

## DATA SET DESCRIPTION

**Content of data set:** homogenized climate data series for 4 stations from 1870

**Name of data set files:**

meta\_<name of the station>.csv:

<name of the station>: name of the settlement of the meteorological station

<parameter>\_h\_< name of the station >\_period.csv

<parameter>: name of the meteorological parameter

h: homogenized

<period>: year of the start and end of the data series

### DATA SET CHARACTERISTICS

**Spatial coverage:** Budapest, Nyíregyháza, Sopron, Szeged

**Temporal coverage:** 1870 - 2021

**Temporal resolution:** daily

**Data formats:**

meta\_< name of the station >.csv:

It contains metadata of data series

<parameter>\_h\_< name of the station >\_period.csv:

1. row: header

2. row: data

1. column: date

2. column: data

**Parameters:**

t: mean temperature (°C)

r: precipitation (mm)

**Uncertainties:**

The quality of the homogenized data series in general depends on the number of stations used for the joint studies, the extent of the lack of data, the accuracy of the measurement of the raw data and, of course, the mathematical methods used in the procedure. With MASH v3.03 software, high-quality homogenized data sets can be generated, including error statistics for

raw data, test statistics for inhomogeneities of series, station network representativity values, test statistics for station history information, and the extent of modifications.

**Data quality information:**

The homogenized data series were adapted to the present measurement conditions, inhomogeneities due to changes in measurement conditions (station relocation, instrument change, measurement time change, etc.) were filtered out, gaps in the raw data set were filled in, and a second data check was performed.

**DATA ORIGIN, METHODOLOGY**

The raw data series from the official database of the Hungarian Meteorological Service are quality controlled, homogenized and completed by the MASH homogenization method.

**CONSIDERATIONS/SUGGESTIONS FOR APPLICATIONS**

Homogenized station data sets are suitable, among other things, for deriving climatic averages and other climatic characteristics, such as various climate indices, and for monitoring climate change over time. However, it should be noted that the measured extremes in the original data sets are not necessarily found in the homogenized data sets, as homogenization always takes into account the current conditions and adjusts the previous periods accordingly. Thus, it may be that an older extreme would not have occurred in the current circumstances, but extremes may also arise during homogenization.

**ADDITIONAL INFORMATION**

According to the plans, the data series will be updated with the data of the previous year by March 31 of each year. Any changes in the database during the updates (data replacement, data verification) may result in changes in the data series. In addition, the range and number of stations taken into account in the calculations could change, and thus the homogenized values may vary.

**POINT OF CONTACT**

[odp@met.hu](mailto:odp@met.hu)