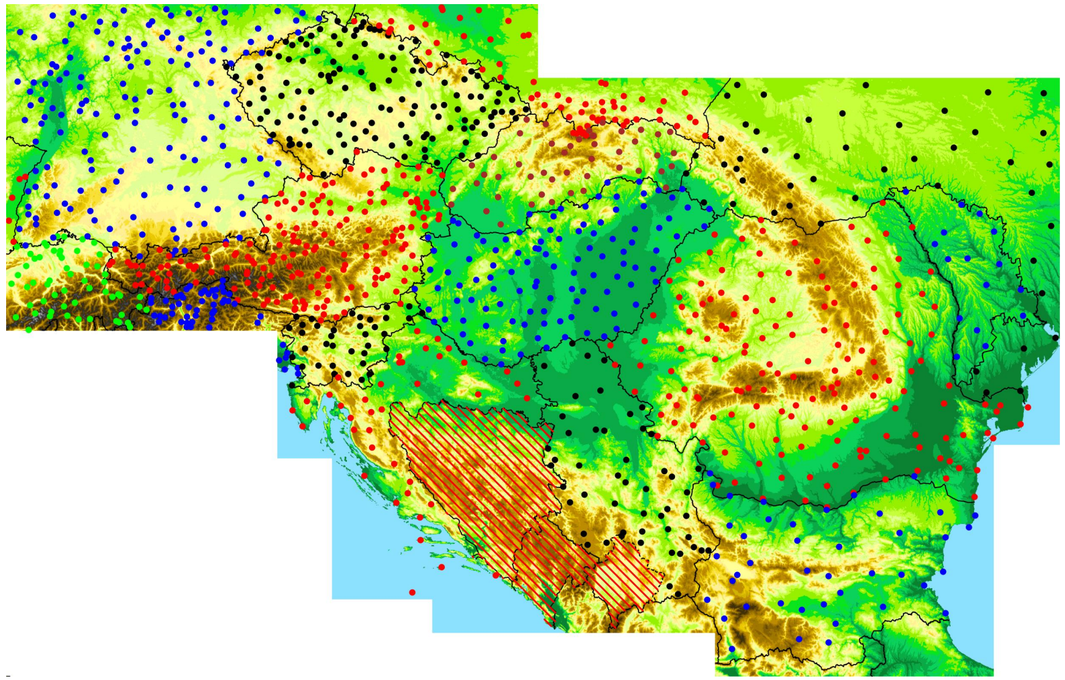
Interreg
Danube Region

Co-funded by
the European Union

Danube-ADAPT

Creation of climate database for the Danube Region: first results



Olivér Szentes[1], Beatrix Izsák[1], Mónika Lakatos[1], Zita Bihari[1], Tamás Szentimrey[2]

[1]Department of Climate Research, HungaroMet Hungarian Meteorological Service, Budapest, Hungary

[2]Varimax Limited Partnership, Budapest, Hungary

Budapest, 7th May 2026

Task:

Creation of a grided database based on meteorological measurements for the Danube region, using the principles applied in the CarpatClim database and freely available statistical climatological methods (MASH, MISH), which can be downloaded free of charge from the project website.

The resolution will be 0.1 degrees for all variables.

Interreg Danube Region
Co-funded by the European Union
Danube-ADAPT

The meteorological elements in the Danube Adapt project

Variable	*Description*	*Units*	*Time period*
Ta	*2 m mean daily air temperature*	*°C*	*1970-2024*
Tmin	*Minimum air temperature*	*°C*	*1970-2024*
Tmax	*Maximum air temperature*	*°C*	*1970-2024*
p	*Accumulated total precipitation*	*mm*	*1970-2024*
VV	*10 m horizontal wind speed*	*m/s*	*2000-2024*
Vmax	*10 m horizontal max. wind speed*	*m/s*	*2000-2024*
Rglobal	*Global radiation*	*J/cm^2*	*2000-2024*
RH	*Relative humidity*	*%*	*1970-2024*
pair	*Surface air pressure*	*hPa*	*1970-2024*

Overview of the main steps for gridded datasets:

- Near border data exchange before homogenization
- Homogenization (MASH)
- Data exchange after homogenization
- Controlling of the cross-border harmonization (MASH)
- Gridding, interpolation (MISH) per countries, sub-regions with exchanged data
- Compilation of gridded series from sub-regions in one file per variable

The main features of MASHv3.03

The software consists of two parts.

Part 1: Homogenization of monthly series:

– It is a relative homogeneity test procedure.

– It is a step-by-step procedure: the role of series (candidate or reference series) changes step by step in the course of the procedure.

– Interactive automatic, artificial intelligence (AI) system (Szentimrey, 2023).

– Additive (e.g. temperature) or multiplicative (e.g. precipitation) model can be used depending on the climate elements.

– Providing the homogeneity of the seasonal and annual series as well.

– Metadata (probable dates of break points) can be used automatically.

– The homogenization and quality control (QC) results can be evaluated on the basis of verification tables generated automatically during the procedure.

Part 2: Homogenization of daily series:

– Homogenization is based on the detected monthly inhomogeneities.

