DOI:10.28974/idojaras.2022.1.1

IDŐJÁRÁS

Quarterly Journal of the Hungarian Meteorological Service
Vol. 126, No. 1, January – March, 2022, pp. 1–26

Creation of a representative climatological database for Hungary from 1870 to 2020

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

[1]*Hungarian Meteorological Service,*
Kitabel Pál Street 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 Department of Meteorology,*
Pázmány Péter st. 1/A, H-1117, Budapest, Hungary

**Corresponding author E-mail: izsak.b@met.hu*

(Manuscript received in final form August 13, 2021)

Abstract— Climate studies, particularly those that are related to climate change, require long, high-quality controlled data sets, which are representative both spatially and temporally. Changing the conditions of measurements, for example relocating the station, or changing the frequency and timing of measurements, or changing the instruments used can cause breaks in the time series. To avoid these problems, data errors and inhomogeneities are eliminated and the data gaps are filled by using the MASH (Multiple Analysis of Series for Homogenization, *Szentimrey,* 1999, 2008) homogenization procedure. The Hungarian meteorological observation network was upgraded significantly in the last decades. Homogenization of the data series raises the question of how to homogenize long and short data series together within the same process. It is possible to solve this with the MASH method due it has solid mathematical foundations, which make it suitable for such purposes. The solution includes the synchronization of the common parts' inhomogeneities within three (or more) different MASH processing of the three (or more) datasets with different lengths depending on the time periods and elements. After the homogenization process, the station data series were interpolated to a 0.1 degree regular grid covering the whole area of Hungary. The MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; *Szentimrey* and *Bihari,* 2007) program system was used for this purpose. The MISH procedure was developed specifically for the interpolation of various meteorological elements. In the case of mean temperature, we also renewed the MISH modeling, as compared to previous years, the number of homogenized stations doubled due to the new work, so it was expedient to model the climate statistical parameters with this extended station system. Time series of daily mean temperature and precipitation sum for the period 1870–2020 for Hungary were used in this study. As a result, the longest ever homogenized, gridded daily data sets became available for Hungary. The method described here can also be applied to produce representative datasets for other meteorological elements.

Key-words: homogenization, interpolation, quality control,climatological database, ANOVA method,gridding

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 (*WMO,* 2020). Since we have data series of different lengths, this is not an easy task, because in order to create a representative database we need to use as much data and as long data series as possible.

Homogenization, quality control and completion of the daily data series were implemented by the method MASH (*Szentimrey,* 1999, 2008). The interpolation or gridding of the daily data series was carried out using the method MISH method (*Szentimrey* and *Bihari,* 2007). A general overview and evaluation of the homogenization methods and the theoretical background can be found in these references (*Venema et al.,* 2012; *WMO,* 2020). While the following references (*Cressie,* 1991; *Szentimrey et al.,* 2011) are suggested for studying the general theoretical questions and methods of spatial interpolation.

2. The software MASH v3.03

Changing measurement conditions, such as station relocation, changing measuring time, or changing the instrument, may result in undue fractures in time series. At the Hungarian Meteorological Service (HMS), data errors and inhomogeneities are eliminated and data gaps are filled 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 modifies 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 MASHv3.03

The advantages of MASHv3.03 in the homogenization of monthly series are:

– 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.

- Either an additive (e.g., temperature) or a multiplicative (e.g., precipitation) model can be used depending on the distribution of the target meteorological element.
- 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.

2.2. The MASH procedure for daily series

The MASH procedure developed for daily series (*Szentimrey*, 2008, 2013) consist of the following steps:

1. preparation of 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.

The procedure repeats steps 2–8 if it is necessary.

2.3. The verification statistics in MASH

The test statistics generated automatically during the procedure are the following:
Test statistics for series inhomogenity:

- test statistics after homogenization,
- test statistics before homogenization,
- statistics for estimated inhomogeneities.

Characterization of inhomogenity:

- 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.

3. The software MISHv1.03

The method MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; *Szentimrey* and *Bihari,* 2007, 2014) was developed specifically for the spatial interpolation of surface meteorological elements. This is a meteorological system not only in respect of the aim but in respect of the tools as well. It means that as well as the predictors, all the valuable meteorological information – climate statistical parameters, supplementary model variables, and possible background information – are included. For that purpose, developing an adequate mathematical background was also necessary of course. The main difference between MISH and the geostatistical interpolation methods can be found in the amount of information used for modeling the necessary statistical parameters (*Szentimrey et al.,* 2011). In general, when using the geostatistical methods built in GIS, the sample elements for modeling are only the predictors, which sample is merely a single realization in time. However, MISH method uses the spatio-temporal data for modeling since long data series form a sample in time and space as well. The long data series is such a specialty of meteorology, which enables efficient modelling of the statistical parameters in question.

The software version MISHv1.03 (*Szentimrey* and *Bihari,* 2014) consists of two units, the modeling and the interpolation systems. The interpolation system can be applied to the results of the modeling system. These two units of the software developed can be summarized as follows.

The modeling subsystem for climate statistical (local and stochastic) parameters:

- This is based on long homogenized data series and supplementary deterministic model variables. The model variables may include such elements as height, topography, distance from the sea, etc. Neighborhood modeling and correlation model are applied to each grid point.
- It is also a benchmark study, a cross-validation test for interpolation error or representativity.
- It should be noted that the modeling procedure must be executed only once before the interpolation applications.

The interpolation subsystem:

- Either additive (e.g., temperature) or multiplicative (e.g., precipitation) model and interpolation formula can be used, depending on the climate elements.
- Daily or monthly values and means from a number of years can be interpolated.
- Only a few predictors are sufficient for the interpolation, and no problem arises if the greater part of daily precipitation predictors is equal to 0.
- Representativity is also modeled.
- Supplementary background information (stochastic variables), e.g., satellite, radar, forecast data can also be used.
- Data series completion, namely, missing value interpolation, completion for monthly or daily station data series is possible.
- Interpolation, the gridding of monthly or daily station data series for given predictand locations is possible. In case of gridding, the predictand locations are the nodes of a relatively dense grid.

For this study, station datasets are interpolated to a 0.1 degree regular grid. We calculate regional and national averages from these gridded data series. In addition, in the case of temperature, modeling of the climate statistical parameters was also renewed, as the number of stations had increased significantly. In the case of precipitation, the number of homogenized data series used for modeling did not change, so here we interpolate with the previous modeling results.

The quality of interpolation can be characterized by the representativity value in the s_0 location (*Szentimrey* and *Bihari,* 2014):

$$\mathrm{REP}(s_0) = 1 - \frac{\mathrm{RMSE}(s_0)}{\mathrm{D}(s_0)} ,$$

where $\mathrm{RMSE}(s_0)$ is the root-mean-square error and $\mathrm{D}(s_0)$ is the standard deviation in the s_0 location. We show the representativity values for the gridpoint databases obtained by gridding the different station systems.

4. ANOVA (analysis of variance) examination

Since the representativity values depend on the distribution of the given meteorological element, the comparison of the gridded data sets is supplemented. We focus on the ANOVA methodology that we used originally in the MISH procedure (*Szentimrey* and *Bihari,* 2014).

Using the basic theorem of ANOVA, the total spatiotemporal variance can be partitioned equivalently as follows.

- Sum of spatial variance of temporal means and spatial mean of temporal variances. The temporal means and temporal variances (or standard deviations) in the space can be visualized by maps.
- Sum of temporal variance of spatial means and temporal mean of spatial variances. The series of spatial means and spatial variances (or standard deviations) can be visualized by graphics.
- The above ANOVA methodology can be used for gridded monthly, seasonal, and annual series calculated from the daily series. Mean series are calculated for temperature, while sum series are calculated for precipitation.

