based on measurements, but the methodology itself can be quite different. Raw data sets contain errors, significant inhomogeneities are found in them, and missing data must be replaced. To get to know the climate of recent periods, we need a representative database in space and time based on measurements (Izsák and Szentimrey, 2020).
In order to have a spatially and temporally representative database, the first step is to homogenize, check, and complete missing values in the station data series. Since we have data series of different lengths this is not such an easy task since in order to create representative database we have to use as much data and as long data series as possible!

2. THE SOFTWARE MASHv3.03

Changing measurement conditions, such as station relocation, change in measuring time, or instrument change, may result in undue fracture in time series. At the Hungarian Meteorological Service (HMS), data errors and inhomogeneities are eliminated and data gaps are filled in using the MASH (Multiple Analysis of Series for Homogenization; Szentimrey, 1999, 2008, 2017) homogenization procedure. What kind of software is employed for homogenization is of great importance, because if not just inhomogeneities are removed from the data series, but also the process unintentionally alters the signal of climate change, the result will be misleading. Thanks to the mathematical model, using the MASH software, it is possible to detect climate change in the homogenized data set.

2.1 THE MAIN PROPERTIES OF THE VERSION MASHV3.03

Advantages of MASHv3.03 in the homogenization of monthly series:

– It is a relative homogeneity test procedure.
– **It is a step-by-step iteration procedure**: the role of series (candidate, reference) changes step by step in the course of the procedure.
– An additive (e.g. temperature) or multiplicative (e.g. precipitation) model can be used depending on the distribution.
– It includes quality control and missing data completion.
– It provides the homogeneity of the seasonal and annual series as well.
– Metadata (probable dates of break points) can be used automatically.
– **The homogenization results can be evaluated on the basis of verification tables generated automatically during the procedure.**

In the homogenization of daily series:

– The procedure is **based on the detected monthly inhomogeneities.**
– It includes quality control and the completion of missing data in daily data.

The MASH procedure for daily series (Szentimrey, 2008, 2013):

1. Monthly series from daily series.
2. MASH homogenization procedure for monthly series, estimation of monthly inhomogeneities.
3. On the basis of estimated monthly inhomogeneities, smooth estimation for daily inhomogeneities.
4. Homogenization of daily series.
5. Quality control for homogenized daily data.
6. Missing daily data completion.
7. Monthly series from homogenized, quality-controlled, completed daily data.
8. Test of homogeneity for the new monthly series by MASH.

(Repeating steps 2-8 if it is necessary.)

2.2 THE VERIFICATION STATISTICS IN MASH

The test statistics generated automatically during the procedure:

TEST STATISTICS FOR SERIES INHOMOGENEITY

- *Test statistics after homogenization*
- *Test statistics before homogenization*
- *Statistics for estimated inhomogeneities*

CHARACTERIZATION OF INHOMOGENEITY

- *Relative estimated inhomogeneities*
- *Relative modification of series*
- *Lower confidence limit for relative residual inhomogeneities*

REPRESENTATIVITY OF STATION NETWORK

EVALUATION OF META DATA

- *Test statistics*
- *Representativity of META data*

2.3 HOMOGENIZATION OF SPATIALLY AND TEMPORALLY EXPANDED STATION SYSTEMS

When the station network is upgraded and we have short data series besides the long series, the common section must be homogeneous together with the long as well as with the short data series, while the two or more systems have to be homogeneous themselves too. MASH is able to fulfill these criteria, as it is based on hypothesis testing and it involves an iteration procedure (Szentimrey, 1999, 2017). Because MASH is an iteration procedure, the series are examined and adjusted many times, therefore the homogenization of the new system can be considered as a continuation of the earlier homogenization procedure. The test of hypothesis and throughout this test, the test statistics enable us to use the former results.

The solution is that we synchronize the common part's inhomogeneities within two or more different MASH processing for the two or more datasets with different length.

We have harmonized two MASH systems in recent years. In case of daily precipitation sum, the new MASH consists of data series of 11 stations starting from 1871 as a result of the digitization effort made for the 150th anniversary. Therefore, we harmonize three MASH systems for the precipitation database. This year, based on the renewed daily average temperature database, the task is to homogenize four MASH systems together. We first note that the selection of station systems is also a difficult task, as we have to look for discontinued stations close to the automatic measuring stations launched a few years ago, or even to find a continuation data series for stations that have been discontinued for decades but had 80-100 years old time series.

In order to update our database annually or to homogenize several MASH systems together, test statistics must be studied at each step and the homogenization can be continued based on these, or it can be decided that the overall homogeneity of the station network is acceptable at a given significance level.

2.4 QUALITY CONTROL FOR DAILY DATA

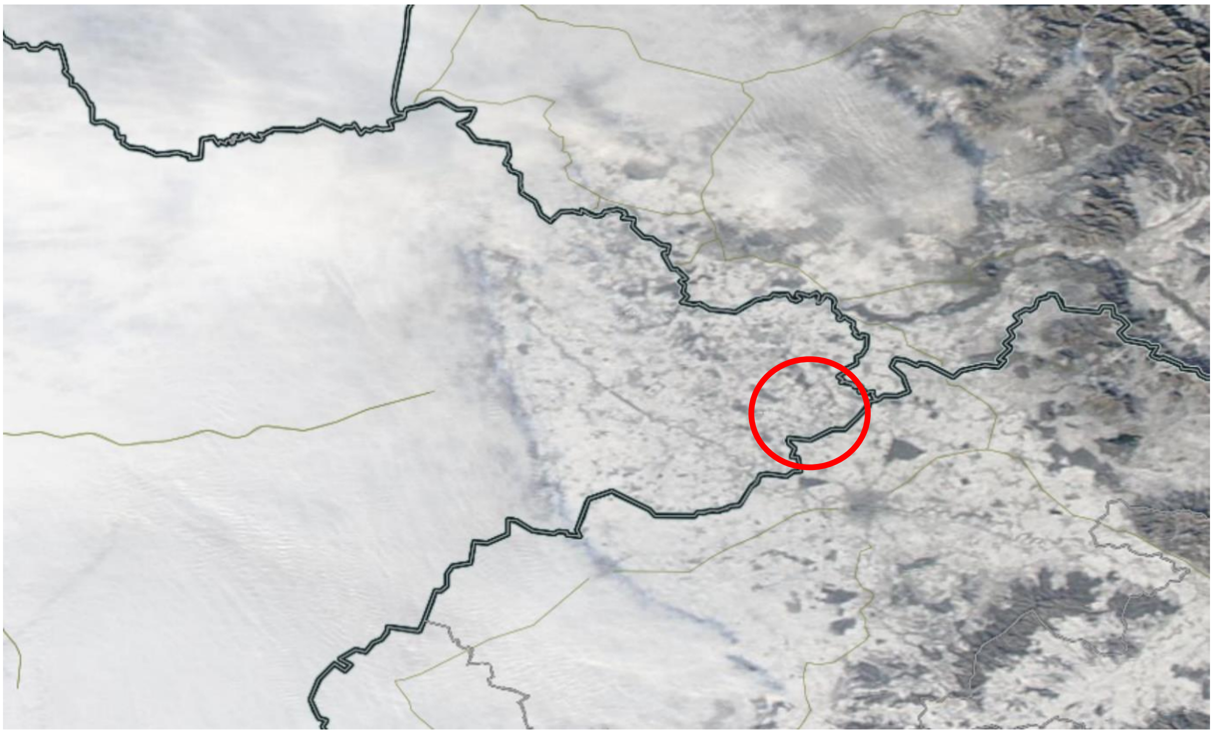


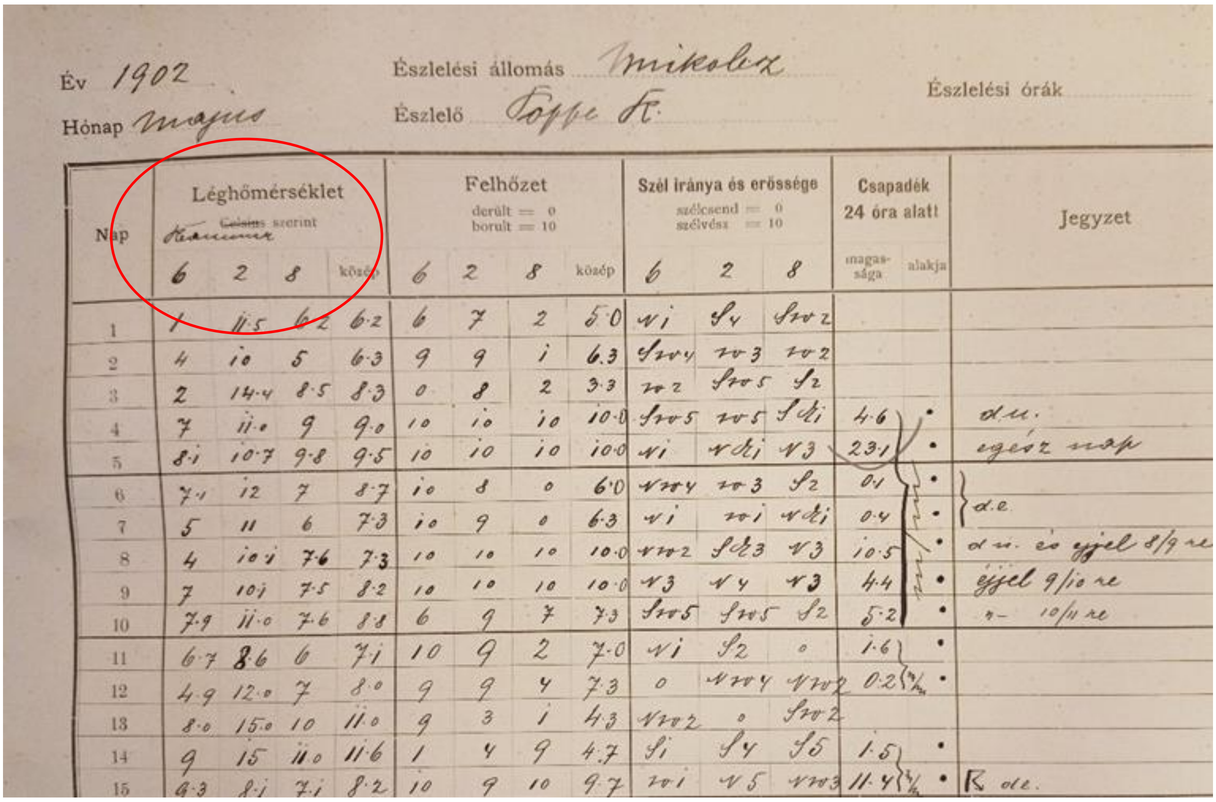
Fig. 1.: Satellite image, 05.12.2019, in the area marked with a red circle, a very low daily average temperature was measured compared to the nearest stations

The data check section of MASH is very useful not only for checking archive data, but also for a comprehensive examination of daily data for the entire period. When checking the data of the past years, we can more easily examine the suspicious data marked in the MASH's automatically generated error.res file, as we have data of radar, satellite data sets or other information. For example, in the case of temperature, if a measurement point has a much lower value than the nearest stations, looking at the satellite image, it is clear that this is possible due to the lack of cloud and snow cover, so the data in error file is actually extreme and not an error (Fig. 1). And there are also cases where the value in the error file is indeed erroneous, from a meteorologist's point of view, nothing justifies a large discrepancy based on the comparison with neighbouring stations. There are many such examples from the present, the more difficult task is to examine the data of 100 and more years ago, as there were no satellites or radars at that time. There are cases where statistical calculation indicates an error but the meteorologist is unable to make a decision due to the small amount of information. For instance, if the nearest stations (distance ≥ 100km, as the station network was even less dense 150 years ago) show 20 ºC and the given station shows 10 ºC, then this is possible in the Carpathian Basin (e.g. due to a cold front), but it may also be a measurement error. (Mathematical procedures cannot distinguish between extreme values and data errors.) Of course, recording errors can also be easily found by MASH, in which case we can correct the data based on the annual books. In several cases, we found errors of varying magnitude and sign for the entire month, and we now know that in all of these cases, data for the previous or next month were recorded. Typical errors are the sign error and the absence of the decimal point. It can also be seen from these examples that professional analyses are needed after the automatic data verification and the two together ensure a comprehensive examination. The advantage of quality control in MASH is definitely that only a few percent of the millions of data need to be subjected to further analysis, and if we have checked past data, we do not have to re-examine them again, only the new year or years have to be examined, which is only less than 10 such examinations per year per meteorological variable.

2.5 QUALITY CONTROL FOR MONTHLY DATA

In the MASH system, errors in the monthly data are displayed as outliers. However, it is also worth exploring large inhomogeneities over several months or even years. Here are a few examples.

As can be seen from the two graphs on page 42 of the MASH manual (Szentimrey, 2017), Miskolc station has high inhomogeneity values for the period 1901-1908. We found an explanation for these about 20 years later, by investigating them in the archives, because at that



Év 1902 Észlelési állomás Miskolcz Észlelési órák

Hónap május Észlelő Poppe K.

Nap	Léghőmérséklet ~~Celsius~~ Reaumur szerint				Felhőzet derült = 0 borult = 10				Szél iránya és erőssége szélcsend = 0 szélvész = 10			Csapadék 24 óra alatt		Jegyzet
	6	2	8	közép	6	2	8	közép	6	2	8	magassága	alakja	
1	1	11.5	6.2	6.2	6	7	2	5.0	d1	é4	ény2			
2	4	10	5	6.3	9	9	1	6.3	ény4	ny3	ny2			
3	2	14.4	8.5	8.3	0	8	2	3.3	ny2	ény5	é2			
4	7	11.0	9	9.0	10	10	10	10.0	ény5	ny5	ék1	4.6	•	d.u.
5	8.1	10.7	9.8	9.5	10	10	10	10.0	d1	dk1	d3	23.1	•	egész nap
6	7.1	12	7	8.7	10	8	0	6.0	dny4	ny3	é2	0.1	•	d.e.
7	5	11	6	7.3	10	9	0	6.3	d1	ny1	dk1	0.4	•	
8	4	10.1	7.6	7.3	10	10	10	10.0	dny2	ék3	d3	10.5	•	d.u. és éjjel 8/9-re
9	7	10.1	7.5	8.2	10	10	10	10.0	d3	d4	d3	4.4	•	éjjel 9/10-re
10	7.9	11.0	7.6	8.8	6	9	7	7.3	ény5	ény5	é2	5.2	•	" 10/11-re
11	6.7	8.6	6	7.1	10	9	2	7.0	d1	é2	0	1.6	•	
12	4.9	12.0	7	8.0	9	9	4	7.3	0	dny4	dny2	0.2	•	
13	8.0	15.0	10	11.0	9	3	1	4.3	dny2	0	ény2			
14	9	15	11.0	11.6	1	4	9	4.7	é1	é4	é5	1.5	•	
15	9.3	8.1	7.1	8.2	10	9	10	9.7	ny1	d5	dny3	11.4	•	K de.

Fig. 2.: The observations of May 1902 at Miskolc station. It is clear that instead of the ° C already widely used at that time, the air temperature values are given on the Réaumur scale. (Source: HMS)

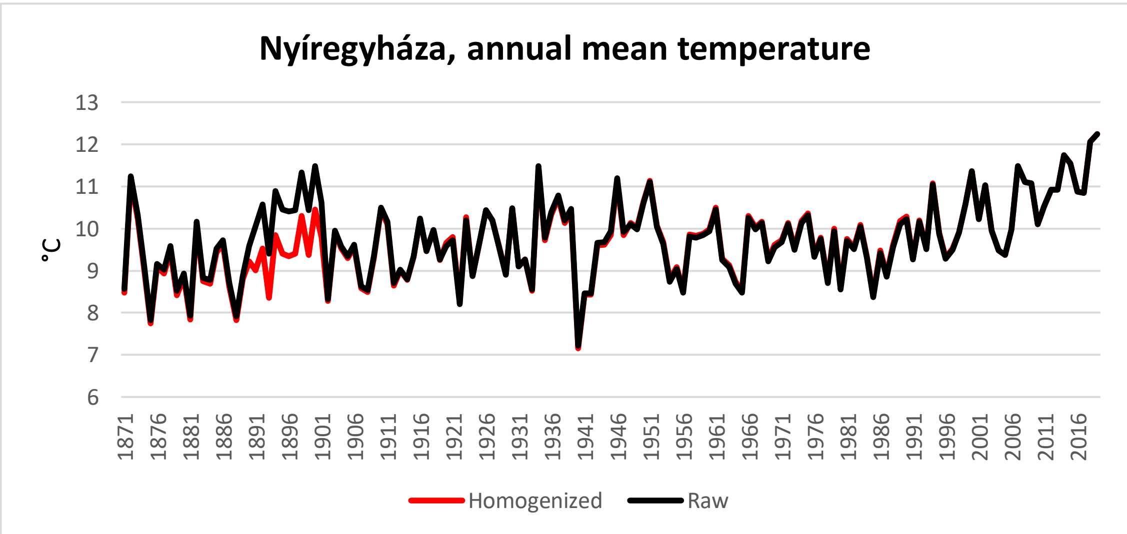


Fig. 3.: Annual average temperature values calculated from raw and homogenized data series at Nyíregyháza station, 1871-2019

time they were not measured in Celsius, but in Réaumur scale (Fig. 2). When checking the

archive precipitation data, we found the same discrepancies due to the fact that the unit of measurement was different, e.g. Paris line or inch. In these cases, converting the basic data to the appropriate unit, we get valid data. The significant monthly and annual inhomogeneities detected at the Nyíregyháza station can also be easily explained, as we found in the annual books that the observation time took place 1 hour later in the morning and 1 hour earlier in the evening between 1890-1901, so the high inhomogeneity found by MASH does not indicate erroneous data (Fig. 3). These are inhomogeneities, but these resulted in much higher daily averages than if they had been detected at the standard time.
It also shows that the models based on the classical mathematical theories that form the basis of the MASH system work very efficiently, but human intervention is also needed!

2.6 AUTOMATIC USE OF META DATA IN MASH

One of the great benefits of MASH is that it handles META (station history information) data automatically, so we created these META files as well. However, we know that these are often incomplete and we also know that not all changes cause inhomogeneity. The decision mechanism in MASH ensures that breakpoints are indeed detected (Szentimrey, 2017), the user can choose between basic, strict and light versions and evaluates the META data with automatically generated statistics (Szentimrey, 2017). In the case of the three longer systems the representativity values of the META data are below 0.4, in the case of the shortest system they are less than 0.5, but even in this case we cannot say that the inhomogeneities themselves can be well explained with this information. However, if the annual breakpoints are searched for and treated as metadata file, the verification statistics will improve significantly (around 0.8), but still there are stations that cannot be explained with META data at all and there are several stations where all breakpoints can be explained with META data.

3. HOMOGENIZATION OF DAILY DATA SERIES BY MASH

3.1 JOINT HOMOGENIZATION OF TIME SERIES WITH UNEQUAL LENGTH IN CASE OF AVERAGE TEMPERATURES

We use the MASH system at the Climate Department of the HMS to homogenize and complete the daily station data sets and it also includes a significant data check. As the database is constantly updated, not only with the data of the past year, but also with the archive data being continuously digitized, it is also necessary to complete and verify them. Our goal is to use as many station data sets as possible for climatological analyses each year. So far, in the case of average temperatures, we used the data of 25 stations from 1901 to 2019, and from 1971 to 2019 for further 33 stations, i.e. a total of 58 stations, to compile our database for climatological purposes.

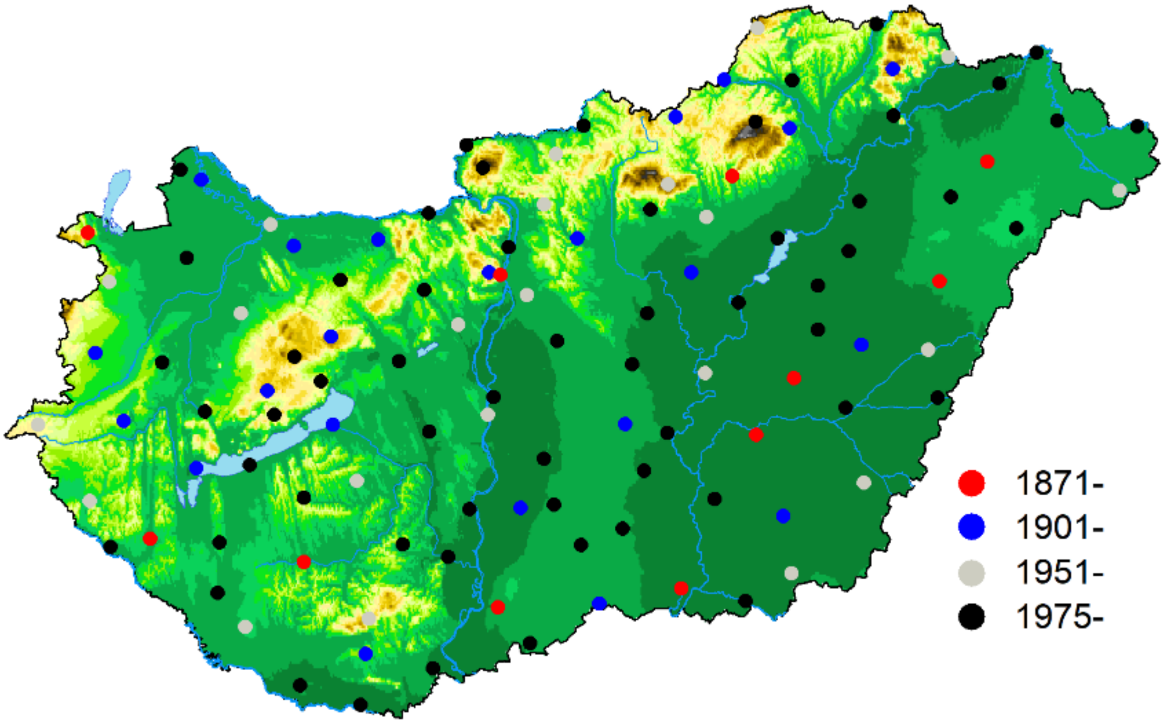


Fig. 4.: Location of the stations in case of temperature

This year, we re-reviewed our station system and significantly expanded it (Fig. 4). From 1871, data sets from 11 stations were checked, completed, and homogenized. For the period starting from 1901, we use the data of 22 stations more, so from 1901 our homogenized, completed and controlled database is based on the long data series of 33 stations. The next time step was chosen to be 1951, when the number of available station data sets was significantly expanded, so we use 22 further stations, i.e. 55 station data sets altogether. The shortest period starts from 1975, when we already use the data series of 110 stations. (Note here that we also use data from 4 other stations, which have too many missing data to be included in the short system, but their geographical location justifies taking their data into account. We will discuss this issue later.) So, our task was to homogenize the data sets of the four MASH systems together, i.e. the common part should be homogeneous in each system.

3.1.1 Homogenization steps using four MASH systems

The step zero is to compile the station systems, since we cannot predict in advance how much missing data the MASH will work with, i.e. when the data series will be dependent. The methodology of the selection of stations to include the processing or not is also a complex task, as data gaps, relocation and closure of stations are common. Thus, in many cases, it is necessary to concatenate the data of a station closed with the data of a station located 10-20 km apart. Such as Zirc and Tés, or Salgótarján and Zabar for instance. All stations are important to include in both examples due to their locations at higher elevation as the station network is sparce in Bakony and in Börzsöny, respectively (the mountainous area where these stations are located). It is necessary to select between the stations in many cases to end up with a coherent system. The stations where the missing data are utmost have to be left out from the process. The systems drawn in Fig. 5 were used in the homogenization process.

The steps of homogenization of temperature station data series are as follows:

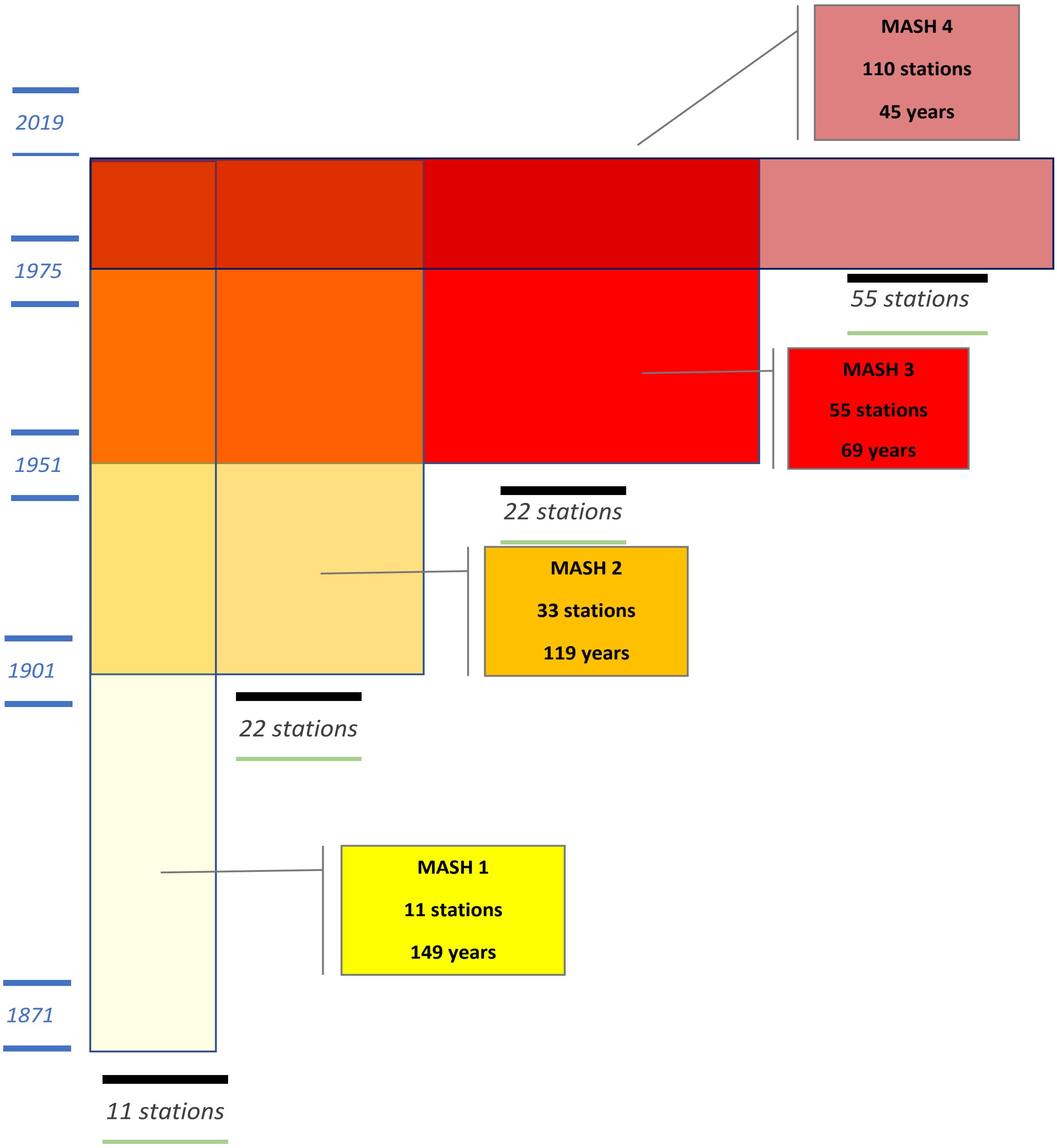


Fig. 5.: The scheme of the four different station systems for the MASH procedure in the case of temperature

1. *MASH1: homogenization of monthly data*
2. *Cut out the inhomogeneities of the common part and insert it into the other three MASHs*
3. *MASH2: homogenization of monthly data*
4. *Cut out the inhomogeneities of the common part and insert them into the other three MASHs*
5. *MASH3: homogenization of monthly data*
6. *Cut out the inhomogeneities of the common part and insert them into the other three MASHs*
7. *MASH4: homogenization of monthly data*
8. *Cut out the inhomogeneities of the common part and insert them into the other three MASHs*
9. *If statistics are acceptable in MASH1: go to point 10, if not, go to step 1.*

10. If statistics are acceptable in MASH2: go to point 11, if not, go to step 3.
11. If statistics are acceptable in MASH3: go to point 12, if not, go to step 5.
12. Homogenization of daily data in MASH1, MASH2, MASH3, MASH4.
13. Gathering the homogenized data sets from the different MASH systems.

Table 1.: The most important verification statistics in case of temperature

MASH	1871-2019	1901-2019	1951-2019	1975-2019
Significance level: 0.05	Critical value: 22.05	Critical value: 21.76	Critical value: 21.31	Critical value: 20.86
Test Statistics Before Homogenization	AVERAGE: 1275.65	AVERAGE: 780.44	AVERAGE: 370.29	AVERAGE: 303.26
Test Statistics After Homogenization	AVERAGE: 31.75	AVERAGE: 26.83	AVERAGE: 22.64	AVERAGE: 19.73
Relative Modification of Series	AVERAGE: 0.46	AVERAGE: 0.53	AVERAGE: 0.41	AVERAGE: 0.39
REPRESENTATIVITY OF STATION NETWORK	AVERAGE: 0.86	AVERAGE: 0.88	AVERAGE: 0.90	AVERAGE: 0.90

Summarizing the steps 1-12: we homogenized the data series from the longest to the shortest ones. In this case, two runs provided good results as after inserting the common inhomogeneities the test statistics no longer increased.
Overall, we can say that the co-homogenization was successful, the data sets can be considered homogeneous at the 0.05 level of significance. The most important verification statistics are summarized in Table 1.

3.1.2 Some examples of manual correction

Here, we note that we took advantage of the possibilities of the interactive program system, since we needed manual correction in many cases. The subroutines and subdirectories built into the MASH provide this option. (This, in turn, requires a thorough study of the mathematical background.) As mentioned earlier, we have added several stations to the new system that have a significant lack of data, in which case the completion of missing values will take place from neighboring stations. The test statistics after homogenization value can be very high without having an actual breakpoint for that data set. In this case, the built-in graphics program, the mashex1-2 subroutines, can help you decide if to look for an additional breakpoint. It may also be necessary to manually exclude a particular station, e.g. which we completed from the candidate series, in order to find the new breakpoint. We have the opportunity to do this with the mashgame program and in many cases we have found the breakpoint. However, you may also need to manually delete a breakpoint with mashcor program. For example, if the test statistics did not change before and after homogenization, however, the relative modification of series value increased significantly. The principle of MASH is to homogenize the station data series so as to make the slightest change in the data. In months when the relative modification of series value is very high on a station, it is always worth looking at what happened.
In our work, we omitted two stations from the older system that were homogenized, but the 80–100-year-old coherent data set has been discontinued for more than 20 years and there is no continuation (e.g. Mencshely and Hárskút). Two more stations only have data sets from the last 10 years but these are also in locations where data are highly needed. Thus, at the very end of the homogenization, these were inserted next to the homogenized data sets and the verification statistics were regenerated. This year, these did not increase and we did not receive any indication of inhomogeneity at the given stations, so we completed them based on the neighbouring data series and thus finally the data set of 114 stations is included in the homogenized station database.

Of course, if the statistics for these stations increase in the coming years, we will have to examine them thoroughly, as a new breakpoint may appear after new data added.

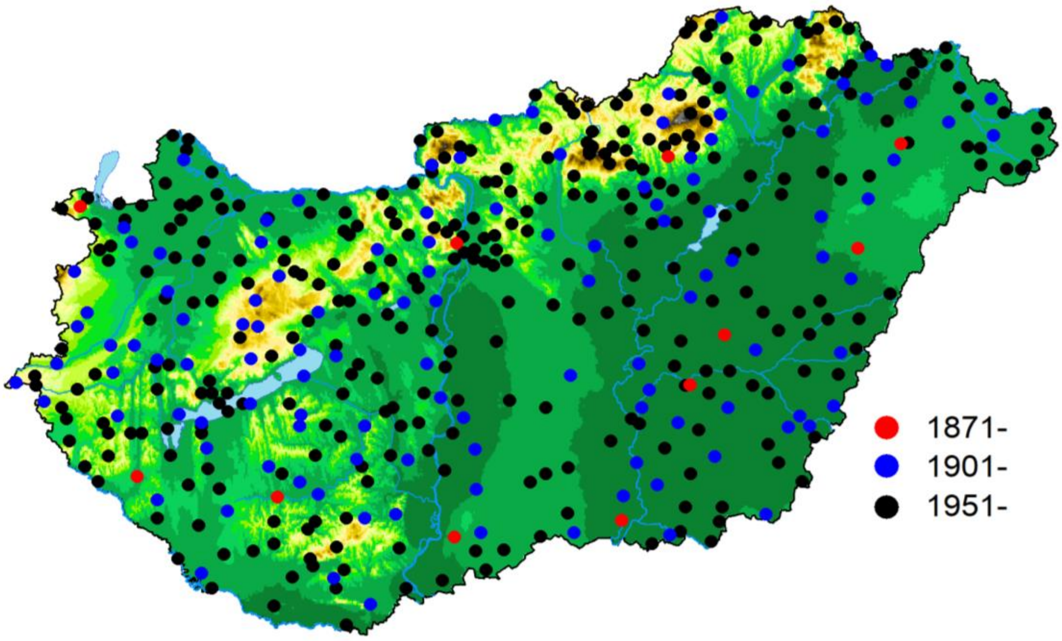


Fig. 6.: Location of the stations in the case of precipitation

3.2 JOINT HOMOGENIZATION OF TIME SERIES WITH UNEQUAL LENGTH IN CASE OF PRECIPITATION

The task to fulfill is much simpler in the case of precipitation than temperature as the list of precipitation stations that could be used for homogenization and gridding remained unchanged compared to previous years (Fig. 6). The only new issue comes from the inclusion of the recently digitized data from 1871-1900 into the homogenization and gridding process. The quality control, data homogenization and data completion are the first steps to execute here as well. Three MASH systems were built and harmonized for precipitation time series. Fig. 7 illustrates the periods and the number of stations that were used in this process.