– Automatic procedures for quality control (QC) and missing data completion of daily data. The quality control results can be evaluated by test tables generated automatically during the procedure.

The main features of MISHv1.03

The software MISHv1.03 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.

Modelling system for climate statistical parameters in space:

– Based on long homogenized data series and supplementary deterministic model variables e.g. topography. Neighborhood modelling of climate statistical parameters.

– cross-validation test for interpolation error or representativity.

– Modelling procedure must be executed only once before the interpolation applications!

Interpolation system:

– 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.

– The interpolation error or representativity is modelled too.

– Capability for application of supplementary background information (stochastic variables) e.g. satellite, radar, forecast data.

– Capability for gridding of data series.

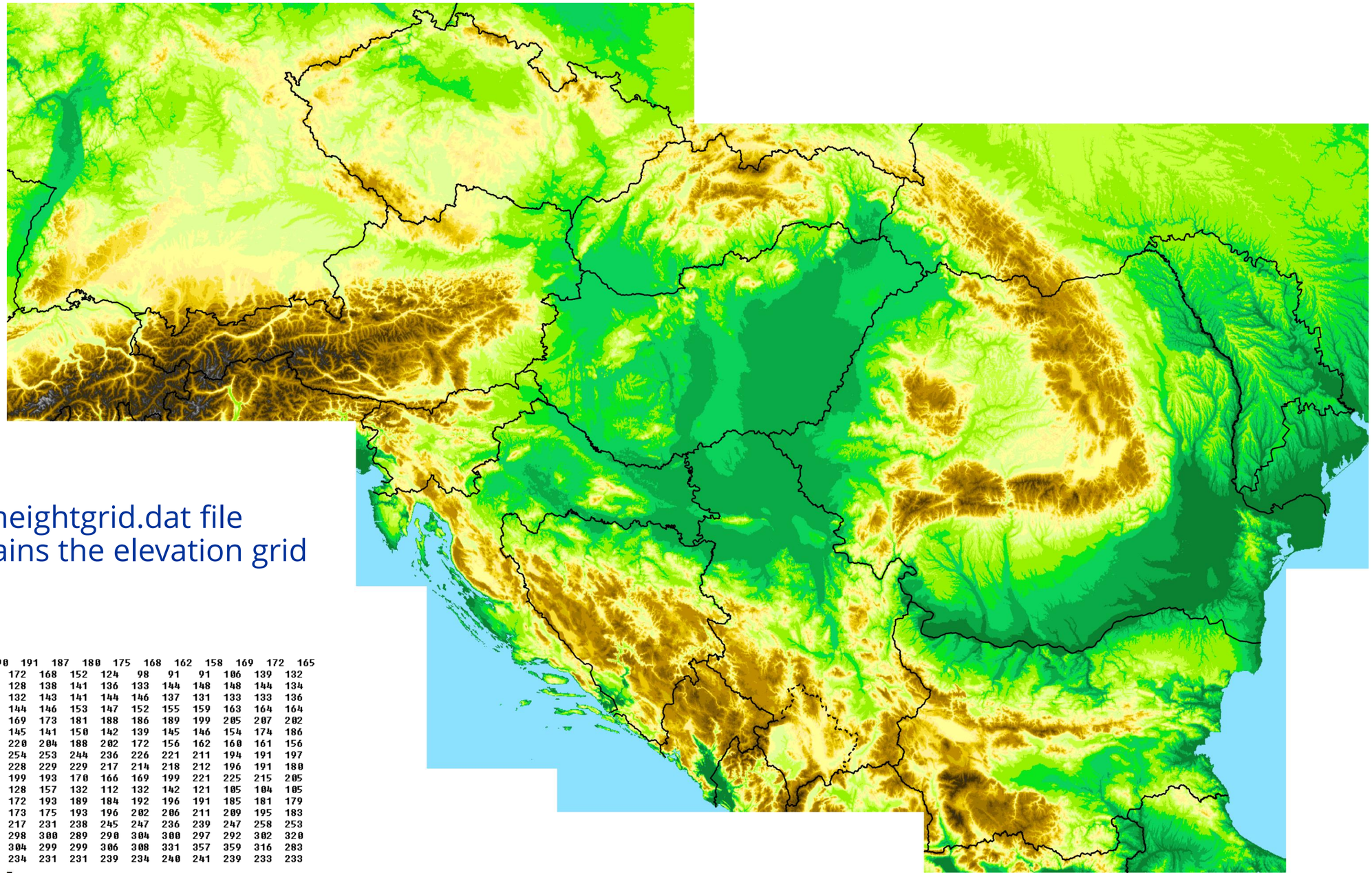
The main features of MISHv1.03

Remark

The interpolation area is covered by a gridded rectangle with half minutes 0.5’x0.5’ resolution, that is defined with the elevation data file HEIGHTGRID.DAT. This grid as a matrix, maximum number of rows is 900, maximum number of columns is 1400. There is further condition, i.e. the maximum number of the predictor stations is 500 within this gridded rectangle. (See details on p. 20 in MISH Manual)