4.1. Mathematical description

In our ANOVA examination, the following mathematical formulas are used:

- data series at s_j location and t time: $Z(s_j,t)\ (j=1,\dots,N;\ t=1,\dots,n)$,
- temporal mean at location s_j: $\hat{E}(s_j)=\frac{1}{n}\sum_{t=1}^{n}Z(s_j,t)$ (j=1,…,N) ,
- temporal standard deviation at location s_j:

$$\hat{D}(s_j)=\sqrt{\frac{1}{n}\sum_{t=1}^{n}(Z(s_j,t)-\hat{E}(s_j))^2}\quad (j=1,\dots,N)$$

,

- spatial mean at moment t: $\hat{E}(t)=\frac{1}{N}\sum_{j=1}^{N}Z(s_j,t)$ (t=1,…,n) ,
- spatial standard deviation at moment t:

$$\hat{D}(t)=\sqrt{\frac{1}{N}\sum_{j=1}^{N}(Z(s_j,t)-\hat{E}(t))^2}\quad (t=1,\dots,n)$$

- total mean: $\hat{E}=\frac{1}{N\cdot n}\sum_{j=1}^{N}\sum_{t=1}^{n}Z(s_j,t)=\frac{1}{N}\sum_{j=1}^{N}\hat{E}(s_j)=\frac{1}{n}\sum_{t=1}^{n}\hat{E}(t)$
- total variance: $\hat{D}^2=\frac{1}{N\cdot n}\sum_{j=1}^{N}\sum_{t=1}^{n}(Z(s_j,t)-\hat{E})^2$

4.2. Partitioning of total variance (Theorem)

Using the formulas defined above, the following equation can be written:

$$\widehat{D}^2 = \frac{1}{N}\sum_{j=1}^{N}(\hat{E}(s_j) - \hat{E})^2 + \frac{1}{N}\sum_{j=1}^{N}\widehat{D}^2(s_j) = \frac{1}{n}\sum_{t=1}^{n}(\hat{E}(t) - \hat{E})^2 + \frac{1}{n}\sum_{t=1}^{n}\widehat{D}^2(t) \ . \quad (1)$$

The analysis of the following terms is recommended to characterize the spatiotemporal variability:

Spatial terms: spatial variance of temporal means: $\frac{1}{N}\sum_{j=1}^{N}(\hat{E}(s_j) - \hat{E})^2$,

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

Temporal terms: spatial mean of temporal variances: $\frac{1}{N}\sum_{j=1}^{N}\widehat{D}^2(s_j)$,

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

We do not show the variances but the standard deviations to make the values easier to interpret, especially in the case of precipitation.

Total standard deviation: $\widehat{D} = \sqrt{\frac{1}{N \cdot n}\sum_{j=1}^{N}\sum_{t=1}^{n}\left(Z(s_j, t) - \hat{E}\right)^2}$.

Spatial standard deviation of temporal means: $\sqrt{\frac{1}{N}\sum_{j=1}^{N}(\hat{E}(s_j) - \hat{E})^2}$.

Root spatial mean of temporal variances: $\sqrt{\frac{1}{N}\sum_{j=1}^{N}\widehat{D}^2(s_j)}$.

Temporal standard deviation of spatial means: $\sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{E}(t) - \hat{E})^2}$.

Root temporal mean of spatial variances: $\sqrt{\frac{1}{n}\sum_{t=1}^{n}\widehat{D}^2(t)}$.

5. Homogenization of spatially and temporally expanded station systems

When the station network is upgraded and we have short data series along with the long series, the common section must be homogeneous together with the long as well as with the short data series, whilst the two or more systems have also to be homogeneous. MASH is able to fulfil 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 inhomogeneities of the common part within two or more different MASH processings for two or more datasets with different length.

We have harmonized two MASH systems in the recent years. However, as a result of digitization undertaken for the 150th anniversary of the HMS, daily precipitation sums and mean temperature data for 11 stations, starting in 1870 are now available. Three and four MASH systems are used to combine the daily precipitation sum and mean daily temperature, respectively, into one homogeneous climate database. Therefore, we harmonize three MASH systems for the precipitation database. This year, on the basis of the larger daily mean temperature database, the task is to homogenize four MASH systems together. We first note that the selection of station systems is 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 an 80- to100-year-long time series.

In order to update our database annually or to homogenize several MASH systems together, test statistics must be studied at each step so that 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.

5.1. Quality control of daily data

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 recent years, we can more easily examine the suspicious data marked in automatically generated error.res file of MASH, as we have additional data from radar or satellite measurements, or from other information sources. 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 could be possible due to the lack of cloud and snow cover, so the data in the error file is actually extreme and not an error (*Fig. 1*). There are also cases where the value in the error file is indeed erroneous, such as when there is a large discrepancy based on comparison with neighboring stations, and from a meteorologist's point of view there is nothing to justify this difference. There are many such examples from the present, while the more difficult task is to examine the data measured 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 ≥ 100 km, as the station network was even less dense 150 years ago) show 20 °C and the given station shows 10 °C, 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.) Recording errors are 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 now we know that in all of these cases, data for the previous or following month were recorded. Typical errors are the sign error and the absence of the decimal point. It can be seen from these examples, that professional analyses are needed after the automatic data verification, and the two together ensure a comprehensive examination. A major advantage of quality control in MASH is 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 does not exceed 10 such examinations per year per meteorological variable.

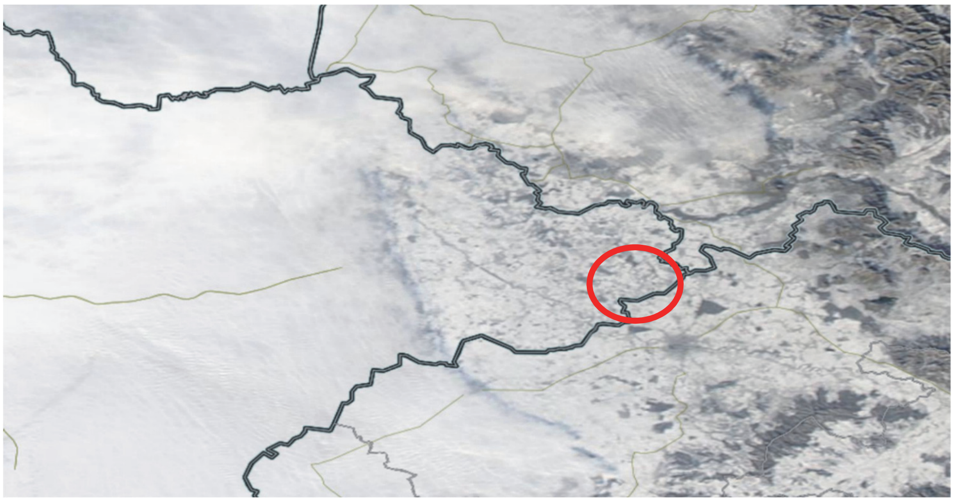

Fig. 1. Satellite image, captured on December 5, 2019. In the area marked with a red circle, a very low daily average temperature was measured compared to the nearest stations.

5.2. Quality control of 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 it can be seen in *Fig. 2*, Miskolc station has high inhomogeneity values for the period 1901–1908. We found an explanation for these about 20 years later, in the paper archives. At that time they were not measured in Celsius, but in Réaumur scale (*Fig. 3*). When checking the precipitation data archive, 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, give us 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 was 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. 4*). These are inhomogeneities, which 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 which forms the basis of the MASH system work very efficiently, but human intervention is also needed!