The steps of homogenization of precipitation station data series are as follows:

1. *MASH1: homogenization of monthly data*
2. *Cut out and insert the inhomogeneities of the common part into MASH2 and MASH3*
3. *MASH2: homogenization of monthly data*
4. *Cut out the inhomogeneities of the common part and insert them into MASH3 and MASH1*
5. *MASH3: homogenization of monthly data*
6. *Cut out the inhomogeneities of the common part and insert them into MASH1 and MASH2*
7. *If statistics are acceptable in MASH1: go to 8, if not, go to step 1.*
8. *If the statistics are acceptable in MASH2: go to 9, if not, go to step 3.*
9. *Homogenization of daily data in MASH1, MASH2 and MASH3*

10. Gathering the homogenized data sets from the different MASH systems.

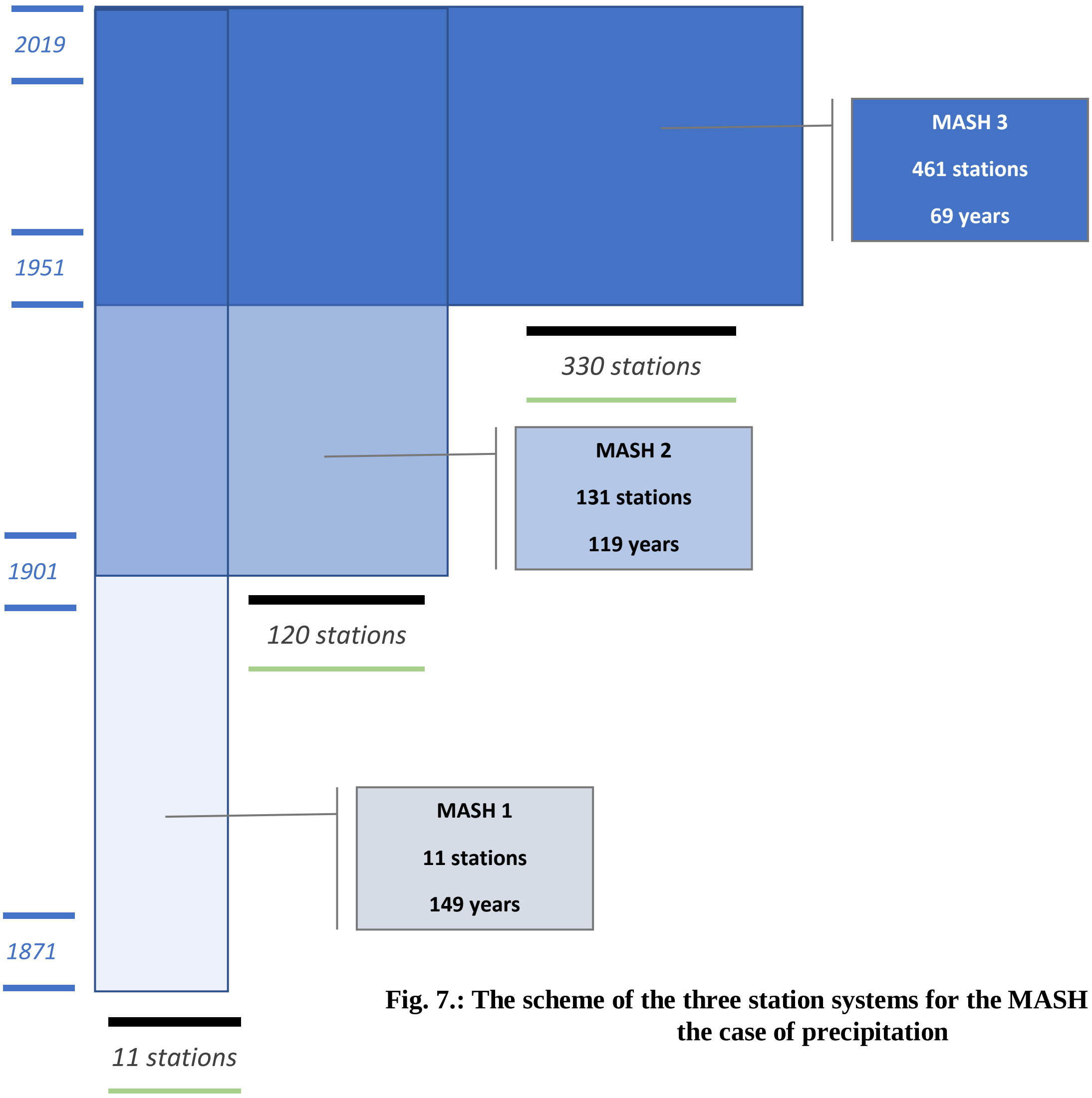


Fig. 7.: The scheme of the three station systems for the MASH procedure, in the case of precipitation

Table 2.: The most important verification statistics for daily precipitation

MASH	1871-2019	1901-2019	1951-2019
Significance level: 0.01	Critical value: 28.00	Critical value: 28.00	Critical value: 30.00
Test Statistics Before Homogenization	AVERAGE: 45.46	AVERAGE: 63.42	AVERAGE: 42.52
Test Statistics After Homogenization	AVERAGE: 18.79	AVERAGE: 27.48	AVERAGE: 30.22
Relative Modification of Series	AVERAGE: 0.23	AVERAGE: 0.19	AVERAGE: 0.09
REPRESENTATIVITY OF STATION NETWORK	AVERAGE: 0.46	AVERAGE: 0.63	AVERAGE: 0.70

The most important verification statistics are listed in Table 2. Obviously, the available data from 461 stations bring an improvement over the results from 131 stations only. Not surprisingly, the representativity values of station system consisting of 11 data series only are very low. The reason is that the precipitation varies more spatially and temporally unlike the average temperature, therefore much denser network of stations is needed to characterize well the distribution of precipitation.
Similarly to temperature, we do not recommend the mere use of automatic algorithms in the case of precipitation either, even though in this case we are much more careful and use a significance level of 0.01. In each case, therefore, it is necessary to study the test statistics and, on this basis, the homogenized station data series is prepared. The great advantage of MASH is that we can correct not only automatically but also manually in almost every step, however of course, the biggest help is the study of verification statistics. This is an opportunity to really talk about updating every year without having to start homogenizing again from the beginning.

4. SUMMARY

Homogenized, completed and quality-controlled station datasets were derived from the daily average temperature and daily precipitation sums for Hungary for different time periods. One of the most important achievements is that 149-years-long climate data series were homogenized for Hungary. Three MASH systems for the three different time intervals were harmonized for precipitation and four MASH systems for the daily mean temperature. In the case of temperature, the daily data of 11 meteorological stations from 1871, 33 from 1901, 55 from 1951, and 114 from 1975 were used in this process. Daily precipitation sums of 11 stations from 1871, 131 from 1901, and 461 from 1951 were quality-controlled, homogenized and completed. Significant quality improvement was achieved by expanding the station system in the case of temperatures. We used unprocessed archived data to creation of the datasets for temperature and precipitation, too. This will allow to study the climate change in Hungary over a longer period of time than earlier.
Our future plan is the updating of the model variables used for interpolation. Thanks to this work that we have done, there is opportunity to renew the model variables used for interpolation of the average temperature time series. Fewer stations were applied for modeling previously, but the modeling can be performed by using long time series of 114 stations by now. The new station system described in this paper is suitable for the regular update of the daily maximum and minimum temperature datasets. The modeling part of MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; Szentimrey and Bihari, 2007, 2014) is suggested to re-run with the updated time series on a regular basis. The methodology of homogenization together with three or more MASH systems detailed here made possible to use as many measurements as possible to produce representative datasets for other meteorological elements as well.
The research is supported by the ÚNKP-20-3 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Found.

ÚNKP Új Nemzeti Kiválóság Program

National Research, Development and Innovation Office Hungary

Innovációs és Technológiai Minisztérium

References

Izsák, B., Szentimrey, T., 2020: To what extent does the detection of climate change in Hungary depend on the choice of statistical methods? Int J Geomath 11, 17 https://doi.org/10.1007/s13137-020-00159-7

Szentimrey, T., 1999: Multiple Analysis of Series for Homogenization (MASH), Proceedings of the Second Seminar for Homogenization of Surface Climatological Data, Budapest, Hungary; WMO, WCDMP-No. 41, pp. 27-46.

Szentimrey, T., Bihari, Z., 2007: Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis), Proceedings of the Conference on Spatial Interpolation in Climatology and Meteorology, Budapest, Hungary, 2004, COST Action 719, COST Office, 2007, pp. 17-27

Szentimrey, T., 2008: Development of MASH homogenization procedure for daily data. Proceedings of the Fifth Seminar for Homogenization and Quality Control in Climatological Databases, Budapest, 2006, WCDMP-No. 71, WMO/TD-NO. 1493, 123–130.

Szentimrey, T., 2013: Theoretical questions of daily data homogenization, Időjárás Vol. 117. No. 1, January-March 2013. pp. 113-122.

Szentimrey, T., Bihari, Z., 2014: Manual of interpolation software MISHv1.03, Hungarian Meteorological Service, p. 60.

Szentimrey, T. 2017: Manual of homogenization software MASHv3.03, Hungarian Meteorological Service, p.71.

Venema, V. K. C. et al., 2012: Benchmarking homogenization algorithms for monthly data, Climate of the Past, 8, 89-115

World Meteorological Organization (WMO), 2020: Guidelines on Homogenization, WMO-No. 1245, p. 63.

MATHEMATICAL QUESTIONS OF SPATIAL INTERPOLATION AND SUMMARY OF MISH

Tamás Szentimrey

Varimax Limited Partnership, Budapest, Hungary
szentimrey.t@gmail.com

Abstract

We focus on the basic mathematical and theoretical questions of spatial interpolation of meteorological elements. Nowadays in meteorology the most often applied procedures for spatial interpolation are the geostatistical interpolation methods built also in GIS software. The mathematical basis of these methods is the geostatistics that is an exact but special part of the mathematical statistics. However special meteorological spatial interpolation methods for climate variables also can be developed on the basis of the mathematical statistical theory. The main difference between the geostatistical and meteorological interpolation methods can be found in the amount of information used for modelling the necessary statistical parameters. In geostatistics the usable information or the sample for modelling is only the predictors, which are a single realization in time. While in meteorology we have spatiotemporal data, namely the long data series which form a sample in time and space as well. The long data series is such a specialty of the meteorology that makes possible to model efficiently the statistical parameters in question. The planned topics to be discussed are as follows.

- Interpolation formulas and loss functions depending on the spatial probability distribution of climate variables.
- Estimation and modelling of climate statistical parameters (e.g.: spatial trend, covariance or variogram) for interpolation formulas using spatiotemporal sample and supplementary model variables (topography). Use of background information (e.g.: dynamical model results, satellite, radar data) for spatial interpolation.

The earlier versions of our method MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; Szentimrey and Bihari) were developed formerly at the Hungarian Meteorological Service. At MISH method we use spatiotemporal data for modelling the climate statistical parameters and the interpolation system is based on these results. The earlier modelling system was elaborated for the monthly and daily expected values and the spatial correlations. At the new version MISHv2.01 the monthly and daily standard deviations and the daily temporal correlations also can be modelled. Consequently the modelling subsystem of MISH is completed for all the first two spatiotemporal moments on monthly and daily scales. If the joint spatiotemporal probability distribution of the given variable is normal then the above spatiotemporal moments determined uniquely this distribution that is the mathematical model of the climate. Another developments are modelling of the interpolation error RMSE (Root Mean Square Error) in order to characterize quantitatively the uncertainty of the interpolation, furthermore real time Quality Control for daily and monthly data. We will present a summary of the method MISH.

1. INTRODUCTION

I retired from the Hungarian Meteorological Service two years ago but I continue my activity in my VARIMAX Limited Partnership. This activity includes the development of advanced mathematics for meteorology as well as the development of efficient software on the basis of the mathematical results.

Concerning our topic we have the following question. What kind of mathematics of spatial interpolation is adequate for meteorology? Nowadays the geostatistical interpolation methods built in GIS software are applied in meteorology, e.g. the various kriging methods. The mathematical basis of these methods is the geostatistics that is an exact but special part of the mathematical statistics. The specialty is connected with the assumption that the data are purely spatial ones. Consequently, as we see it, the geostatistical methods cannot efficiently use the spatiotemporal data like the meteorological data series while the data series make possible to obtain the necessary climate information for the interpolation in meteorology.

Modelling of the climate statistical parameters is a key issue to the interpolation of meteorological elements and modelling can be based on the long data series. However the data series are usually affected by inhomogeneities (artificial shifts), due to changes in the measurement conditions (relocations, instrumentation) therefore the direct analysis of the raw data series can lead to wrong conclusions. In order to deal with this crucial problem many statistical homogenization procedures have been developed for detection and correction of these inhomogeneities. Similarly to the connection of interpolation and homogenization in our conception the meteorological questions cannot be treated separately. Therefore we present a block diagram (*Fig. 1*) to illustrate the possible connection between various important meteorological topics.

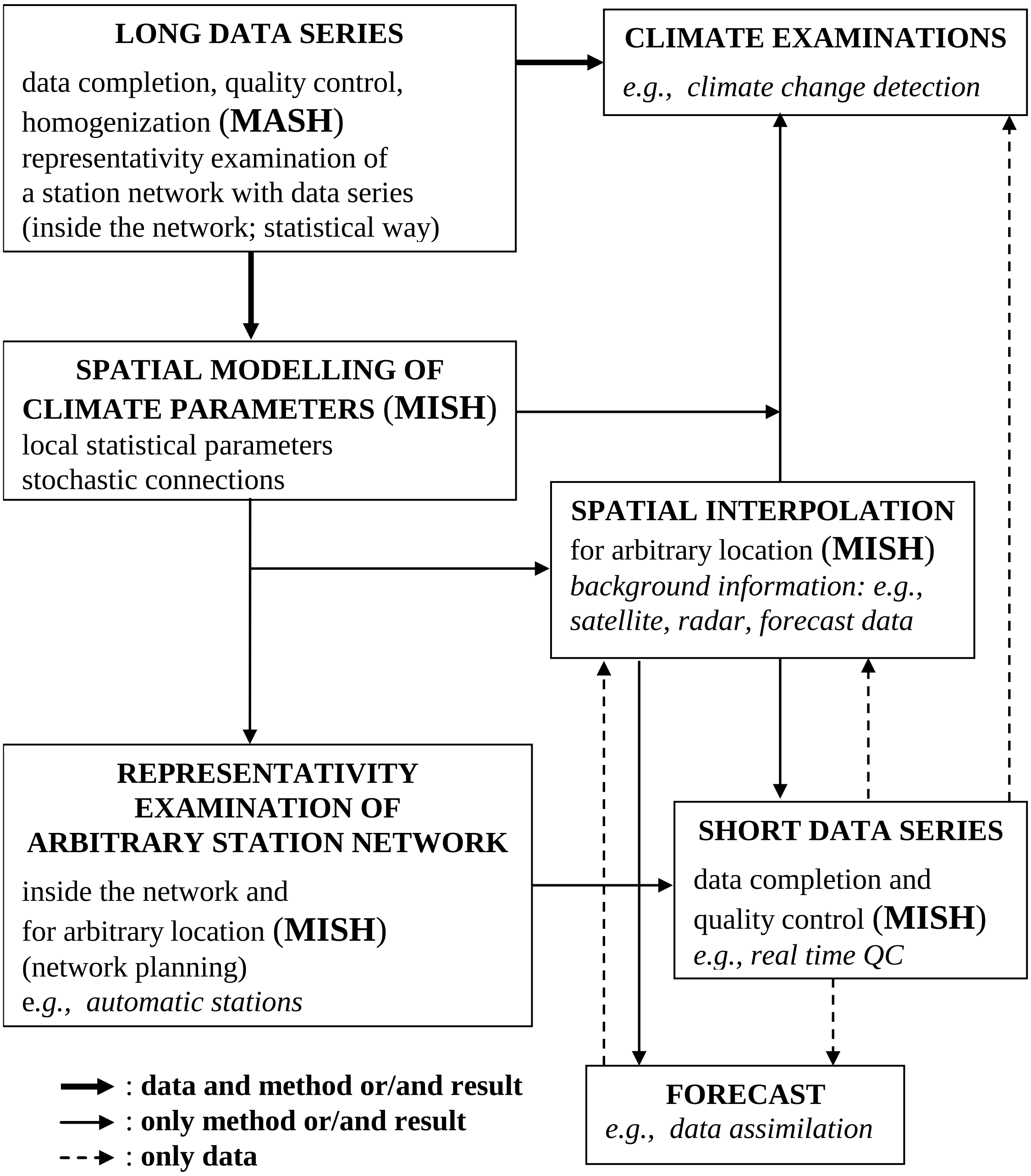


Fig. 1.: Block diagram for the possible connection between various basic meteorological topics and systems.

The software MASH (Multiple Analysis of Series for Homogenization; *Szentimrey*, 1999, 2017b) and MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; *Szentimrey* and *Bihari*, 2007, 2014, 2017a) were developed by us. These software were applied also in CARPATCLIM project (*Szentimrey et al.*, 2012a,b). We plan to share the new versions MASHv4.01 and MISHv2.01 on the website of VARIMAX this year. Our paper is summary of our conception on spatial interpolation and of the method MISH.

2. MATHEMATICAL OVERVIEW OF SPATIAL INTERPOLATION PROBLEM IN METEOROLOGY

According to the interpolation problem the unknown predictand $Z(\mathbf{s}_0,t)$ is estimated by use of the known predictors $Z(\mathbf{s}_i,t)$ $(i=1,...,M)$ where the location vectors $\mathbf{s}$ are the elements of the given space domain D and t is the time.

2.1 ADDITIVE MODEL OF SPATIAL INTERPOLATION

The type of the adequate interpolation formula depends on the probability distribution of the meteorological variable. Assuming normal distribution (e.g. temperature) the additive (linear) formula is adequate.

2.1.1 Statistical parameters

In general the interpolation formulas have some unknown interpolation parameters which are known functions of certain statistical parameters. At the additive interpolation formulas the basic statistical parameters can be divided into two groups such as the local and the stochastic parameters. The local parameters are the expected values $\mathrm{E}(Z(\mathbf{s}_i,t))$ $(i=0,...,M)$. The stochastic parameters are the covariances belonging to the predictand and predictors such as,

$\mathbf{c}$: predictand-predictors covariance vector,

$\mathbf{C}$: predictors-predictors covariance matrix.

This covariance system is equivalent with the standard deviations $D(\mathbf{s}_i)=\mathrm{D}(Z(\mathbf{s}_i,t))$ $(i=0,...,M)$ and the correlation system as,

$\mathbf{r}$: predictand-predictors correlation vector,

$\mathbf{R}$: predictors-predictors correlation matrix.

2.1.2 Linear meteorological model for expected values

At the statistical modelling of the meteorological elements we have to assume that the expected values of the variables are changing in space and in time alike. The spatial change means that the climate is different in the regions. The temporal change is the result of the possible global climate change. Consequently in case of linear modelling of expected values we assume that

$$\mathrm{E}(Z(\mathbf{s}_i,t))=\mu(t)+E(\mathbf{s}_i) \quad (i=0,..,M) \tag{1}$$

where $\mu(t)$ is the temporal trend or the climate change signal and $E(\mathbf{s})$ is the spatial trend.

2.1.3 Additive (Linear) Interpolation Formula

Assuming the linear model (1) and normal distribution the appropriate additive meteorological interpolation formula is as follows,

$$\hat{Z}(\mathbf{s}_0,t)=\lambda_0+\sum_{i=1}^{M}\lambda_i\cdot Z(\mathbf{s}_i,t) \tag{2}$$

where $\sum_{i=1}^{M}\lambda_i=1$ because of unknown $\mu(t)$.

The quality of interpolation can be characterized by the root-mean-square error,

$$RMSE(\mathbf{s}_0)=\sqrt{\mathrm{E}\left(\left(Z(\mathbf{s}_0,t)-\hat{Z}(\mathbf{s}_0,t)\right)^2\right)}, \tag{3}$$

and by the representativity value: $REP(\mathbf{s}_0)=1-\dfrac{RMSE(\mathbf{s}_0)}{D(\mathbf{s}_0)}$. (4)

2.1.4 The optimal interpolation

The optimal interpolation parameters $\lambda_0,\ \lambda_i\ (i=1,...,M)$ minimize the root-mean-square error and these are known functions of the statistical parameters! The optimal interpolation is $\hat{Z}_{opt}(\mathbf{s}_0,t)$ when we use the optimal parameters.

The optimal constant term is: $\lambda_0=\sum_{i=1}^{M}\lambda_i\left(E(\mathbf{s}_0)-E(\mathbf{s}_i)\right)$ (5)

The vector of optimal weighting factors $\boldsymbol{\lambda}=[\lambda_1,..,\lambda_M]^{\mathrm{T}}$ can be written in covariance form,

$$\boldsymbol{\lambda}=\mathbf{C}^{-1}\left(\mathbf{c}+\frac{\left(1-\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{c}\right)}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}\mathbf{1}\right), \tag{6}$$

and it is known function of the parameters: $D(\mathbf{s}_0)/D(\mathbf{s}_i)\ (i=1,...,M),\ \mathbf{r},\ \mathbf{R}$.

Consequently the unknown statistical parameters are the spatial trend differences $E(\mathbf{s}_0)-E(\mathbf{s}_i)(i=1,...,M)$, the standard deviation ratios $D(\mathbf{s}_0)/D(\mathbf{s}_i)\ (i=1,...,M)$ and the correlation system $\mathbf{r},\ \mathbf{R}$. In essence these parameters are climate parameters which in fact means we could interpolate optimally if we knew the climate.

2.1.5 The optimal interpolation error RMSE

The uncertainty of the optimal interpolation can be characterized quantitatively by the optimal RMSE that can be expressed as,

$$RMSE_{opt}(\mathbf{s}_0)=\sqrt{\left(D^2(\mathbf{s}_0)-\mathbf{c}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{c}\right)+\left(1-\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{c}\right)^2\cdot\frac{1}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}} \tag{7}$$

in the case of optimum interpolation parameters, and it can be written as function of the parameters: $D(\mathbf{s}_i)\ (i=0,..,M),\mathbf{r},\mathbf{R}$.

It can be proved that then the optimal representativity value $REP_{opt}(\mathbf{s}_0) = 1 - \dfrac{RMSE_{opt}(\mathbf{s}_0)}{D(\mathbf{s}_0)}$ depends on the following parameters only, $D(\mathbf{s}_0)/D(\mathbf{s}_i)$ $(i = 1,...,M)$, $\mathbf{r}$, $\mathbf{R}$.
Moreover if $D(\mathbf{s}_0)/D(\mathbf{s}_i) = 1$ $(i = 1,...,M)$ then,

$$REP_{opt}(\mathbf{s}_0) = 1 - \sqrt{\left(1 - \mathbf{r}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r}\right) + \left(1 - \mathbf{1}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{r}\right)^2 \cdot \frac{1}{\mathbf{1}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{1}}} \tag{8}$$

Consequently the necessary statistical parameters for calculation of $\hat{Z}_{opt}(\mathbf{s}_0)$, $RMSE_{opt}(\mathbf{s}_0)$ and $REP_{opt}(\mathbf{s}_0)$ are as follows: $E(\mathbf{s}_i)$, $D(\mathbf{s}_i)$ $(i = 0,...,M)$, $\mathbf{r}$, $\mathbf{R}$.
The question is how do we get know them?

3. THEORETICAL BACKGROUND OF MISH SYSTEM

Our method MISH (Meteorological Interpolation based on Surface Homogenized Data Basis) for the spatial interpolation of surface meteorological elements was developed (*Szentimrey* and *Bihari*, 2007, 2014, 2017a) according to the mathematical background that is outlined in Section 2. This is a meteorological system not only in respect of the aim but in respect of the tools as well. It means that using all the valuable meteorological information – e.g. climate and possible background information – is required.

3.1 THE BASIC MODELLING SYSTEM IN MISH

3.1.1 Modelling of monthly, daily spatiotemporal statistical parameters in MISH

The necessary unknown climate parameters can be modelled. These modelled monthly, and daily spatiotemporal statistical parameters in MISH system are:

i. Spatial expected values (spatial trend) $E(\mathbf{s})$
ii. Spatial standard deviations $D(\mathbf{s})$
iii. Spatial correlations $r(\mathbf{s}_1, \mathbf{s}_2)$
iv. Temporal autocorrelations $r(t_1, t_2)$

Consequently the first two spatiotemporal moments can be modelled for daily and monthly data by the MISH procedure! An example is presented on Fig. 2. Modelling is based on long station data series that means sample in space and in time! There is a substantial difference between geostatistics and meteorology namely the amount of information for modelling the statistical parameters. Consequently we should know the present climate well, not the future climate only! For this purpose special advanced mathematics is needed of course! “There is no royal road!” (Archimedes)

3.1.2 Possibility for modelling of unknown statistical parameters in Meteorology

The special possibility in meteorology is to use the long meteorological data series for modelling of the climate statistical parameters in question. The data series make possible to know the climate in accordance with the fundament of statistical climatology!

The main difference between geostatistics and meteorology can be found in the amount of information being usable for modelling the statistical parameters. In geostatistics the usable information or the sample for modelling is only the actual predictors $Z(\mathbf{s}_i, t)$ $(i = 1,...,M)$ which belong to a fixed instant of time, that is a single realization in time. While in meteorology we have spatiotemporal data, namely the long data series which form a sample in time and space as well and make possible to model the climate statistical parameters in question. If the meteorological stations $\mathbf{S}_k$ $(k = 1,..,K)$ $(\mathbf{s} \in D)$ have long data series then the spatial trend differences $E(\mathbf{S}_k) - E(\mathbf{S}_l)$ $(k,l = 1,...,K)$ as well as the covariances $\mathrm{cov}(Z(\mathbf{S}_k), Z(\mathbf{S}_l))$ $(k,l = 1,...,K)$ can be estimated statistically. Consequently these parameters are essentially known and provide much more information for modelling than the predictors $Z(\mathbf{s}_i, t)$ $(i = 1,...,M)$ only. However nowadays unfortunately the geostatistical interpolation methods built in GIS software are applied in meteorology mostly.

Example

Mean temperature data in September for 10 arbitrary locations somewhere in Hungary.

Input: the location coordinates only without any temperature data.

Output: modelled climate statistical parameters

Location indices:

	1	2	3	4	5	6	7	8	9	10
Monthly Expected Values:										
	14.59	14.99	14.95	15.06	15.16	15.16	15.13	15.08	15.01	15.05
Daily Expected Values:										
	14.59	14.99	14.95	15.06	15.16	15.16	15.13	15.08	15.01	15.05
Monthly Standard Deviations:										
	1.34	1.62	1.68	1.67	1.68	1.66	1.72	1.66	1.61	1.64
Daily Standard Deviations:										
	2.84	3.44	3.47	3.46	3.47	3.60	3.73	3.58	3.48	3.46
Temporal Daily Autocorrelations:										
	0.74	0.74	0.75	0.75	0.75	0.73	0.73	0.73	0.73	0.74
Matrix of Spatial Autocorrelations:										
	1.00	0.99	0.99	0.98	0.97	0.96	0.97	0.97	0.98	0.98
	0.99	1.00	0.99	0.99	0.98	0.95	0.96	0.96	0.97	0.98
	0.99	0.99	1.00	0.99	0.99	0.94	0.95	0.95	0.96	0.97
	0.98	0.99	0.99	1.00	0.99	0.91	0.93	0.93	0.95	0.96
	0.97	0.98	0.99	0.99	1.00	0.90	0.91	0.91	0.93	0.94
	0.96	0.95	0.94	0.91	0.90	1.00	0.99	0.99	0.98	0.98
	0.97	0.96	0.95	0.93	0.91	0.99	1.00	0.99	0.99	0.98
	0.97	0.96	0.95	0.93	0.91	0.99	0.99	1.00	0.99	0.99
	0.98	0.97	0.96	0.95	0.93	0.98	0.99	0.99	1.00	0.99
	0.98	0.98	0.97	0.96	0.94	0.98	0.98	0.99	0.99	1.00

Fig. 2.: Modelling of Present Climate by MISH.

3.2 INTERPOLATION APPLICATIONS FOR MONTHLY AND DAILY DATA IN MISH

The optimal interpolation $\hat{Z}_{opt}(\mathbf{s}_0)$, the optimal error $RMSE_{opt}(\mathbf{s}_0)$ and the optimal representativity $REP_{opt}(\mathbf{s}_0)$ can be calculated from the above modelled parameters (3.1.1), according to the mathematical formulas, equations (2),(3),(4),(5),(6),(7) defined in Section 2. The uncertainty of the interpolation can be characterized by $RMSE_{opt}(\mathbf{s}_0)$ or $REP_{opt}(\mathbf{s}_0)$ that make possible the representativity examination of any station network.

4. FURTHER DEVELOPMENTS IN MISH SYSTEM

4.1 AUTOMATED REAL TIME QUALITY CONTROL FOR DAILY AND MONTHLY DATA

The principle of the Test Statistics (*TS*) of Quality Control procedure in case of additive, normal distribution model is as follows. If the predictand $Z(\mathbf{s}_0)$ to be controlled and the predictors are correct, then

$$TS = \frac{Z(\mathbf{s}_0,t) - \hat{Z}_{opt}(\mathbf{s}_0,t)}{RMSE_{opt}(\mathbf{s}_0)} \in N(0,1) \text{ (=standard normal distribution)} \tag{9}$$

where $\hat{Z}_{opt}(\mathbf{s}_0,t)$ is the optimal interpolated value and $RMSE_{opt}(\mathbf{s}_0)$ is the optimal interpolation error calculated from the modelled parameters. During the automatic real time QC procedure in MISH multiple spatial comparison is tested similarly to the automatic QC procedure built in our MASH method (*Szentimrey*, 2017b) for station data series.

4.2 MULTIPLICATIVE MODEL OF SPATIAL INTERPOLATION

In this paper only the linear or additive model was described in detail which is appropriate in case of normal probability distribution. However perhaps it is worthwhile to remark that for case of a quasi lognormal distribution (e.g. precipitation sum) we deduced a mixed additive multiplicative formula which is used also in our MISH system and it can be written in the following form,

$$\hat{Z}(\mathbf{s}_0,t) = \vartheta \cdot \left(\prod_{q_i \cdot Z(\mathbf{s}_i,t) \geq \vartheta} \left(\frac{q_i \cdot Z(\mathbf{s}_i,t)}{\vartheta} \right)^{\lambda_i} \right) \cdot \left(\sum_{q_i \cdot Z(\mathbf{s}_i,t) \geq \vartheta} \lambda_i + \sum_{q_i \cdot Z(\mathbf{s}_i,t) < \vartheta} \lambda_i \cdot \left(\frac{q_i \cdot Z(\mathbf{s}_i,t)}{\vartheta} \right) \right) \tag{10}$$

where the interpolation parameters are $\lambda_i \geq 0 \ (i = 1,...,M)$, $\sum_{i=1}^{M} \lambda_i = 1$ and $\vartheta = m(\mathbf{s}_0)$, $q_i = m(\mathbf{s}_0)/m(\mathbf{s}_i)$, where $m(\mathbf{s}_i)$ ($i = 0,...,M$) are the spatial median values.

4.3 GRIDDING, TRANSFORMATION OF GRID-POINT DATASETS FOR GRID-BOX AVERAGE DATASETS

MISH has capability for interpolation, gridding of monthly or daily station data series, as grid-point values. For example the CarpatClim datasets (*Szentimrey et al.*, 2012b) were developed for grid-points, i.e. grid- point values were interpolated by MISH, while other datasets e.g. E-OBS were constructed as grid-box average values. For comparability of the datasets (*Szentimrey*, 2019b) a transformation procedure was developed at MISH for calculation of grid-box averages (*Szentimrey*, 2019a). The basis of the conception was the MISH specialty that the necessary statistical parameters - like spatial expected values, medians, st. deviations and correlation structure - are modelled for a very dense half minutes grid and saved. We developed a mathematical procedure and applied it for the gridded series using these saved parameters. The transformation formulas are as follows.

i, Additive model: $\hat{Z}_{average}(\mathbf{s}_0,t) = \left(\bar{m}(\mathbf{s}_0) - m(\mathbf{s}_0)\right) + \hat{Z}(\mathbf{s}_0,t)$

where $\hat{Z}(\mathbf{s}_0,t)$ is the grid-point interpolation according to the formula (2), $m(\mathbf{s}) = E(\mathbf{s})$ are the spatial expected values or medians, $\bar{m}(\mathbf{s}_0)$ is the grid-box average of medians and $\hat{Z}_{average}(\mathbf{s}_0,t)$ is grid-box average of point interpolated values.

ii, Multiplicative model: $\hat{Z}_{average}(\mathbf{s}_0,t) = \dfrac{\bar{m}(\mathbf{s}_0)}{m(\mathbf{s}_0)} \cdot \hat{Z}(\mathbf{s}_0,t)$

where $\hat{Z}(\mathbf{s}_0,t)$ is the grid-point interpolation according to the formula (10), $m(\mathbf{s})$ are the spatial medians, $\bar{m}(\mathbf{s}_0)$ is the grid-box average of medians and $\hat{Z}_{average}(\mathbf{s}_0,t)$ is grid-box average of point interpolated values.

