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], Olivér Szentes[1]

[1] Hungarian Meteorological Service, Unit of Climatology, Hungary

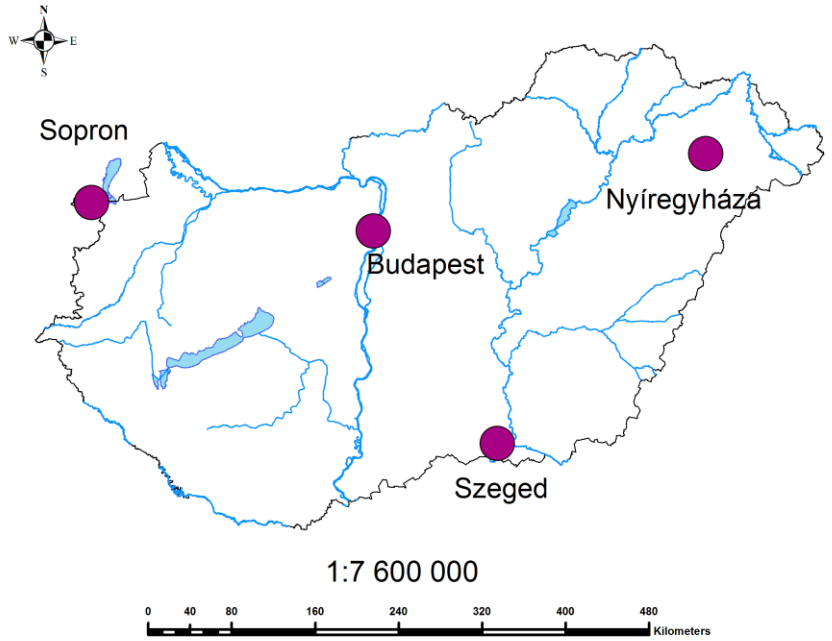
[2] ELTE Faculty of Science, Doctoral School of Earth Sciences Budapest, Hungary

[3] Varimax Limited Partnership Budapest, Hungary

[4] ELTE Institute of Geography and Earth Sciences, Department of Meteorology Budapest, Hungary

The Hungarian Meteorological Service (OMSZ) is celebrating its 150th birthday this year.

On the homepage of the Hungarian Meteorological Service we have published 4 stations data series from 1870 until 2020. We used MASH for quality control of the archive daily data series.



ORSZÁGOS METEOROLÓGIAI SZOLGÁLAT
1870-2020

Creating a representative database for the study of Hungary's climate

Creating the representative database we have to use as much data and as long data series as possible!

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.

Innovációs és Technológiai Minisztérium

Új Nemzeti Kiválóság Program

National Research, Development and Innovation Office Hungary

A spatially and temporally representative climatological database is created using MASH and MISH software:

MASHv3.03

(Multiple Analysis of Series for Homogenization; *Szentimrey, T.*)

For homogenization, quality control and missing value completion of station daily data series

MISHv1.03

(Meteorological Interpolation based on Surface Homogenized Data Basis; *Szentimrey, T.and Bihari, Z.*)

For gridding (interpolation) of homogenized daily data series

http://www.met.hu/en/omsz/rendezvenyek/homogenization_interpolation/software/

ORSZÁGOS METEOROLÓGIAI SZOLGÁLAT
1870-2020

The main features of MASHv3.03

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.
- Additive (e.g. temperature) or multiplicative (e.g. precipitation) model can be used depending on the climate elements.
- Including Quality Control and missing data completion.
- Providing the homogeneity of the seasonal and annual series as well.
- Meta data (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.**

Homogenization of daily series:

- Based on the detected monthly inhomogeneities.
  Including Quality Control and missing data completion for daily data.
  The quality control results can be evaluated by test tables generated automatically during the procedure.

Problem: how do we homogenize 3 time series of different lengths together?

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 three 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 three different MASH processing for the three datasets with different length.