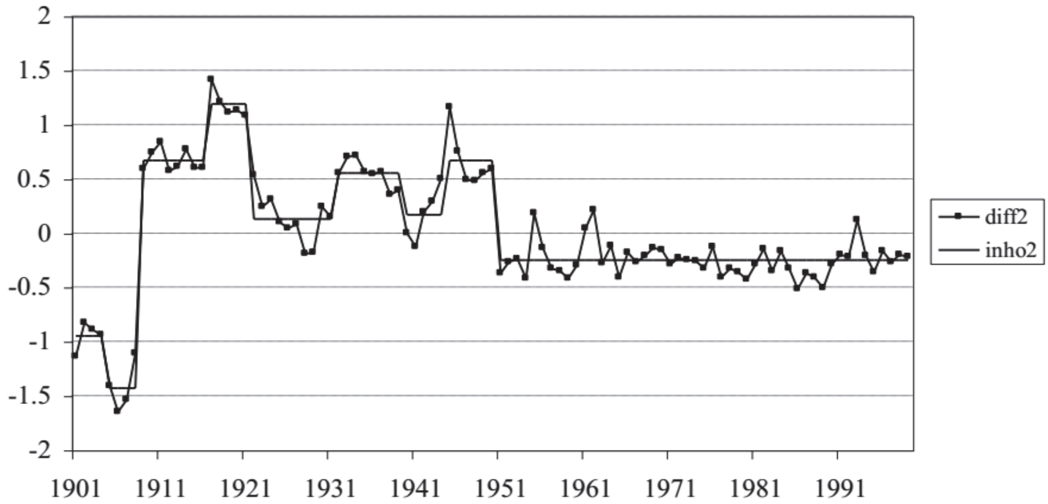

Fig. 2. The original graph of the MASH manual (*Szentimrey,* 2017): the difference series for Miskolc station and the estimated inhomogeneity.

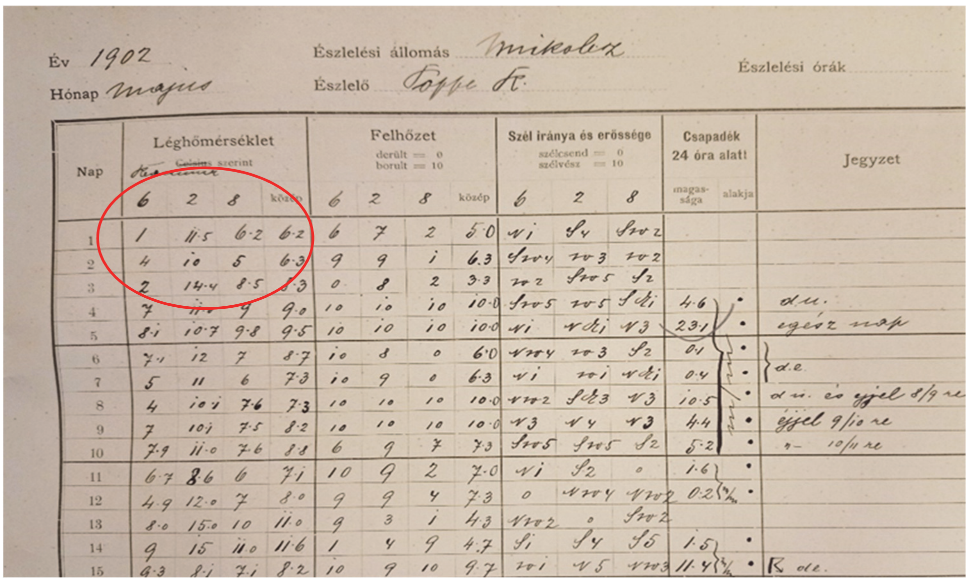

Fig. 3. The observations of May 1902 at Miskolc station. It is clear that instead of the °C already widely used at that time, the air temperature values are given in Réaumur scale. (Source: HMS)

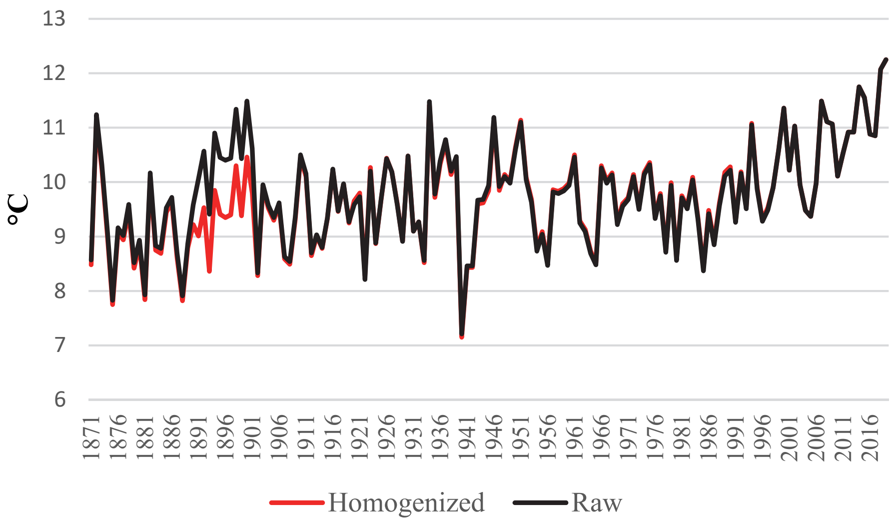

Fig. 4. Spatial mean series of annual mean temperature calculated from raw and homogenized data series at Nyíregyháza station, 1871–2019.

5.3. *Automatic use of metadata in MASH*

One of the great benefits of MASH is that it handles metadata (i.e., station history information) 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 detected (*Szentimrey,* 2017), the user can choose between basic, strict, and light versions, and MASH evaluates the metadata with automatically generated statistics (*Szentimrey,* 2017). In the case of the three longer systems the representativity values of the metadata 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, the verification statistics will improve significantly (around 0.8), but still there are stations that cannot be explained with metadata at all, and there are several stations where all breakpoints can be explained with metadata.

6. Creation of a representative climatological database

6.1. *Joint homogenization of time series with unequal length in the case of mean temperatures*

We use the MASH system at the Climate Department of the Hungarian Meteorological Service to check, homogenize, and complete the daily station data sets. As the database is constantly updated, it is necessary to complete and verify them, not only with the data of the past year, but also with the archive data being continuously digitized. Our goal is to use as many station data sets as possible for climatological analyses each year. So far, in the case of mean temperatures, we used the data of 25 stations from 1901 to 2020, and from 1971 to 2020 a further 33 stations were added, i.e., a total of 58 stations were included, to compile our database for climatological purposes.

This year, we reviewed our station system and significantly expanded it (*Fig. 5*). From 1870, data sets from 11 stations were checked, homogenized, and completed. For the period starting in 1901, we also use the data of an additional 22 stations, so from 1901 our homogenized, completed, and quality 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, with a further 22 stations, i.e., 55 station data sets altogether. The shortest period starts from 1975 and includes 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.

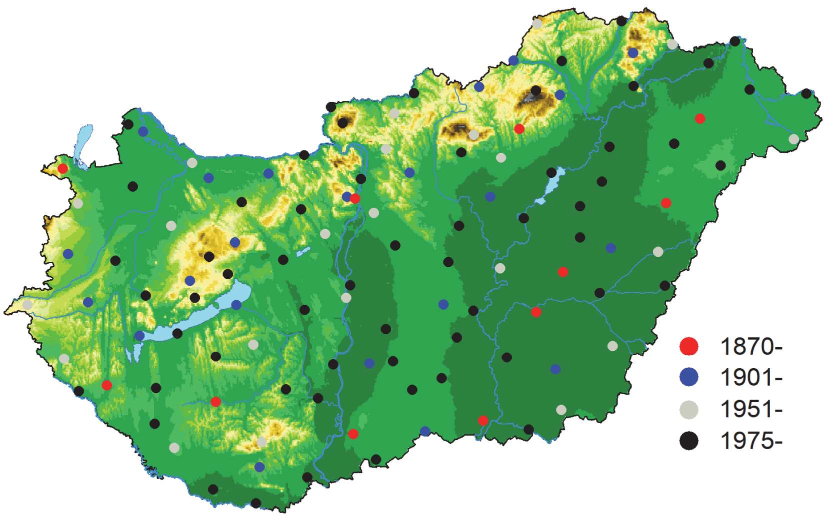

Fig. 5. Location of the stations in case of temperature.

6.1.1. Homogenization steps using four MASH systems

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 for station selection is 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. These station pairs are both at high elevations in the data-sparse mountainous regions of Bakony and Börzsöny, respectively, so it is important to included them in the analysis. Station selection is very important in order to end up with a coherent dataset. Stations where missing data are utmost have to be left out from the process. The 4 systems used in the homogenization process are presented in *Fig. 6*.