Consequently now two versions of gridded datasets can be constructed by MISH, namely grid-point and grid-box average datasets.

4.4 INTERPOLATION WITH BACKGROUND INFORMATION

The background information e.g. forecast, satellite, radar data can be efficiently used to decrease the interpolation error. In this paper only the interpolation based on additive model or normal distribution is presented.

According to the Section 2. let us assume that,

$Z(\mathbf{s}_0,t)$: predictand,

$\hat{Z}(\mathbf{s}_0,t) = \lambda_0 + \sum_{i=1}^{M} \lambda_i Z(\mathbf{s}_i,t)$: interpolated predictand,

moreover there is given,

$\mathbf{G} = \{G(\mathbf{s},t) \mid \mathbf{s} \in D\}$: background information on a dense grid.

4.4.1 The principle of interpolation with background information

The interpolated predictand given $\mathbf{G}$ can be expressed as,

$$\hat{Z}_G(\mathbf{s}_0,t) = \hat{Z}(\mathbf{s}_0,t) + \mathrm{E}\left(Z(\mathbf{s}_0,t) - \hat{Z}(\mathbf{s}_0,t) \mid \mathbf{G} \right) \tag{11}$$

where $\mathrm{E}\left(Z(\mathbf{s}_0,t) - \hat{Z}(\mathbf{s}_0,t) \mid \mathbf{G} \right)$ is the conditional expectation of $Z(\mathbf{s}_0,t) - \hat{Z}(\mathbf{s}_0,t)$, given $\mathbf{G}$.

4.4.2 Reanalysis data, Data Assimilation

The reanalysis data are based on the data assimilation which procedure is in strong relationship with the methodology of interpolation with background information. The Bayes estimation theory is the mathematical background of the data assimilation and the following variational cost function has to be minimized in order to estimate the analysis field,

$$J(\mathbf{z}) = (\mathbf{z}-\mathbf{g})^{\mathrm{T}} \mathbf{Q}^{-1} (\mathbf{z}-\mathbf{g}) + (\mathbf{y}_0 - \mathbf{F}\mathbf{z})^{\mathrm{T}} \mathbf{P}^{-1} (\mathbf{y}_0 - \mathbf{F}\mathbf{z}) \tag{12}$$

$\mathbf{z}$: analysis field, predictand (grid),

$\mathbf{g}$: given background field (forecast),

$\mathbf{y}_0$: given observations, predictors; $\mathbf{Fz} = \mathrm{E}(\mathbf{y}_0|\mathbf{z})$,

$\mathbf{Q}$: background error covariance matrix,
$\mathbf{P}$: observation error covariance matrix.

It can be proved that this procedure is essentially an interpolation with background information including a quality control part for the predictors.
However there are several problems (*Szentimrey*, 2016) with the reanalysis data in the practice:

i, Inhomogeneous predictor station data series are used.

ii, Few stations are used with little spatial representativity.

iii, There are also some problems with the data assimilation formula (12):

- Lack of good climate statistical parameters in matrix **Q.**
- Formula (12) includes an implicit, unjustifiable assumption of $\mathrm{E}(\mathbf{z}|\mathbf{g}) = \mathbf{g}$.

5. THE MAIN FEATURES OF SOFTWARE MISHV2.01

The new software version MISHv2.01 consists of two units that are the modelling and the interpolation systems. The interpolation system can be operated on the results of the modelling system. We summarize briefly the most important facts about these two units of the developed software.

Modelling subsystem for climate statistical (local and stochastic) parameters:

– Modelling of all the first two spatiotemporal moments for daily and monthly data. (expected values, standard deviations, spatiotemporal correlations)
– Based on long homogenized data series and supplementary deterministic model variables. The model variables may be such as height, topography, distance from the sea etc.. Neighbourhood modelling, correlation model for each grid point, dense half minutes grid.
– Benchmark study, cross-validation test for interpolation error or representativity.
– Modelling procedure must be executed only once before the interpolation applications.
Totally different principle from the other methods!

Interpolation subsystem:

– Additive (e.g. temperature) or multiplicative (e.g. precipitation) model and interpolation formula can be used depending on the climate elements.
– Daily, monthly values and many years' means can be interpolated.
– Few predictors are also sufficient for the interpolation and no problem if the greater part of daily precipitation predictors is equal to 0.
– The expected interpolation error RMSE is modelled too, representativity examination of arbitrary station network.
– Real time Quality Control for daily and monthly data (additive model).
– Capability for application of supplementary background information (stochastic variables) e.g. satellite, radar, forecast data. (with QC: data assimilation)
– Data series completion that is missing value interpolation, completion for monthly or daily station data series.
– Capability for interpolation, gridding of monthly or daily station data series, as grid-point and grid-box average datasets alike.

The elder version of MISH-MASH software can be downloaded from:
http://www.met.hu/en/omsz/rendezvenyek/homogenization_and_interpolation/software/
We plan to share the new version MISHv2.01 this year 2021.

References

Cressie, N., 1991: Statistics for Spatial Data., Wiley, New York, p. 900.

Szentimrey, T., 1999: Multiple Analysis of Series for Homogenization (MASH), Proceedings of the Second Seminar for Homogenization of Surface Climatological Data, Budapest, Hungary; WMO, WCDMP-No. 41, pp. 27-46.

Szentimrey, T., Bihari, Z., 2007: Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis), Proceedings of the Conference on Spatial Interpolation in Climatology and Meteorology, Budapest, Hungary, 2004, COST Action 719, COST Office, 2007, pp. 17-27

Szentimrey, T, Bihari, Z., Lakatos, M., Szalai,S., 2011: Mathematical, methodological questions concerning the spatial interpolation of climate elements. Proceedings of the Second Conference on Spatial Interpolation in Climatology and Meteorology, Budapest, Hungary, 2009, Időjárás 115, 1-2, 1-11

Szentimrey T, Lakatos M, Bihari Z, Kovacs T, Nemeth A, Szalai S, Auer I, Hiebl J, Milkovic J, Štěpánek P, Zahradnicek P, Tolasz R, Kilar P, Pyrc R, Limanowka D, Cheval S, Matei M, Kajaba P, Ivanakova G, Bochnicek O, Nejedlik P, Statsny P, Mihic D, Petrovic P, Savic T, Skrynyk O, Nabyvanets Y, Gnatiuk N, 2012a: Final report on quality control and data homogenization measures applied per country, including QC protocols and measures to determine the achieved increase in data quality. Carpatclim Project, Deliverable D1.12, http://www.carpatclim-eu.org/docs/deliverables/D1_12.pdf.

Szentimrey T, Lakatos M, Bihari Z, Kovacs T, Nemeth A, Szalai S, Auer I, Hiebl J, Milkovic J, Štěpánek P, Zahradnicek P, Tolasz R, Kilar P, Pyrc R, Limanowka D, Cheval S, Matei M, Kajaba P, Ivanakova G, Bochnicek O, Nejedlik P, Statsny P, Mihic D, Petrovic P, Savic T, Skrynyk O, Nabyvanets Y, Gnatiuk N, 2012b: Final report on the creation of national gridded datasets, per country. Carpatclim Project, Deliverable D2.9, http://www.carpatclim-eu.org/docs/deliverables/D2_9.pdf.

Szentimrey,T., Bihari, Z., 2014: Manual of interpolation software MISHv1.03, Hungarian Meteorological Service, p. 60.

Szentimrey T., Bihari Z., Lakatos M., 2015: Mathematical questions of spatial interpolation of climate variables, Proceedings of the 8th Seminar for Homogenization and Quality Control in Climatological Databases and 3rd Conference on Spatial Interpolation Techniques in Climatology and Meteorology, Budapest, Hungary, 2014, WCDMP-No. 84, pp. 107-114

Szentimrey, T. (2016): Analysis of the data assimilation methods from the mathematical point of view. In: Bátkai, A., Csomós, P., Faragó, I., Horányi, A., Szépszó, G. (ed.): *Mathematical Problems in Meteorological Modelling*, Springer International Publishing, Switzerland, 193–205, DOI:10.1007/978-3-319-40157-7_10.

Szentimrey, T., 2017a: New developments of interpolation method MISH: modelling of interpolation error RMSE, automated real time quality control, Proceedings of the 9th Seminar for Homogenization and Quality Control in Climatological Databases and 4th Conference on Spatial Interpolation Techniques in Climatology and Meteorology (Ed. Szentimrey T, Lakatos M, Hoffmann L), Budapest, Hungary, 2017, WCDMP-No. 85, pp. 115-124

Szentimrey, T.,2017b: Manual of homogenization software MASHv3.03, Hungarian Meteorological Service,p.71.

Szentimrey, T.,2019a: Transformation of CarpatClim datasets for grid-box average datasets, Study for Hungarian Meteorological Service (COPERNICUS C3Surf project (WP4)), Varimax Limited Partnership, Budapest, Hungary, p. 7.

Szentimrey, T.,2019b: Mathematical methodology and software for comparison of gridded datasets, Study for Hungarian Meteorological Service (COPERNICUS C3Surf project (WP4)), Varimax Limited Partnership, Budapest, Hungary, p. 10.

TRANSFORMATION OF CARPATCLIM DATASETS TO GRID-BOX AVERAGE DATASETS

Beatrix Izsák[1, 2], Tamás Szentimrey[3], Mónika Lakatos[1], Zita Bihari[1], Andrea Kircsi[1]

[1] Hungarian Meteorological Service (OMSZ), Unit of Climatology, Hungary
[2] ELTE Faculty of Science, Doctoral School of Earth Sciences Budapest, Hungary
[3] Varimax Limited Partnership Budapest, Hungary
izsak.b@met.hu

Abstract

The CarpatClim datasets were developed for grid points, i.e. the meteorological variables were interpolated to grid points, while the E-OBS datasets were constructed as grid-box averages.
For comparability we have transformed the CarpatClim datasets for grid-box averages. For this purpose, beside the gridded values with 0.1 x 0.1-degree resolution we used also certain modelled climate statistical parameters. These statistical parameters were modelled during the construction of CarpatClim datasets and they were also outputs of our MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; Szentimrey and Bihari) procedure applied for gridding. There is a MISH specialty that the necessary statistical parameters - like spatial trend and correlation structure - are modelled for a very dense half minutes grid and saved. We developed a mathematical procedure and applied it for the gridded series using these saved parameters. Now we have two versions of CarpatClim gridded datasets for temperature (Tx, Tn) and precipitation, namely grid-point and grid-box average datasets. Comparison of CarpatClim grid-point and CarpatClim grid-box datasets are presented too.

1. INTRODUCTION

1.1 THE CARPATCLIM DATASET

CARPATCLIM project was a consortium of ten organizations founded for a tender published by Joint Research Centre. The objective of the project was to investigate the detailed temporal and spatial structure of the climate of Carpathian Region using unified methods.
The main aim was to produce gridded climatological database for this region. The grids cover the area between latitudes 44°N and 50°N, and longitudes 17°E and 27°E. Daily values of more than ten meteorological variables were calculated on a 0.1° spatial resolution grid for the period 1961-2010. Climate statistics (monthly and annual values) and different climate indices were also determined from the daily grids.
For ensuring the usage of largest possible station density the necessary work phases were implemented on national level but by the same methods and software. The commonly used methods and software were the method MASH (Multiple Analysis of Series for Homogenization; *Szentimrey*, 1999, 2008, 2017a) for homogenization, quality control, completion of the observed daily data series, and the method MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; *Szentimrey and Bihari*, 2007, 2014, 2017b) for gridding of homogenized daily data series. Besides the common software, the harmonization of the results across country borders is promoted also by near border data exchange.

1.2 THE E-OBS DATASET

Recently a new version of E-OBS dataset was developed. This new version of E-OBS dataset was tested against CARPATLIM and against other regional datasets in the framework of the COPERNICUS C3S_311a_Lot4 project. We have done such a comparison between the E-OBS and CARPATLIM datasets before (*Lakatos et al.*, 2017), but we had to face the following methodological problem. The CarpatClim dataset was developed for grid points, i.e. grid point values were interpolated, while the E-OBS dataset was constructed as grid-box averages. However this comparison problem can be solved since the CarpatClim dataset can be transformed also to grid-box averages as a consequence of some useful properties of the applied interpolation procedure MISH.

1.3 THE SOFTWARE MASH AND MISH APPLIED AT CARPATCLIM PROJECT

The homogenization, quality control and completion of the daily data series were implemented by the method MASH (*Szentimrey et al.*, 2012a). The interpolation or gridding of the daily data series was implemented by the method MISH (*Szentimrey et al.*, 2012b,c). The CarpatClim dataset was developed for grid points. The transformation procedure to calculate the grid-box averages can be based on some properties of the method MISH.

1.3.1 The properties of the version MISHv1.03

The applied software version MISHv1.03 (*Szentimrey and Bihari*, 2014) consists of two units that are the modelling and the interpolation systems. The interpolation system can be operated on the results of the modelling system. We summarize briefly the most important facts about these two units of the developed software.

Modelling subsystem for climate statistical (local and stochastic) parameters:

- Based on long homogenized data series and supplementary deterministic model variables. The model variables may be such as height, topography, distance from the sea etc.. Neighbourhood modelling, correlation model for each grid point.
- Benchmark study, cross-validation test for interpolation error or representativity.
- Modelling procedure must be executed only once before the interpolation applications!

Interpolation subsystem:

- Additive (e.g. temperature) or multiplicative (e.g. precipitation) model and interpolation formula can be used depending on the climate elements.
- Daily, monthly values and many years' means can be interpolated.
- Few predictors are also sufficient for the interpolation and no problem if the greater part of daily precipitation predictors is equal to 0.
- The representativity is modeled too.
- Capability for application of supplementary background information (stochastic variables) e.g. satellite, radar, forecast data.
- Data series completion that is missing value interpolation, completion for monthly or daily station data series.
- Interpolation, gridding of monthly or daily station data series for given predictand locations. In case of gridding the predictand locations are the nodes of a relatively dense grid.

Our MISH-MASH software can be downloaded from:
http://www.met.hu/en/omsz/rendezvenyek/homogenizationand_interpolation/software/
According to the modelling subsystem there is a MISH specialty that the necessary statistical parameters - like spatial trend and correlation structure - are modelled for a very dense half minutes grid and saved. These statistical parameters were also modelled during the construction of CarpatClim datasets and they were also outputs of our MISH procedure applied for gridding. Using these parameters we can transform the grid-point values to the grid-box averages and this transformed dataset can be compared with the E-OBS dataset. The mathematical background of this transformation is detailed in the following section.

2. MATHEMATICAL BACKGROUND OF TRANSFORMATION OF GRIDPOINT DATASETS FOR GRID-BOX AVERAGE DATASETS

The mathematics of the transformation in MISH was developed for additive (temperature) and multiplicative (precipitation) model alike (*Szentimrey*, 2019a).

2.1 ADDITIVE (LINEAR) INTERPOLATION (TEMPERATURE)

2.1.1 Mathematical model

According to the interpolation problem the unknown predictand $Z(\mathbf{s}_0)$ is estimated by use of the known predictors $Z(\mathbf{s}_i)$ $(i=1,...,M)$ where the location vectors $\mathbf{s}$ are the elements of the given space domain D.

Assuming the linear model – in case of normal distribution - the appropriate additive meteorological interpolation formula is as follows,

$$\hat{Z}(\mathbf{s}_0)=\lambda_0+\sum_{i=1}^{M}\lambda_i\cdot Z(\mathbf{s}_i) \tag{1}$$

where $\sum_{i=1}^{M}\lambda_i=1$ because of unknown climate change.

The optimal interpolation parameters λ_0, λ_i $(i=1,...,M)$ minimize the root-mean-square error and these are known functions of some climate statistical parameters!

The optimal constant term is,

$$\lambda_0=\sum_{i=1}^{M}\lambda_i\left(E(\mathbf{s}_0)-E(\mathbf{s}_i)\right) \tag{2}$$

where $E(\mathbf{s}_i)=\mathrm{E}(Z(\mathbf{s}_i))$ $(i=0,...,M)$ are the expected values or spatial trend values.

The vector of optimal weighting factors $\boldsymbol{\lambda}^{\mathrm{T}}=[\lambda_1,..,\lambda_M]$ written in covariance form,

$$\boldsymbol{\lambda}^{\mathrm{T}}=\left(\mathbf{c}^{\mathrm{T}}+\mathbf{1}^{\mathrm{T}}\frac{\left(1-\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{c}\right)}{\mathbf{1}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{1}}\right)\mathbf{C}^{-1},$$

where $\mathbf{c}$ is the predictand-predictors covariance vector, $\mathbf{C}$ is the predictors-predictors covariance matrix and they are known functions of the climate statistical parameters as, standard deviations $D(\mathbf{s}_i)=\mathrm{D}(Z(\mathbf{s}_i))$ $(i=0,...,M)$ predictand-predictors correlation vector $\mathbf{r}$ and predictors-predictors correlation matrix $\mathbf{R}$.

The climate statistical parameters must be modelled and the special possibility in meteorology is to use the long meteorological data series for modelling of the climate statistical parameters in question. The data series make possible to know the climate in accordance with the fundament of statistical climatology!

According to this principle we developed the method MISH for the spatial interpolation of surface meteorological elements and there is a MISH specialty that the necessary statistical parameters - like spatial trend and correlation structure - are modelled for a very dense half minutes grid and saved. These statistical parameters were modelled during also the construction of CarpatClim datasets and they were also outputs of our MISH procedure applied for gridding. Using these saved parameters the transformation of CarpatClim datasets for grid-box average datasets is possible.

2.1.2 Mathematical background of transformation of gridpoint datasets for grid-box average datasets in case of additive model

According to the formulas (1), (2) the interpolated gridpoint values can be written as,

$$\hat{Z}(\mathbf{s}_0) = \sum_{i=1}^{M} \lambda_i \left(E(\mathbf{s}_0) - E(\mathbf{s}_i) \right) + \sum_{i=1}^{M} \lambda_i Z(\mathbf{s}_i)$$

Let $\mathbf{s}_{0,k} \in B(\mathbf{s}_0)$, $(k = 1,...,K)$ where $B(\mathbf{s}_0)$ is a gridbox round $\mathbf{s}_0$.
Then interpolated values within the gridbox are,

$$\hat{Z}(\mathbf{s}_{0,k}) = \sum_{i=1}^{M} \lambda_{i,k} \left(E(\mathbf{s}_{0,k}) - E(\mathbf{s}_i) \right) + \sum_{i=1}^{M} \lambda_{i,k} Z(\mathbf{s}_i) \quad (k = 1,...,K)$$

Therefore the grid-box average is,

$$\hat{Z}_{Average}(\mathbf{s}_0) = \frac{1}{K} \sum_{k=1}^{K} \hat{Z}(\mathbf{s}_{0,k}) = \frac{1}{K} \sum_{k=1}^{K} \left(\sum_{i=1}^{M} \lambda_{i,k} \left(E(\mathbf{s}_{0,k}) - E(\mathbf{s}_i) \right) + \sum_{i=1}^{M} \lambda_{i,k} Z(\mathbf{s}_i) \right) \approx$$

as a consequence of the similar stochastic connection of the predictands within the grid-box with the predictors,

$$\approx \frac{1}{K} \sum_{k=1}^{K} \left(\sum_{i=1}^{M} \lambda_i \left(E(\mathbf{s}_{0,k}) - E(\mathbf{s}_i) \right) + \sum_{i=1}^{M} \lambda_i Z(\mathbf{s}_i) \right) = \sum_{i=1}^{M} \frac{1}{K} \sum_{k=1}^{K} \lambda_i \left(E(\mathbf{s}_{0,k}) - E(\mathbf{s}_0) \right) + \hat{Z}(\mathbf{s}_0) =$$

$$= \left(\bar{E}(\mathbf{s}_0) - E(\mathbf{s}_0) \right) + \hat{Z}(\mathbf{s}_0) = \left(\bar{m}(\mathbf{s}_0) - m(\mathbf{s}_0) \right) + \hat{Z}(\mathbf{s}_0) \qquad (3)$$

The spatial trend or median $m(\mathbf{s}) = E(\mathbf{s})$ were modelled for a very dense half minutes grid therefore we can transform the temperature CarpatClim gridpoint datasets for grid-box average datasets by using the formula (3).

2.2 MULTIPLICATIVE INTERPOLATION (PRECIPITATION)

2.2.1 Mathematical model

The linear or additive model is appropriate in case of normal probability distribution. However in case of a quasi lognormal distribution (e.g. precipitation sum) we deduced a mixed additive multiplicative formula which is used also in our MISH system and it can be written in the following form,

$$\hat{Z}(\mathbf{s}_0) = \vartheta \cdot \left(\prod_{q_i \cdot Z(\mathbf{s}_i) \geq \vartheta} \left(\frac{q_i \cdot Z(\mathbf{s}_i)}{\vartheta} \right)^{\lambda_i} \right) \cdot \left(\sum_{q_i \cdot Z(\mathbf{s}_i) \geq \vartheta} \lambda_i + \sum_{q_i \cdot Z(\mathbf{s}_i) < \vartheta} \lambda_i \cdot \left(\frac{q_i \cdot Z(\mathbf{s}_i)}{\vartheta} \right) \right) \qquad (4)$$

where the interpolation parameters are $\vartheta > 0$, $q_i > 0$, $\lambda_i \geq 0$ $(i = 1,...,M)$ and $\sum_{i=1}^{M} \lambda_i = 1$.

During the construction of CarpatClim datasets we applied this interpolation formula with interpolation paramaters:

$$\vartheta = m(\mathbf{s}_0), \quad q_i = m(\mathbf{s}_0)/m(\mathbf{s}_i) \tag{5}$$

where $m(\mathbf{s}_i)$ ($i = 0,...,M$) are the spatial median values.

2.2.2 Mathematical background of transformation of gridpoint datasets for grid-box average datasets in case of multiplicative model

According to the formulas (4), (5) the interpolated gridpoint values can be written as,

$$\hat{Z}(\mathbf{s}_0) = m(\mathbf{s}_0)\cdot\left(\prod_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)^{\lambda_i}\right)\cdot\left(\sum_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\lambda_i + \sum_{Z(\mathbf{s}_i)<m(\mathbf{s}_i)}\lambda_i\cdot\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)\right)$$

Let $\mathbf{s}_{0,k} \in B(\mathbf{s}_0)$, $(k = 1,...,K)$ where $B(\mathbf{s}_0)$ is a gridbox round $\mathbf{s}_0$.
Then interpolated values within the gridbox are,

$$\hat{Z}(\mathbf{s}_{0.k}) = m(\mathbf{s}_{0.k})\cdot\left(\prod_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)^{\lambda_{i,k}}\right)\cdot\left(\sum_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\lambda_{i,k} + \sum_{Z(\mathbf{s}_i)<m(\mathbf{s}_i)}\lambda_{i,k}\cdot\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)\right)$$

Therefore the grid-box average is,

$$\hat{Z}_{Average}(\mathbf{s}_0) = \frac{1}{K}\sum_{k=1}^{K}\hat{Z}(\mathbf{s}_{0,k}) =$$

$$= \frac{1}{K}\sum_{k=1}^{K} m(\mathbf{s}_{0.k})\cdot\left(\prod_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)^{\lambda_{i,k}}\right)\cdot\left(\sum_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\lambda_{i,k} + \sum_{Z(\mathbf{s}_i)<m(\mathbf{s}_i)}\lambda_{i,k}\cdot\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)\right) \approx$$

as a consequence of the similar stochastic connection of the predictands within the grid-box with the predictors,

$$\approx \frac{1}{K}\sum_{k=1}^{K} m(\mathbf{s}_{0.k})\cdot\left(\prod_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)^{\lambda_i}\right)\cdot\left(\sum_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\lambda_i + \sum_{Z(\mathbf{s}_i)<m(\mathbf{s}_i)}\lambda_i\cdot\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)\right) =$$

$$= \left(\frac{1}{K}\sum_{k=1}^{K} m(\mathbf{s}_{0.k})\right)\cdot\left(\prod_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)^{\lambda_i}\right)\cdot\left(\sum_{Z(\mathbf{s}_i)\ge m(\mathbf{s}_i)}\lambda_i + \sum_{Z(\mathbf{s}_i)<m(\mathbf{s}_i)}\lambda_i\cdot\left(\frac{Z(\mathbf{s}_i)}{m(\mathbf{s}_i)}\right)\right) =$$

$$= \frac{\frac{1}{K}\sum_{k=1}^{K} m(\mathbf{s}_{0.k})}{m(\mathbf{s}_0)}\cdot\hat{Z}(\mathbf{s}_0) = \frac{\overline{m}(\mathbf{s}_0)}{m(\mathbf{s}_0)}\cdot\hat{Z}(\mathbf{s}_0) \tag{6}$$

The spatial median $m(\mathbf{s})$ were modelled for a very dense half minutes grid therefore we can transform the precipitation CarpatClim grid-point datasets for grid-box average datasets by using the formula (6).

3. COMPARISON OF THE CARPATCLIM GRID-POINT AND GRID-BOX AVERAGE DATASETS

According to the above transformation procedures we calculated the grid-box average datasets for maximum temperature, minimum temperature and precipitation. In this section we show the difference between the two types of CarpatClim gridded datasets.

3.1 MATHEMATICAL BACKGROUND OF THE COMPARISON

The applied methodology is part of that was developed for the comparison of the CarpatClim and E-OBS datasets (Szentimrey, 2019b) in the framework of the COPERNICUS C3S_311a_Lot4 project.

3.1.1 ANOVA (Analysis Of Variance) examination

We applied the ANOVA methodology originally at the modelling unit of the MISH procedure (*Szentimrey and Bihari*, 2014) and later for the comparison of the E-OBS and CarpatClim datasets (*Lakatos et al.*, 2017).
Using the basic theorem of ANOVA the total spatiotemporal variance can be partitioned (7) equivalently as follows.
- Sum of spatial variance of temporal means and spatial mean of temporal variances.
  The temporal means and temporal variances (or st. deviations) in the space can be visualised by maps.
- Or sum of temporal variance of spatial means and temporal mean of spatial variances.
  The series of spatial means and spatial variances (or st. deviations) can be visualised by graphics.

The above ANOVA methodology can be used for the gridded monthly, seasonal and annual series calculated from the daily series. Mean series for temperature, while sum series for precipitation. The calculated maps and graphics for various datasets can be compared.

3.1.1.1 Mathematical description of ANOVA methodology

$Z(\mathbf{s}_j,t)\ (j=1,..,N;\ t=1,...,n)$ – data series $(\mathbf{s}_j$: location; t : time$)$

$\hat{E}(\mathbf{s}_j)=\frac{1}{n}\sum_{t=1}^{n}Z(\mathbf{s}_j,t)\quad (j=1,..,N)$ – temporal mean at location $\mathbf{s}_j$

$\hat{D}(\mathbf{s}_j)=\left(\frac{1}{n}\sum_{t=1}^{n}\left(Z(\mathbf{s}_j,t)-\hat{E}(\mathbf{s}_j)\right)^2\right)^{\frac{1}{2}}\quad (j=1,..,N)$ – temporal st. deviation at location $\mathbf{s}_j$

$\hat{E}(t)=\frac{1}{N}\sum_{j=1}^{N}Z(\mathbf{s}_j,t)\quad (t=1,..,n)$ – spatial mean at moment t

$\hat{D}(t)=\left(\frac{1}{N}\sum_{j=1}^{N}\left(Z(\mathbf{s}_j,t)-\hat{E}(t)\right)^2\right)^{\frac{1}{2}}\quad (t=1,..,n)$ – spatial st. deviation at moment t

$$\hat{E} = \frac{1}{N \cdot n}\sum_{j=1}^{N}\sum_{t=1}^{n} Z(\mathbf{s}_j, t) = \frac{1}{N}\sum_{j=1}^{N}\hat{E}(\mathbf{s}_j) = \frac{1}{n}\sum_{t=1}^{n}\hat{E}(t) \text{ – total mean}$$

$$\hat{D}^2 = \frac{1}{N \cdot n}\sum_{j=1}^{N}\sum_{t=1}^{n}\left(Z(\mathbf{s}_j, t) - \hat{E}\right)^2 \text{ – total variance}$$

Partitioning of Total Variance (Theorem)

$$\hat{D}^2 = \frac{1}{N}\sum_{j=1}^{N}\left(\hat{E}(\mathbf{s}_j) - \hat{E}\right)^2 + \frac{1}{N}\sum_{j=1}^{N}\hat{D}^2(\mathbf{s}_j) = \frac{1}{n}\sum_{t=1}^{n}\left(\hat{E}(t) - \hat{E}\right)^2 + \frac{1}{n}\sum_{t=1}^{n}\hat{D}^2(t) \tag{7}$$

The analysis of these terms is recommended to characterize the spatiotemporal variability.
Spatial terms: spatial variance of temporal means $\frac{1}{N}\sum_{j=1}^{N}\left(\hat{E}(\mathbf{s}_j) - \hat{E}\right)^2$

and temporal mean of spatial variances $\frac{1}{n}\sum_{t=1}^{n}\hat{D}^2(t)$.

Temporal terms: spatial mean of temporal variances $\frac{1}{N}\sum_{j=1}^{N}\hat{D}^2(\mathbf{s}_j)$

and temporal variance of spatial means $\frac{1}{n}\sum_{t=1}^{n}\left(\hat{E}(t) - \hat{E}\right)^2$.

3.1.2 Scores, examination of differences of gridded datasets

The applied methodology of scores is also part of that was developed for the comparison of the CarpatClim and E-OBS datasets (*Szentimrey*, 2019b) in the framework of the COPERNICUS C3S_311a_Lot4 project
Scores RMSE, MSESS were calculated for the difference of CarpatClim grid point and grid-box average data series. This procedure was performed for monthly data and the scores were determined for the months and year. The scores on the grid can be visualised by maps.

3.1.2.1 Mathematical description

$Z(y,m)$, $Z_1(y,m)$, $Z_2(y,m)$ (y : year, m : month) – monthly series

Monthly variance for the months: $V(Z(m)) = \frac{1}{n_y}\sum_{y=1}^{n_y}\left(Z(y,m) - \frac{1}{n_y}\sum_{y=1}^{n_y}Z(y,m)\right)^2$

Monthly variance for the year: $V(Z) = \frac{1}{12}\sum_{m=1}^{12}V(Z(m))$

Monthly MSE for the months: $MSE(Z_{1,2}(m)) = \frac{1}{n_y}\sum_{y=1}^{n_y}\left(Z_1(y,m) - Z_2(y,m)\right)^2$

Monthly MSE for the year: $MSE(Z_{1,2}) = \frac{1}{12}\sum_{m=1}^{12}MSE(Z_{1,2}(m))$

Monthly RMSE for the months: $RMSE(Z_{1,2}(m)) = \sqrt{MSE(Z_{1,2}(m))}$

Monthly RMSE for the year: $RMSE(Z_{1,2}) = \sqrt{MSE(Z_{1,2})}$ $(m = 1,..,12)$

Monthly MSESS for the months: $MSESS\left(Z_{1,2}\left(m\right)\right)=1-MSE\left(Z_{1,2}\left(m\right)\right)/V\left(Z_{2}\left(m\right)\right)$

Monthly MSESS for the year: $MSESS\left(Z_{1,2}\right)=1-MSE\left(Z_{1,2}\right)/V\left(Z_{2}\right)$

4. COMPARISON RESULTS

We created the daily maximum temperature, daily minimum temperature and daily precipitation sum grid-box databases from the CarpatClim grid-point database. For the entire area of the CarpatClim and for the entire period from 1961 to 2010. The goal was to make CarpatClim comparable to other grid-box databases, however, many questioned what the relationship was between the two CarpatClim datasets. One possible way to do this, for example, is to present the ANOVA results.

4.1 ANOVA RESULTS FOR MAXIMUM TEMPERATURE

In the Fig. 1 we can see that the total mean is the same, while with the averaging the total variance decreases in case of CarpatClim_BOX. The values of spatial mean of temporal variances and temporal variance of spatial means are the same for the whole area and period, the values of spatial variance of temporal means and temporal mean of spatial variances are smaller for the box average. It is worth seeing in Fig. 2 that the spatial mean series of annual mean values for the entire CarpatClim area are almost identical. Fig. 3, on the other hand, shows that the standard deviations decreased slightly.