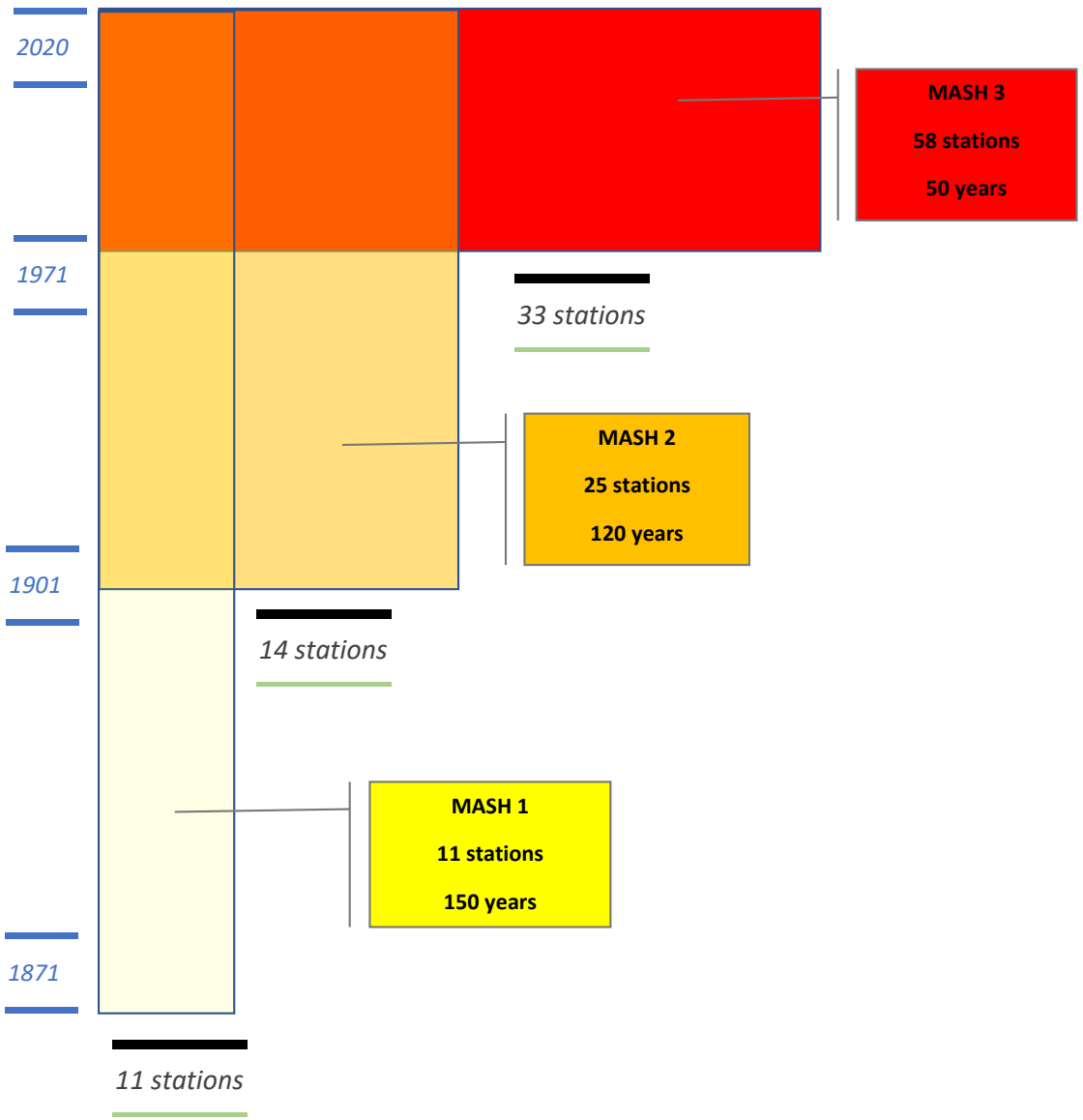
MASH is an iteration procedure, the series are examined and adjusted many times, therefore the homogenization of the new system can be considered as a continuation of the earlier homogenization procedure.

The **test of hypothesis** and throughout this test, the **test statistics** enable us to use the former results.