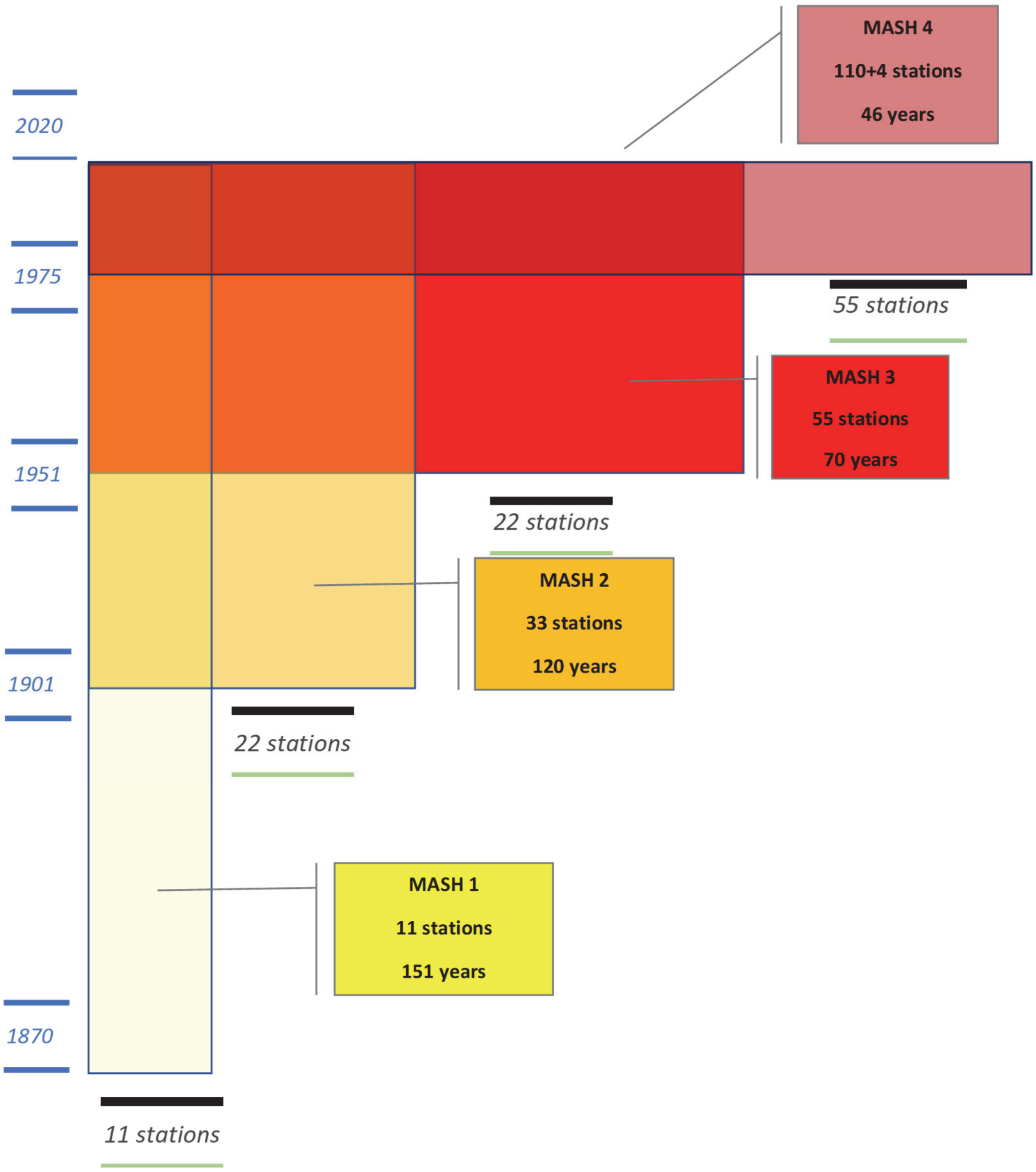

Fig. 6. The scheme of the four different station systems for the MASH procedure in the case of temperature.

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

1. MASH1: homogenization of monthly data;
2. Cutting out the inhomogeneities of the common part and inserting it into the other three MASHs (i.e., MI1, ...MI12 in the SAM directory);
3. MASH2: homogenization of monthly data;
4. Cutting out the inhomogeneities of the common part and inserting them into the other three MASHs;
5. MASH3: homogenization of monthly data;
6. Cutting out the inhomogeneities of the common part and inserting them into the other three MASHs;
7. MASH4: homogenization of monthly data;
8. Cutting out the inhomogeneities of the common part and inserting them into the other three MASHs;
9. If statistics are acceptable in MASH1: go to step 10, if not, go to step 1;
10. If statistics are acceptable in MASH2: go to step 11, if not, go to step 3;
11. If statistics are acceptable in MASH3: go to step 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.

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*.

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

MASH	MASH1	MASH2	MASH3	MASH4
Significance level: 0.05	Critical value: 22.05	Critical value: 21.76	Critical value: 21.31	Critical value: 20.86
Test statistics before homogenization	1253.08	877.95	357.2	257.49
Test statistics after homogenization	44.89	29.1	24.83	21.81
Relative modification of series	0.43	0.48	0.39	0.38
Representativity of station network	0.83	0.87	0.89	0.89

6.1.2. 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 takes 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, *mashex1-2* subroutines, can help us to decide whether to look for an additional breakpoint. It may also be necessary to manually exclude a particular station (e.g., one 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 a breakpoint. However, we may also need to manually delete a breakpoint with the *mashcor* program. For example, if the test statistics did not change before and after homogenization, despite the relative modification of series, value increased significantly. The principle of MASH is to homogenize the station data series with the slightest change in the data series. In months where the relative modification of the series value is very high for 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 (Mencshely and Hárskút). Two more stations have data sets only 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 neighboring 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 is added.

6.2. Gridding in the case of mean temperature and the ANOVA results

Through homogenization and missing data completion, the station data set was completed and finalized. These homogenized station data series formed inputs to the MISH interpolation procedure (*Szentimrey* and *Bihari,* 2007, 2014), which is applied for gridding. Because far fewer stations were considered in the previous modeling procedure, we updated the results of the MISH model using the new homogenized data sets. These results were applied to gridding in all four cases. The resulting gridded dataset has acceptable quality, it is representative spatially as well as temporally, therefore, it can be used for climate change studies. Since during the MISH modeling the resolution is set to a very fine 0.5′ grid, the information obtained in this way can be applied to any location within Hungary in the future. The gridded datasets from the four different station networks (*Figs. 5, 7*) were compared. Considering the representativity values for these data sets, a significant improvement is obvious, especially in the northern counties (*Fig. 8*). However, the representativity values themselves are not yet sufficient to characterize the created database. The ANOVA methodology (see Section 4), among others, enables us to compare the grid point databases. We looked at the ANOVA results for the common period 1975–2020. Temporal mean values and temporal standard deviation values of annual mean temperature for the four different station systems are shown in *Fig. 9*. The difference between the temporal means of the systems, which consists of stations 11 or 114, is illustrated in *Fig. 10.* Lower annual mean temperatures appear in extended regions with the denser station system. Higher values were obtained mainly in the central and southern areas of Transdanubia. Regarding the time series of spatial statistics, while the spatial means are almost identical, the spatial standard deviations are slightly different for the four different systems (*Fig. 11*). From this, we can conclude that as a result of MISH modeling, indeed, few predictors are sufficient to calculate a national average. In order to analyze the climatic conditions of a smaller area, it is no longer sufficient to use a less dense station system, but in this case the system of 33 and 55 stations also returns the spatial variability obtained from the system of 114 stations. Further ANOVA results are listed in *Table 2*.

Finally, the four gridded datasets are merged to a common database by taking grid point values from the dataset, which consists of most stations in the given time period (*Fig. 7*). The spatial mean series of annual mean temperature obtained with the merged dataset are shown in *Fig. 12*.