CarpatClim gridBOX

Total mean: 13.85
Total variance: 6.10

Spatial variance of temporal means:	5.32
Spatial mean of temporal variances:	0.78
Temporal variance of spatial means:	0.69
Temporal mean of spatial variances:	5.41

CarpatClim gridPoint

Total mean: 13.83
Total variance: 6.90

Spatial variance of temporal means:	6.13
Spatial mean of temporal variances:	0.78
Temporal variance of spatial means:	0.69
Temporal mean of spatial variances:	6.21

Fig. 1.: The most important ANOVA statistics for annual mean maximum temperature

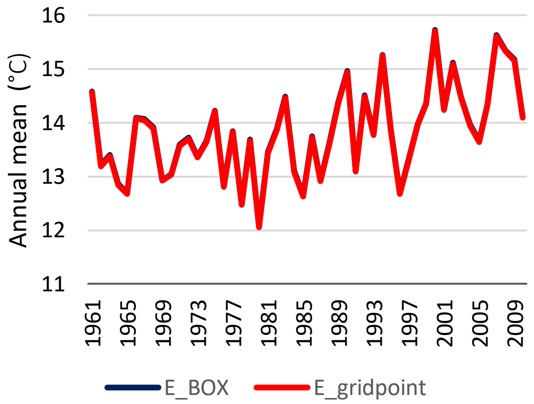


Fig. 2.: Spatial mean series of annual mean maximum temperature for the period 1961-2010

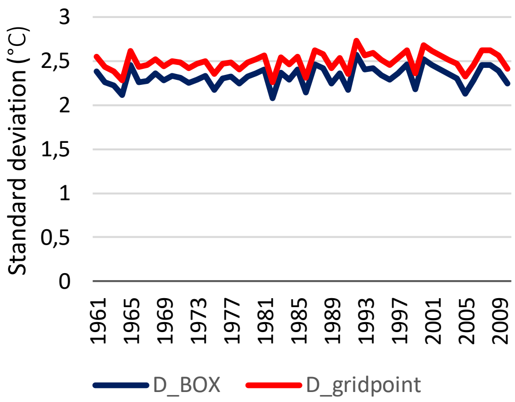


Fig.3.: Spatial standard deviation series of annual mean maximum temperature for the period 1961-2010

We plotted the temporal mean values on a map (Fig. 4-5) and the difference between the two CarpatClim datasets in Fig. 6. Although we found that the values are the same on average, it is clear that where the surface is more variable, the topography is different within a given cell, there are more differences. Typically, the Carpathian ranges emerge. In the same way, Figures 7-8 show the temporal standard deviations and Figure 9 shows the difference between the two Carpatclim datasets. The two maps are exactly the same, where there is a difference, even there this value is less than 0.01 Celsius.

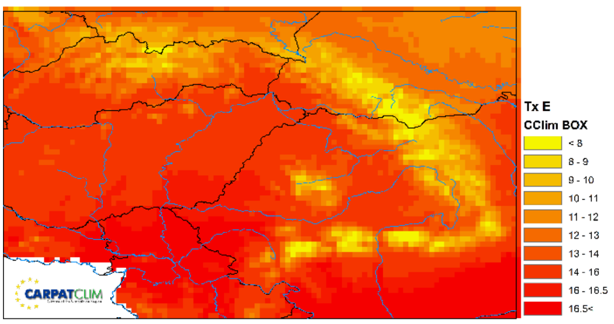


Fig. 4.: CarpatClim_BOX: temporal mean values of annual mean maximum temperature for the period 1961-2010

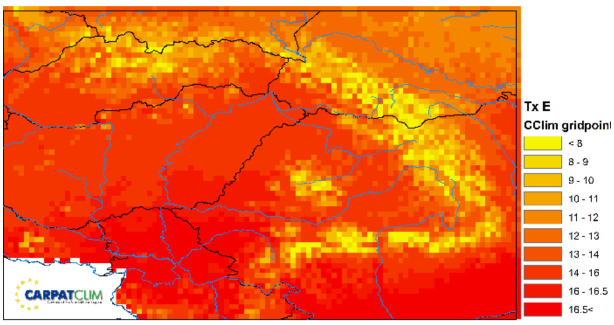


Fig. 5.: CarpatClim_gridpoint: temporal mean values of annual mean maximum temperature for the period 1961-2010

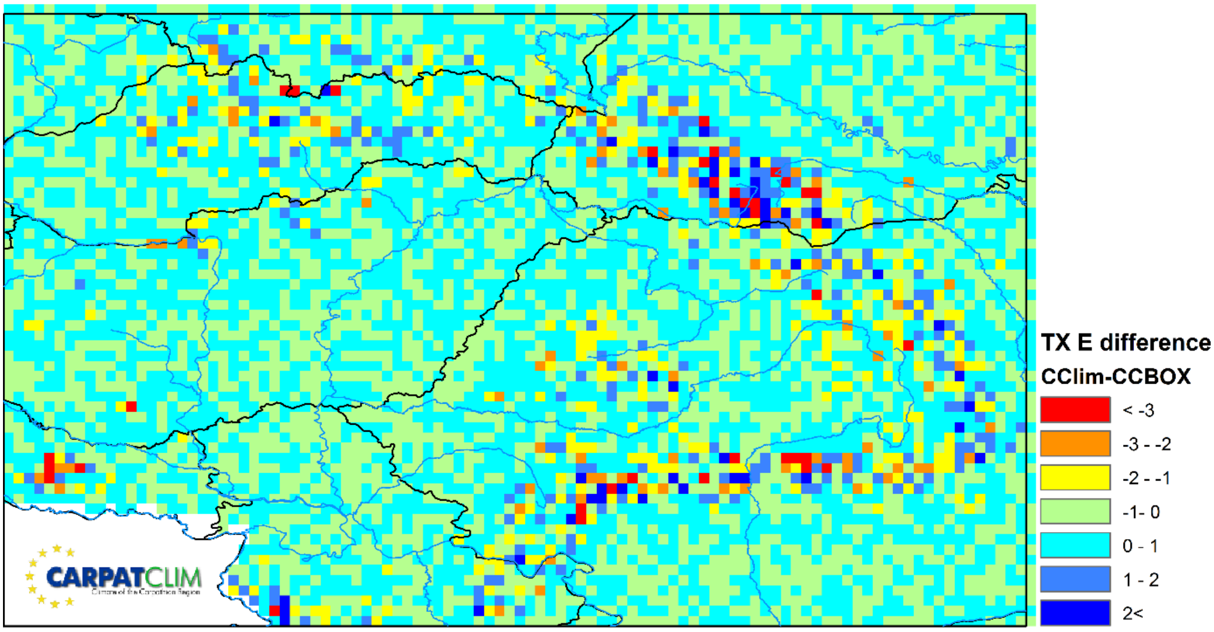


Fig. 6.: The difference between the two CarpatClim, Maximum temperature annual mean, 1961-2010

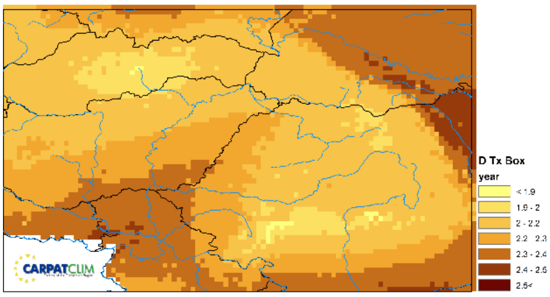


Fig. 7.: CarpatClim_BOX: temporal standard deviation values of annual mean maximum temperature for the period 1961-2010

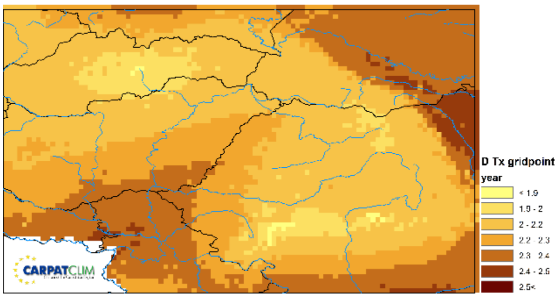


Fig. 8.: CarpatClim_ gridpoint: temporal standard deviation values of annual mean maximum temperature for the period 1961-2010

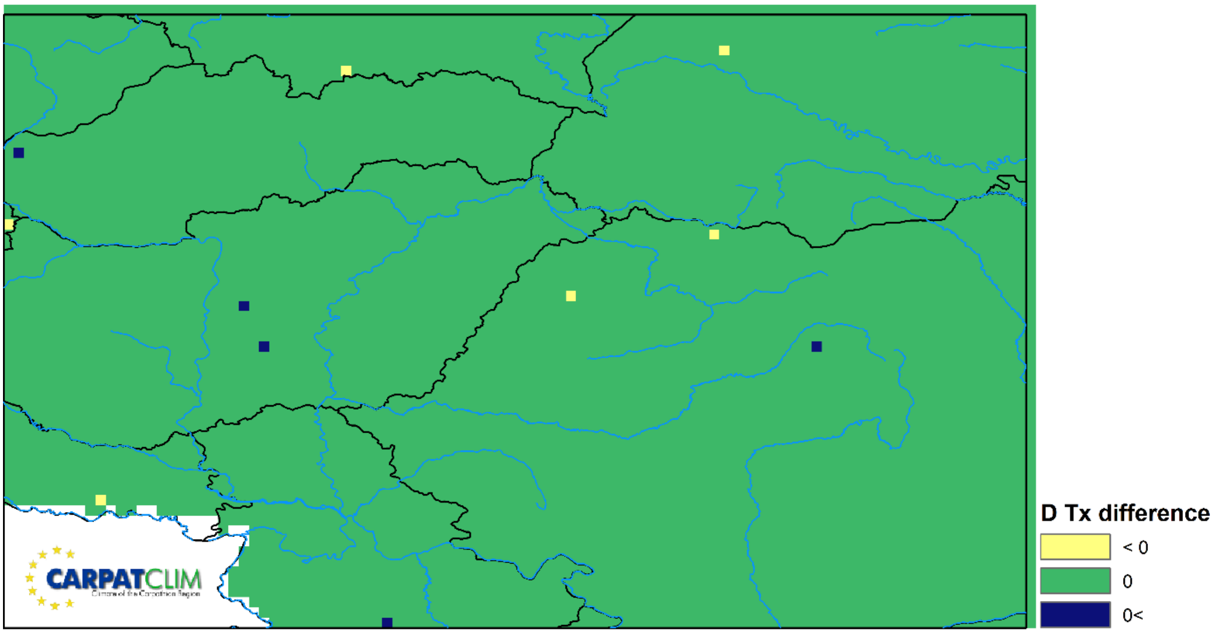


Fig. 9.: The difference between the two CarpatClim, temporal standard deviation, 1961-2010

4.2 SCORE RESULTS FOR MAXIMUM TEMPERATURE

The monthly RMSE and MSESS values for the year are shown in Fig. 10 and 11. Here, too, we can say that where the surface is variable, there are greater differences. Here we would note that these values are still smaller than if we compared any CarpatClim values with E-OBS or ERA5 datasets. For example, spatial average of RMSE values is 0.44 °C between the two CarpatClim datasets, while for E-OBS Carpatclim_box 0.67 °C and for E-OBs CarpatClim_gridpoint 0.85 °C.

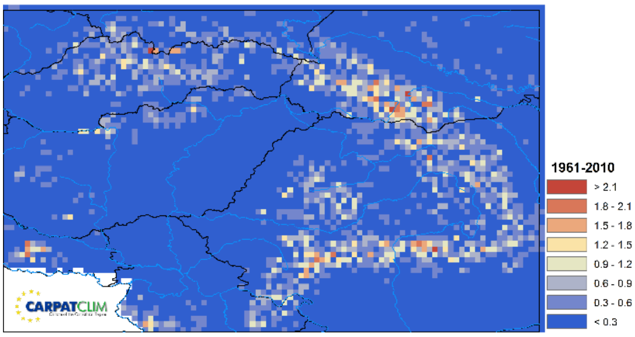


Fig. 10.: Monthly RMSE for the year, Maximum temperature, 1961-2010

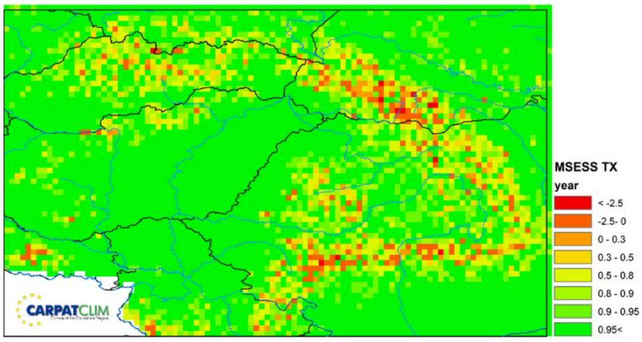


Fig. 11.: Monthly MSESS for the year, Maximum temperature, 1961-2010

4.3 ANOVA RESULTS FOR PRECIPITATION SUM

The most important ANOVA statistics for precipitation (Fig. 12) are similar to the results obtained for temperature values, the standard deviation decreased with averaging here as well.

CarpatClim gridBOX

Total mean: 700.96
Total variance: 39178.73

Spatial variance of temporal means: 23225.56
Spatial mean of temporal variances: 15953.17
Temporal variance of spatial means: 8293.03
Temporal mean of spatial variances: 30885.27

CarpatClim gridpoint

Total mean: 701.21
Total variance: 40565.66

Spatial variance of temporal means: 24571.22
Spatial mean of temporal variances: 15994.44
Temporal variance of spatial means: 8294.19
Temporal mean of spatial variances: 32271.01

Fig. 12.: The most important ANOVA statistics, annual precipitation sum

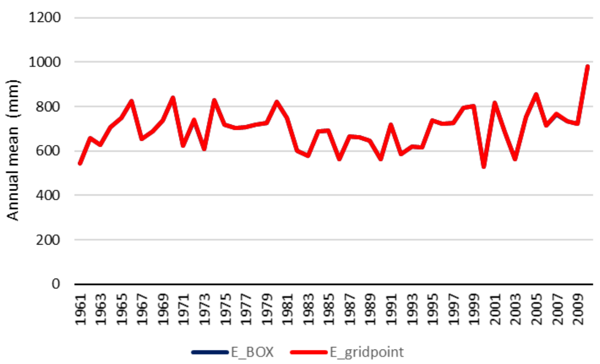


Fig. 13.: Spatial mean series of annual precipitation sum for the period 1961-2010

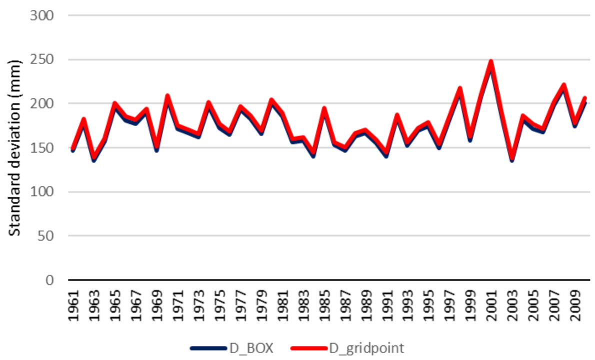


Fig. 14.: Spatial standard deviation series of annual precipitation sum for the period 1961-2010

Even for daily precipitation sum, the annual mean series values are the same (Fig. 13), while the standard deviations (Fig. 14) are slightly lower.

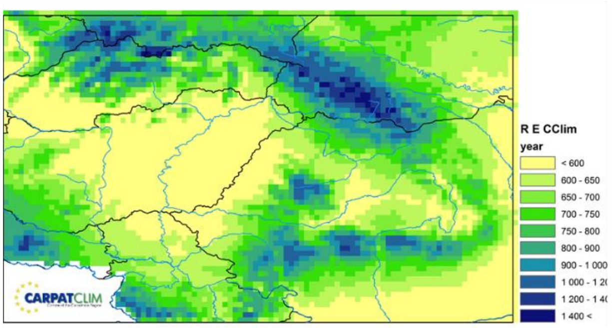


Fig. 15.: CarpatClim_gridpoint: temporal mean values of annual precipitation sum for the period 1961-2010

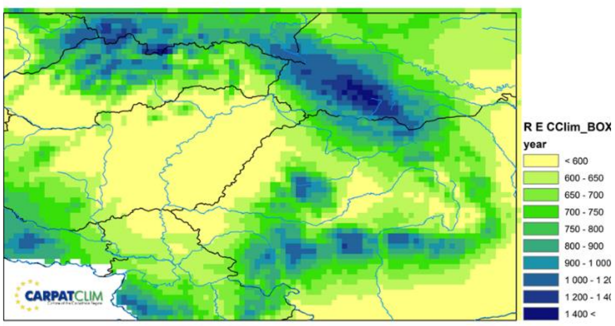


Fig. 16.: CarpatClim_BOX: temporal mean values of annual precipitation sum for the period 1961-2010

We displayed the temporal mean values on a map (Fig. 15-16) and the difference between the two CarpatClim in Fig. 17. Although we found that the values look like the same on average, it is clear that where the surface is more variable, the topography is different within a given cell, there are more differences. The greatest differences are in the case of mountains and valleys, even with the opposite sign.

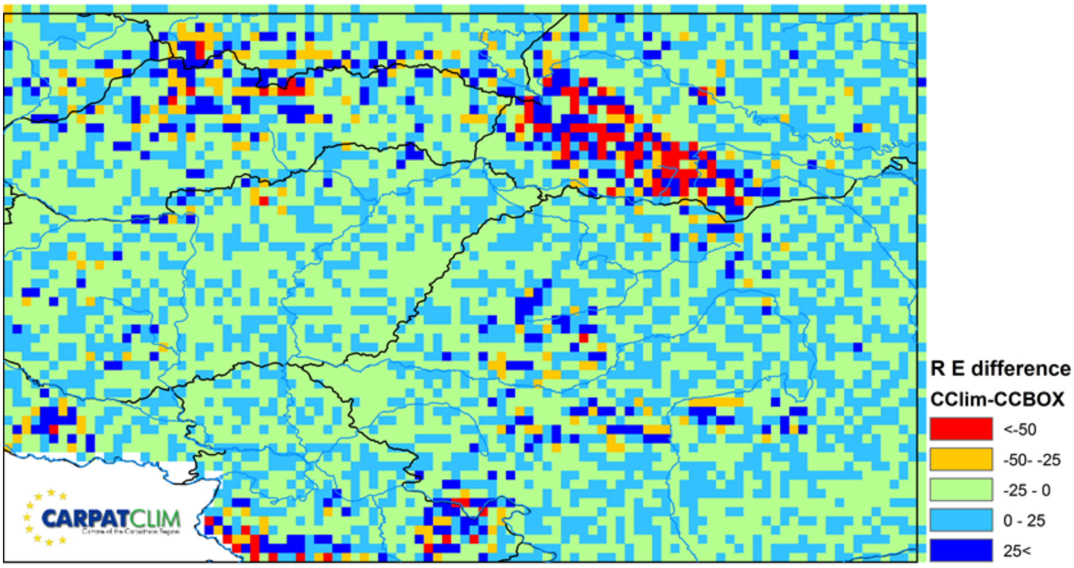


Fig. 17.: The difference between the two CarpatClim, Precipitation annual mean, 1961-2010

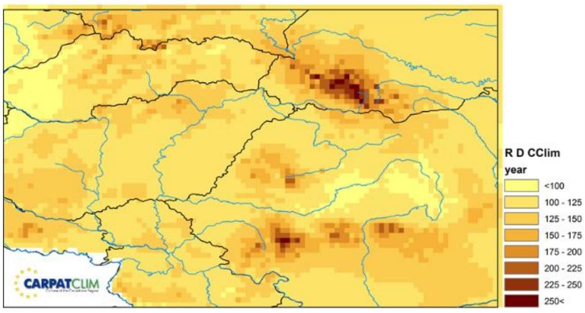


Fig. 18.: CarpatClim_gridpoint: temporal standard deviation values of annual precipitation sum for the period 1961-2010

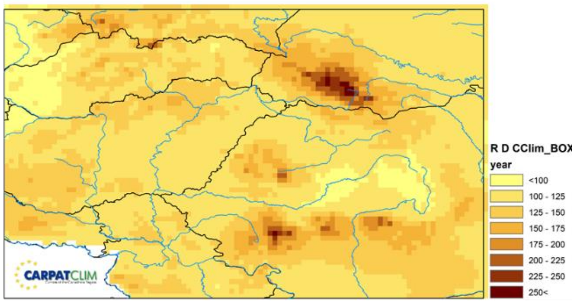


Fig. 19.: CarpatClim_BOX: temporal standard deviation values of annual precipitation sum for the period 1961-2010

In the same way, Figures 18-19. show the temporal standard deviations and Fig. 20 shows the difference between the two CarpatClim.

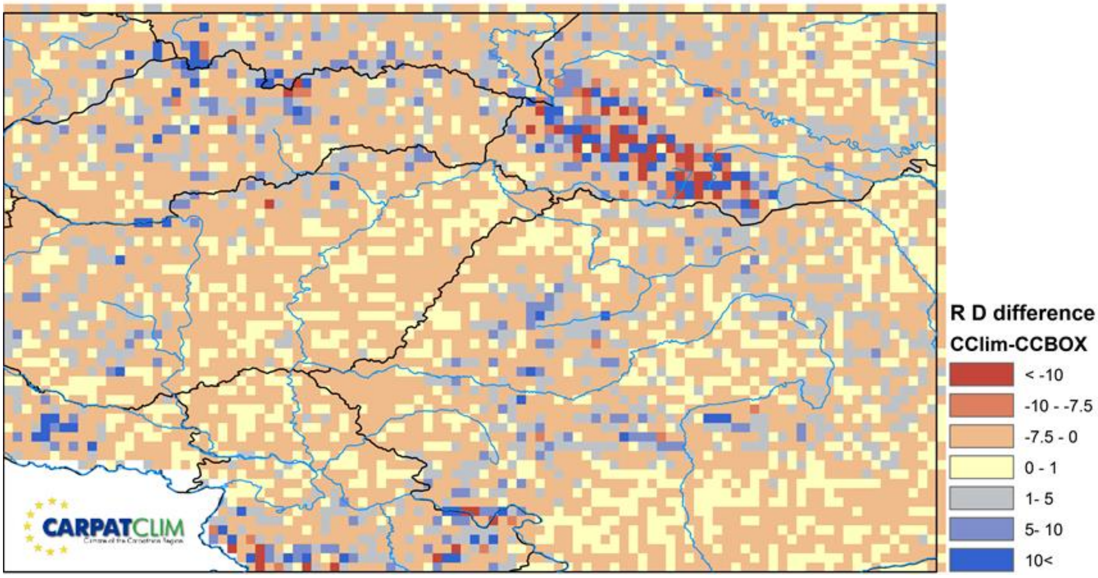


Fig. 20.: The difference between the two CarpatClim, temporal standard deviation, 1961-2010

4.4 SCORE RESULTS FOR PRECIPITATION SUM

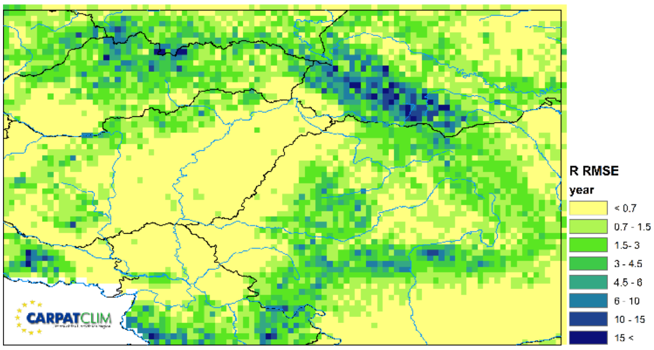


Fig. 21.: Monthly RMSE for the year, Precipitation, 1961-2010

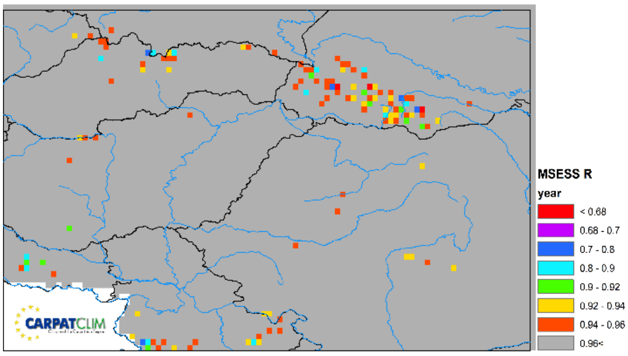


Fig. 22.: Monthly MSESS for the year Precipitation, 1961-2010

The monthly RMSE and MSESS values for the year are shown in Fig. 21 and 22. Here, too, we can say that where the surface is variable, there are greater differences and the values are the same on the flat surface. We note here that these values are significantly smaller than if we compared any CarpatClim values with E-OBS or ERA5 datasets.

5. CONCLUSION

At present we have two versions of CarpatClim gridded datasets for temperature (Tx, Tn) and precipitation, namely grid-point and grid-box average datasets. Comparison of these datasets was also happened, and according to the results, there is more difference between them where the surface is more complex: in the mountains, in the valleys, slopes and other surface forms.

For climatological studies, we need extremes, which can be inevitably lost by averaging. Therefore, we recommend the grid-point datasets for such applications. The methodology of climate modelling needs smoother data, so we recommend the grid-box average datasets, just as the output of the climate models is also cell average, so this version is more adequate for verification.

As regards the future this transformation procedure can be applied also for the other variables of the CarpatClim database by the software MISH and using the saved modelled statistical parameters. In addition the new version MISHv2.01 is planned to be able for gridding of data series, as grid-point and grid-box average series alike.

References

Haylock, M. R., N. Hofstra, A. M. G. Klein Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950{2006, J. Geophys. Res., 113 (D20), D20,119, doi:10.1029/2008JD010201.

Lakatos, M., Szentimrey, T., Izsák, B., Hoffmann, L., 2017: Comparison of the E-OBS and the CARPATCLIM gridded datasets for minimum temperatures, maximum temperatures and precipitation by the analysis of variance (ANOVA) method, Proceedings of the 9th Seminar for Homogenization and Quality Control in Climatological Databases and 4th Conference on Spatial Interpolation Techniques in Climatology and Meteorology (Ed. Szentimrey T, Lakatos M, Hoffmann L), Budapest, Hungary, 2017, WCDMP-No. 85, pp. 151-1664

Szentimrey, T., 1999: Multiple Analysis of Series for Homogenization (MASH), Proceedings of the Second Seminar for Homogenization of Surface Climatological Data, Budapest, Hungary; WMO, WCDMP-No. 41, pp. 27-46.

Szentimrey, T., Bihari, Z., 2007: Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis), Proceedings of the Conference on Spatial Interpolation in Climatology and Meteorology, Budapest, Hungary, 2004, COST Action 719, COST Office, 2007, pp. 17-27

Szentimrey, T., 2008: Development of MASH homogenization procedure for daily data. Proceedings of the Fifth Seminar for Homogenization and Quality Control in Climatological Databases, Budapest, 2006, WCDMP-No. 71, WMO/TD-NO. 1493, 123–130.

Szentimrey, T, Bihari, Z., Lakatos, M., Szalai,S., 2011: Mathematical, methodological questions concerning the spatial interpolation of climate elements. Proceedings of the Second Conference on Spatial Interpolation in Climatology and Meteorology, Budapest, Hungary, 2009, Időjárás 115, 1-2, 1-11

Szentimrey T, Lakatos M, Bihari Z, Kovacs T, Nemeth A, Szalai S, Auer I, Hiebl J, Milkovic J, Štěpánek P, Zahradnicek P, Tolasz R, Kilar P, Pyrc R, Limanowka D, Cheval S, Matei M, Kajaba P, Ivanakova G, Bochnicek O, Nejedlik P, Statsny P, Mihic D, Petrovic P, Savic T, Skrynyk O, Nabyvanets Y, Gnatiuk N, 2012a: Final report on quality control and data homogenization measures applied per country, including QC protocols and measures to determine the achieved increase in data quality. Carpatclim Project, Deliverable D1.12, http://www.carpatclim-eu.org/docs/deliverables/D1_12.pdf

Szentimrey T, Lakatos M, Bihari Z, Kovacs T, Nemeth A, Szalai S, Auer I, Hiebl J, Milkovic J, Štěpánek P, Zahradnicek P, Tolasz R, Kilar P, Pyrc R, Limanowka D, Cheval S, Matei M, Kajaba P, Ivanakova G, Bochnicek O, Nejedlik P, Statsny P, Mihic D, Petrovic P, Savic T, Skrynyk O, Nabyvanets Y, Gnatiuk N, 2012b: Final version of gridded datasets of harmonized and spatially interpolated meteorological parameters, per country. Carpatclim Project, Deliverable D2.8, http://www.carpatclim-eu.org/docs/deliverables/D2_8.pdf

Szentimrey T, Lakatos M, Bihari Z, Kovacs T, Nemeth A, Szalai S, Auer I, Hiebl J, Milkovic J, Štěpánek P, Zahradnicek P, Tolasz R, Kilar P, Pyrc R, Limanowka D, Cheval S, Matei M, Kajaba P, Ivanakova G, Bochnicek O, Nejedlik P, Statsny P, Mihic D, Petrovic P, Savic T, Skrynyk O, Nabyvanets Y, Gnatiuk N, 2012c: Final report on the creation of national gridded datasets, per country. Carpatclim Project, Deliverable D2.9, http://www.carpatclim-eu.org/docs/deliverables/D2_9.pdf

Szentimrey,T., Bihari, Z., 2014: Manual of interpolation software MISHv1.03, Hungarian Meteorological Service, p. 60.

Szentimrey, T., 2017a: Manual of homogenization software MASHv3.03, Hungarian Meteorological Service,p.71.

Szentimrey, T., 2017b: New developments of interpolation method MISH: modelling of interpolation error RMSE, automated real time Quality Control. Proceedings of the 9th Seminar for Homogenization and Quality Control in Climatological Databases and 4th Conference on Spatial Interpolation Techniques in Climatology and Meteorology, Budapest, Hungary, 2017, WCDMP-No. 85, pp. 115-124

Szentimrey, T.,2019a: Transformation of CarpatClim datasets for grid-box average datasets, Study for Hungarian Meteorological Service (COPERNICUS C3Surf project (WP4)), Varimax Limited Partnership, Budapest, Hungary, p. 7.

Szentimrey, T.,2019b: Mathematical methodology and software for comparison of gridded datasets, Study for Hungarian Meteorological Service (COPERNICUS C3Surf project (WP4)), Varimax Limited Partnership, Budapest, Hungary, p. 10.

COMPARATIVE STUDY OF CARPATCLIM, E-OBS AND ERA5 DATASET

[1]Mónika Lakatos, [2]Tamás Szentimrey, [1]Beatrix Izsák, [1]Olivér Szentes, [1]Lilla Hoffmann, [1]Andrea Kircsi, [1]Zita Bihari

[1]Hungarian Meteorological Service, Budapest, Hungary
[2]Varimax Limited Partnership, Budapest, Hungary
lakatos.m@met.hu

Abstract

Recently the pan-European observational dataset E-OBS has been considered as a reference for several European climate analyses. Moreover, the usage of the newly available global reanalysis ERA5 is increasing for climate change studies. CARPATCLIM is a regional climate dataset for the Carpathian region, which is situated in central-eastern Europe. The E-OBS and ERA5 dataset were tested against CARPATLIM and against other regional datasets in the framework of the COPERNICUS C3S_311a_Lot4 project. The common time period of E-OBS, ERA5 and CARPATCLIM is the period of 1979-2010. Different measures, evaluation statistics were computed for comparison of the gridded Tx, Tn and precipitation fields for this period. Analysis of Variance (ANOVA) method was applied for instance, which is an adequate statistical method to explore the statistical structure of different datasets. ANOVA can be used effectively for the characterization of the spatiotemporal statistical properties of CARPATCLIM, E-OBS and ERA5. In addition, different evaluation scores, yearly cycle, absolute and monthly extremes, quantiles, wet days frequency, several climate indices for temperature were computed and reported in the COPERNICUS C3S_311a_Lot4 project. Trend analysis (exponential trend model for precipitation and linear trend model for temperature) and homogeneity test for the gridded data were applied too. The differences between the datasets come from the station density behind the grids and also the methods used for homogenization and gridding determine the results. The main outcomes of this comparative study are presented on graphs and maps in this paper.

1. INTRODUCTION

The pan-European observational dataset E-OBS (Haylock et al., 2008) has been widely used as a reference for climate analysis in Europe. After a major update an evaluation is necessary to find out about the suitability of E-OBS. In order to do so, the updated E-OBS dataset is compared to several sub-regional high-resolution observational climate datasets in the framework of the COPERNICUS C3Surf (C3S_311a_Lot4): Climate monitoring products for Europe based on surface in-situ observations project in Work Package 4 (WP4). Moreover, the newly available global reanalysis ERA5 was tested and documented against sub-regional datasets and then compared to E-OBS too in the WP4:Sub-regional datasets. To examine the performance of E-OBS and the reanalysis datasets under different circumstances, three European subregions are considered: the Carpathian region, Fennoscandia and the Alpine region in WP4 of C3S_311a_Lot4 initiative. The detailed comparison of daily precipitation, daily maximum, minimum and average temperatures can be found in the project report (Bandhauer, et al., 2020). In addition to that, the evaluation results for precipitation are summarized in an article submitted to the International Journal of Climatology (Bandhauer, et al., 2021).