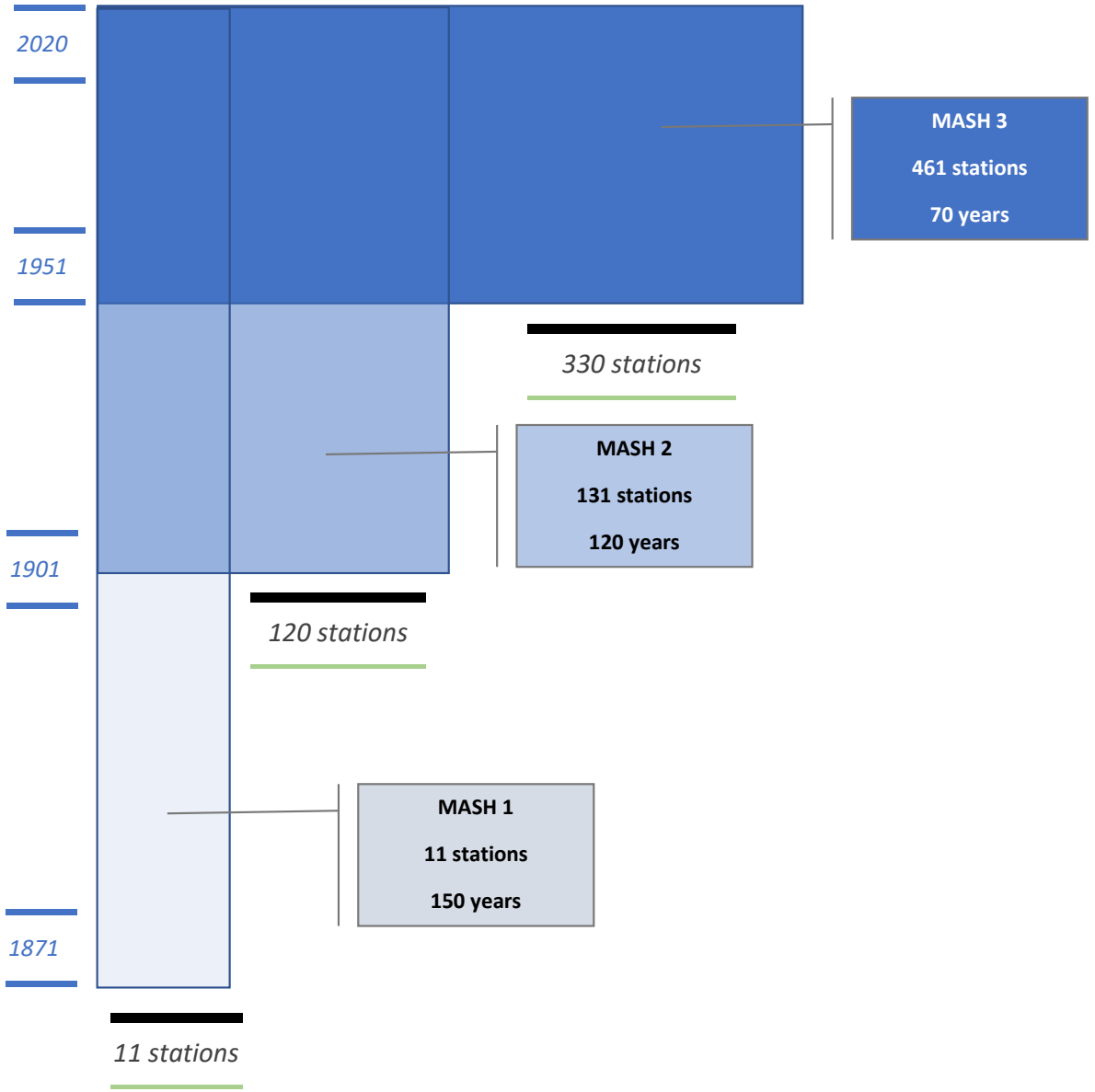
Verification of homogenization

- 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

Daily mean temperature



Daily precipitation



Homogenization with 3 MASH systems

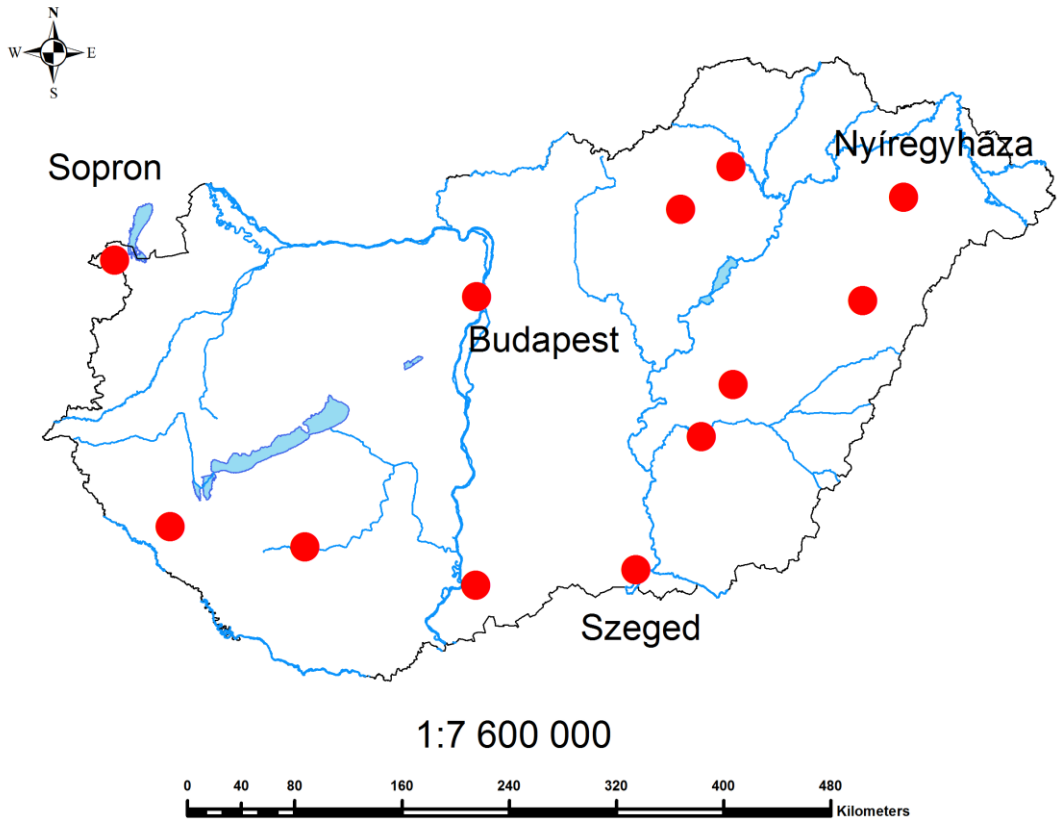
Steps:

1. -MASH1: monthly data homogenization
2. -Cut out the inhomogeneities of the common parts and insert them into MASH2, MASH3
3. -MASH2: monthly data homogenization
4. -Cut out the inhomogeneities of the common parts and insert them into MASH1, MASH3
5. -MASH3: monthly data homogenization
6. -Cut out the inhomogeneities of the common parts and insert them into MASH1, MASH2
7. -If the statistics in MASH1 are acceptable go to step 8
   - If not, go to step 1
8. - If the statistics in MASH2 are acceptable go to step 9
   - If not, go to step 3
9. - Homogenization of daily data in MASH1, MASH2, MASH3
10. - Compilation of homogenized daily data from MASH1, MASH2, MASH3
11. –End

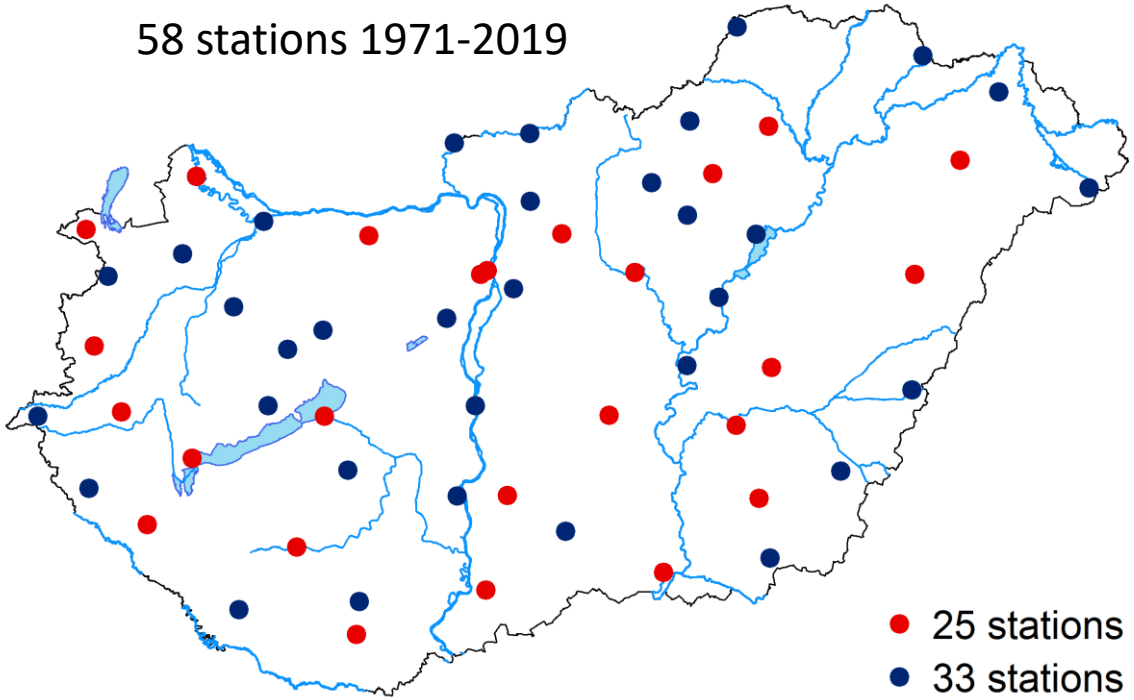
ORSZÁGOS METEOROLÓGIAI SZOLGÁLAT
1870-2020