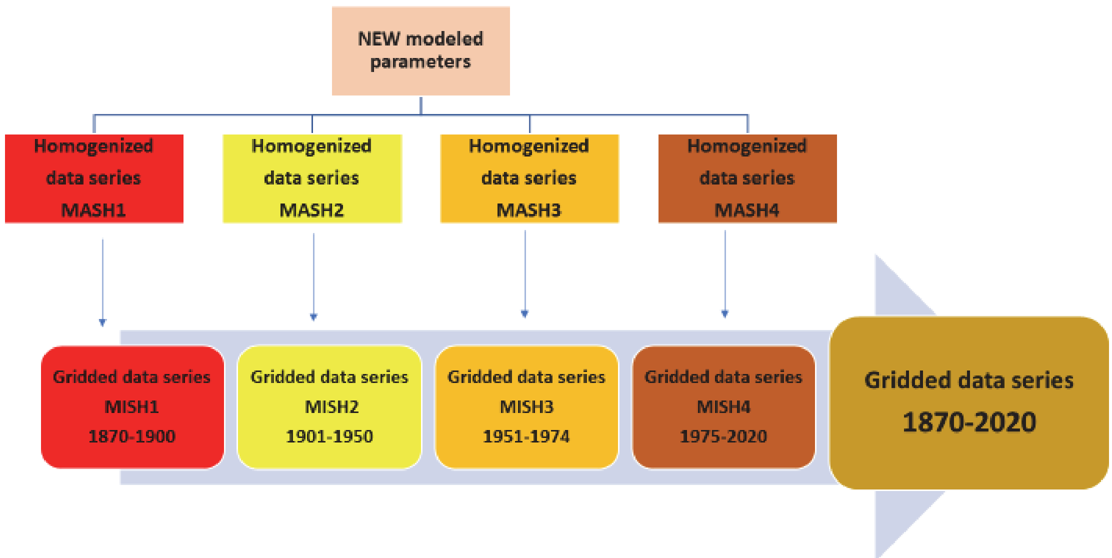

Fig. 7. The shematic diagram of creating the gridded database.

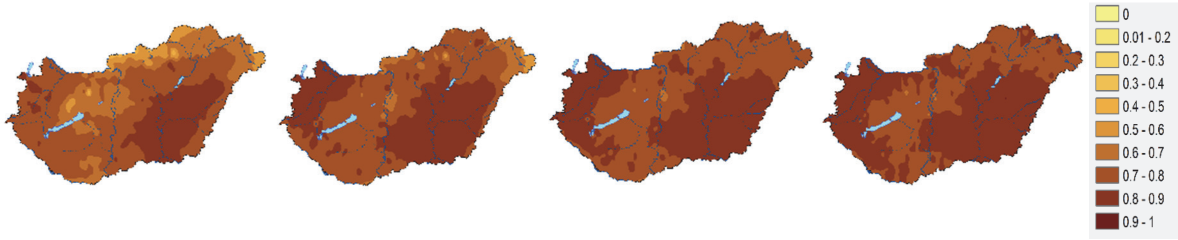

Fig. 8. Representativity values for the year for the four different station systems.

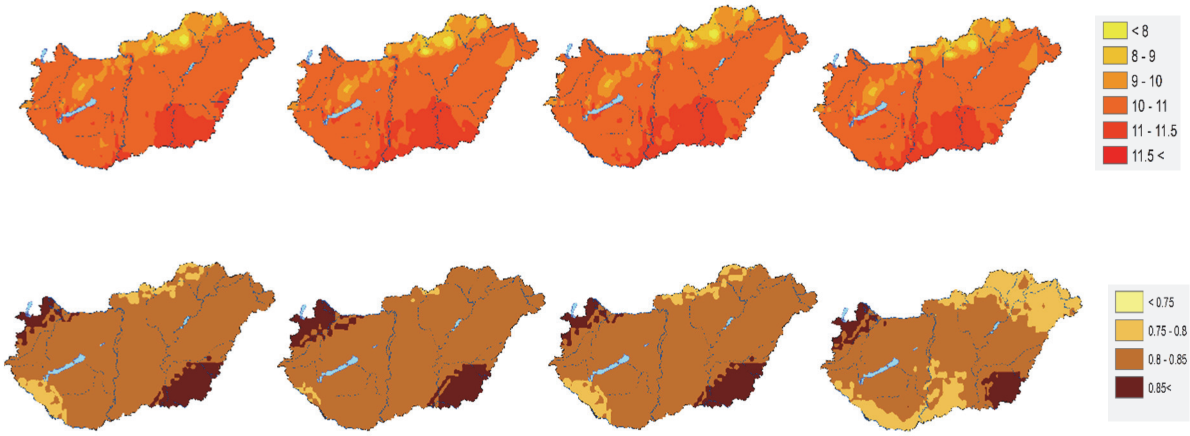

Fig. 9. Temporal mean values of annual mean temperature (top, in °C) and temporal standard deviation values of annual mean temperature (bottom, in °C) for the period 1975–2020 for the four different station systems.

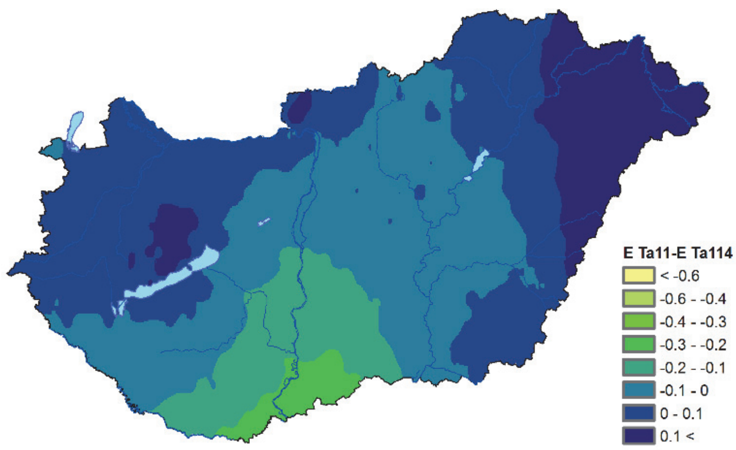

Fig. 10. The difference between the temporal mean values of annual mean temperature based on 114 or 11 gridded stations for the common period 1975–2020 (in °C).

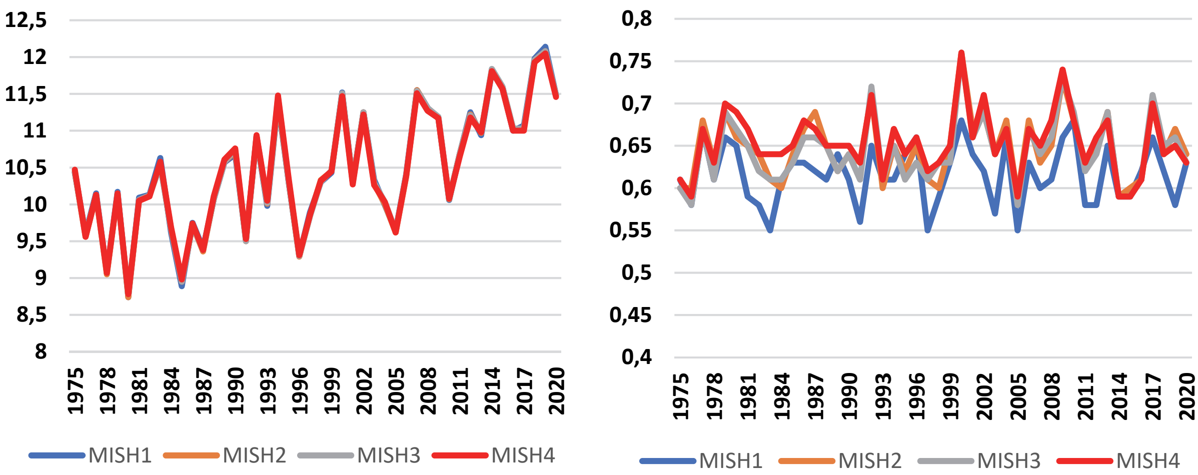

Fig. 11. Spatial mean series (left) and spatial standard deviation series (right) of annual mean temperature for the different systems (in °C).