The updated E-OBS dataset and the newly derived ERA5 reanalyses dataset is examined with the help of several statistical tests. These statistical tests include ANOVA (Analysis of Variance) show the general, seasonal and extreme behaviour and the deviations between the examined datasets. After an overview of the CARPATCLIM dataset, the examined statistical measures of maximum and minimum temperature and daily precipitation are presented here. Moreover, the comparison of trend and the residual inhomogeneity in the datasets are illustrated here as well.

2. DATA

The CARPATCLIM is a regional dataset available for the Carpathian Region. It covers approximately 500 000 km^2 in Europe. The CARPATCLIM consortium (led by the Hungarian Meteorological Service) consists of nine participant countries in the Larger Carpathian Region created a homogenised, harmonized, gridded dataset to support the investigation of the climate change in the region. CARPATCLIM dataset encompassing gridded daily observations for the Carpathian region. It is available on a 0.1°(~10 km×10 km) grid and includes homogenized, gridded daily time series of various meteorological parameters from 1961 to 2010. It consists of several ECVs: temperature (maximum, minimum, average), precipitation, 10 and 2 meter wind speed, 10 meter wind direction, maximum 10 meter wind speed, sunshine duration, cloud cover, global radiation, relative humidity, air pressure, vapour pressure, snow depth (modelled), and an extensive number of climate indices (37 indices, including 7 drought indices). The gridded fields are based on a dense station network consisting of 415 climate stations and 904 precipitation stations. The method and software used in CARPATCLIM project (www.carpatclim-eu.org.) for data quality control, homogenization and data completion was the MASH (Szentimrey, 1999, 2008, 2017). Interpolation of the homogenized time series was carried out by applying the MISH (Szentimrey and Bihari, 2007, 2014) procedure. The grids are publically available on the CARPATCLIM webpage: http://www.carpatclim-eu.org/pages/home/

Conversion of CARPATCLIM dataset was necessary before comparisons. It was re-gridded onto the 0.1° x 0.1° regular latitude-longitude grid of E-OBS to make them comparable. During the re-gridding of CARPATCLIM to the E-OBS grid, a 0.05° shift between the grid systems was found. In order to use the same grid points in all analyses, hence the E-OBS was re-gridded onto the CARPATCLIM grid by applying the nearest neighbour method. Regarding the ERA5, the data was downscaled from 0.25° to 0.1° with a bilinear interpolation before comparison work. Beside this, the CARPATCLIM dataset includes interpolated data for gridpoints, while E-OBS includes gridbox averages. Thanks to the special methodology of MISH interpolation method the CARPATCLIM data, which are valid at grid points can be converted to grid box average (Szentimrey, 2019a). Notably the speciality of MISH that the statistical parameters - like spatial trend and correlation structure - are modelled and saved for further computations.

3. EVALUATION OF DAILY MAXIMUM AND DAILY MINIMUM TEMPERATURES

The methods and formulas were used in comparison are detailed in Szentimrey 2019b and also in the Annex of the WP4 project report. The ANOVA results are shown at first here for daily maximum and minimum temperatures. The ANOVA can be used effectively for the characterization of the spatio-temporal properties of CARPATCLIM, E-OBS and ERA5 datasets (Lakatos et al., 2017). The main principle of the ANOVA is that the total variance can be partitioning as the sum of the spatial variance of the temporal means and the spatial mean of the temporal variances on one hand; and the sum of the temporal variance of the spatial means and the temporal mean of the spatial variances on the other hand. The magnitude and spatial distribution of the specific components of the total variance can be analyzed for different time periods, years and seasons for instance. When a dataset can capture the spatio-temporal variability of a given climate parameter well, that can be settled as a dataset with better quality in general. This methodology is built into the modelling part of method MISH in order to evaluate the modelling results automatically.

The spatial means and spatial variances at the moment *t* are illustrated on graphs and temporal means and temporal variances at a specific location *s* are illustrated on maps and compared.

Notations:

Et (s) - temporal mean at location s
Es (t) - spatial mean at time t
Dt (s) - temporal st. deviation at location s
Ds (t) - spatial st. deviation at time t

The Et (s), Dt (s) are illustrated on maps, and Es (t), Ds (t) are shown on graphs.

The derived statistics are as follows:

E - total mean
Dt - spatial mean of temporal st. deviations
Ds - temporal mean of spatial st. deviations
DEt - spatial st. deviation of temporal means
DEs - temporal st. deviation of spatial means

Besides the yearly characteristics, the spring and summer means and variances were analysed in this study. The spring and summer have been chosen from the seasons for presenting here, as the preliminary analysis showed greater differences between datasets in those seasons.

The daily maxima in ERA5 can be characterized with lower spatial variances than the observational datasets (Figure 1.). The Figure 1. indicates that the lowest daily minima and the smallest variances appear in CARPATCLIM and the highest ones in ERA5.

The Figure 2. shows the time series of the "Es (t)-spatial mean" and "Ds (t)- spatial st. deviation" from 1979 to 2010. The CARPATCLIM and E-OBS yearly, spring and summer means of the daily maximum temperatures are running closely, but the curves based on ERA5 are going bottommost. The largest differences between CARPATCLIM and ERA5 can be seen in summer, 0.63°C on average for the period 1979-2010. The spatial standard deviation characterizing E-OBS and ERA5 remains below CARPATCLIM in the whole period, particularly in spring. There is a large jump in spatial standard deviation of the spring average daily maximum temperatures in E-OBS in 2005, reaching 2.79°C (Figure 2.). The yearly and seasonal means derived from E-OBS and CARPATCLIM are close in the case of daily minima. The mean and standard deviation series run together from 1979-2010, the ERA5 is running at the uppermost (Figure 3.).

The maps in the (Figure 4. and 5.) show the spatial distribution of the yearly means and standard deviations of the daily maximum temperatures for the CARPATCLIM region. Lower mean values at higher elevation are obvious in all three dataset. The warmest area goes up in the Great Hungarian plain to a larger extent in CARPATCLIM than it is appearing in E-OBS and in ERA5 as well. Smallest standard deviation of the yearly average Tx shows up in the mountainous area at all three dataset, mentionable that the standard deviation of the yearly average Tx is low in ERA5 almost in the whole territory of Romania (Figure 5.). Patches with relative higher standard deviation values to the surroundings turn out on the yearly map in the foreground of the Eastern Carpathians in Romania in E-OBS (Figure 5.).

The features of the topography can be explored on the map of the yearly mean minima (Figure 6.). Higher values than in CARPATCLIM (>7°C yearly averages) are standing out between the Lake Balaton and the Danube River in E-OBS in Hungary and in extended regions in Serbia in ERA5 too. The standard deviation of the yearly average daily minima is decreasing from north to south across the region (Figure 7.). The patches with higher standard deviation on the yearly map of E-OBS tracing out the locations of the measuring stations which are involved in E-OBS.

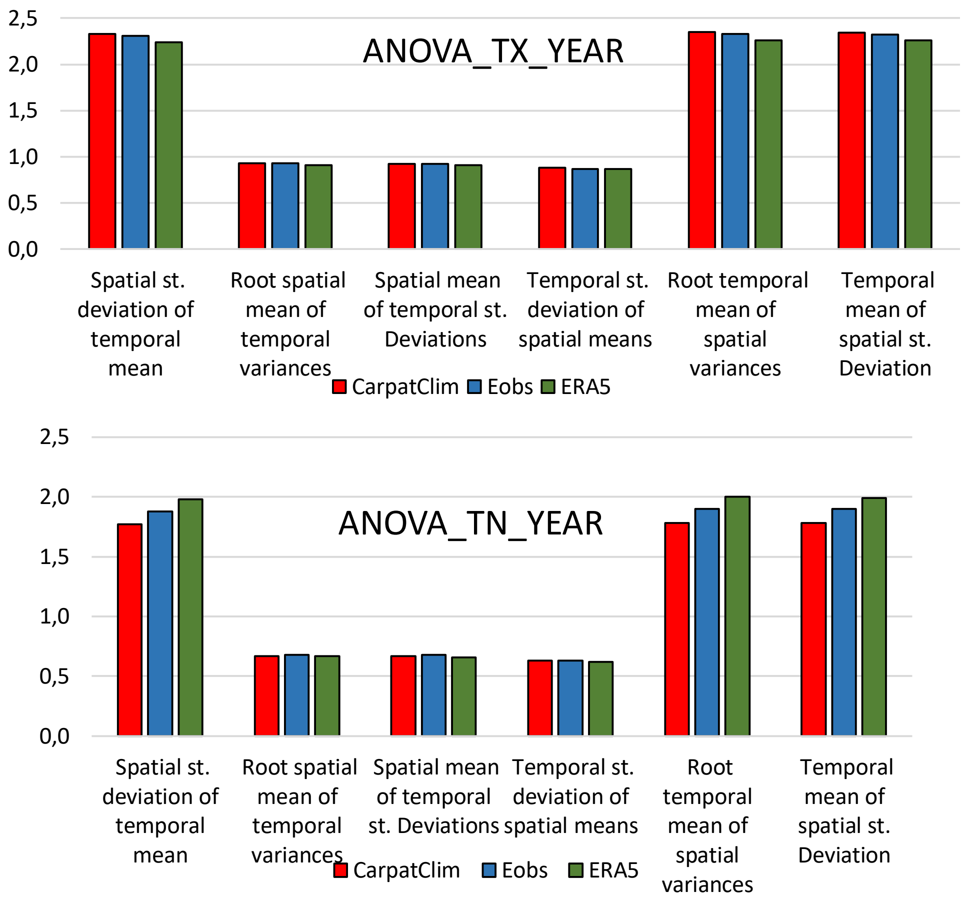
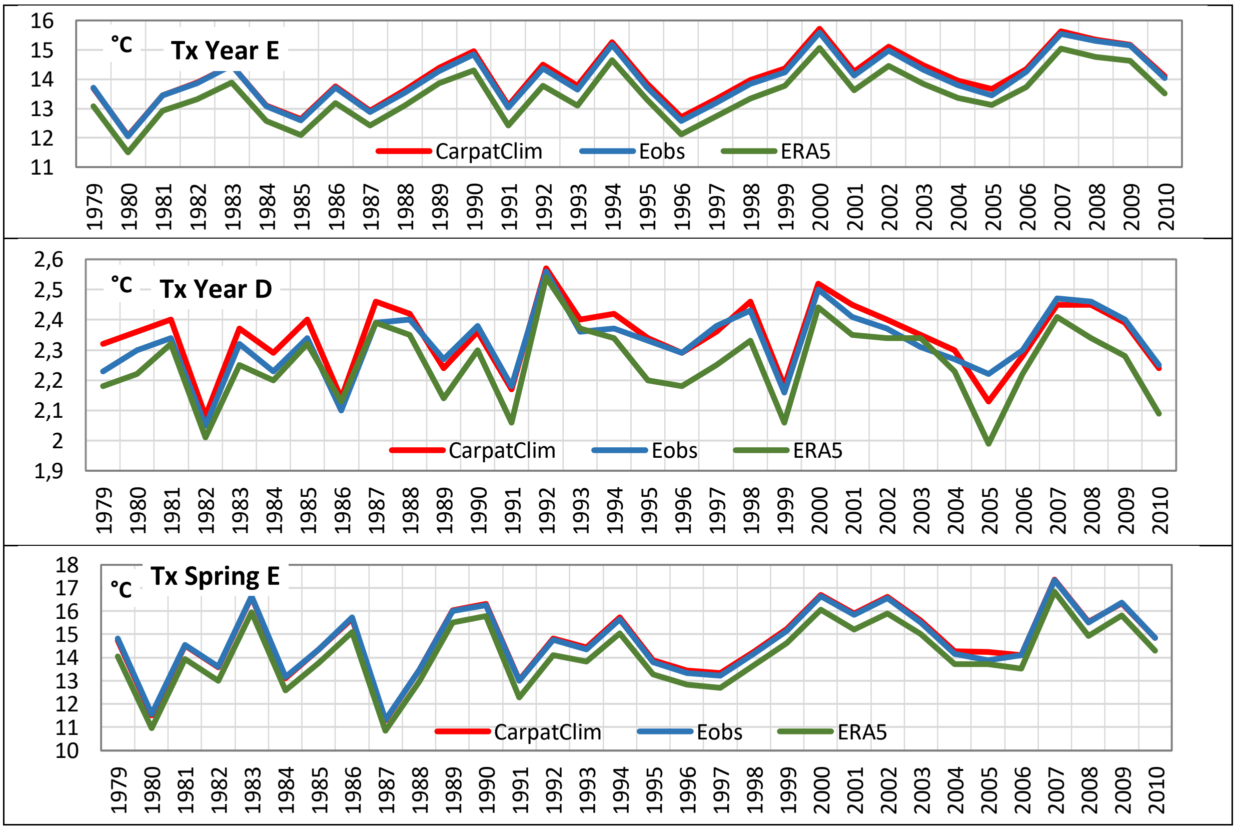


Fig. 1.: The output statistics of ANOVA of daily maximum and minimum temperatures for CARPATCLIM, E-OBS and ERA5 for the period 1979-2010.



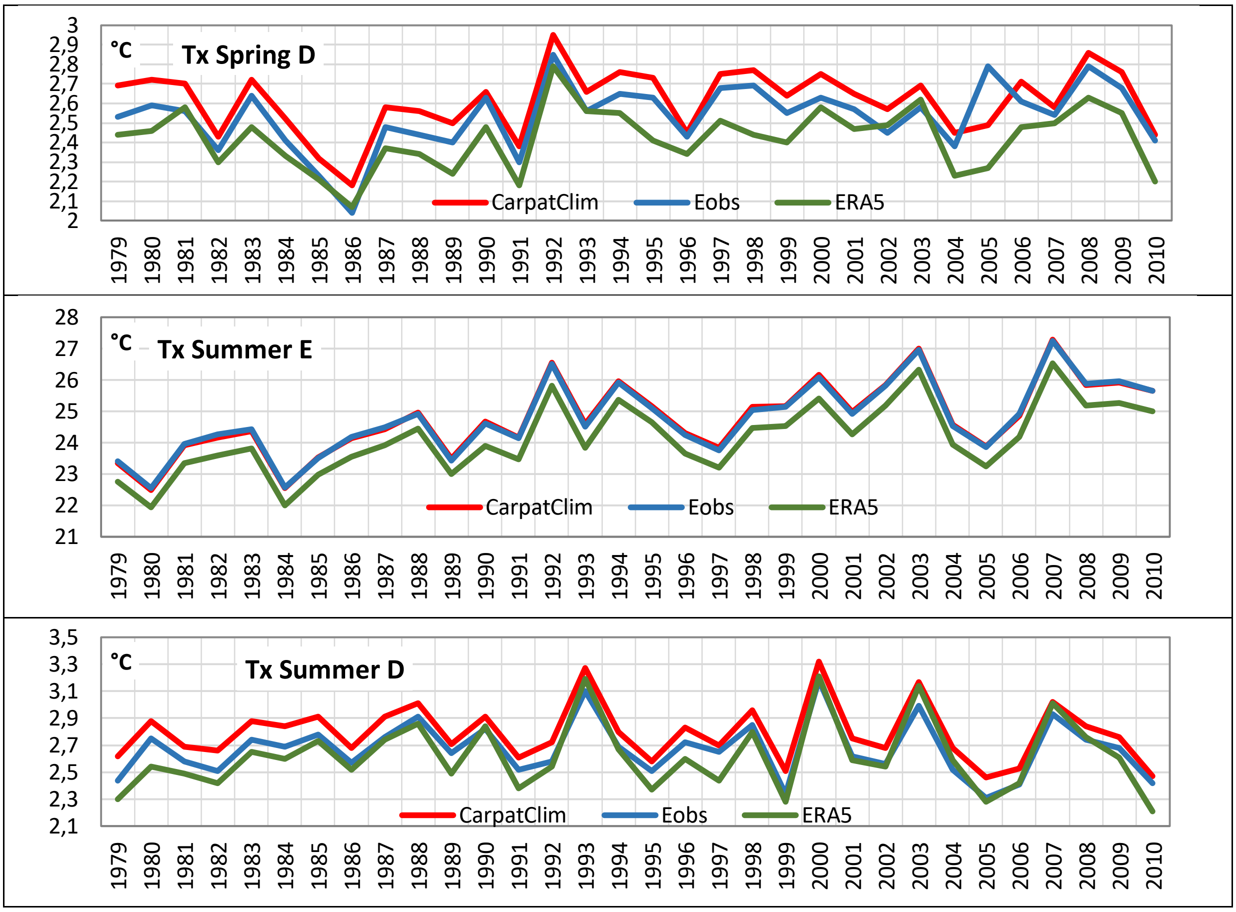
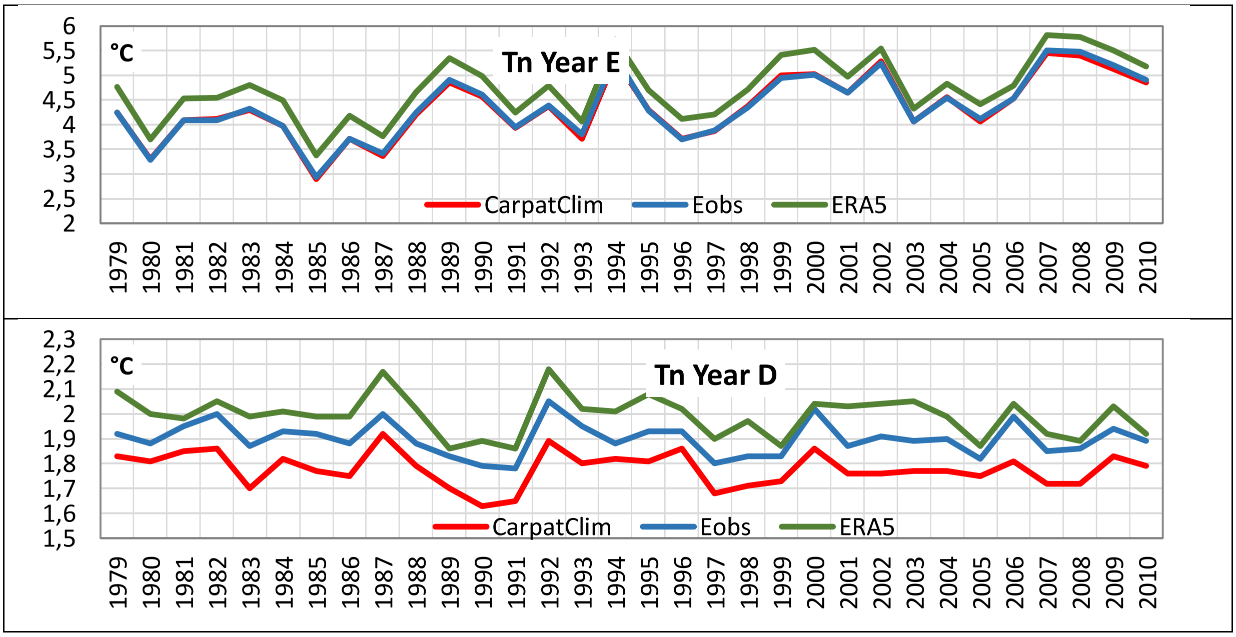


Fig. 2.: Yearly, spring and summer Es (t)-spatial mean and Ds (t)- spatial st. deviation from 1979-2010 for CARPATCLIM, E-OBS and ERA5.



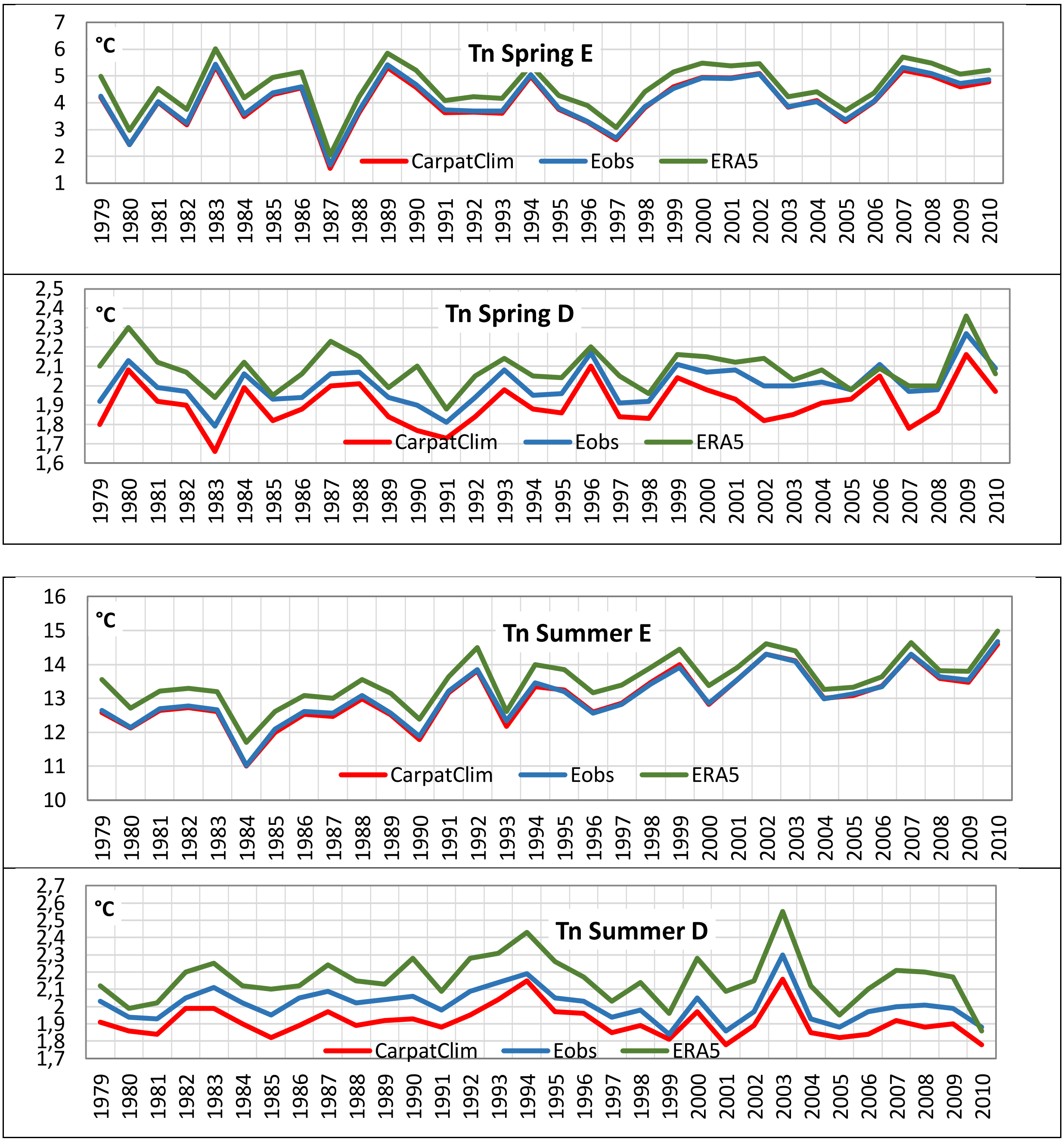


Fig. 3.: Yearly, spring and summer Es (t)-spatial mean and Ds (t)- spatial st. deviation for CARPATCLIM, E-OBS and ERA5 dataset in the period 1979-2010.

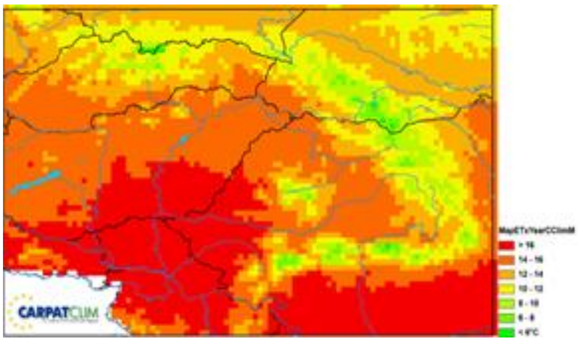
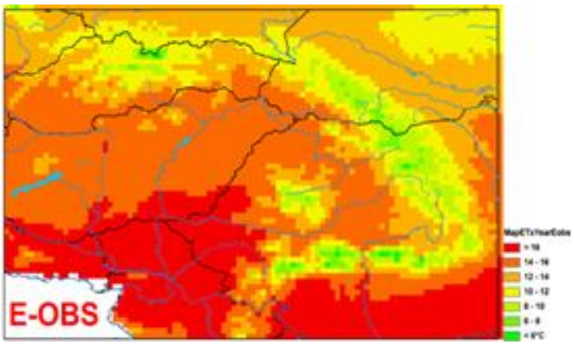
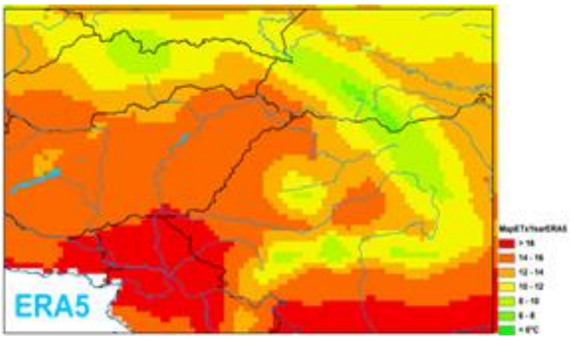
Tx CARPATCLIM EOBS, ERA5 E year	
	
	

Fig. 4.: Et (s)-temporal mean of the yearly average daily maximum temperatures for CARPATCLIM, E-OBS and ERA5 datasets in the period 1979-2010.

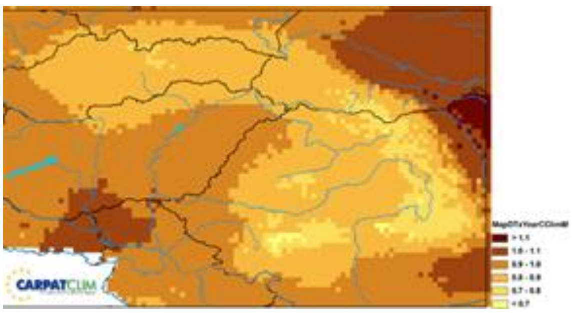
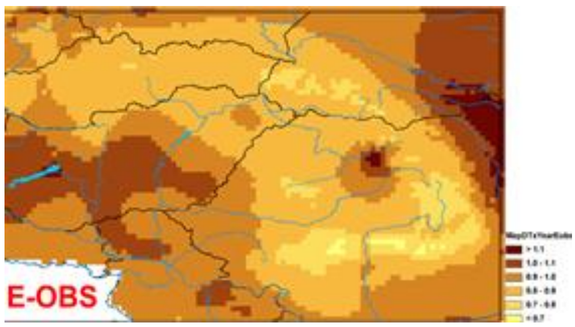
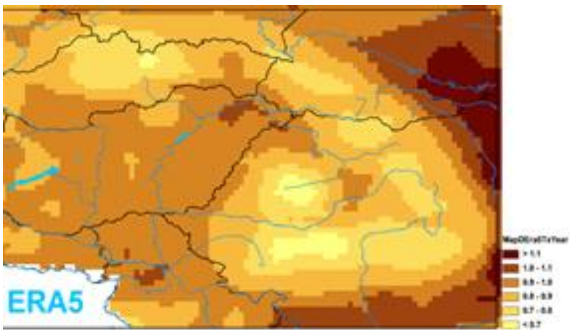
Tx CARPATCLIM EOBS, ERA5 D year	
	
	

Fig. 5.: Dt (s)-temporal st. deviation of the yearly average daily maximum temperatures for CARPATCLIM, E-OBS and ERA5 datasets in the period 1979-2010.

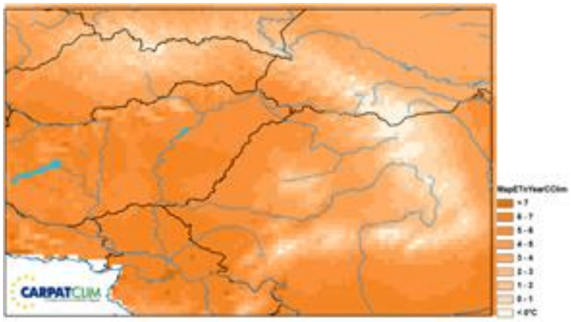
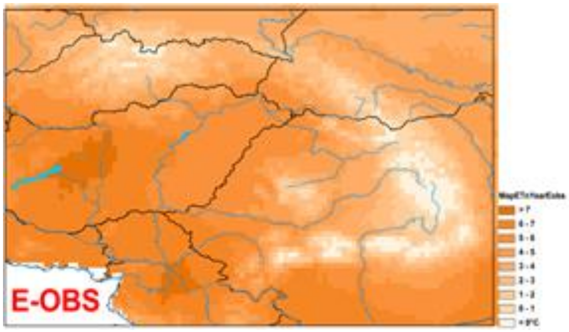
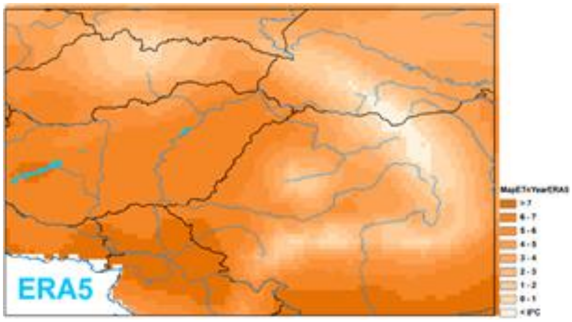
Tn CARPATCLIM EOBS, ERA5 E Year	
	
	

Fig 6.: Yearly average (Et (s)-temporal mean) daily minimum temperatures for CARPATCLIM, E-OBS and ERA5 datasets in the period 1979-2010.

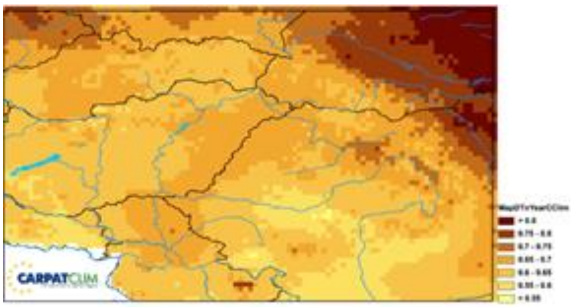
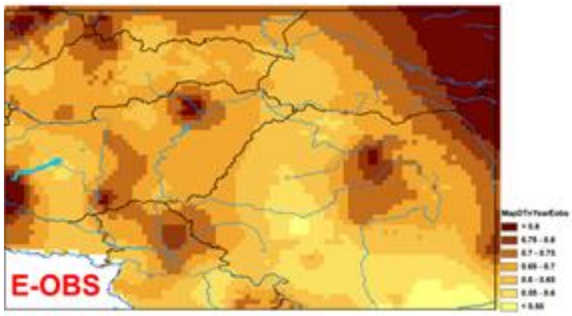
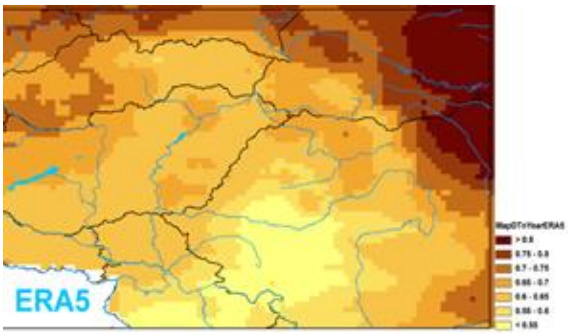
Tn CARPATCLIM EOBS, ERA5 D Year	
	
	

Fig. 7.: Dt (s)-temporal st. deviation of the yearly average daily minimum temperatures for CARPATCLIM, E-OBS and ERA5 datasets in the period 1979-2010.

3.1 TRENDS

The trend maps in Figure 8-10. are remarkably different. The spatial distribution of the change of the daily maximum temperature from 1979 to 2010 is unexpectedly heterogeneous in E-OBS. Very high change, exceeding 2.5 °C comes up in mid-Hungary with the largest values at the Lake Balaton, while the change is less than 0.4 °C to the north-east in Hungary near the Slovakian border in the environment of Miskolc weather station (Figure 8. top right). The probable explanation of the high variability of the trend of the daily maxima in E-OBS is the less effective performance of QC and homogenization procedure. The trend fitted to ERA5 gridded time series indicates lower climate change signal in the Carpathian region as a whole (Figure 8., bottom left), though inexplicable high change appears where the Slovakian, Hungarian and Ukraine border meet.

The influence of the continentality can be explored on the CARPATCLIM summer change map (Figure 9, top left). The map that illustrates the E-OBS summer change is highly variable in the region and causelessly spotted, especially in the Carpathian Basin. For ERA5 the spatial distribution of the summer change is similar to yearly with the larger trend along the Slovak and Hungarian border, what is missing from CARPATCLIM.