Location of the temperature stations

11 stations 1871-2019

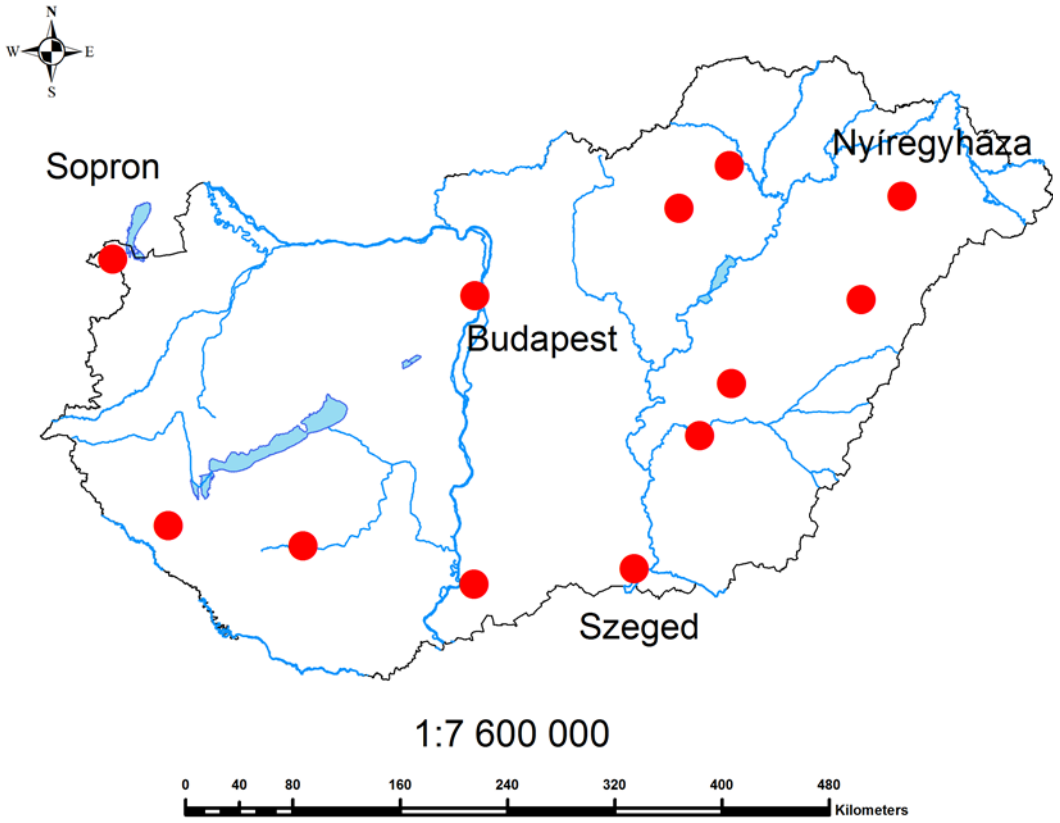


25 stations 1901-2019
58 stations 1971-2019

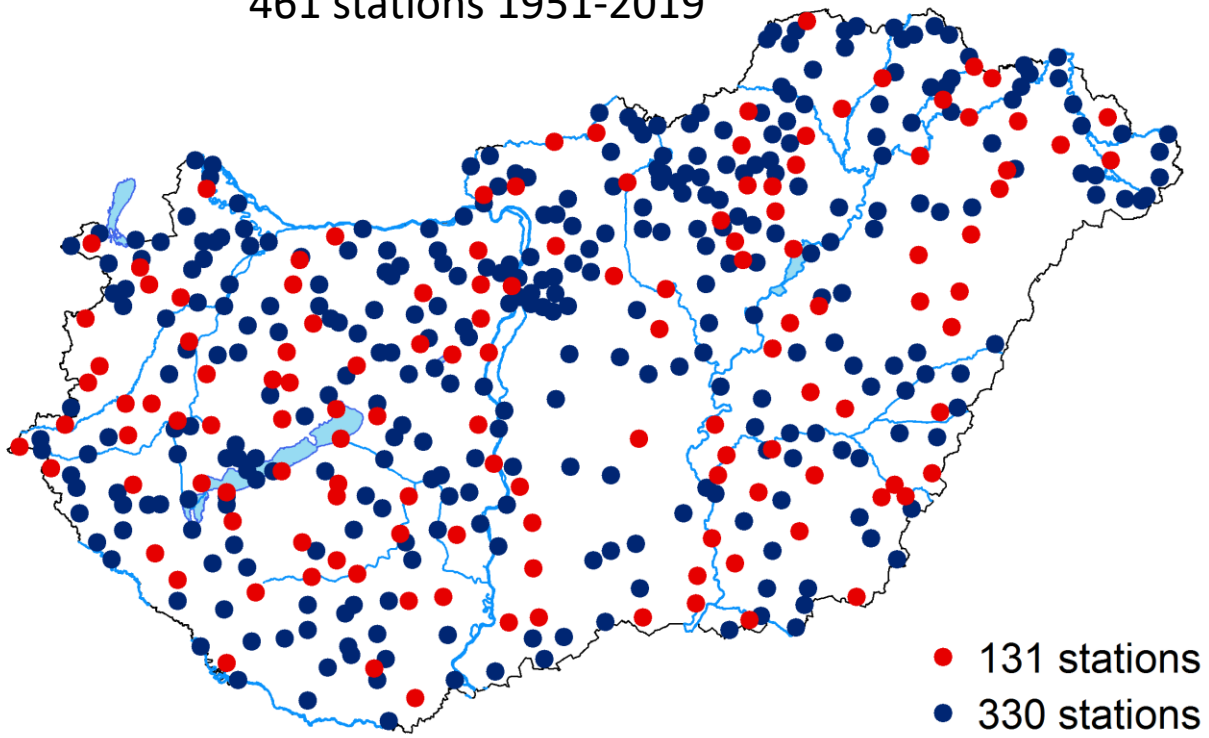


Location of the precipitation stations

11 stations 1871-2019



131 stations 1901-2019
461 stations 1951-2019



Some preliminary results: v17

MASH1

Mean temperature, annual values

Statistics of the 11 station series from 1871 to 2019:

Test Statistics Before Homogenization

Series	Index	TSB	Series	Index	TSB	Series	Index	TSB
63413	10	2042.24	56312	8	1864.82	55808	7	1194.04
44121	4	851.32	58116	9	478.47	48101	5	302.67
64711	11	285.2	27815	3	203.58	17809	2	177.6
53215	6	118.82	13704	1	116.05			

AVERAGE: 694.07

Null hypothesis: the examined series are homogeneous.
Critical value (significance level 0.05): 22.05
Test statistics (TS) can be compared to the critical value.

Some preliminary results: v17

MASH2

Mean temperature, annual values 25 stations, 1901-2019

1. Test Statistics After Homogenization

Series	Index	TSA	Series	Index	TSA	Series	Index	TSA
17809	4	68.56	44121	12	57.04	34234	8	54.3
34808	9	51.14	64704	25	45.97	63411	24	42.53
47107	16	38.1	36100	10	35.27	27815	7	34.97
48101	17	34.16	58113	23	33.55	16414	3	32.15
55808	20	29.09	15310	2	25.38	44214	13	23.83
26505	6	23.62	13704	1	23.17	56300	21	22.65
57311	22	22	52744	18	21.05	39113	11	19.75
44400	14	19.18	46400	15	18.84	23201	5	17.88
53215	19	15.17						
AVERAGE:	**32.37**							

2. Test Statistics Before Homogenization

Series	Index	TSB	Series	Index	TSB	Series	Index	TSB