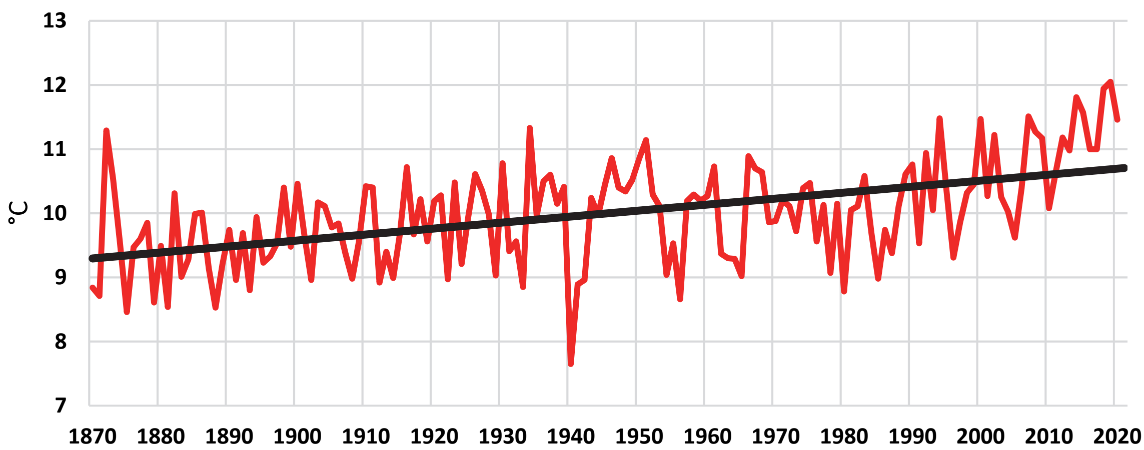

Fig. 12: Spatial mean series of annual mean temperature for Hungary from 1870 to 2020 with the estimated linear trend.

Table 2. The most important ANOVA results for the gridded annual mean temperature series for the different station systems computed for the common time period 1975–2020

	MISH1	MISH2	MISH3	MISH4
Total mean	10.47	10.46	10.47	10.46
Total standard deviation	3.25	3.23	3.24	3.23
Spatial standard deviation of temporal means	1.04	1.10	1.07	1.08
Root spatial mean of temporal variances	0.62	0.72	0.66	0.69
Temporal standard deviation of spatial means	0.83	0.84	0.84	0.83
Root temporal mean of spatial variances	0.82	0.83	0.83	0.82

6.3. Joint homogenization of time series with unequal length in the case of precipitation

The task is much simpler in the case of precipitation than in case of temperature, as the list of precipitation stations which could be used for homogenization and gridding remained unchanged compared to previous years (*Fig. 13*). The only new issue comes from the inclusion of the recently digitized data from 1870–1900 into the homogenization and gridding process. 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. 14* illustrates the periods and the number of stations that were used in this process.

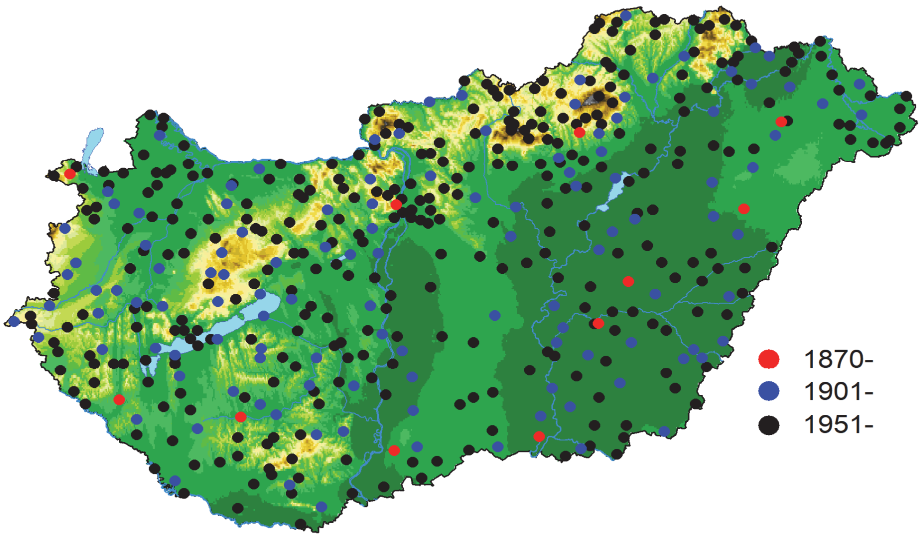

Fig. 13. Location of the stations in the case of precipitation.

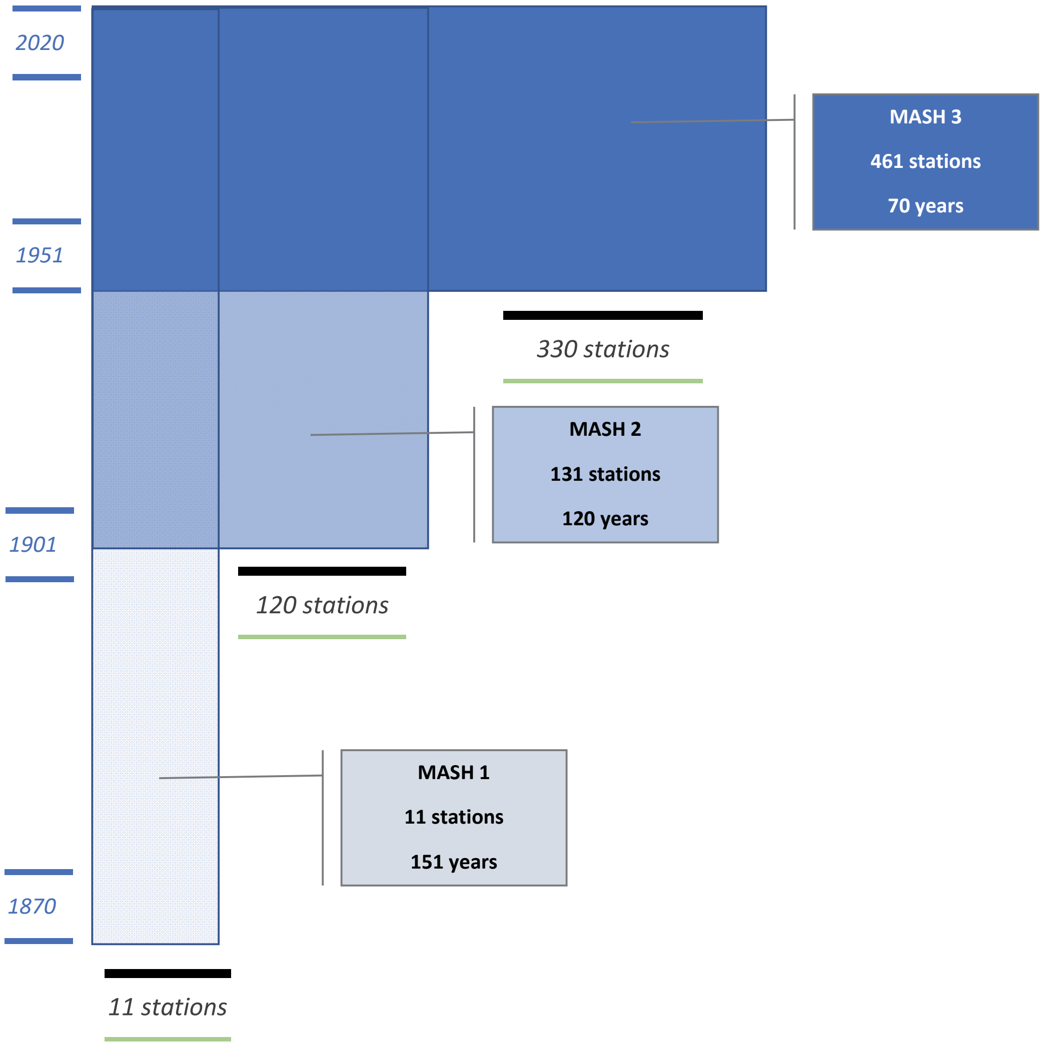

Fig. 14. The scheme of the three station systems for the MASH procedure, in the case of precipitation.