Extremely high spatial variability turns up on the trend map representing the changes of the daily minimum temperatures for E-OBS (Figure 10., top right). Regions marked with very big and very low increase appear in Transdanubia in small a distance and in Slovakia, moreover E-OBS gives decreasing signal of minimum temperature in Romania narrow regions. The ERA5 underestimate the change of the minimum temperature in wider region south from the midline of the domain.

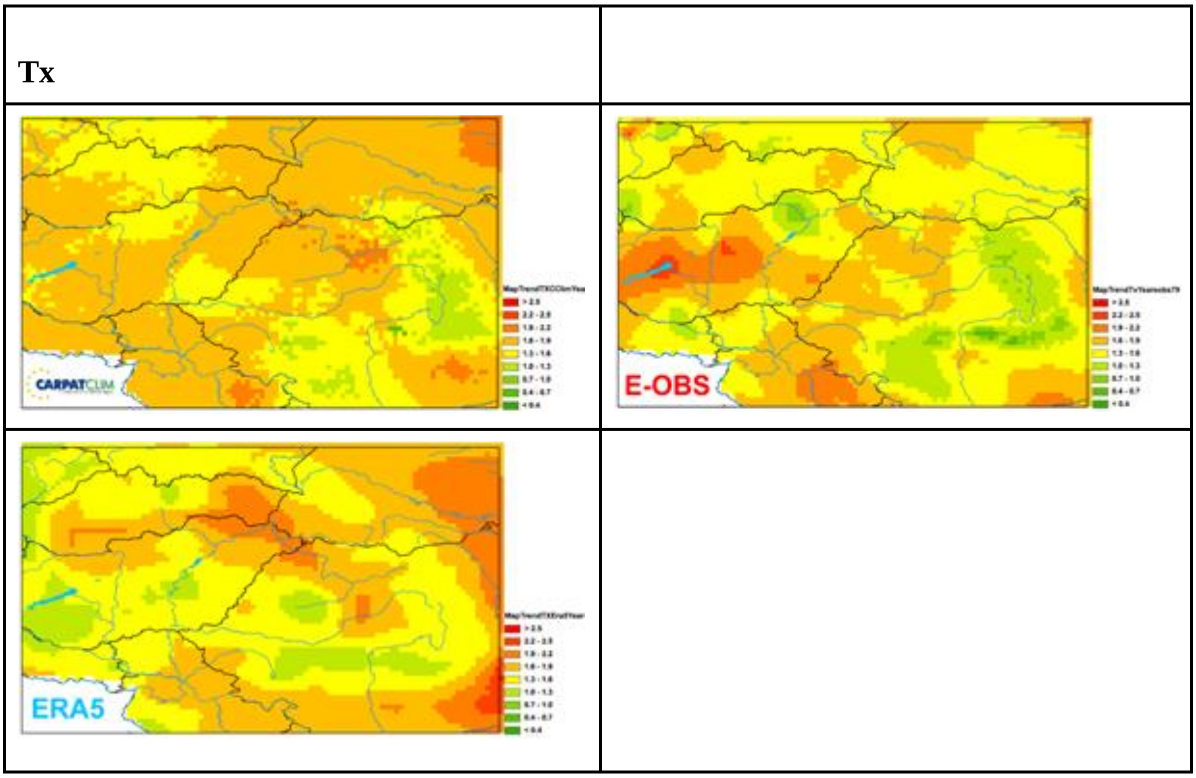


Fig. 8.: Linear trend for average annual Tx temperature in degC/32year. for CARPATCLIM (top left), EOBS (top right) and ERA5 (bottom left) over the time period 1979-2010.

Tx summer

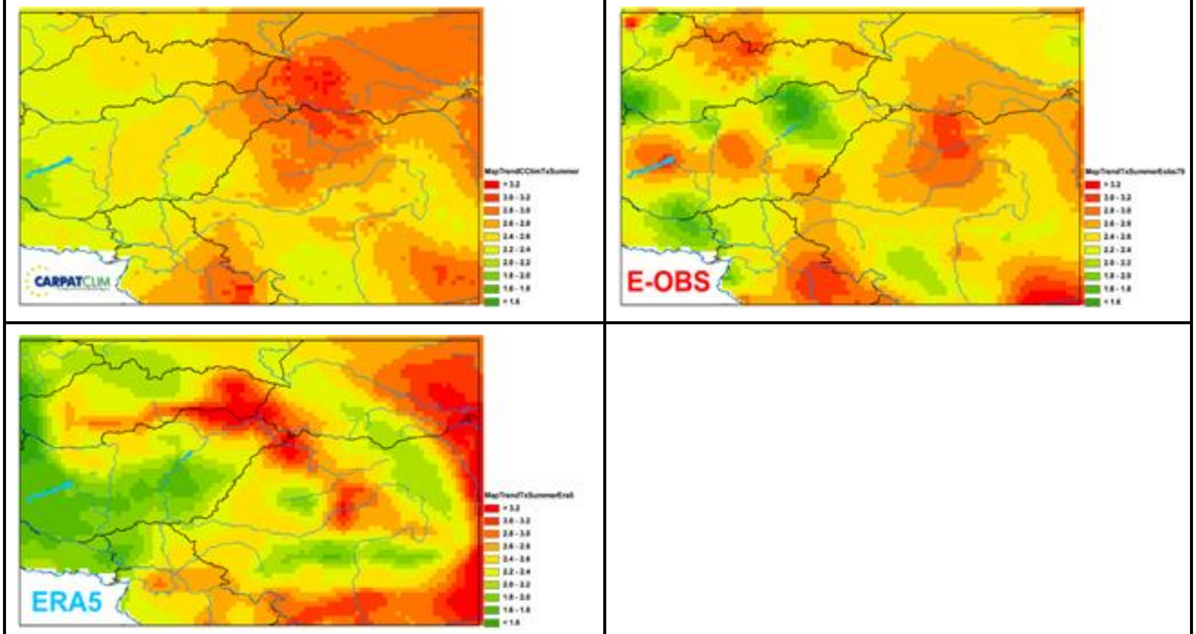


Fig. 9.: Linear trend for average summer Tx temperature in degC/32year. for CARPATCLIM (top left), EOBS (top right) and ERA5 (bottom left) over the time period 1979-2010.

Tn

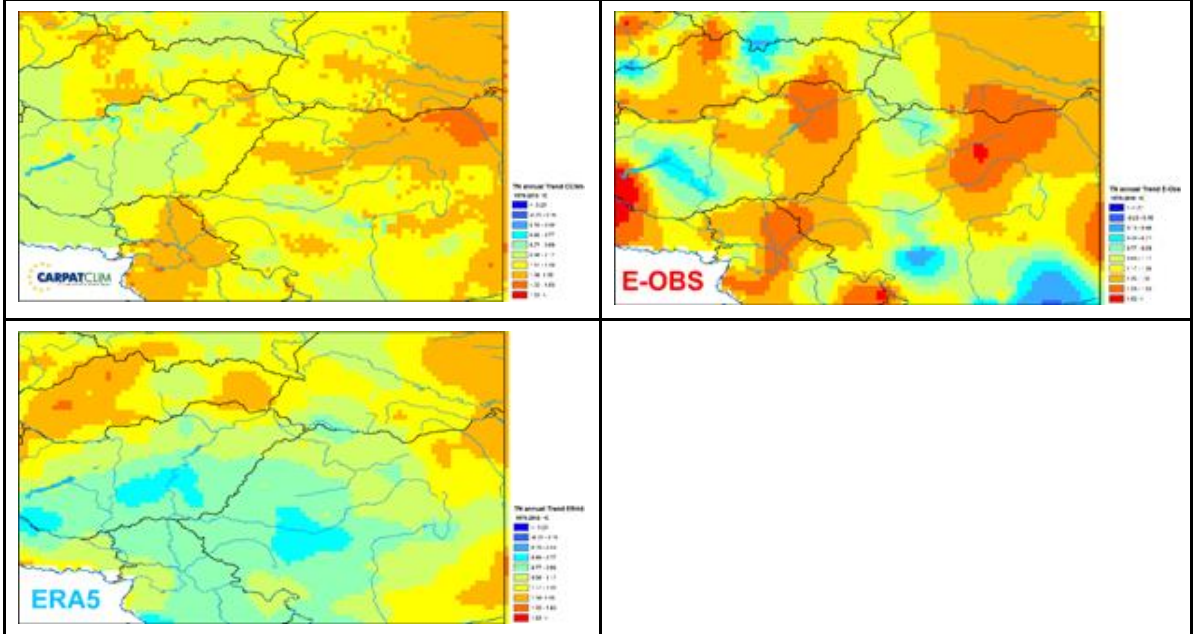


Fig. 10.: Linear trend for average annual Tn temperature in degC/32year. for CARPATCLIM (top left), EOBS (top right) and ERA5 (bottom left) over the time period 1979-2010.

3.2 HOMOGENEITY TEST FOR TEMPERATURE

However, the E-OBS gridded series were derived from homogenized station data series there is no information on the quality of the gridded series. The MASH (Multiple Analysis of Series for Homogenization, Szentimrey; 2014) software system for homogenization consists of functions for testing the residual inhomogeneity in any dataset. For making this test the closest grid points to 51 Hungarian meteorological stations (Figure 11.) were selected from CARPATCLIM and also from E-OBS, then the gridded daily maximum and minimum temperature series from 1961-2010 (the longest overlap between the datasets) were tested by MASH homogenization method.

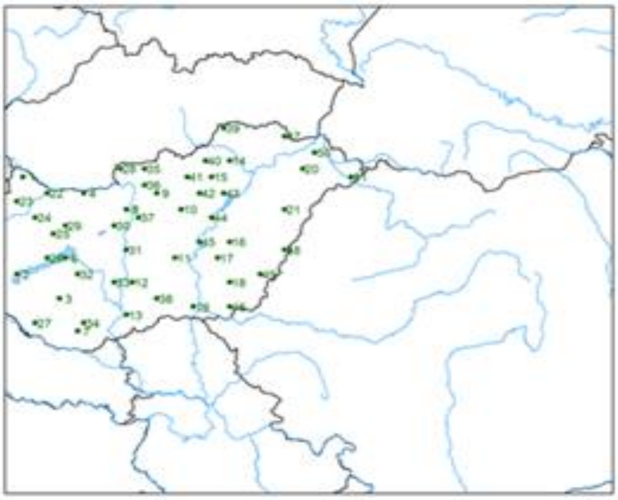


Fig. 11.: The location of the stations which were used in the homogeneity test.

The test statistics for inhomogeneity of the gridded series are listed in tables below. The null hypothesis is that the examined gridded series are homogeneous. The critical value related to significance level 0.05 comes to 20.86 Test statistics (TSA) can be compared to the critical value. The larger TSA values are the more suspicious. The highest TSA values are of the order of hundreds in many cases in E-OBS and appear at grid-points near the stations where inhomogeneities were found and eliminated by MASH during preparation of CARPATCLIM.

3.2.1. Homogeneity test for TN

Test Statistics After Homogenization **for TN in E-OBS**

Table 1.: The test statistics in decreasing order for E-OBS (after homogenization).

Series	TSA	Series	TSA	Series	TSA
37	529.82	6	417.25	13	366.72
36	342.52	11	322.14	15	291.76
14	265.76	38	261.02	41	182.02
22	176.71	20	170.18	24	163.54
26	152.44	5	150.96	8	150.96
34	148.13	9	128.54	33	124.65
40	106.52	1	103.2	23	99.75
27	96.67	49	84.07	10	81.79
47	68.82	31	63.15	25	60.87
51	60.76	43	58.93	17	48.55
21	48.25	2	48.13	46	45.73
7	45.64	16	45.49	39	44.83
29	44.81	4	39.63	50	36.39
32	34.76	12	33.59	3	31.32

18	30.68	19	26.91	42	26.48
45	26.06	35	25.69	44	25.12
30	23.55	48	23.47	28	21.75

AVERAGE: 117.77

Test Statistics After Homogenization **for TN in CARPATCLIM**

Table 2.: The test statistics in decreasing order for E-OBS (after homogenization).

Series	TSA	Series	TSA	Series	TSA
32	50.61	22	49.95	26	44.76
30	44.52	16	44.49	14	39.18
38	38.36	19	37.46	12	36.28
28	33.64	11	32.76	21	31.34
36	28.60	24	27.90	41	27.75
48	27.68	45	25.34	10	24.17
3	22.76	18	22.23	33	22.13
23	21.34	4	21.15	35	20.99
43	20.22	37	20.12	17	18.96
7	18.86	5	17.85	8	17.85
1	17.63	20	16.47	27	16.01
40	15.91	49	14.09	42	13.29
6	13.08	50	12.99	46	12.01
2	11.72	31	11.31	9	11.23
34	11.04	47	10.39	29	9.87
44	9.60	25	9.18	15	8.93
51	8.66	13	8.65	39	8.12

AVERAGE: 22.34

Possibly the residual inhomogeneities and erroneous data cause the extremely variable trend for E-OBS compared to CARPATCLIM. The largest TSA values appear grid points near Miskolc (15), Siófok (6), Baja (13) where the Figure 12. is patchy. The MASH found these stations inhomogeneous. During MASH QC and homogenization these errors were filtered and the series were adjusted in the CARPATCLIM project. Inexplicable large and low changes can be seen in E-OBS. Note that the trends are analysed and compared for CARPATCLIM, E-OBS and ERA5 in the previous chapter for a shorter overlapping period from 1979-2010.

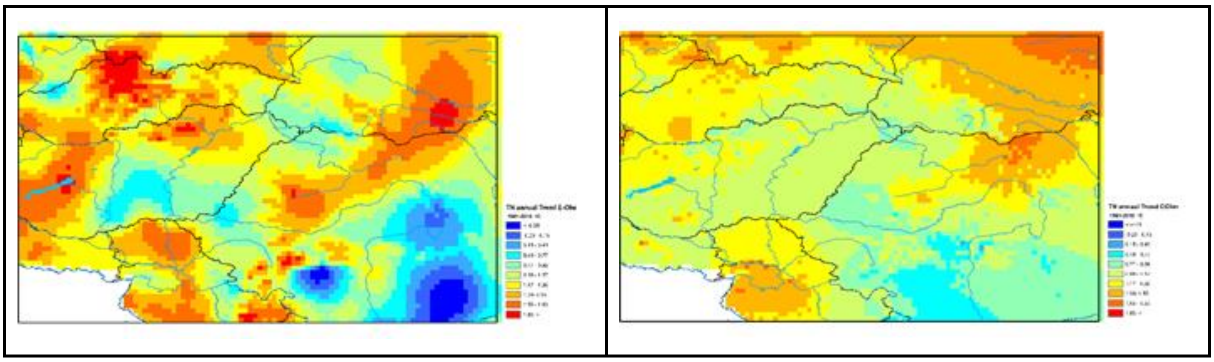


Fig. 12.: Linear trend (change in the whole period) for average annual average TN temperature in degC for E-OBS (left) and CARPATCLIM (right) over the time period 1961-2010, the longest overlapping period.

4. EVALUATION OF THE PRECIPITATION

The ANOVA results, the comparison of trends and the homogeneity test results are presented here as a complementary analysis of the article submitted to International Journal of Climatology (Brandhauer, 2021). Some of the yearly measures are similar for all three dataset (Figure 13.), Larger differences turn up in the case of "Spatial st. deviation of temporal mean" and in the "Root temporal mean of spatial variances" and in the „Temporal mean of spatial st. Deviation" between datasets, with the largest values in ERA5, CARPATCLIM and E-OBS respectively. To make these measures more expressive the spatial means and spatial variances at the moment *t* are illustrated on graphs. The Figure 14. shows the time series of the "Es (t)-spatial mean" and "Ds (t)- spatial st. deviation" from 1979 to 2010. The CARPATCLIM and the E-OBS yearly spatial means are running parallel, with larger amounts representing the CARPATCLIM than the E-OBS during the whole analysed period. The curves representing the ERA5 spatial means are going bottommost far above CARPATCLIM. Larger overestimation (165 mm or higher) can be found in ERA5 yearly spatial means, mainly in the beginning of the period in 1988, 1985, 1987, 1981, 1989, 1996, 1980 and 1981. The spatial standard deviation of E-OBS and ERA5 running together until 1991, thereafter the E-OBS curve is breaking down on the yearly graph. There is an increase in CARPATCLIM and a decrease in E-OBS between 1979 and 2010.

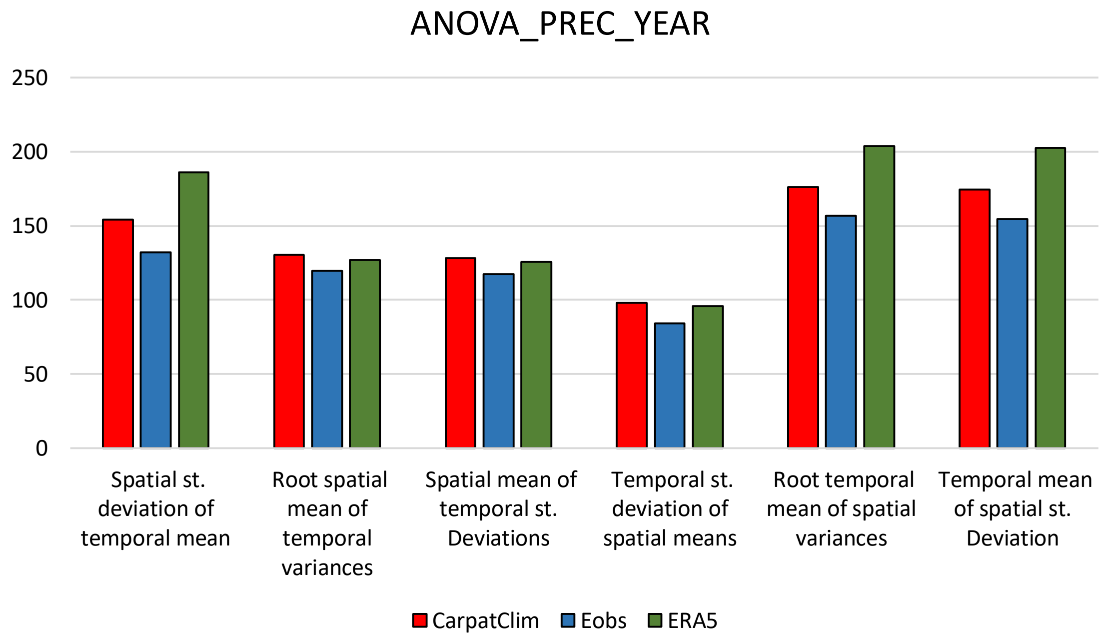
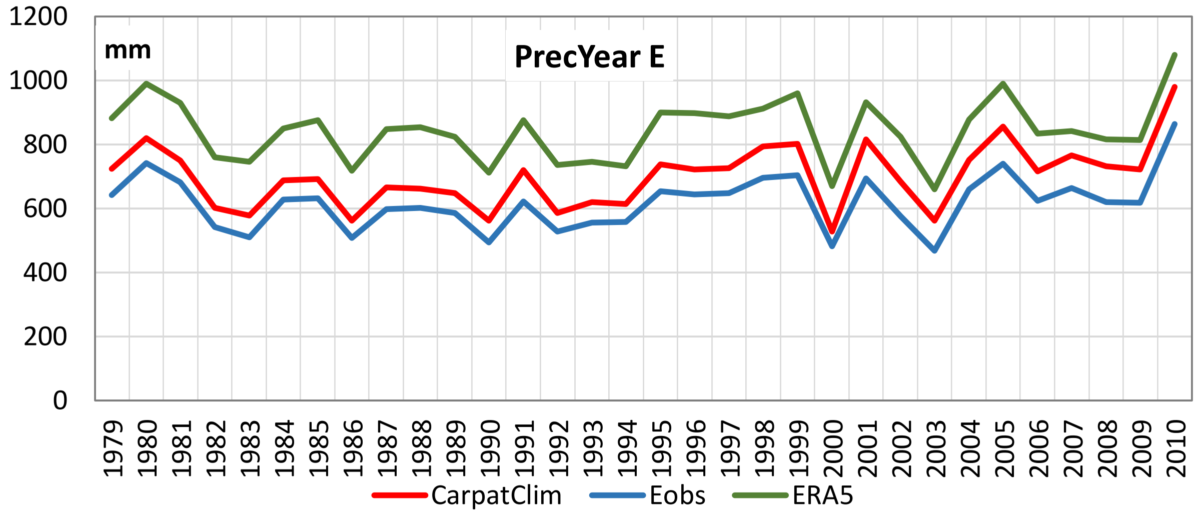


Fig. 13.: Some of the main statistics listed in Table 3.1.2. for yearly precipitation



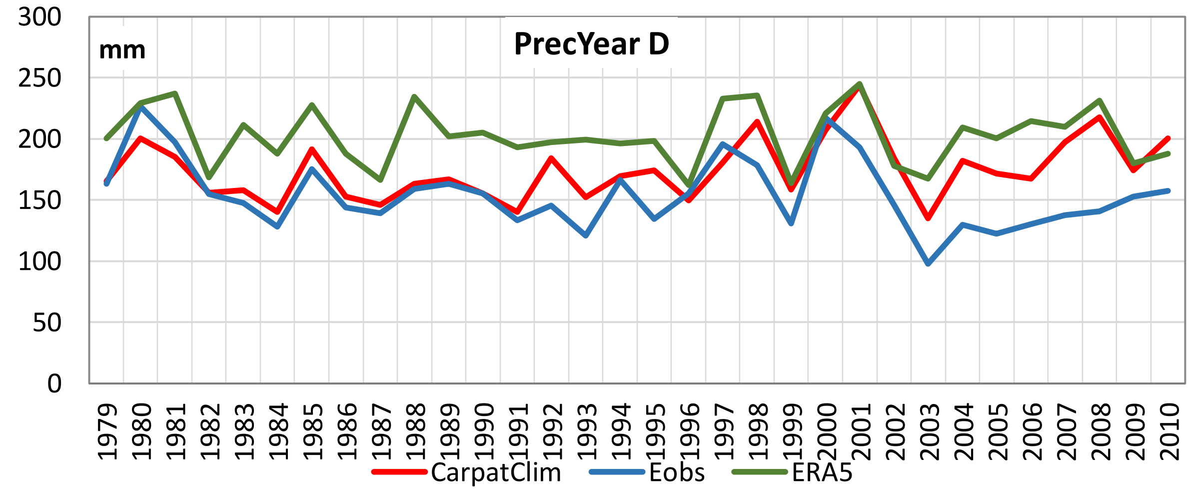


Fig. 14.: Yearly Es (t)-spatial mean (top) and Ds (t)- spatial st. deviation (bottom) from 1979-2010 for CARPATCLIM, E-OBS and ERA5.

4.1. TREND

An exponential trend model was fitted to the annual precipitation sums at each grid points in this comparative study. The estimated changes over the whole period (1979-2010) in % can be seen in the Figure 15. for the different datasets. The spatial pattern of the precipitation changes remarkably diverge. An extended region with decreasing precipitation emerges in Ukraine, Transcarpatia in E-OBS. Possibly the less station data applied for gridding in E-OBS resulted in this unreasonable decreasing trend. ERA5 produces 10% or even major precipitation decrease in the eastern part of the Carpathian region, which is completely missing from CARPATCLIM. The precipitation increase is moderate in ERA5, 20% in the large part of the region, while there are around 30-40% increase in Polish Carpathians and in Tatras in CARPATCLIM.

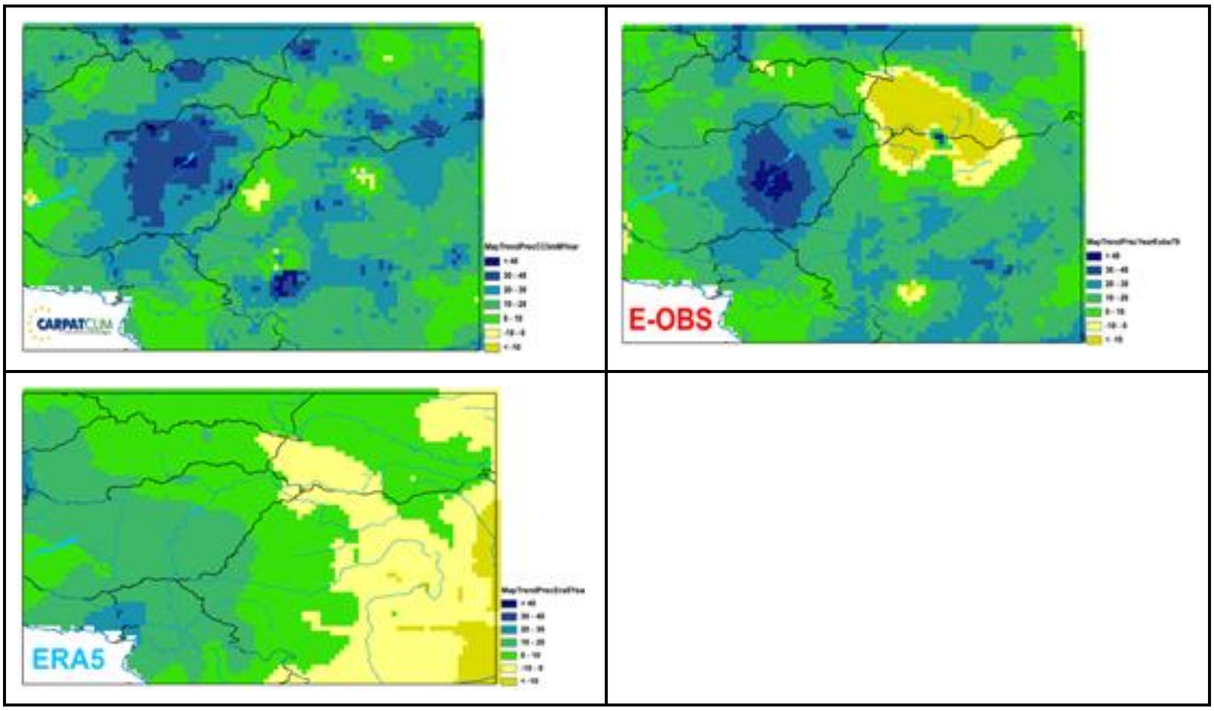


Fig. 15: Result of exponential trend fitting for yearly precipitation in %/ 32 year. for CARPATCLIM (top left), EOBS (top right) and the ERA5 (bottom) over the time period 1979-2010.

4.2. HOMOGENEITY TEST

The MASH (Multiple Analysis of Series for Homogenization, Szentimrey; 2014) software system for homogenization consists of functions for testing the residual inhomogeneity in any dataset. For doing this the closest grid points to 51 Hungarian meteorological stations (Figure 16.) were selected from CARPATCLIM and also from E-OBS, then the gridded daily precipitation series from 1961-2010 were tested by MASH homogenization method.

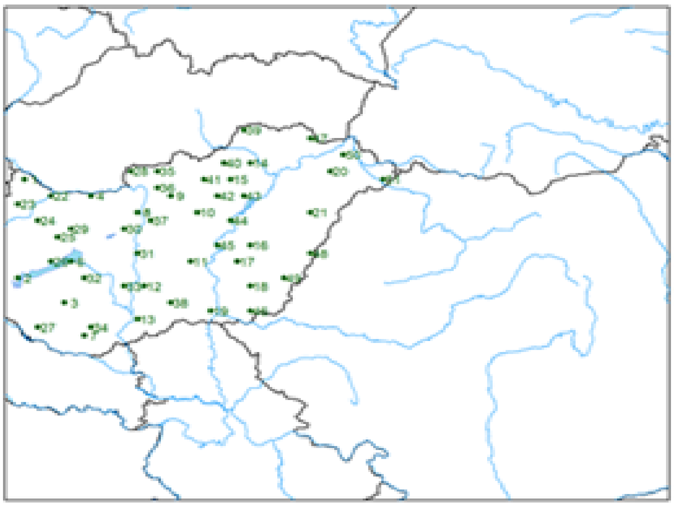


Fig. 16.: The location of the stations which were used in the homogeneity test.

The test statistics for inhomogeneity of the gridded series are listed in the Table 3. and Table 4. The null hypothesis is that the examined gridded series are homogeneous. The critical value related to significance level 0.01 comes to 31. Test statistics (TSA) can be compared to the critical value, the larger TSA values are the more suspicious.

Test Statistics After Homogenization for E-OBS

Table 3.: The test statistics in decreasing order for E-OBS (after homogenization).

Series	TSA	Series	TSA	Series	TSA
51	95.70	20	64.53	13	58.61
44	56.85	41	54.93	6	49.78
7	48.51	32	48.20	45	47.46
29	43.23	49	42.48	50	34.40
42	33.92	10	33.89	4	33.55
11	33.24	23	32.80	31	31.21
2	24.52	14	24.30	37	24.02
19	21.04	9	20.53	47	19.75
21	19.46	48	19.18	40	18.21
5	17.05	8	17.05	3	16.50
34	16.38	24	15.47	26	14.63
16	14.06	39	13.39	30	12.79
33	12.12	35	11.80	12	11.20
15	11.19	22	11.17	43	10.55
18	10.50	36	9.95	46	9.37
25	9.09	28	8.17	38	7.51
27	7.11	17	6.83	1	6.21

AVERAGE: 25.77

Test Statistics After Homogenization for CARPATCLIM

Table 4.: The test statistics in decreasing order for CARPATCLIM (after homogenization)

Series	TSA	Series	TSA	Series	TSA
42	50.53	2	36.04	6	36.03
37	34.07	49	32.05	46	30.12
50	29.21	12	28.13	29	27.51
44	27.06	5	25.00	8	25.00
40	23.11	17	22.71	20	20.79

16	20.55	43	19.92	3	19.49
21	18.67	1	17.75	14	17.37
48	16.86	39	16.71	15	16.65
32	15.81	31	15.62	13	15.28
4	14.80	9	14.68	33	14.64
35	14.32	51	14.20	41	14.05
18	13.45	38	13.37	23	13.27
22	12.71	25	11.78	7	11.49
11	10.71	47	9.86	26	9.14
19	9.07	30	8.91	45	8.17
36	7.89	27	7.24	24	6.11
34	5.30	10	4.58	28	4.30

AVERAGE: 17.88

The largest test statistics is 95.7 regarding E-OBS (Table 3.), while 50.53 in CARPATCLIM (Table 4.) taking the analysed grid point series. The average of the statistics is smaller in CARPATCLIM (17.88) than in E-OBS (25.77) what suggest that the E-OBS is bothered with residual inhomogeneity. Beyond doubt, the climate change signal strongly depends on the QC and the homogeneity procedure were applied.

5. CONCLUSION

The evaluation of the updated E-OBS dataset and the new global reanalysis ERA5 against regional datasets has been performed by calculating different statistical measures considering the variables of daily precipitation and temperature in the COPERNICUS C3S_311a_Lot4 project. When considering general temperature distribution of E-OBS, a remarkable agreement between E-OBS and the CARPATCLIM dataset. Regarding ERA5 it has lower maximum and higher minimum values, especially in colder regions. Furthermore, a general underestimation of precipitation magnitude is visible in the whole Carpathian region. The performance of E-OBS seems to rely strongly on the data availability. The calculation of long-term trends should be avoided or treated with caution because of the temporal inconsistencies caused by residual inhomogeneities in the station network. The major drawback of ERA5 is the constant overestimation of the precipitation amounts.

Acknowledgement
This study was supported by the COPERNICUS C3Surf (C3S_311a_Lot4): Climate monitoring products for Europe based on surface in-situ observations project.

References

Bandhauer, M., Isotta, F., Lussana, C., Tveito, O.E., Båserud, L., Lakatos, M., Szentimrey, T., Izsák, B., Szentes, O., Hoffmann, L., Kircsi, A., Bihari, Z., Frei, C. (2020): Report for Deliverable 4 (D4): Scientific report on evaluation of new E-OBS data set.

Bandhauer, M., Isotta, F., Lussana, C., Tveito, O.E., Båserud, L., Lakatos, M., Izsák, B., Szentes, O., Frei, C. (2021): Evaluation of daily precipitation analyses in E-OBS (v19.0) and ERA5 by comparison to regional high-resolution datasets in European regions. (submitted)

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., & New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. Journal of Geophysical Research, 114, D20119. https://doi.org/10.1029/2009JD011799

Lakatos, M., Szentimrey, T., Izsák, B., Hoffmann, L. (2017a): Comparison of the E-OBS and the CARPATCLIM gridded datasets for minimum temperatures, maximum temperatures and precipitation by the analysis of variance (ANOVA) method, Proceedings of the 9th Seminar for Homogenization and Quality Control in Climatological Databases and 4th Conference on Spatial Interpolation Techniques in Climatology and Meteorology (Ed. Szentimrey T, Lakatos M, Hoffmann L), Budapest, Hungary, 2017, WCDMP-No. 85, pp. 151-1664

Szentimrey, T. (1999): Multiple Analysis of Series for Homogenization (MASH). Proceedings of the Second Seminar for Homogenization of Surface Climatological Data, Budapest, Hungary, WMO, WCDMP-No. 41, 27–46.

Szentimrey, T. (2008): Development of MASH homogenization procedure for daily data. Proceedings of the Fifth Seminar for Homogenization and Quality Control in Climatological Databases, Budapest, 2006, WCDMP-No. 71, WMO/TD-NO. 1493, 123–130.