56300	21	2014.2	55808	20	1485.39	36100	10	1023.77
26505	6	895.57	58113	23	876.65	47107	16	671.48
52744	18	607.99	48101	17	573.09	34234	8	492.8
46400	15	447.74	39113	11	422.55	64704	25	421.06
44214	13	357.59	57311	22	303.46	44400	14	301.43
23201	5	239.63	44121	12	228.59	17809	4	214.64
27815	7	192.58	15310	2	188.04	16414	3	157.46
13704	1	154.99	63411	24	118.2	34808	9	107.52
53215	19	93.73						
AVERAGE:	**503.61**							

Critical value (significance level 0.05): 21.76

Some preliminary results: v17

MASH3

Mean temperature, annual values, 58 stations, 1971-2019

1. Test Statistics After Homogenization

Series	Index	TSA	Series	Index	TSA	Series	Index	TSA
24709	31	70.53	28700	34	68.48	44121	12	63.67
24120	30	60.88	46400	15	48.43	64704	25	47.93
52744	18	44.24	55507	52	41.98	58300	53	40.52
17809	4	37.54	26505	6	35.62	58113	23	34.38
15310	2	34.22	25408	32	33.51	23703	29	33.14
72103	57	32.97	56300	21	30.85	36407	38	29.62
66519	56	29.52	53521	49	27.83	34234	8	27.45
17306	28	27.14	36100	10	26.56	14707	26	26.48
57311	22	25	42602	42	24.95	26306	33	24.84
61709	54	24.52	27815	7	24.28	44214	13	23.3
73313	58	23.08	63411	24	22.98	34808	9	22.75
35116	36	21.82	52517	47	21.29	44400	14	21.09
55808	20	20.33	36500	39	19.62	54309	50	19.02
53215	19	18.81	51705	46	18.1	43613	43	17.57
16414	3	17.09	47606	45	16.85	48101	17	16.5
39113	11	14.65	13704	1	14.09	16204	27	13.99
44527	44	13.6	54600	51	13.45	37402	40	13.32
65700	55	13.28	35418	37	12.76	38605	41	12.24
47107	16	11.72	32805	35	11.54	53101	48	7.88
23201	5	5.57						
AVERAGE:	**26.82**							