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

1. MASH1: homogenization of monthly data;
2. Cutting out and inserting the inhomogeneities of the common part into MASH2 and MASH3;
3. MASH2: homogenization of monthly data;
4. Cutting out the inhomogeneities of the common part and inserting them into MASH3 and MASH1;
5. MASH3: homogenization of monthly data;
6. Cutting out the inhomogeneities of the common part and inserting them into MASH1 and MASH2;

7. If statistics are acceptable in MASH1: go to step 8, if not, go to step 1;
8. If the statistics are acceptable in MASH2: go to step 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.

The most important verification statistics are listed in *Table 3*. Obviously, using data from 461 stations brings an improvement over using only 131 stations. Not surprisingly the representativity values of the station system consisting of 11 data series are very low. The reason for this is that precipitation varies more both spatially and temporally than the mean temperature, therefore, a 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, 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 at almost every step, however of course, the biggest help is the study of verification statistics. The advantage of this is that the database can be updated every year without having to restart homogenizing again from the beginning.

Table 3. The most important verification statistics for annual precipitation sum

MASH	**MASH1**	**MASH2**	**MASH3**
Significance level: 0.01	Critical value: 28.00	Critical value: 28.00	Critical value: 30.00
Test statistics before homogenization	45.46	63.42	42.19
Test statistics after homogenization	18.79	27.91	26.75
Relative modification of series	0.23	0.18	0.11
Representativity of station network	0.46	0.63	0.7

6.4. Gridding in the case of precipitation and the ANOVA results

Using the previous modeling results in MISH, the grid point data sets in all three station networks were created by applying MISH interpolation procedure (*Figs. 13, 15*). The representativity values for January and July can be seen in the maps of *Fig. 16*. It is clear that the representativity values are smaller in July than in January. As the station density increases, the spatial pattern is more accurate.

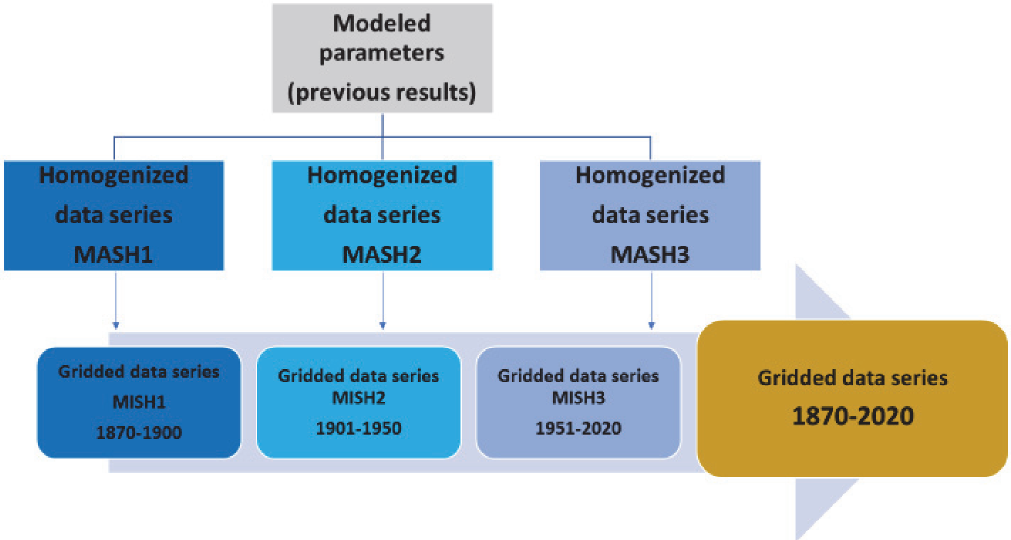

Fig 15. The shematic diagram of creating the gridded database.

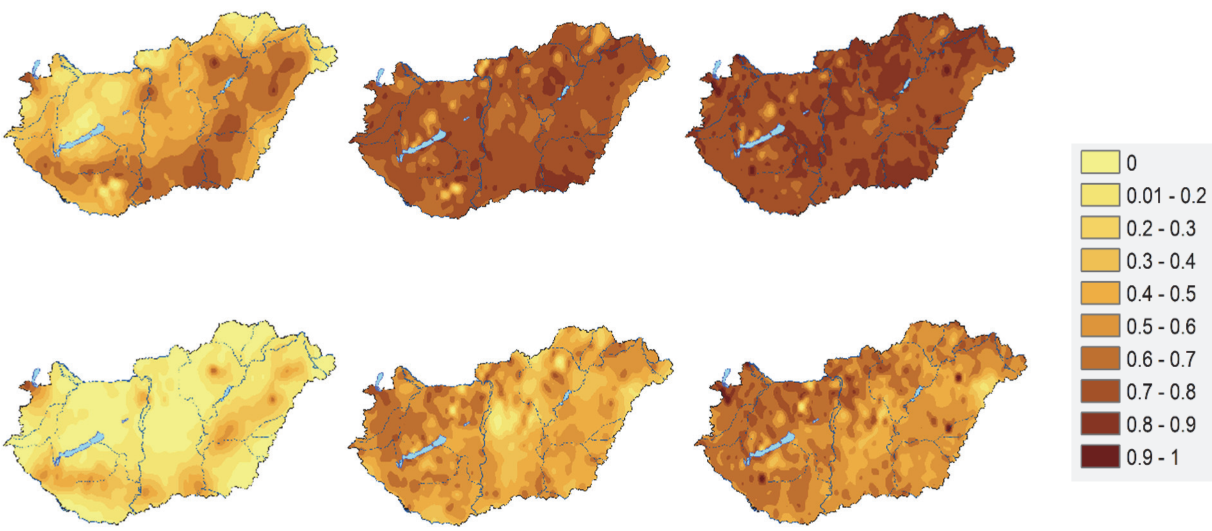

Fig. 16. January (top) and July (bottom) representativity values for the different station systems.

The ANOVA (see Section 4) results for the three precipitation station systems were compared (*Table 4*). The temporal means and standard deviations based on the common period of all three datasets are visualized in the maps of *Fig. 17*. The maps in *Fig. 18* enable us to identify the areas where more or less precipitation sums appear depending on the station systems. The spatial means are almost identical, the spatial standard deviations are slightly different for the three different systems (*Fig. 19*).

Table 4: The most important ANOVA results for the gridded annual precipitation sum series for the different station systems

	MISH1	**MISH2**	**MISH3**
Total mean	591.42	603.74	601.65
Total standard deviation	132.51	135.02	135.48
Spatial standard deviation of temporal means	68.85	66.23	65.46
Root spatial mean of temporal variances	113.22	117.66	118.62
Temporal standard deviation of spatial means	99.05	101.38	101.07
Root temporal mean of spatial variances	88.02	89.17	90.22