Szentimrey, T. and Bihari, Z. (2007) Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis). In: Proceedings from the Conference on Spatial Interpolation in Climatology and Meteorology, Budapest, Hungary, 2004. COST Action 719, COST Office, pp. 17–27

Szentimrey,T., Bihari, Z. (2014): Manual of interpolation software MISHv1.03, Hungarian Meteorological Service, p. 60.

Szentimrey, T. (2017): Manual of homogenization software MASHv3.03, Hungarian Meteorological Service,p.71.

Szentimrey, T. (2019a): Transformation of CarpatClim datasets for grid-box average datasets, Study for Hungarian Meteorological Service (COPERNICUS C3Surf project (WP4)), Varimax Limited Partnership, Budapest, Hungary, p. 7.

Szentimrey, T. (2019b): Mathematical methodology and software for comparison of gridded datasets, Study for Hungarian Meteorological Service (COPERNICUS C3Surf project (WP4)), Varimax Limited Partnership, Budapest, Hungary, p. 10.

DEVELOPMENT OF HIGH RESOLUTION GRIDDED DATASETS OF MONTHLY TEMPERATURE SINCE 1916 FOR SPAIN

Andrés CHAZARRA, Belinda LORENZO, Roser BOTEY

Spanish State Meteorological Agency (AEMET)
Leonardo Prieto Castro 8, 28071 Madrid, Spain
achazarrab@aemet.es

Abstract

This article describes the methodology used in the Spanish State Meteorological Agency (AEMET) for obtaining gridded datasets of monthly minimum, maximum and mean temperature with 1 × 1 km spatial resolution for Spain, covering the period 1916-2018. These datasets have been created for climate analysis and monitoring, and will be updated periodically to extend the time coverage. The data used to produce the grids have undergone a quality control process in order to remove or correct erroneous data. The spatial interpolation method consists on a multiple linear regression with ordinary kriging of the regression residuals, using elevation, easting, northing and distance to the coast as independent variables in the regression. The performance of the interpolation method and the accuracy of the grids are evaluated using a cross-validation approach to estimate the errors. Some examples of derived products are shown, as well as a temperature analysis over the 1916-2018 period in Spain based on the gridded datasets.

Key words: temperature, variability, climate change, geographic information system, grid.

1. INTRODUCTION

Recently, 1 x 1 km gridded datasets of monthly temperature for Spain covering the period 1961–2018 have been created and published by the Spanish State Meteorological Agency for climate analysis and monitoring (Chazarra et al., 2020).
These datasets are continuously updated to extend the time coverage to the present. Every month the new monthly provisional grids are generated using the temperature data available, and six months later the definitive grids are created using the complete network data.
New mean temperature series in Spain for the 1981-2010 reference period based on these gridded datasets have replaced the previous reference series used for climate monitoring, which were based on a set of 42 stations distributed along mainland Spain and the Balearic Islands.
In addition, some studies have been done to extend the gridded dataset back in time as long as possible using the same methodology. At present, temperature gridded datasets have been generated for mainland Spain and the Balearic Islands since 1916.
In this article, the methodology used for creating the grids since 1916 is described and the resulting 1916-2018 gridded datasets are analysed.

2. METHODOLOGY AND DATA

2.1. STUDY AREAS AND DATA

Two study areas were selected: the first one covers mainland Spain, the Balearic Islands and the autonomous cities of Ceuta and Melilla (mainland Spain area), whereas the second one covers the Canary Islands (Canary Islands area), as shown in Figure 1.

The coordinate reference systems used were ETRS89 / UTM zone 30N (EPSG: 25830) for the mainland Spain area and REGCAN95/UTM zone 28N (EPSG: 4083) for the Canary Islands area.

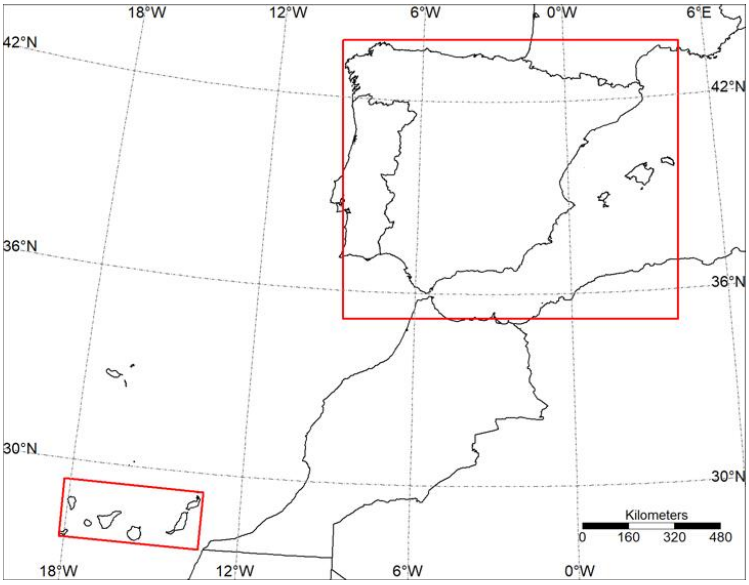


Fig. 1.: Study areas: mainland Spain area and Canary Islands area (in red color).

The variables considered for creating the grids were the monthly mean daily maximum temperature, the monthly mean daily minimum temperature and the monthly mean temperature. The monthly mean daily maximum and minimum temperature grids were spatially interpolated from the data, whereas the monthly mean temperature grids were derived by calculating the mean of the maximum and minimum temperature grids, in order to ensure the consistency of the three variable grids.

It should be noted that the station data were not homogenized before the spatial interpolation and, therefore, the resulting grids are not suitable for studying the climate variability in specific points. Nevertheless, they can be useful for climate monitoring in large areas such as provinces, autonomous regions or the whole country, where the effects of possible local inhomogeneities are expected to cancel each other out to a certain extent.

2.2. SELECTION OF THE STUDY PERIOD

Before selecting the time period for the study, all the monthly temperature data recorded in the Spanish National Climate Database since the beginning of the observations were analysed. The number of temperature stations with valid monthly temperature data for the two study areas are shown in figure 2.

In the mainland Spain area, the number of stations with monthly temperature data was fewer than 20 until the late 1870s. During the following decades, the number of stations increased slowly, and after 1911, the number of stations began to grow quickly and reached more than 100 in the second half of the 1910s. In the 1920s, the number of stations did not change significantly, but in the early 1930s the number rapidly rose, reaching 250 stations in 1936. During the Spanish Civil War (1936-1939) almost half of the stations stopped measuring data. After the war, the number of stations experienced a steady increase for several decades, reaching 1400 stations in 1975. Since then, the number of stations have been oscilating around 1600.

In the Canary Islands, the number of stations with valid monthly temperature data was fewer than 5 stations until 1945. During the late 1940s and the 1951-1960 decade the number of station oscilated between 15 and 30, and only after 1961 the number of valid data was constantly higher than 20.

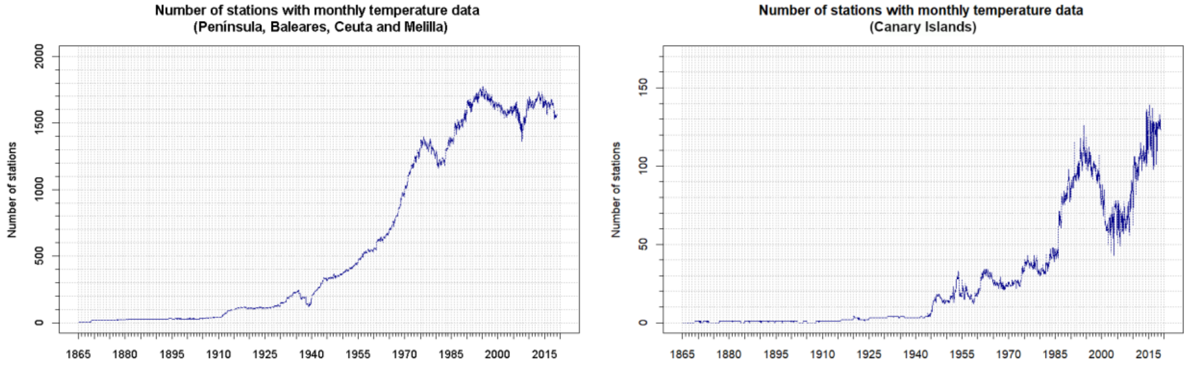


Fig. 2.: Number of stations with valid monthly temperature data in the two study areas since 1865.

After the analysis of the monthly temperature data avalaible, some spatial interpolations were carried out, by using the same interpolation method that will be used for creating the definitive grids, to identify the minimum number of data needed to generate valid grids, rejecting those grids that showed artifacts or noise due to a low data density. The visual analysis of the resulting grids since 1865 led to the conclusion that it is possible to generate high enough quality grids with 100 or more stations evenly distributed in the mainland Spain study area, and at least 20 stations in the Canary Islands area. These minimum numbers of stations were available after the first half of the 1910s in the mainland Spain area, and after 1961 in the Canary Islands area.

On the other hand, it is well known the existence of inhomogeneities in the historical temperature series of many countries due to the use, in the past, of meteorological shelters that differ from the current models (Parker, 1994). In Spain, during the XIXth century and the first years of the XXth century different screen models were used before the Stevenson screen became a standard. The most frequent pre-Stevenson shelter were the Montsouris and Glaisher open wooden stands, often called *facistoles* (lecterns) in Spanish because of their similar appearance to a choir lectern (Giménez, 1992).

According to some studies carried out in Spain (Brunet et al., 2004), in which temperatures simultaneously registered in both Stevenson and Montsouris shelters were compared, the daily maximum temperatures recorded under Montsouris stands are considerably overestimated (between 0.14 and 0.28 ºC in average, depending on the season and the location), whereas the daily minimum temperatures are slightly subestimated. Therefore, temperature data recorded under the old open stands should no be used for climate monitoring climate and especially for temperature trend analysis without a previous bias correction.

The Stevenson screen introduction in the Spanish meteorological network was gradual: although the replacement took place in the Astronomic Observatory of Madrid station in 1894, in other principal stations the Montsouris stands were not substituted by Stevenson shelters until the mid 1910s (Observatorio Central Meteorológico, 1918).

Taking all the above into account, January 1916 was finally chosen as the starting date for the study period in the mainland Spain area, since the number os stations was over the minimum number previously specified to obtain grids with enough quality, and the introduction of the Stevenson shelters had been already taken place in the main meteorological stations.

In the Canary Islands, the study was not finally extended back in time due to the low density of stations before the 1960s.

2.3. DATA QUALITY CONTROL

At present, several data quality control procedures are applied in the Spanish National Climate Database according to the World Meteorological Organization guidelines (WMO, 2018).

However, as these processes have evolved over time, it is possible to find in the historical temperature series anomalous data that have not undergone the present quality control procedures, so it is recommendable to incorporate additional data controls to improve the grid quality.
For this purpose, an automatic validation process has been applied to test the spatial consistency of the data. Every temperature data is compared to the estimated value obtained in that point by interpolating the neighbour data using the same method that will be used for creating the grids, and those data that significantly differ from the estimated value are rejected (Chazarra et al., 2020). Similar automatic control processes are currently applied in other European meteorological services for producing historical climate grids (Hollis et al., 2019).

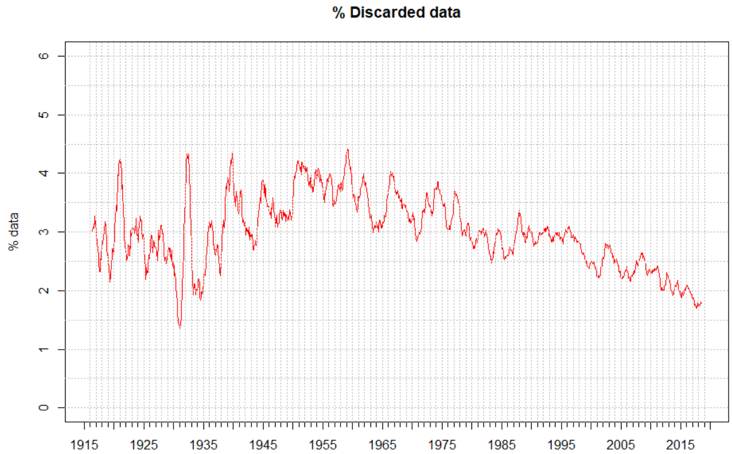


Fig. 3.: Percentage of monthly temperature data discarded in the automatic validation process in the mainland Spain area.

As shown in figure 3, the percentage of the discarded data was relatively small during the first two decades of the study period, highly oscillating around 3 %. In the late 1930s, there was a rise in that percentage into line with the increased number of stations, reaching 4 % in the early 1950s. Since then, the percentage of discarded data has constantly declined until the end of the study period, being fewer than 2 % in the recent years.
This analysis shows that the automatic control process works properly when there is a high density of stations so that it is possible to compare effectively each temperature data with its neighbour data. Therefore, the continuous decrease of the discarded data from 1960 up to now indicates a progressive increase in the quality of the temperature data of the Spanish National the National Climate Database. On the contrary, the relatively low percentage of the discarded data and the high variability of this figure in the first decades can be attributed to a small density of stations that makes difficult to find representative neighbourg data to compare. Consequently, it is to be expected a significantly lower quality in the first half of the study period grids, associated to the lower number of data and the reduced effectiveness of the automatic control process to detect anomalous data.

2.4. GRID INTERPOLATION AND ERROR ESTIMATION

The monthly gridded temperature datasets were created by spatial interpolation of the temperature data that have passed the quality control process, with 1 × 1 km spatial resolution.
The monthly minimum and maximum temperature data interpolation consisted on a multiple linear regression with ordinary kriging of the regression residuals, using elevation, easting, northing and distance to the coast as independent variables in the regression. An exponential model was used for the variogram adjustment.

The performance of the spatial interpolation method and the accuracy of each grid were estimated by leave-one-out cross validation, calculating for each gridded dataset the root mean square error (RMSE) and the mean absolute error (MAE), the most widely reported average-error measures (Willmott & Matsuura, 2006).
The monthly mean gridded datasets were derived by calculating the mean of the maximum and minimum temperature grids in each grid point. Aditionally, the error statistics were derived from the leave-one-out cross validation of the corresponding maximum and minimum temperature grids.

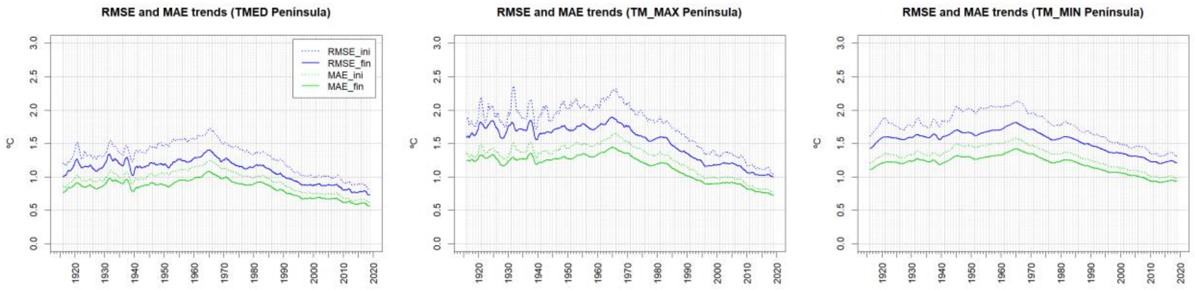


Fig. 4.: Trends of the root mean square error (RMSE) and the mean absolute error (MAE) for the mean, maximum and minimum temperatures before and after the automatic quality control process (mainland Spain area).

The trends of the root mean square error and the mean absolute error for the three considered variables, before and after the data quality control process, are shown in figure 4. As can be seen in that figure, the quality control slightly reduced the errors during the first two decades of the study period, when the density of the data is low. From 1940 to now, the reduction of the errors is much more evident. On the other hand, the magnitude of the errors was approximately constant or even showed a small increase until the mid 1960s, whereas from 1965 until the end of the study period the errors constantly decreased, indicating a continuous improvement of the data and grid quality since then.
The software used for creating the gridded dataset, including the automatic control process and the estimation of the errors, was the open source geographic information system SAGA GIS, version 6.3.0 (Conrad et al., 2015). All the procedures were programmed in R language (R Core Team, 2017) using several packages. These include, among others, RSAGA (Brenning et al., 2018), tmap (Tennekes, 2018) and raster (Hijmans, 2017).

3. RESULTS

Some examples of derived products are shown, as well as a temperature analysis over the 1916-2018 period in the mainland Spain area based on the gridded datasets.

3.1. GRIDDED DATASETS OF MONTHLY, ANNUAL AND SEASONAL TEMPERATURES 1916-2018

The gridded datasets of monthly mean, maximum and minimum temperatures for the study period 1916-2018 were created by following the procedure described above. As an example, the gridded datasets of April 1916 are shown in figure 5.

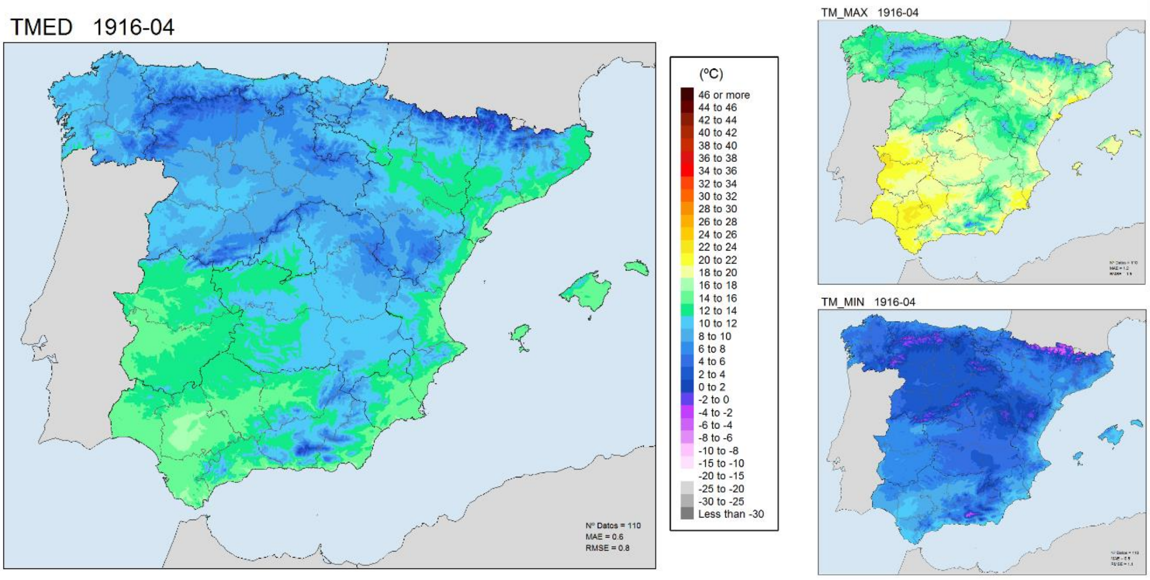


Fig. 5.: Gridded monthly mean (left), maximum and minimum temperatures (right) of April 1916.

Aditionally, annual and seasonal gridded datasets for the study period were calculated based on the monthly gridded datasets, as well as mean gridded datasets for each decade and each 30-year reference period in the study period for climate monitoring in Spain.

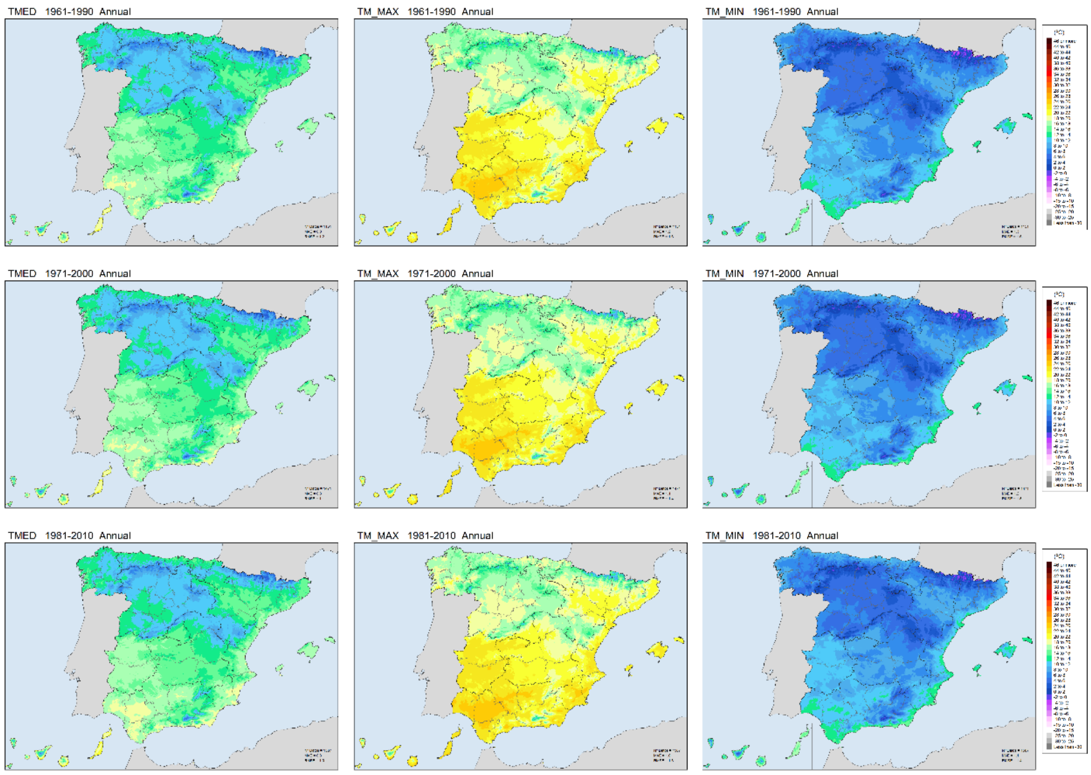


Fig. 6.: Example of several 30-year normal temperature grids.

These gridded temperature datasets have also been used to calculate the temperature anomaly series in mainland Spain for different reference periods, as shown in Figure 7.

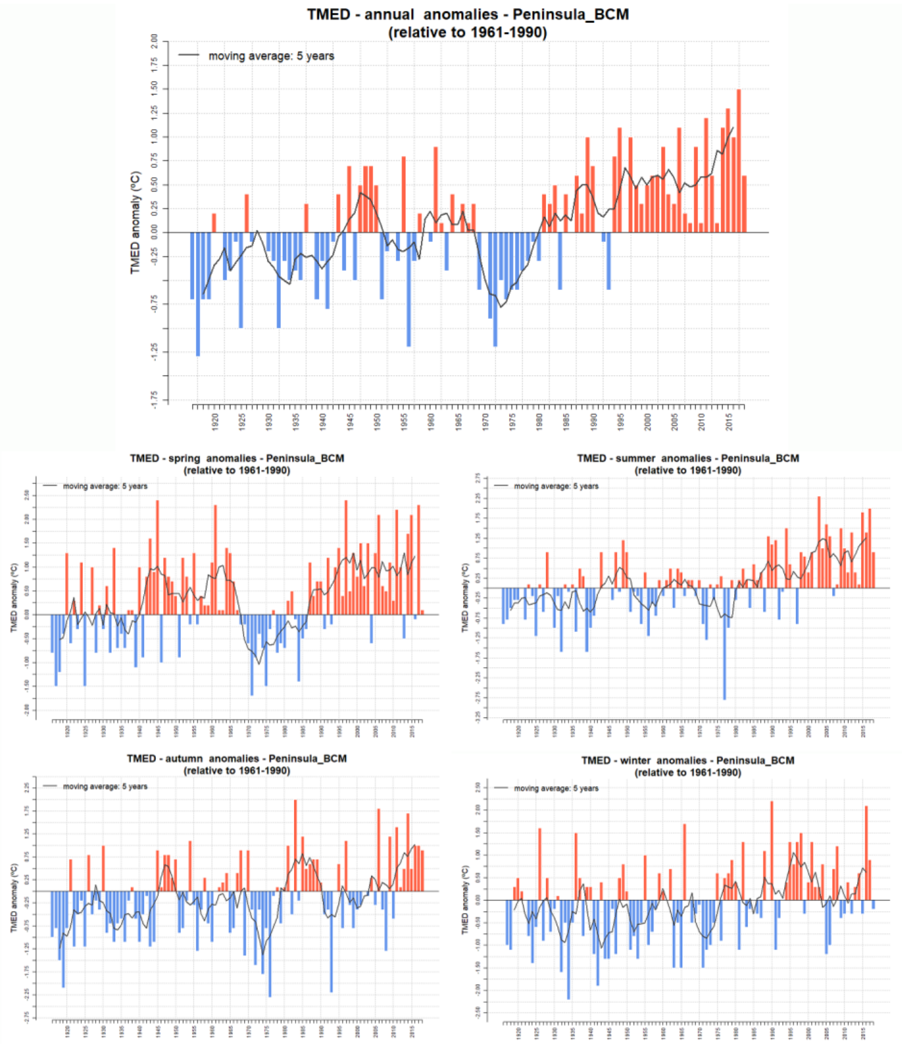


Fig. 7.: Annual and seasonal mean temperature anomaly series relative to 1961-1990 mean for the mainland Spain area.

3.2. TREND ANALYSIS OF THE ANNUAL AND SEASONAL TEMPERATURES IN THE MAINLAND SPAIN AREA 1916-2018

The annual and seasonal mean temperature series in the mainland Spain study area during the 1916-2018 period are shown in figure 8 (blue color), as well as the linear trends calculated by the method of least squares (red color). Additionally, the corresponding series and trends for the maximum and minimum temperatures are shown in figure 9. All the annual and seasonal trends for the 1916-2018 period were statistically significant with a significant level of 0.01 (table 1).

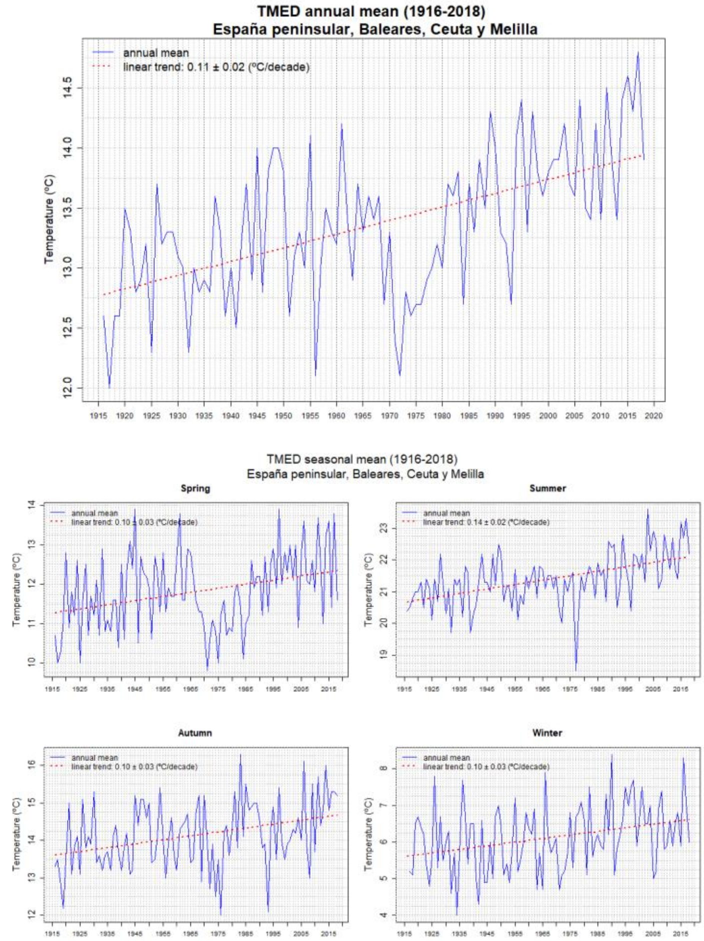


Fig. 8.: Annual and seasonal mean temperature series and trends in the mainland Spain area during the 1916-2018 period.

At first glance, there was a growing trend in the mean annual temperature series since the beginning of the study period until the early 1960s. A sudden cooling was observed during the late 1960s and the early 1970s, with a relative minimum in 1972. After that cold period, mean annual temperatures increased again, showing a noticeable warming in the mid and late 1970s and during the 1980 decade. In the early 1990s a short-term cold period of about three years occurred, which has been attributed to the Pinatubo eruption (González-Hidalgo et al., 2016). During the second half of the 1990s and the 2000 decade a slowdown in the warming was observed: although the mean annual temperatures were relatively high, they did not show any apparent growth trend. Finally, in the 2010s the warming trend has grown again, especially during the second half of the decade, in which the highest temperature values of the study period have been recorded.

According to the least square adjustment results, the mean annual warming in Spain has increased at 0.11 ± 0.02 ºC/decade during the 1916-2018 period. Maximum temperatures have increased at greater rates (0.16 ± 0.02 ºC/decade) than minimum temperatures (0.07 ± 0.02 ºC/decade), with a subsequent rise in the global diurnal temperature range of 0.09 ºC/decade. These results are similar to those of Brunet et al. (2007), which estimated from a series based on a set of stations a mean annual warming of 0.13 ± 0.03 ºC/decade for the mainland Spain in the 1901-2005 period, with larger rates of change for the maximum temperatures (0.17 ± 0.04 ºC/decade) than for the minimum temperatures (0.09 ± 0.03 ºC/decade).

Summer is the season that has contributed most to the annual warming (0.14 ± 0.02 ºC/decade), with a larger rate for the maximum temperatures (0.20 ± 0.03 ºC/decade) than for the minimum temperatures (0.08 ± 0.02 ºC/decade). Spring maximum temperatures has the second higher rate (0.17 ± 0.04 ºC/decade). The lowest warming rates were observed in the spring and winter minimum temperatures (0.04 ± 0.03 ºC/decade and 0.06 ± 0.04 ºC/decade, respectively).

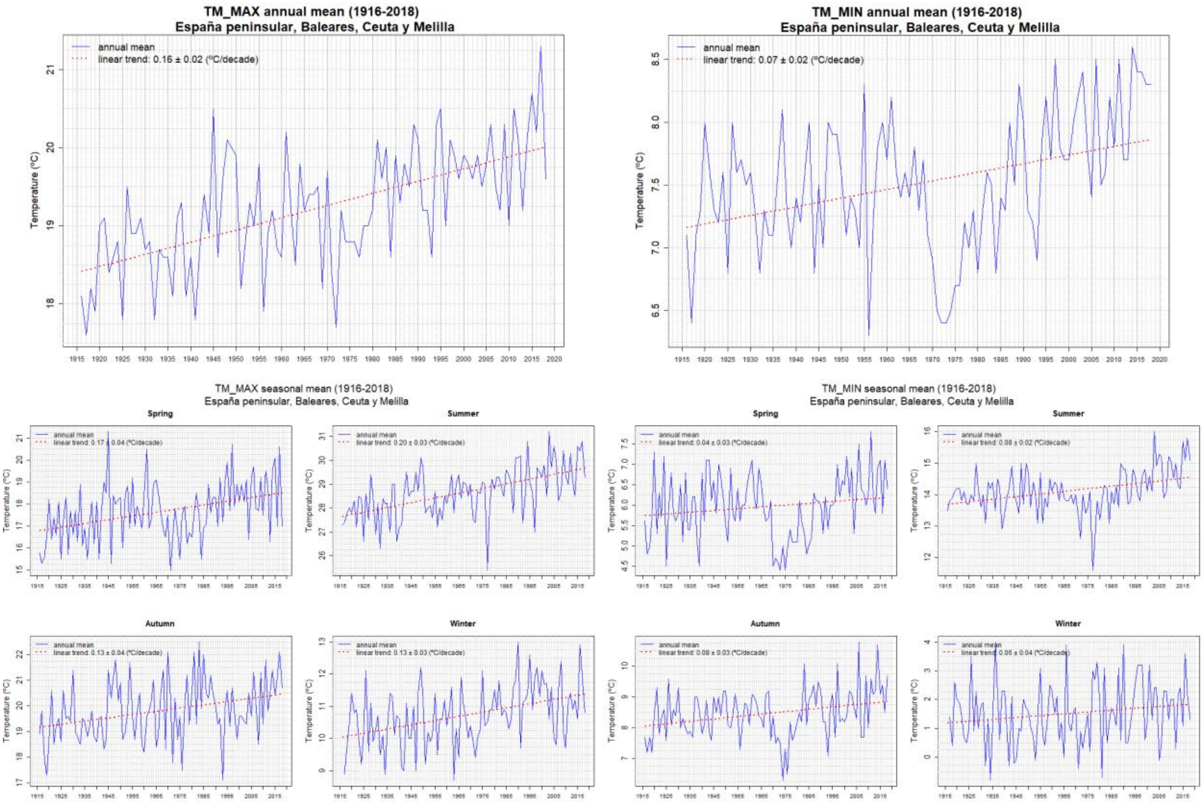


Fig. 9.: Annual and seasonal maximum and minimum temperature series and trends in the mainland Spain area during the 1916-2018 period.

Table 1.: Temperature trends in the 1916-2018 period and in 30-year periods (in bold, tendencies statistically significant at the 0.01 level; in bold and italics, tendencies statistically significant at the 0.05 level).