Source of elevation data: CGIAR SRTM data from diva-gis.org

The heightgrid of the Danube Adapt area



The heightgrid.dat file contains the elevation grid

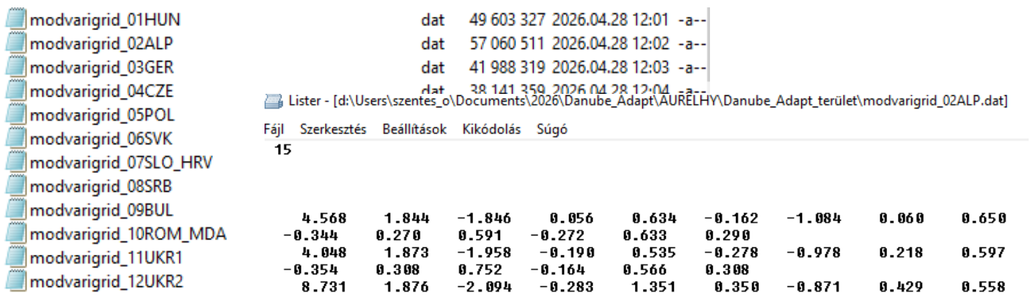
```
ncols:
854
nrows:
326
xllcorner:
11.946666565
yllcorner:
48.444999694
cellsize:
0.008333333
  182  186  191  192  190  190  191  187  180  175  168  162  158  169  172  165
152  141  166  178  180  178  172  168  152  124   98   91   91  106  139  132
125  146  145  133  130  128  128  138  141  136  133  144  148  148  144  134
138  151  151  149  146  142  132  143  141  144  146  137  131  133  133  136
131  127  132  145  153  151  144  146  153  147  152  155  159  163  164  164
164  164  164  165  171  167  169  173  181  188  186  189  199  205  207  202
199  190  180  172  158  139  145  141  150  142  139  145  146  154  174  186
195  209  216  210  213  219  220  204  188  202  172  156  162  160  161  156
187  218  222  234  245  253  254  253  244  236  226  221  211  194  191  197
200  201  199  183  196  214  228  229  229  217  214  218  212  196  191  180
180  183  192  173  175  197  199  193  170  166  169  199  221  225  215  205
201  194  204  218  219  178  128  157  132  112  132  142  121  105  104  105
105  107  114  122  132  146  172  193  189  184  192  196  191  185  181  179
180  184  180  178  185  179  173  175  193  196  202  206  211  209  195  183
189  193  198  196  195  209  217  231  238  245  247  236  239  247  258  253
261  269  265  269  276  290  298  300  289  290  304  300  297  292  302  320
312  312  329  350  338  318  304  299  299  306  308  331  357  359  316  283
282  270  257  251  245  243  234  231  231  239  234  240  241  239  233  233
```

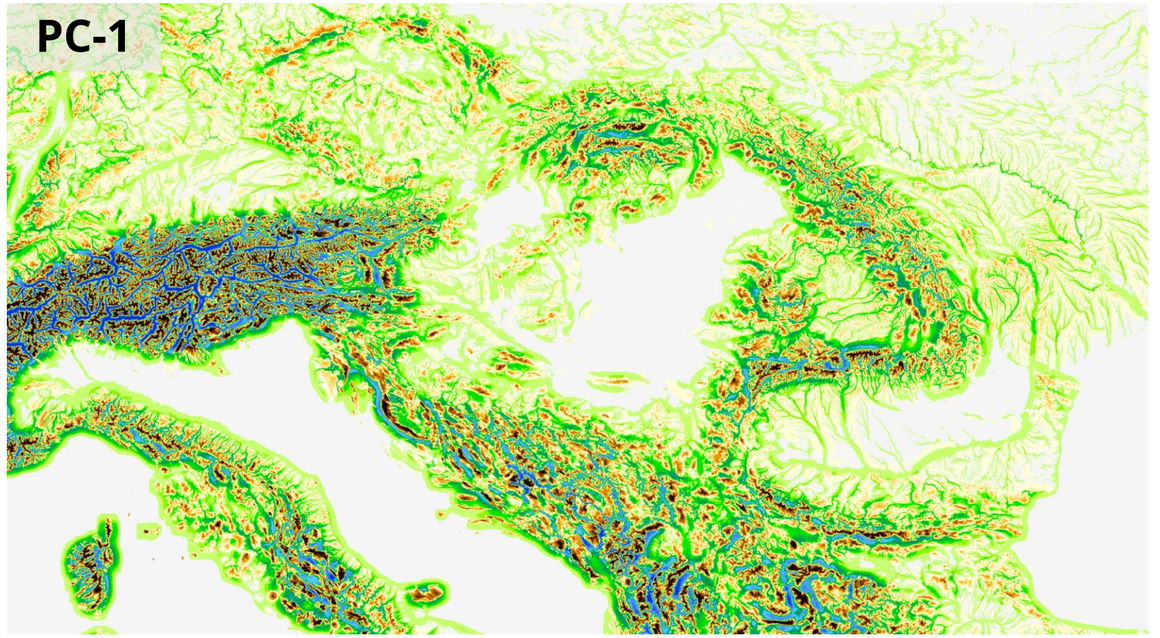
AURELHY principal components at the MISH modeling