Critical value (significance level 0.05): 20.86

2. Test Statistics Before Homogenization

Series	Index	TSB	Series	Index	TSB	Series	Index	TSB
43613	43	1764.94	48101	17	522.25	26505	6	452.45
14707	26	422.65	72103	57	415.65	52744	18	372.61
42602	42	372.58	61709	54	364.53	35418	37	327.96
66519	56	326.88	57311	22	310.23	44214	13	287.73
55507	52	285.89	65700	55	267.57	58113	23	264.69
36100	10	264.62	55808	20	248.52	24120	30	235.25
32805	35	212.4	36407	38	203.49	35116	36	199.69
56300	21	191.67	24709	31	180.63	34808	9	178.75
44121	12	164.74	15310	2	150.98	38605	41	146.45
52517	47	139.71	27815	7	130.12	64704	25	129.89
47606	45	128.6	44400	14	111.78	58300	53	110.26
44527	44	106.99	46400	15	97.1	53521	49	95.27
73313	58	92.13	51705	46	75.51	54309	50	72.6
63411	24	66.47	37402	40	65.75	47107	16	64.86
34234	8	57.33	54600	51	56.74	23703	29	53.04
53215	19	51.71	23201	5	51.63	16414	3	47.7
13704	1	45.25	25408	32	44.91	28700	34	36.89
36500	39	35.54	26306	33	28.81	39113	11	27.86
17306	28	27.53	17809	4	20.26	16204	27	16.47
53101	48	13.26						

AVERAGE: 93.75

ORSZÁGOS METEOROLÓGIAI SZOLGÁLAT
1870-2020

MASH3

Critical value (significance level 0.05): 20.86

Test statistics: PRECIPITATION

MASH1: 11 stations 1871-2019

- Critical value (significance level 0.01): 28.00
- Test Statistics Before Homogenization AVERAGE: 51.10

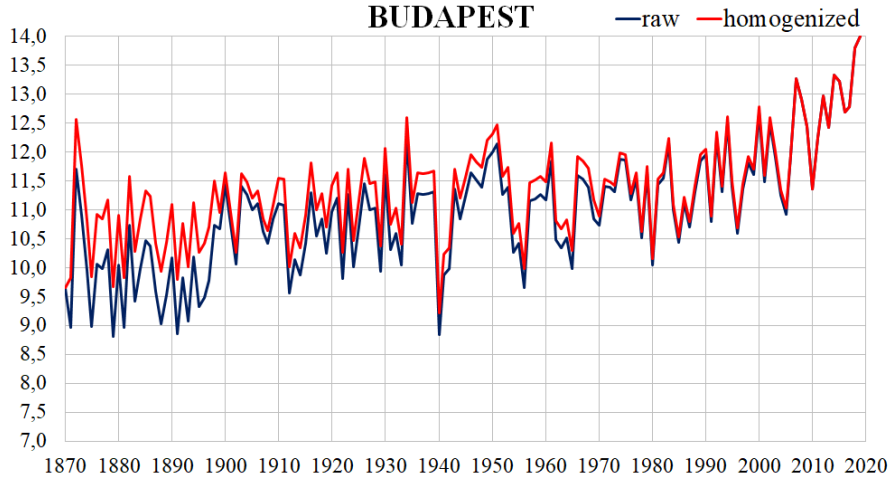
MASH2: 131 stations 1901-2019

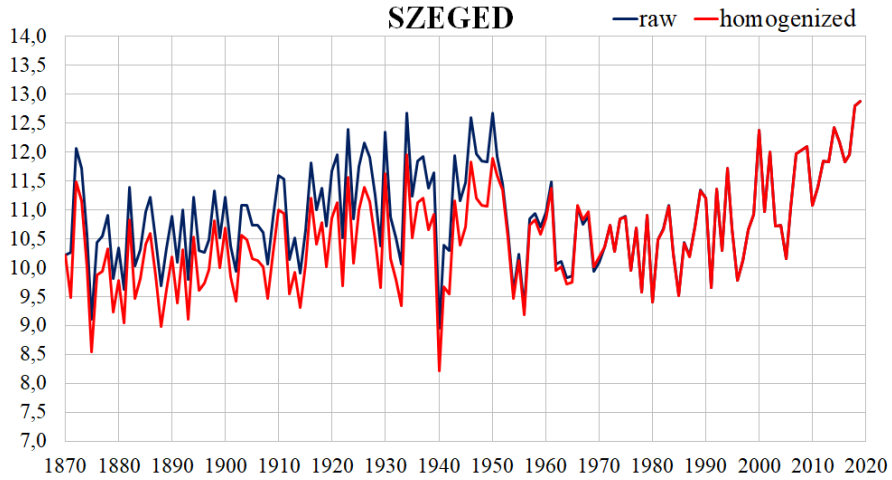
- Critical value (significance level 0.01): 28.00
- Test Statistics Before Homogenization AVERAGE: 63.24
- Test Statistics After Homogenization AVERAGE: 27.56