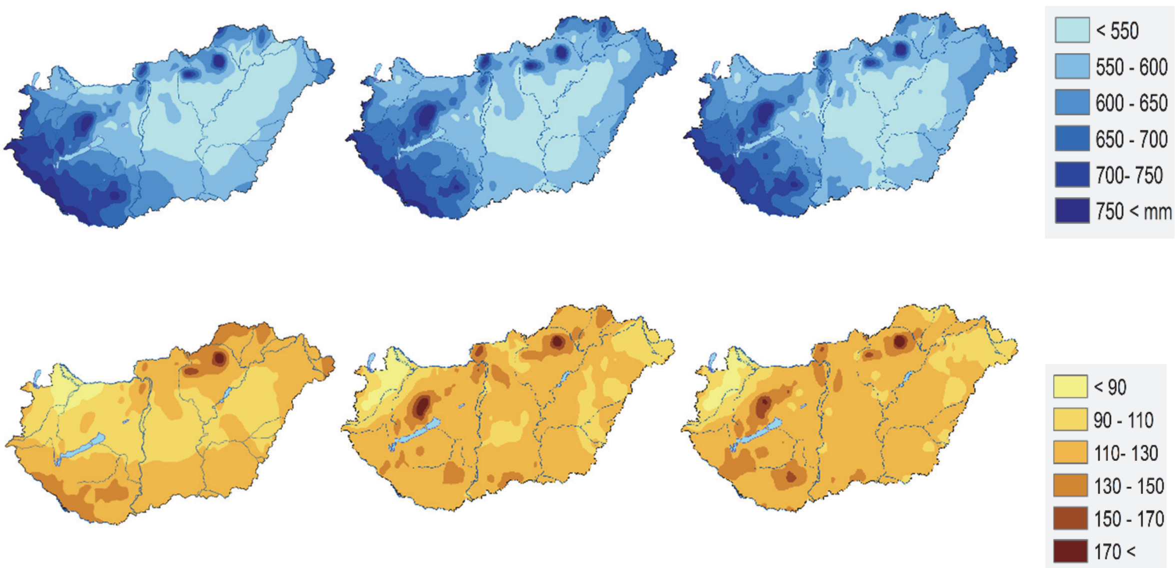

Fig. 17. Temporal mean values of annual precipitation sum (top, in mm) and temporal standard deviation values of annual precipitation sum (bottom, in mm) for the period 1975–2020 for the different station systems.

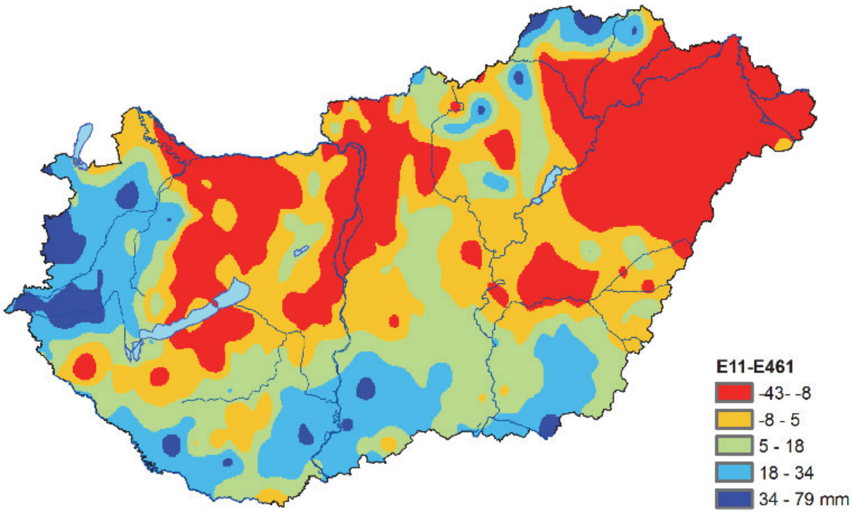

Fig. 18. The difference between the temporal mean values of annual precipitation sum based on 461 or 11 gridded stations for the common period 1975–2020 (in mm).

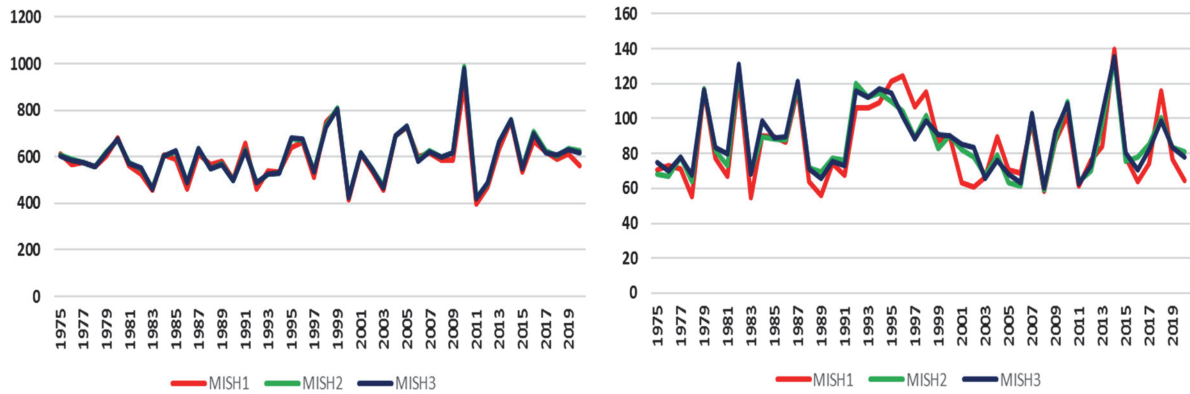

Fig. 19. Spatial mean series (left) and spatial standard deviation series (right) of annual precipitation sum for the different systems (in mm).

The three gridded precipitation datasets can be merged by taking the values from the most dense station systems for specific time periods (*Fig. 15*). Finally, the spatial mean series obtained with the merged dataset can be seen in *Fig. 20*.

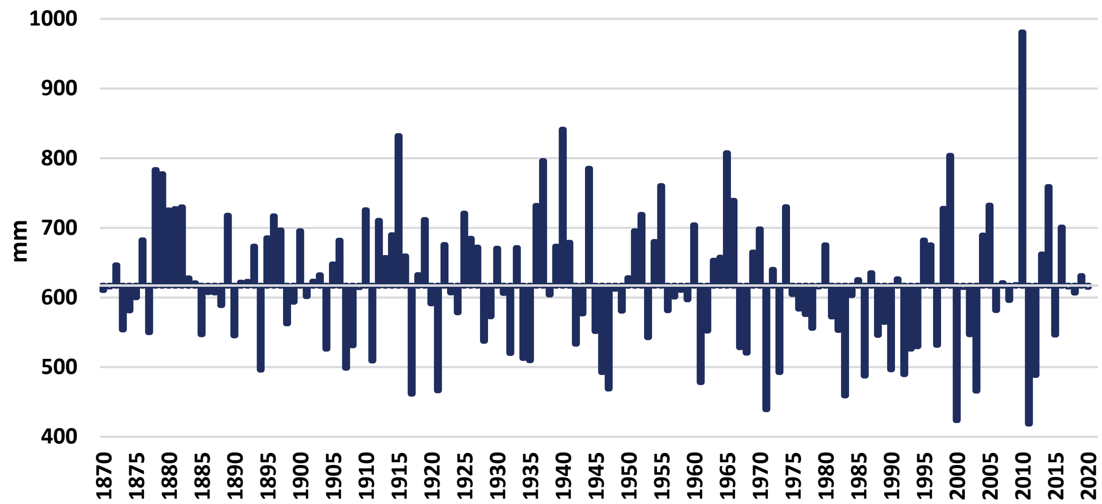

Fig 20. Spatial mean series of annual precipitation sum for Hungary, relative to the 1991-2020 normal, from 1870 to 2020.

We can conclude that we were able to harmonize station data series of different lengths, gaining as much information as possible from the measurements in order to characterize the precipitation. However, the station density for the period 1870–1900 is too sparse for quality checks, data completion, and homogenization. Thereby the quality of the time series from 1870–1900 is lower than the quality of data of the other two systems.

7. Summary

Homogenized, completed, and quality-controlled station datasets were derived from the daily mean temperature and daily precipitation sums for Hungary for different time periods. One of the most important achievements is that 151-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 daily mean temperature. In the case of temperature, the daily data of 11 meteorological stations from 1870, 33 from 1901, 55 from 1951, and 114 from 1975 were used in this process. Daily precipitation sums of 11 stations from 1870, 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 temperature. We used unprocessed archived data for the creation of the datasets for temperature and precipitation. This work will allow us to study the climate change in Hungary over a longer period of time than it was previously possible.

Acknowledgement: 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.

References

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

Izsák, B. and *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., 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 117*, 113–122.

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

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), 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. and *Bihari, Z.*, 2014: Manual of interpolation software MISHv1.03, Hungarian Meteorological Service, p. 60.

Szentimrey, T, Bihari, Z., Lakatos, M., and *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*, 1–11.

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

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