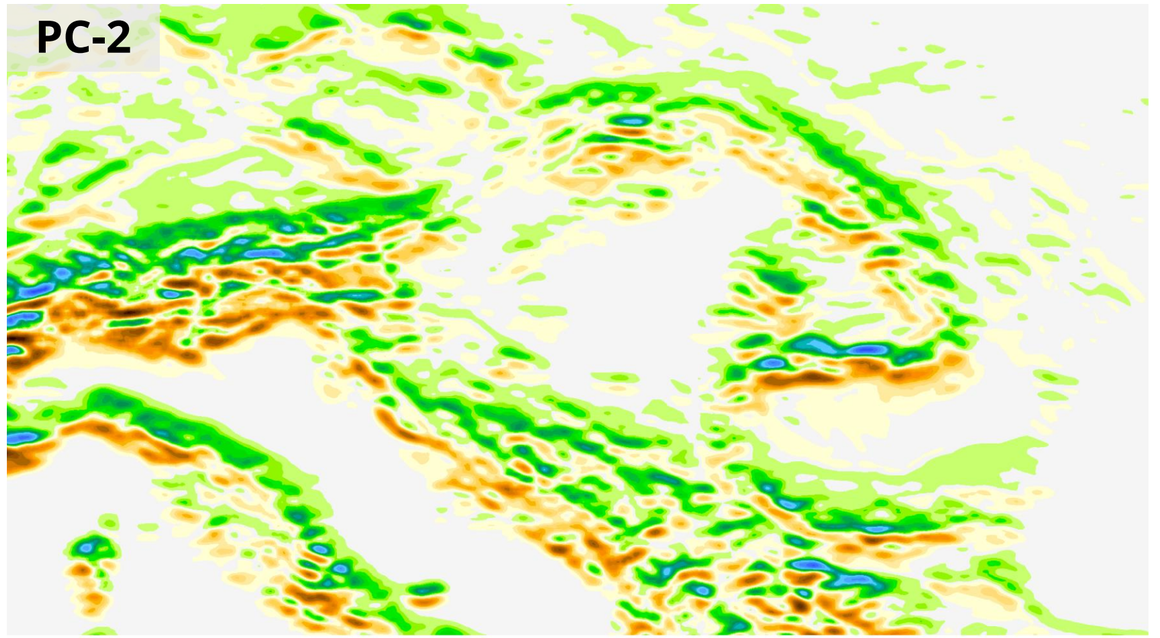
For the modeling climatic statistical parameters, in addition to the homogenized time series and elevation, we also use the first 15 Aurelhy principal components.

The modvarigrid.dat file contains the AURELHY components (it will be uploaded with the heightgrid.dat files to the FTP server soon).

The first five principal components (PCs):
PC-1 indicates peaks and valleys, PC-2 north–south slopes , PC-3 east–west slopes, PC-4 north–south saddles and PC-5 northeast–southwest saddles







Step 1: Data collection and definition of the sub-regions

The following countries homogenize and interpolate their data series themselves: Slovakia, Romania, Ukraine

For Slovakia, Romania, and Ukraine, we perform cross-border data exchange.

Here at HungaroMet, we homogenize and interpolate data originating from these countries:
Austria, Switzerland, Italy, France, Germany, Czechia, Poland, Slovenia, Croatia, Serbia, Bulgaria

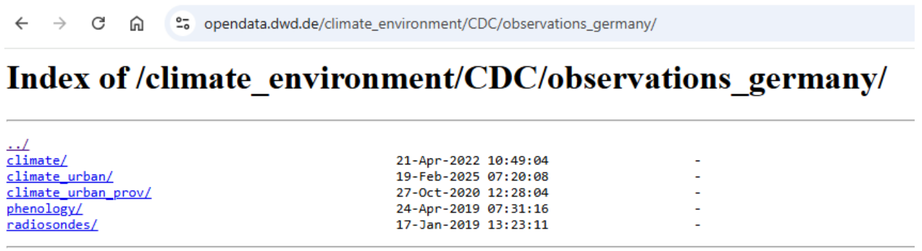
Source of the data: open data portals of national meteorological Institutes/Serveces (e.g. Geosphere Austria, Meteo Schweiz, Deutscher Wetterdienst)

The following countries send us the data: Czechia, Slovenia, Croatia, Bulgaria

We purchased the Serbian data, which was rather difficult.

Not included in the project: Bosnia and Herzegovina, Montenegro, Kosovo

The large Danube Adapt area must be divided into smaller subregions.



Sub-regions of the Danube Adapt area

We divided the entire area into 12 subregions.

Due to the harmonization of the gridded data, there are overlaps between the subregions.