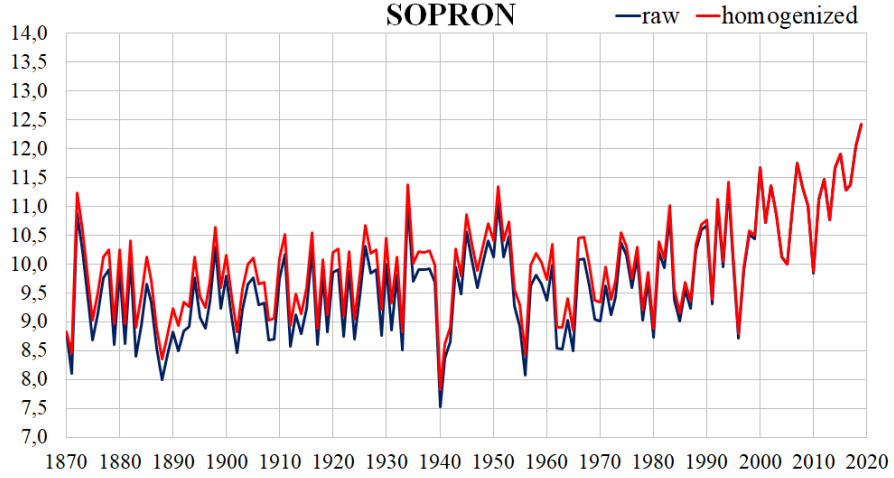
MASH3: 461 stations 1951-2019

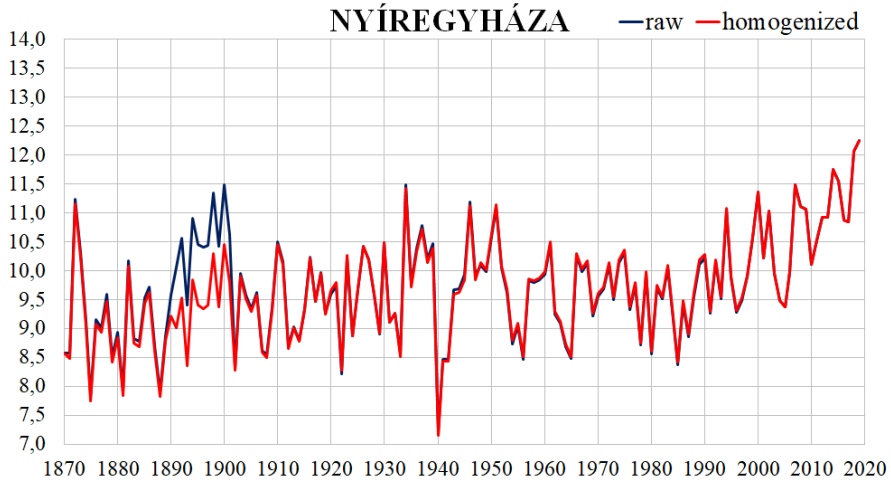
- Critical value (significance level 0.01): 30.00
- Test Statistics Before Homogenization AVERAGE: 42.73
- Test Statistics After Homogenization AVERAGE: 26.79

Homogenized and raw annual average temperature time series from 1870 to 2019, the values are in °C









Values of change (° C) over the total period (1870-2019) obtained by linear trend estimation based on homogenized and raw station data sets
(values are significant in all cases on the significance level of 0.1)

	SOPRON	BUDAPEST	SZEGED	NYÍREGYHÁZA
Homogenized	1.43	1.69	1.48	1.50
Raw	1.82	2.53	0.60	1.10
Difference	–0.39	–0.84	0.88	0.40

Izsák, B., Szentimrey, T. **To what extent does the detection of climate change in Hungary depend on the choice of statistical methods?.** *Int J Geomath* **11,** 17 (2020). https://doi.org/10.1007/s13137-020-00154-y

Thank you for your attention!

ORSZÁGOS METEOROLÓGIAI SZOLGÁLAT
1870-2020