TMED (ºC)

	Spring	Summer	Autumn	Winter	Annual
1916-2018	*0.10 ± 0.03*	*0.14 ± 0.02*	*0.10 ± 0.03*	*0.10 ± 0.03*	*0.11 ± 0.02*
1921-1950	0.35 ± 0.19	0.28 ± 0.15	0.21 ± 0.15	0.00 ± 0.20	0.23 ± 0.09
1931-1960	0.22 ± 0.19	0.08 ± 0.15	0.18 ± 0.15	0.14 ± 0.19	0.16 ± 0.11
1941-1970	-0.13 ± 0.19	-0.02 ± 0.12	0.02 ± 0.16	0.27 ± 0.18	0.01 ± 0.11
1951-1980	***-0.54 ± 0.17***	-0.09 ± 0.15	-0.13 ± 0.17	0.23 ± 0.18	-0.15 ± 0.10
1961-1990	-0.17 ± 0.19	0.12 ± 0.16	0.32 ± 0.20	0.34 ± 0.19	0.14 ± 0.11
1971-2000	***0.81 ± 0.13***	***0.46 ± 0.17***	0.24 ± 0.21	**0.47 ± 0.18**	***0.49 ± 0.09***
1981-2010	**0.39 ± 0.16**	0.29 ± 0.16	-0.25 ± 0.19	0.07 ± 0.19	0.11 ± 0.09
1991-2018*	0.12 ± 0.21	0.36 ± 0.18	***0.60 ± 0.19***	0.05 ± 0.20	**0.29 ± 0.10**

TM_MAX (ºC)

	Spring	Summer	Autumn	Winter	Annual
1869-2018	*0.17 ± 0.04*	*0.20 ± 0.03*	*0.13 ± 0.04*	*0.13 ± 0.03*	*0.16 ± 0.02*
1921-1950	0.51 ± 0.27	***0.48 ± 0.19***	0.35 ± 0.20	0.12 ± 0.19	0.34 ± 0.13
1931-1960	0.33 ± 0.28	0.14 ± 0.19	0.25 ± 0.22	0.15 ± 0.18	0.20 ± 0.14
1941-1970	-0.04 ± 0.28	0.04 ± 0.16	0.03 ± 0.25	0.17 ± 0.19	0.04 ± 0.15
1951-1980	**-0.52 ± 0.24**	0.13 ± 0.18	0.10 ± 0.25	0.23 ± 0.17	-0.02 ± 0.12
1961-1990	-0.17 ± 0.25	0.21 ± 0.19	0.41 ± 0.26	**0.44 ± 0.17**	0.21 ± 0.13
1971-2000	***0.95 ± 0.21***	**0.42 ± 0.22**	-0.01 ± 0.27	***0.55 ± 0.15***	***0.47 ± 0.11***
1981-2010	0.38 ± 0.23	0.21 ± 0.20	-0.41 ± 0.23	-0.08 ± 0.19	0.02 ± 0.11
1991-2018*	0.01 ± 0.29	0.36 ± 0.24	***0.76 ± 0.21***	-0.02 ± 0.21	0.29 ± 0.13

TM_MIN (ºC)

	Spring	Summer	Autumn	Winter	Annual
1869-2018	*0.04 ± 0.03*	*0.08 ± 0.02*	*0.08 ± 0.03*	*0.06 ± 0.04*	*0.07 ± 0.02*
1921-1950	0.18 ± 0.16	0.12 ± 0.11	0.08 ± 0.12	-0.10 ± 0.23	0.07 ± 0.08
1931-1960	0.12 ± 0.15	0.03 ± 0.13	0.11 ± 0.11	0.13 ± 0.24	0.10 ± 0.10
1941-1970	-0.20 ± 0.13	-0.10 ± 0.11	0.02 ± 0.11	0.34 ± 0.20	-0.01 ± 0.10
1951-1980	***-0.54 ± 0.14***	**-0.32 ± 0.13**	**-0.37 ± 0.14**	0.22 ± 0.23	**-0.27 ± 0.10**
1961-1990	-0.19 ± 0.15	0.04 ± 0.15	0.20 ± 0.18	0.27 ± 0.25	0.06 ± 0.11
1971-2000	***0.66 ± 0.08***	***0.51 ± 0.13***	***0.48 ± 0.17***	0.41 ± 0.25	***0.51 ± 0.08***
1981-2010	***0.40 ± 0.11***	***0.37 ± 0.12***	-0.06 ± 0.18	0.18 ± 0.25	**0.21 ± 0.09**
1991-2018*	0.26 ± 0.14	**0.36 ± 0.13**	**0.47 ± 0.19**	0.09 ± 0.24	**0.29 ± 0.09**

4. CONCLUSIONS

At present, high resolution gridded datasets of monthly temperature for the mainland Spain area since 1916 are available, and for the Canary Islands since 1961. These datasets are continuously updated to extend the time coverage to the present.
The error analysis of the grids shows that the quality of the grids is lower during the first decades of the study period, and it progressively increases after the 1960s.
Due to possible local inhomogeneities, these datasets are not suitable for studying the climate variability in specific points, but they can be useful for climate monitoring in whole country or big areas during the last century. The results of the trend analysis of the annual and seasonal temperatures in the mainland Spain area in the period 1916-2018 are consistent with previous studies based on a set of stations.

References

Brenning, A., Bangs, D., & Marc, B. (2018). RSAGA: SAGA Geoprocessing and Terrain Analysis. R package version 1.1.0. https://CRAN.R-project.org/package=RSAGA

Brunet M., Bañón M., García F., Aguilar E., Saladié O., Sigró J., Asín J., López D. 2004. Una aproximación experimental tendente a la minimización del sesgo artificial asociado al tipo de garita meteorológica a través de la observación dual de la temperatura del aire. In Actas de las XXVIII Jornadas Científicas de la AME, Badajoz, 11–13 February 2004, Pub. Asociación Meteorológica Española y Ministerio de Medio Ambiente: La Meteorología y el Clima Atlánticos, Badajoz: 13.

Brunet, M., Jones, P., Sigró, J., Saladié, O., Aguilar, E., Moberg, A., . . . López, D. (2007). Temporal and spatial temperature variability and change over Spain during 1850–2005. J. Geophys. Res., 112, D12117. doi:10.1029/2006JD008249

Chazarra, A., Lorenzo, B., Rodríguez, C., Botey, R. (2020). Análisis de las temperaturas en España en el periodo 1961-2018. Madrid: Agencia Estatal de Meteorología (AEMET). doi: https://doi.org/10.31978/666-20-003-4

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., . . . Böhner, J. (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geosci. Model Dev., 8, 1991-2007. doi:10.5194/gmd-8-1991-2015

Giménez, J.M (1992). La Meteorología en el Observatorio Astronómico de Madrid. En “Doscientos años del Observatorio Astronómico de Madrid” (ISBN: 978-84-88405-00-6), editado por la Asociación de amigos del Observatorio Astronómico de Madrid.

González-Hidalgo, J., Peña-Angulo, D., Brunetti, M., & Cortesi, N. (2016). Recent trend in temperature evolution in Spanish mainland (1951–2010): from warming to hiatus. Int. J. Climatol., 36, 2405-2416. doi:10.1002/joc.4519.

Hijmans, R. J. (2017). raster: Geographic Data Analysis and Modeling. R package version 2.6-7. https://CRAN.R-project.org/package=raster

Hijmans, R. J. (2017). raster: Geographic Data Analysis and Modeling. R package version 2.6-7. https://CRAN.R-project.org/package=raster

Hollis, D., McCarthy, M., Kendon, M., Legg, T., & Simpson, I. (2019). HadUK-Grid—A new UK dataset of gridded climate observations. Geosci Data J.(6), 151-159. doi: https://doi.org/10.1002/gdj3.78

Observatorio Central Meteorológico (1918). Resumen de las observaciones meteorológicas efectuadas en las estaciones del Servicio Meteorológico Español durante el año 1916.

Parker, D. E. (1994), Effects of changing exposure of thermometers at land stations, Int. J. Climatol., 14, 1 – 31.

R Core Team. (2017). R: A language and enviroment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Tennekes, M. (2018). tmap: Thematic Maps. R package version 1.11-1. https://CRAN.Rproject.org/package=tmap

Willmott, C. J., & Matsuura, K. (2006). On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. International Journal of Geographical Information Science, 20(1), 89-102. doi:10.1080/13658810500286976

World Meteorological Organization (2018). Guide to Climatological Practices. OMM-nº100. 2018 edition.

ACTUALIZATION OF NATIONAL CLIMATE CLASSIFICATION MAP OF PERU

Anabel Castro, Cristina Davila, Wil Laura, Felix Cubas, Grinia Avalos, Carlos López, Marcia Valdéz, Donna Villena, Lourdes Menis, Julio Urbiola, Irene Trebejo

National Meteorology and Hydrology Service of Peru (SENAMHI), Lima, Perú
spc2@senamhi.gob.pe

Abstract

The National Meteorology and Hydrology Service of Peru (SENAMHI) elaborated the new national climate classification map in order to share an estimate of the climate sources of different regions of the country, with the purpose of guiding the government planning processes of the economic, environmental and social aspects.
The first climate classification map of Peru was developed and published by SENAMHI in 1988, which provided an amount of 27 types of climate. It was prepared using the climate classification system of Warren Thornthwaite, 1931. For this, twenty year meteorological data between 1965 and 1984 from 250 meteorological stations of the national network were used.
For the preparation of the current Map, for comparative purposes, the same climate classification system of Thornthwaite was used. 30 years of meteorological data were used between the period of 1981 and 2010, from 483 meteorological stations of the national network, in addition with meteorological data from country neighbors: Ecuador (11), Colombia (3) and Bolivia (4), following the recommendation of the World Meteorology Organization (WMO).
For the elaboration of this new map, a new spatial interpolation technique was used, linear regression with the weighted inverse distance error adjustment using satellite information and involves a validation of the statistical model through cross validation. This technique provides information on the relationship between the regional geographic reality and the climate. In addition, a national validation process was carried out with the intervention of national experts familiar with the subject through face to face and virtual workshops throughout the country.
The updated climate classification map presents 38 types of climate in Peru with a better representation of the climates in the region. This product provides national climate information in a graphic and synthetic way that allows connection with natural ecosystems and main human activities.

1. INTRODUCTION

The tropical geographic location of Perú, the position of mountain Andes, the maritime influence of Humboldt's current in cold waters and the warm current from north, as well as the stationality from climate drivers as the Bolivian High System, High Pressure Systems in south pacific and south atlantic and others drivers, module the weather conditions in Peru (thermal and publiometric regimes, among others). This allows that the country presents a greater biologic diversification (MINAM, 2016), associated to different climate types, by the Thornthwaite method.

The Thornthwaite classification system of 1931 takes account values of precipitation and air temperature for the calculation of three indexes: Precipitation effectivity Index (IPE), Seasonal moisture concentration Index (ICEH) and Thermal Efficiency Index (IPE). However, this climate classification was adapted for geographic features and climate conditions that are present in the peruvian territory.

The spatial interpolation technique of multiple linear regression with error adjusted by inverse distance weighting was applied to the climate classification indexes of Thornthwaite, calculated per each meteorology conventional station. In this technique we use like a predictor variable the satellite information and natural regions to estimate the Thornthwaite indexes spatially.

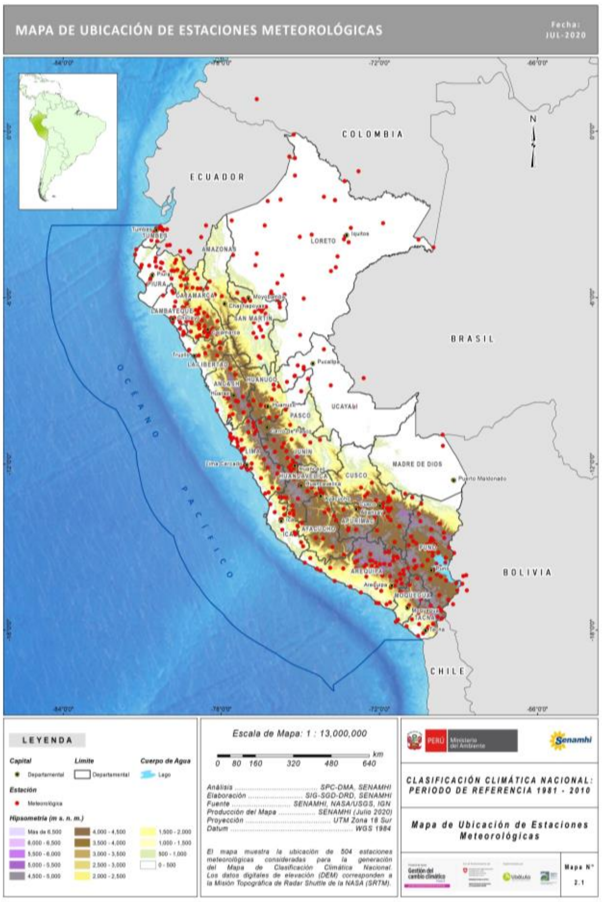
The Thornthwaite classification system was adapted by SENAMHI in 1988, where the ICEH taken into account two more weather stations (autumn and spring) to analyze, since the initial methodology used only the summer and winter stations and the relative humidity variable was introduced with the objective of characterize the climate into “very dry”, “dry”, “humed” and “very humid”. In 2020, SENAMHI keeped the change in the ICEH, however, It was retired the relative humidity since the index evaluates the humidity percentage for each weather station. In the same way, it was retired the subcategories “semi-warm” and “semi-cold”, besides, it was changed the terminology “polar” by “glacier” in the Thermal Efficiency Index.

2. METHODS

2.1. DATA

Climate information was used with quality control of meteorological variables of precipitation, maximum and minimum air temperature, for the period 1981-2010. The conventional meteorological stations of the National Meteorology and Hydrology Service of Peru and other National Meteorological Services[1] were used.

The geographical location of the meteorological station is presented on Map 1.



Map 1. Geographic location of meteorology stations used in the actualization of the climate classification map of Peru.

[1] National Institute of Meteorology and Hydrology Ecuador (INAMHI), Institute of Hydrology, Meteorology and Environmental Studies of Colombia (IDEAM) and National Service of Meteorology and Hydrology Bolivia (SENAMHI- Bolivia).

2.2. METHODOLOGY

The update of the climatic classification in Peru was carried out through an adaptation of the Thornthwaite classification system of 1931, in 2020 by SENAMHI. For each conventional meteorological station the Thornthwaite indices were calculated. Subsequently, the spatial interpolation of the Thornthwaite indices was performed using a standardized multiple linear regression statistical model and the error adjusted by inverse distance weighting (IDW). Finally, national workshops were held with specialists from different areas, who have sufficient experience to validate the results obtained, forhence, the map representative was for each study region.

To develop the climatic classification, It was necessary to calculate the climatic normals of maximum and minimum air temperature (OMM, 2017). This calculation consists of making an average of the daily values of each month, and then averaging the monthly values, during a reference period of 30 years, 1981 - 2010.

Using the climatic normals, the indexes of Thornthwaite were calculated: Precipitation effectivity Index, Seasonal moisture concentration Index and Thermal Efficiency Index. The first of these indexes explain the fraction of precipitation that is used for plants, which is express in terms of precipitation and temperature, through an empiric equation, see Eq.1, where $P$is monthly precipitation in inches and $T$is the monthly air temperature in Fahrenheit degrees.

$$IPE=\sum_{n=1}^{12} 115*\left(\frac{P}{T-10}\right)_n^{10/9} \quad \text{Eq.1}$$

The second index, Seasonal Moisture Concentration, was calculated with respect to precipitation effectivity and its purpose is to differentiate dry from wet seasons. It is classified as a dry month when the IPE value is less than 0.83, therefore to categorize the season as dry, it has to be for at least two dry months.

The third index, Thermal efficiency, expresses the stimulating and limitans effects of low and high air temperatures in the plants growing, through an empiric equation, see Eq 2. expressed in positive values of monthly mean temperatures (above the freezing dot of water), where T is the temperature in Fahrenheit degrees.

$$IET=\sum_{n=1}^{12}\left(\frac{T-32}{4}\right)_n \quad \text{Eq. 2}$$

The results of each Thornthwaite index was clasificated in the following categories:

Table 1.: Precipitation effectivity index classification.

ANUAL VALUE (IPE)	CLIMATE	SYMBOL	REFERENCE ZONES
>4.86	Very rainy	A	Jungle
4.17 a 4.85	rainy	B	Forest
3.50 a 4.16	Medium dry	C	Pastureland
2.84 a 3.49	Semiarid	D	Steppe
Under 2.84	Arid	E	Desert

Table 2.: Thermal efficiency index classification.

ANUAL VALUE (IET)	CLIMATE CHARACTER	SYMBOL	REFERENCE ZONES
>127	Warm	A’	Tropical
64 a 127	Tempered	B’	Mesothermal
32 a 63	Cold	C’	Microthermal
16 a 31	Semifrigid	D’	Taiga
1 a 15	Frigid	E’	Tundra
0	glacier	F’	Perennial ice

Table 3.: Seasonal moisture concentration index classification.

SYMBOL	MEANING
r:	Abundant humidity in all seasons of the year.
i:	With dry Winter.
p:	With dry spring.
v:	With dry summer.
o:	With dry autumn.
d:	Moisture deficiency in all seasons of the year.

The spatial interpolation of Thornthwaite indexes was made using the statistical technique of multiple linear regression and error adjusted by inverse distance weighting; this model establishes an linear relation between dependent variables (Precipitation effectivity and Thermal Efficiency) and independent variables (altitude, latitude, longitude and natural region), it means that geographic and satellite information is used to estimate climates through Thornthwaite indexes, see Eq.3 and Eq.4, finally it was applied an error adjust by IDW (Ninyerola, Pons, et al,. 2007a; Ninyerola, Pons, et al,. 2007b).

$$Lon(IPE+1)=\theta_0+\theta_1(Altitud)+\theta_2(Latitud)+\theta_3(Longitud)+\theta_4(Reg.Natural) \quad (3)$$

$$IET=\theta_0+\theta_1(Altitud)+\theta_2(Latitud)+\theta_3(Longitud) \quad (4)$$

For the realization of the model, four assumptions were evaluated: normal distribution of the variables, linear relationship between independent and dependent variables, homoscedasticity and absence of multicollinearity (Poole and Farrel, 1970). The reliability of the results was determined and was obtained by standardizing the variables used (Osborne and Water, 2002). Finally, a cross validation was carried out using the k-fold method (Refaeilzadeh et al., 2009), to evaluate the bonded adjustment measures, which determine the significance of the model.

Table 4.: Goodness-of-fit measures after applying cross-validation to the interpolation statistical model.

MONTH	RMSE	R^2	MAE
JANUARY	0.47	0.79	0.34
FEBRUARY	0.49	0.76	0.37
MARCH	0.48	0.77	0.37
APRIL	0.45	0.77	0.35
MAY	0.4	0.8	0.31

JUNE	0.37	0.79	0.27
JULY	0.38	0.75	0.28
AUGUST	0.37	0.74	0.28
SEPTEMBER	0.44	0.72	0.34
OCTOBER	0.51	0.73	0.41
NOVEMBER	0.5	0.75	0.38
DECEMBER	0.48	0.79	0.35

The map algebra technique was used to obtain the preliminary climate classification map. This technique applied the parameters obtained from the statistical model, using as grid information the following variables: altitude, latitude, longitude and natural region. The parameters were applied to each grid cell, abarquing the national territory, with the target of estimates the Thornthwaite indexes over all the country. Then, each layer of index information (IPE, IET e ICEH) was converted to vectorial format to superpose it. In this way, each polygon presents the climate classification; this process was developed using a geographic information system software.

Finally, A validation of the climate classification map of Peru was made presenting the diversity and distribution of climates to local experts over the country through face-to-face and virtual technique workshops.

3. RESULTS

As a result of technique work carried out by SENAMHI from 2017 to 2020, the actualization of the climate classification map of Peru, provides us 38 climate types over the country. See Figure 2, elaborate it to a scale of 1:400 000 in centimeters.

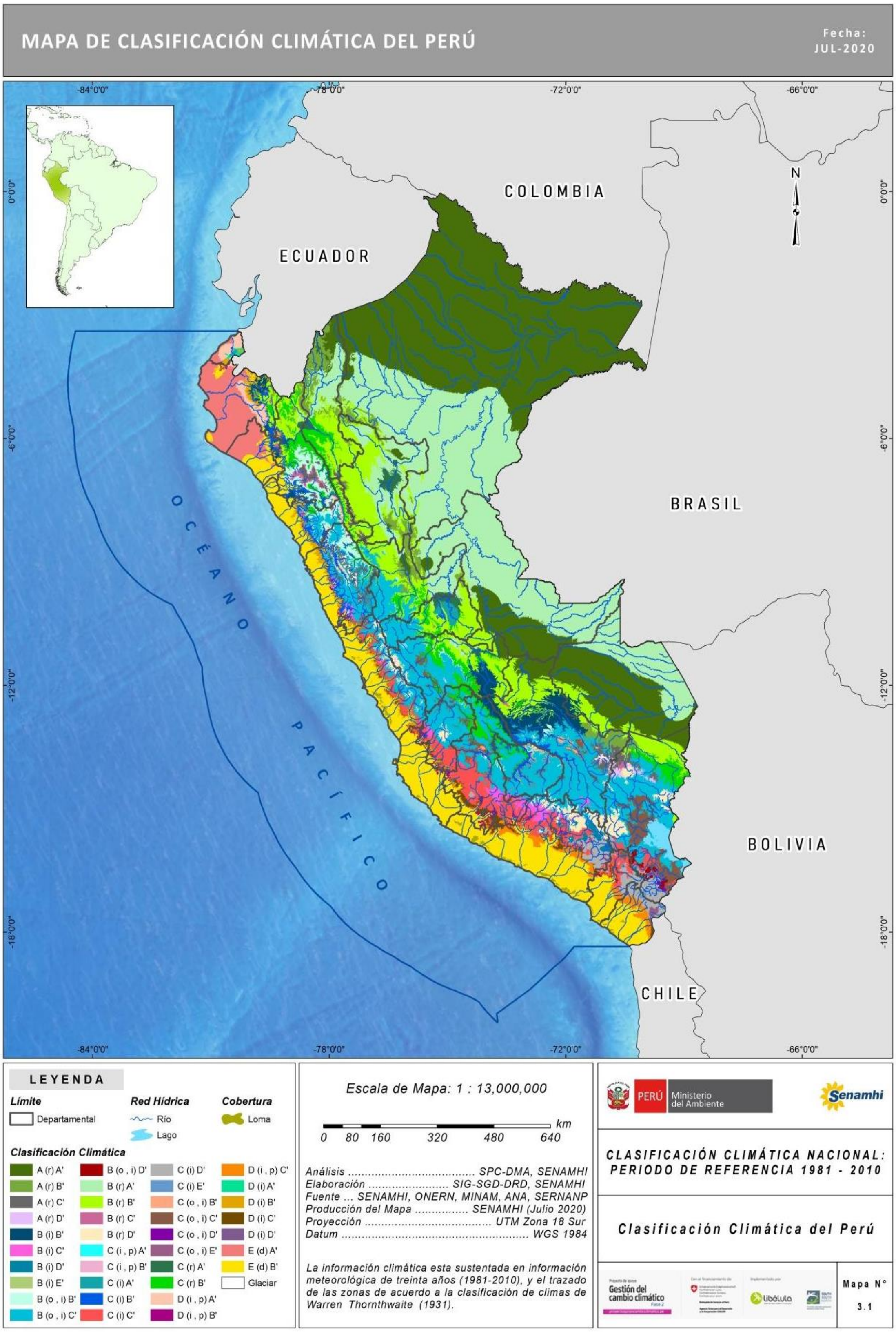


Fig. 2.: Climate classification of Peru.

4. DISCUSSION

Of the 38 climatic types found in Peru, the warm, very rainy and humid climate throughout the year occupied 26.4% of the national territory, the warm, rainy and humid climate throughout the year occupied 20.3% of the country, the rainy and cold with autumn and winter dry occupied 10.2% of the territory and the arid and dry temperate climate throughout the year covers 10.1% of the national area. The two most prevalent climates are found in the jungle, the third climate in the mountains and the fourth on the coast.

Regarding to the map version of 1988, the actualization provide us eleven mores climates, this difference is related to an enhance in the amount of climatic information used for the 2020 version (504 weather stations, almost the twice as the before version), besides, it used the satellite information as the digital elevation model of 90 meters and better techniques to optimize the spatial estimation like the multiple linear regression model with error adjusted by IDW. To improve this methodologie, we added the validation of the map by specialists through face to face and virtual workshops with a national approach. It should be noted that the 1988 version map used 20 year of information (1965 - 1984) and the 2020 version used 30 years (1981-2020), following the OMM recommendations.

The major increment of climates was ubicated in the west side of mountain chain andes, since SENAMHI has a greater weather station net in that area. If we compared the quantity of medium dry and semiarid climates regarding the version of 1988, the actualization has six more in each climate.

The results of this study provide the importance of obtaining a large network of climate information about the country to transform data into climate products and services for citizens and determine the public value chain in climate information.

5. CONCLUSIONS

Using an adaptation of the Thornthwaite methodology to develop the climatic classification and the multiple linear regression statistical technique with inverse distance weighting error adjustment to perform the spatial estimation of the climate over the country, a better representation is obtained according to the climatic reality. Being validated by statistical tests and by the analysis of specialists from various regions of the country.

The updated national climate classification map presents 38 climates: four very rainy, eleven rainy, thirteen semi-dry, seven semi-arid, two arid, and one glacier. The former are found in the jungle region; the second, in the mountains and high jungle; most of the rest, in the highlands of the Pacific slope and desert coast. The current map has 11 additional climates compared to the previous one.

Knowledge of the climate classification map will precisely guide the planning and management process of economic, socio-environmental and territorial intervention activities, necessary for the development of Peru and its regions.

Acknowledgement

A special recognition is expressed to the different public institutions and professional specialists from the different regions of the country, who collaborated with their knowledge and experiences to make possible the faithful elaboration of the new Climate Map of Peru. Special consideration and gratitude, to the National Institute of Meteorology and Hydrology of Ecuador - INAMHI and the Institute of Hydrology, Meteorology and Environmental Studies of Colombia - IDEAM, for providing valuable meteorological information of their stations located in the Amazon area near the border with the Peru and the climate change management support project.

References

Ninyerola, M., Pons, X. y Roure, J.M. (2007a) Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. Theoretical and Applied Climatology, 89(3-4): 195-209.

Ninyerola, M., Pons, X. y Roure, J. M. (2007b) Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. Int J Climatol, 27(9): 1231-1242.

WMO, 2017. Guidelines on the Calculation of Climate Normals. World Meteorological Organization Nº 1203. Available in: https://library.wmo.int/doc_num.php?explnum_id=4167

Osborne, J; Waters, E. (2002) Four assumptions of multiple regression that researchers should always test. Practical Assessment, research and evaluation. Vol 8 Part 2.

Poole, M; Farrel, N. (1970) The assumptions of the linear regression model. Available in: https://pdfs.semanticscholar.org/bc11/336a6f637a44918d708b56efcc5fe9a2e3c8.pdf

Refaeilzadeh, P; Tang, L; Liu, H. (2009) Cross Validation. Encyclopedia of database systems.

Thornthwaite, C. (1931) The Climates of North America According to a New Classification. American Geographical Society. Geographical Review, Vol. 21, No. 4, pp. 633-655.

PROGRAMME

Budapest, Hungary
12-14 October 2020

MONDAY, 12 OCTOBER
13:00-17:00

13:00 – 15:00 OPENING AND PLENARY

Opening addresses by **Kornélia Radics** President of the Hungarian Meteorological Service; Peer Hechler, WMO representative and organizers

Statement by **Peer Hechler**, Scientific Officer, WMO

Session Chair: Peer Hechler

Blair Trewin, Enric Aguilar, Ingeborg Auer, Jose Antonio Guijarro, Peer Hechler, Mónika Lakatos, Matthew Menne, Ghulam Rasul, Clara Rojas, Tamás Szentimrey, Victor Venema, Xiaolan Wang: Implementing homogenization in national meteorological services – the WMO Task Team on Homogenization guidance

Victor Venema: The deleted chapter on future research needs

Tamás Szentimrey: Mathematical questions of homogenization and summary of MASH

15:00 – 15:20 Coffee break

15:20 – 17:00 PLENARY

Session Chair: Tamás Szentimrey

Ralf Lindau and **Victor Venema**: Relative statistical homogenization of observational networks with a low signal to noise ratio

Peter Domonkos: Capacity of Acmantv4 for Homogenizing Climatic Datasets of National Meteorological Services

J.A. Guijarro and E. Aguilar: Quality control and homogenization of the daily series of the ECA&D database under the INDECIS project

Magnus Joelsson, Christophe Sturm, Johan Södling, and Erik Engström: Birth of Bart: Automation and evaluation of the interactive mode of the homogenisation software HOMER

TUESDAY, 13 OCTOBER
13:00-17:00

13:00 – 15:00 PLENARY

Session Chair: Jose A. Guijarro

Athanassios Argiriou, Anna Mamara and Panagiotis Ioannidis: Analysis of parallel measurements of daily maximum and minimum temperatures in Greece

Johan Södling: A method for creating realistic temporal gaps in time series data

Ciara Ryan, Mary Curley, Conor Murphy and Seamus Walsh: Developing a high quality, long term rainfall network for the Island of Ireland 1900-2018.

Yizhak Yosef; Enric Aguilar; Pinhas Alpert: Long-term trends in extreme temperature and precipitation indices for Israel based on a new daily homogenized database

Elinah Khasandi Kuya, Herdis Motrøen Gjelten & Ole Einar Tveito: Homogenization of Norway's monthly temperature and precipitation series

15:00 – 15:20 Coffee break

15:20 – 17:00 PLENARY

Session Chair: Victor Venema

Beatrix Izsák, Mónika Lakatos, Rita Pongrácz, Tamás Szentimrey, Olivér Szentes: Joint homogenization of time series with unequal length by applying the MASH procedure

Roeland Van Malderen, E. Pottiaux, A. Klos, P. Domonkos, M. Elias, T. Ning, O. Bock, J. Guijarro, F. Alshawaf, M. Hoseini, A. Quarello, E. Lebarbier, B. Chimani, V. Tornatore, S. Zengin Kazancı, and J. Bogusz: Break detection in integrated water vapour benchmark datasets

Moritz Buchmann: Evaluating the robustness of snow climate indicators using a unique set of parallel snow measurement series

Gernot Resch, Barbara Chimani, Roland Koch, Wolfgang Schöner, Christoph Marty: Homogenization of long-term snow observations

WEDNESDAY, 14 OCTOBER
13:00-17:00

13:00 – 15:00 PLENARY

Session Chair: Barbara Chimani

Tamás Szentimrey: Mathematical questions of spatial interpolation and summary of MISH

Beatrix Izsák, Tamás Szentimrey, Mónika Lakatos, Zita Bihari, Andrea Kircsi: Transformation of CarpatClim datasets for grid-box average datasets

Moritz Bandhauer, Francesco Isotta, Monika Lakatos, Beatrix Izsák, Olivér Szentes, Cristian Lussana, Ole Einar Tveito and Christoph Frei: Evaluation of the precipitation climate E-OBS and ERA5 with high-resolution grid datasets in European regions

Irene Garcia Marti, **Gerard van der Schrier**, Else van den Besselaar, Cristina Rojas Labanda, Fidel Gonzalez-Rouco: Development of the E-OBS wind strength dataset

Mónika Lakatos, Tamás Szentimrey, Beatrix Izsák, Olivér Szentes, Lilla Hoffmann, Andrea Kircsi, Zita Bihari: Comparative study of CARPATCLIM, E-OBS and ERA5 dataset

15:00 – 15:20 Coffee break

15:20 – 16:10 PLENARY

Session Chair: Ole Einar Tveito

Andrés Chazarra: Development of high resolution gridded datasets of monthly temperature since 1916 for Spain

Wil Laura: Actualization of National Climate Classification Map of Peru

16:10 – 17:00 Poster session

Peter Kajaba, Katarína Mikulová, Maroš Turňa, Jakub Ridzoň: Climatic characteristics used in the design roadway

Svetlana Aniskevich, **Viesturs Zandersons**: Change point detection in monthly mean air temperature observations in Latvia

Elke Rustemeier; M. Ziese; A. Rauthe-Schöch; A. Becker; P. Finger; U. Schneider: Evaluation of interpolation scheme and extreme value indices based on GPCC‘s Full Data Daily

Miroslav Trnka, **Petr Štěpánek**, Zdeněk Žalud, Pavel Zahradníček, Martin Možný, Jan Balek, Daniela Semerádová, Monika Bláhová, Eva Svobodová: Climate monitoring products for farmers in the Czech Republic

For more information, please contact:

World Meteorological Organization

Services Department

7 bis, avenue de la Paix – P.O. Box 2300 – CH 1211 Geneva 2 – Switzerland

Tel.: +41 (0) 22 730 82 68 – Fax: +41 (0) 22 730 80 21

Email: wcdmp@wmo.int

public.wmo.int