Sub-regions of the project area:

01: HUN – Hungary

02: ALP – Alps (Austria & Switzerland)

03: GER – Germany

04: CZE – Czech Republic

05: POL – Poland

06: SVK – Slovakia

07: SLO_HRV – Slovenia & Croatia

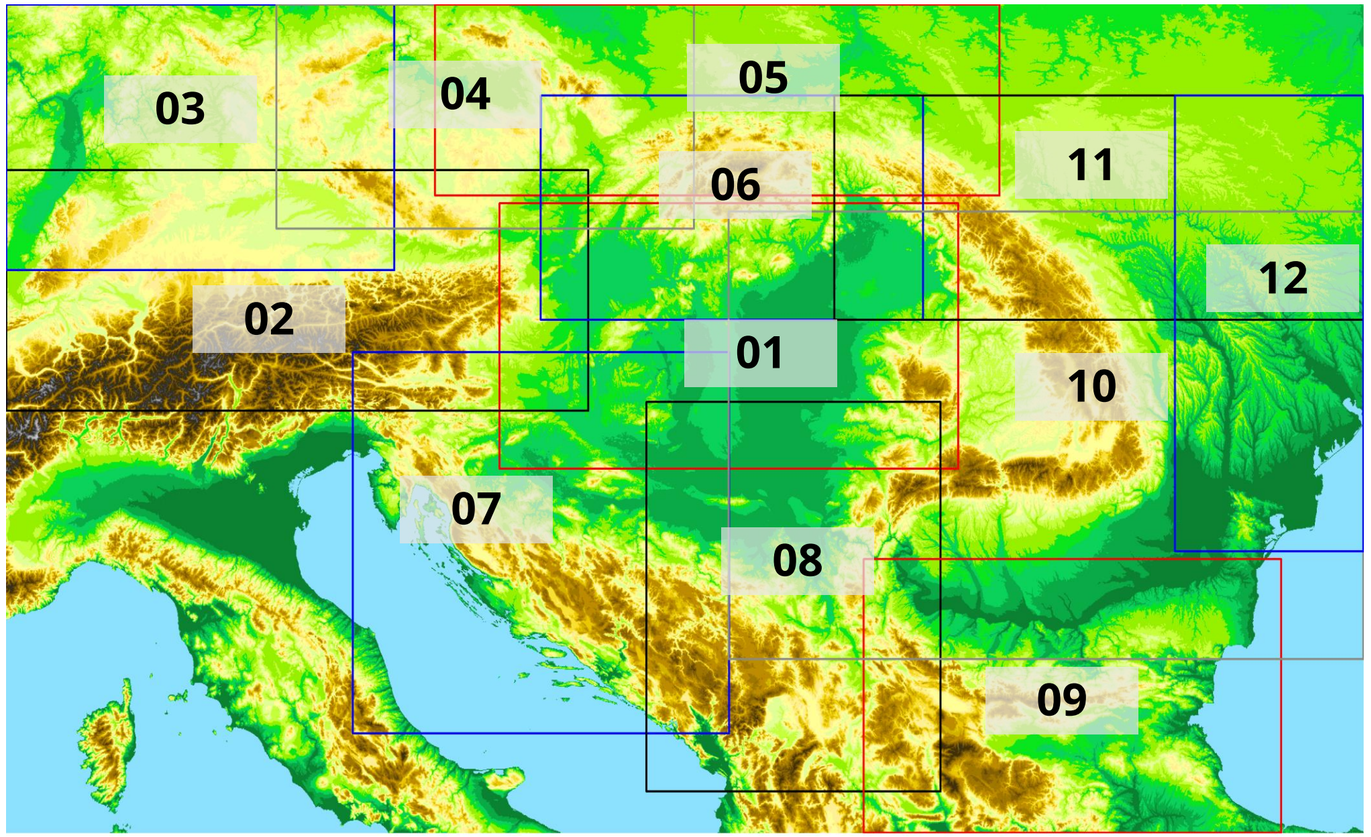
08: SRB – Serbia

09: BUL – Bulgaria

10: ROM_MDA – Romania & Moldova

11: UKR1 – Ukraine (north part)

12: UKR2 – Ukraine (east part)



Interreg Danube Region

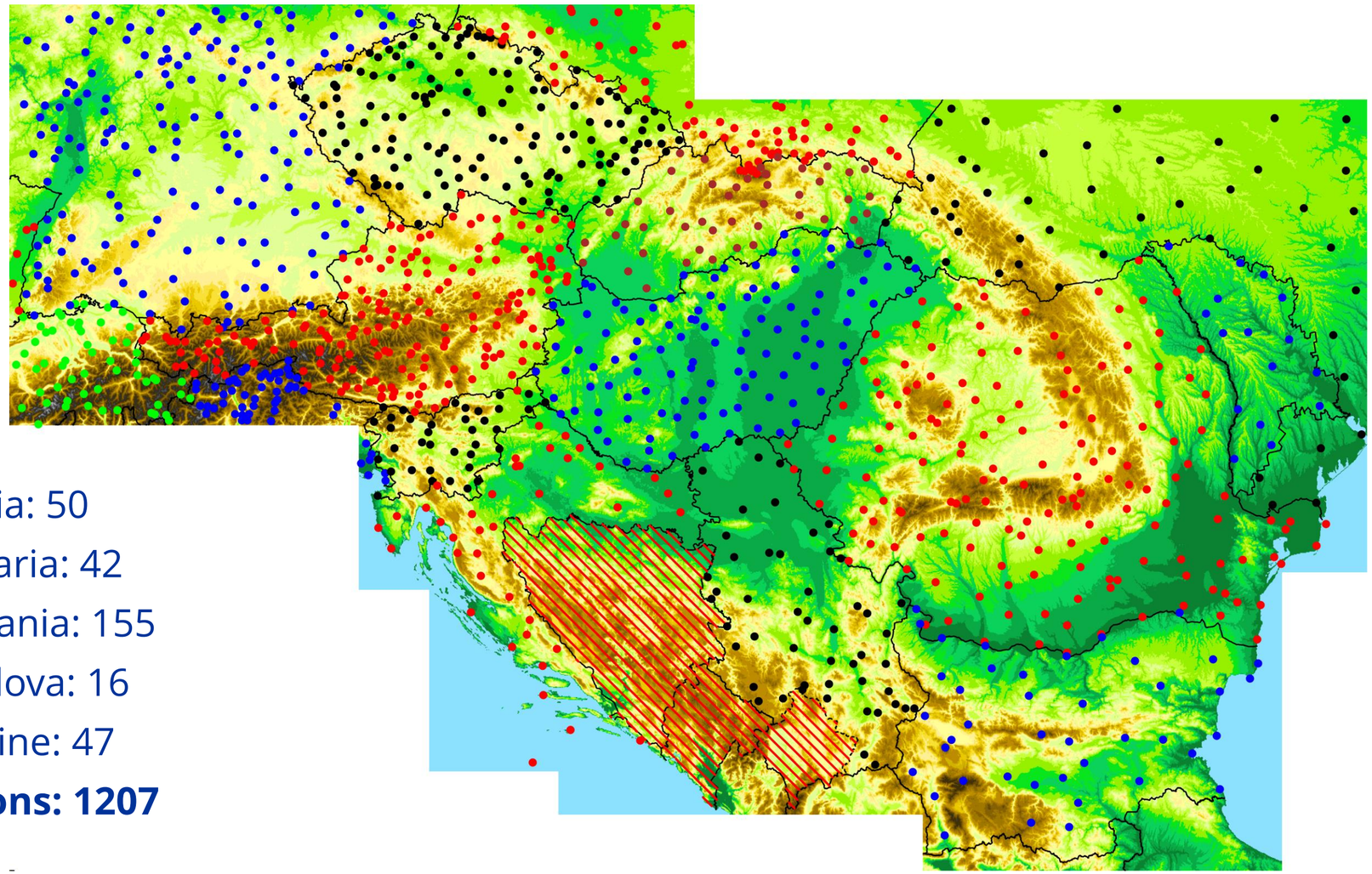
Co-funded by the European Union

Danube-ADAPT

First meteorological element: temperature (mean, min, max)

The project has begun with the homogenization of temperature data

The location of the stations with temperature data series:



Number of stations:

Hungary: 110

Austria: 181

Italy: 53

Switzerland: 57

France: 5

Germany: 161

Czechia: 137

Poland: 69

Slovakia: 47

Slovenia: 37

Croatia: 40

Serbia: 50

Bulgaria: 42

Romania: 155

Moldova: 16

Ukraine: 47

Total number of stations: 1207

Sub-region: 02–ALP (Austria & Switzerland)

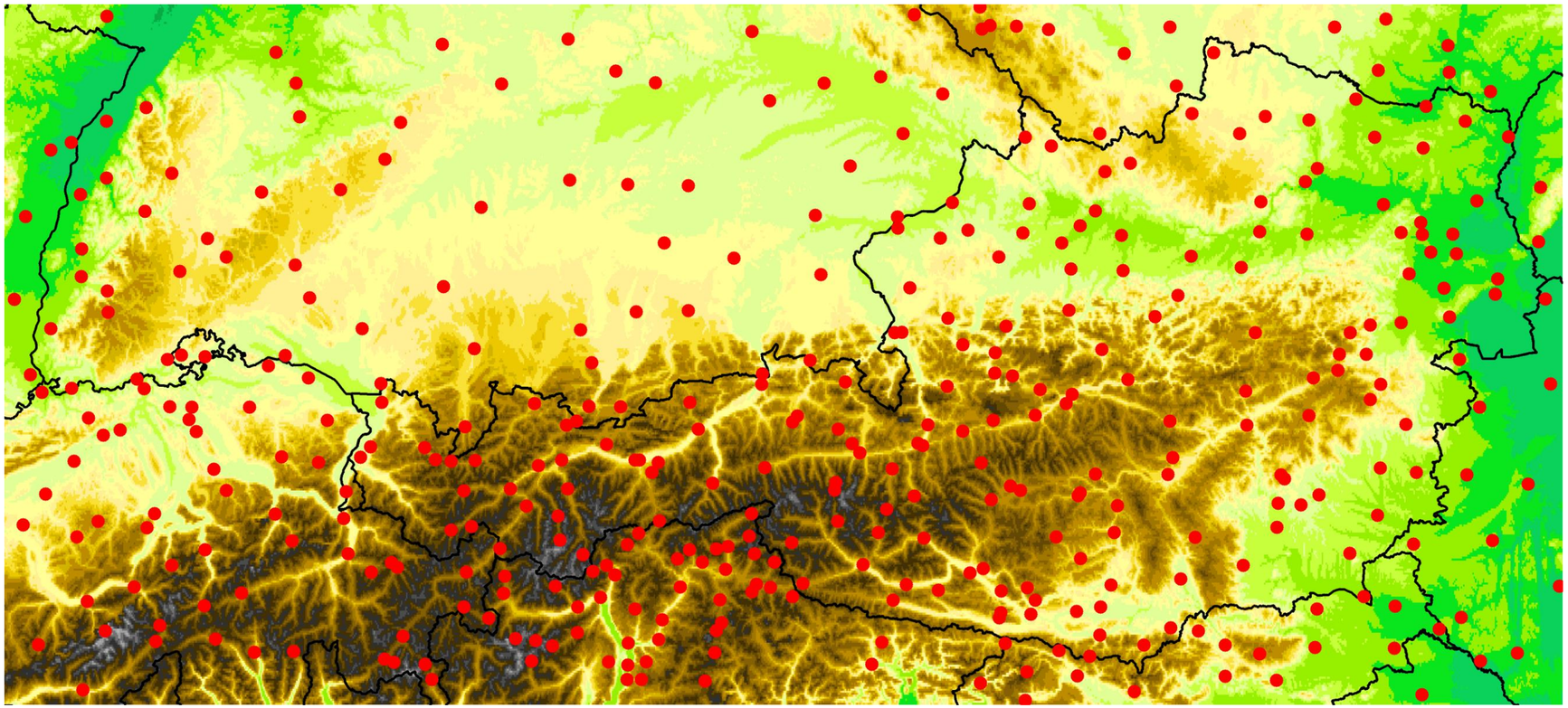
The boundaries of the subregion:

latitude: 46.3° – 49.1°

longitude: 7.4° – 17.2°

number of grids (resolution: 0.1°): 2871

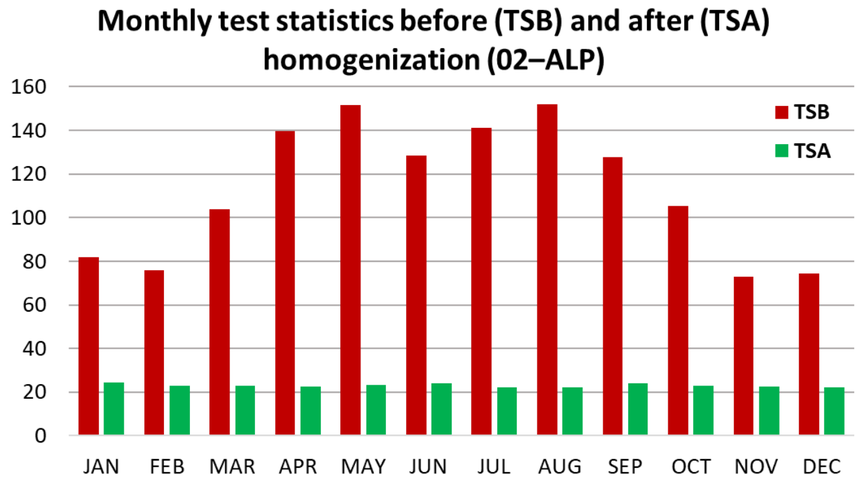
Number of stations: 399

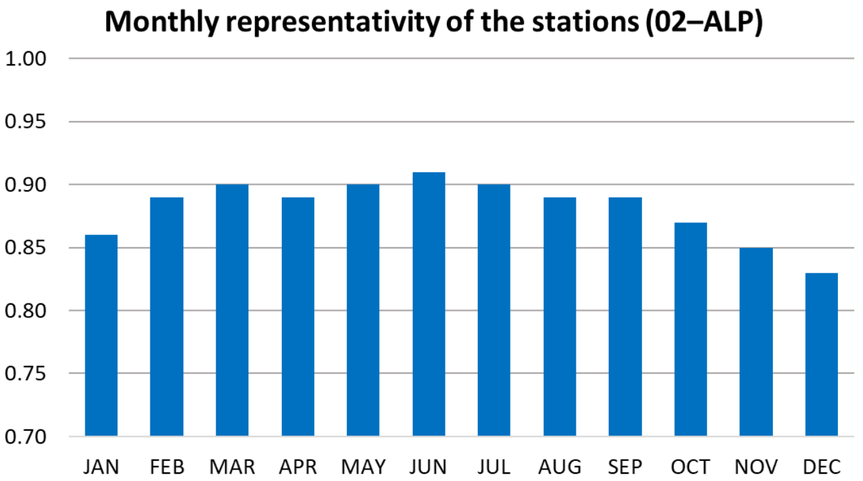


Verification statistics: 02–ALP (Austria & Switzerland)

The most important verification statistics at mean temperature:
We used the additive model with a 0.05 significance level,
where the critical value is: 20.91

	Winter	Spring	Summer	Autumn	Year
Test Statistics Before Homogenization	133.00	191.65	206.76	141.69	278.39
Test Statistics After Homogenization	24.05	22.42	23.07	22.29	25.42
Relative Modification of Series	0.17	0.19	0.21	0.21	0.24
Representativity of station network	0.88	0.90	0.91	0.87	0.91





Sub-region: 03–GER (Germany)

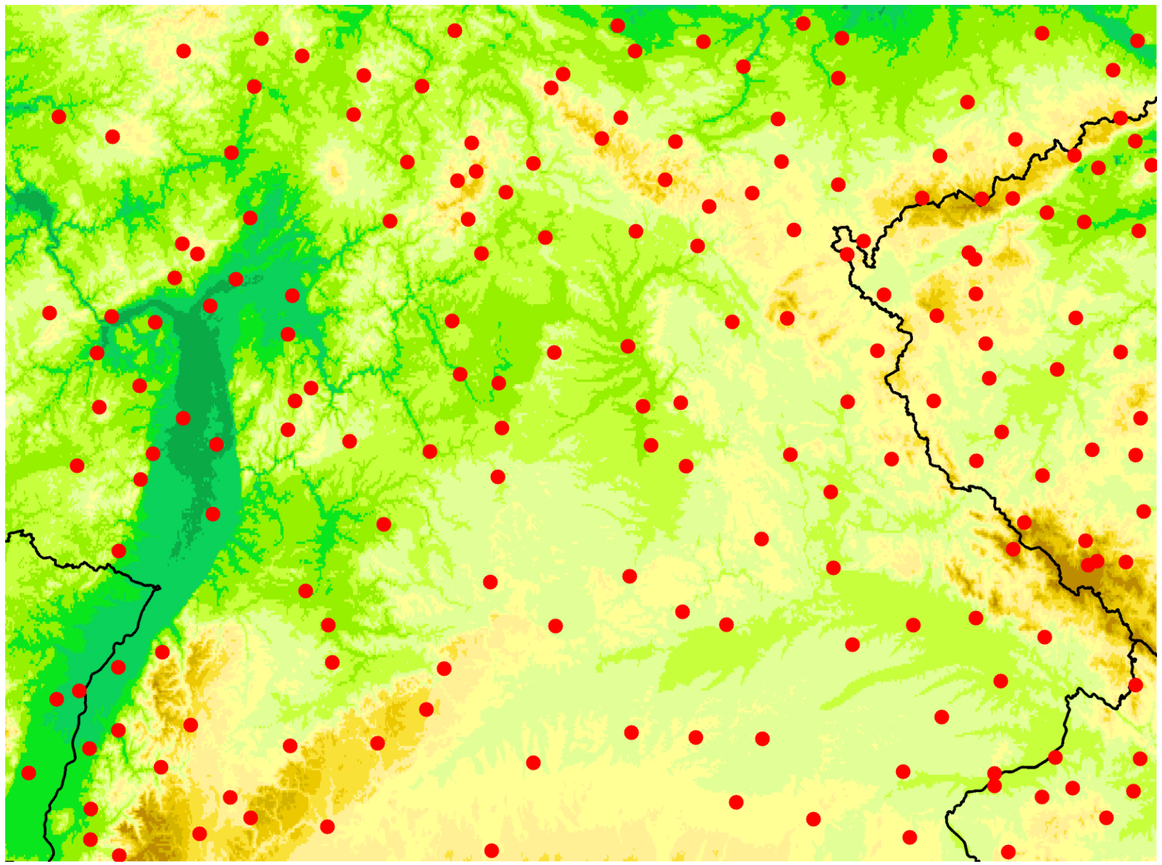
The boundaries of the subregion:

latitude: 48.0° – 51.1°

longitude: 7.4° – 13.9°

number of grids (resolution: 0.1°): 2112

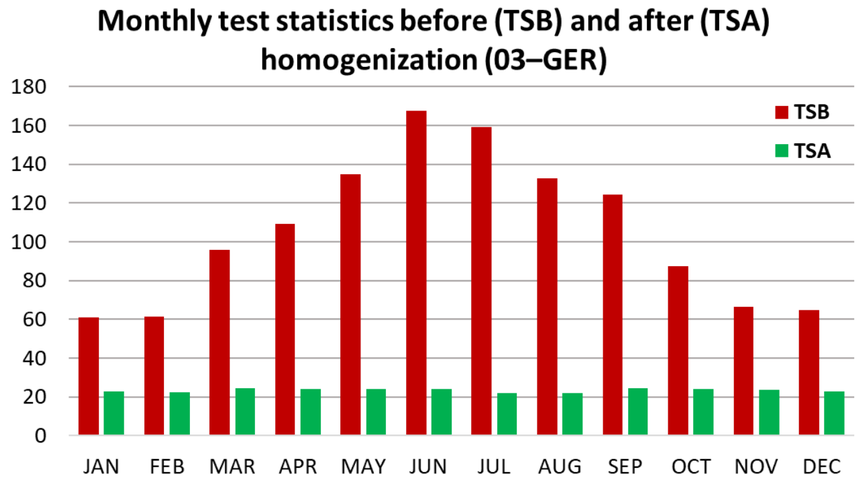
Number of stations: 185

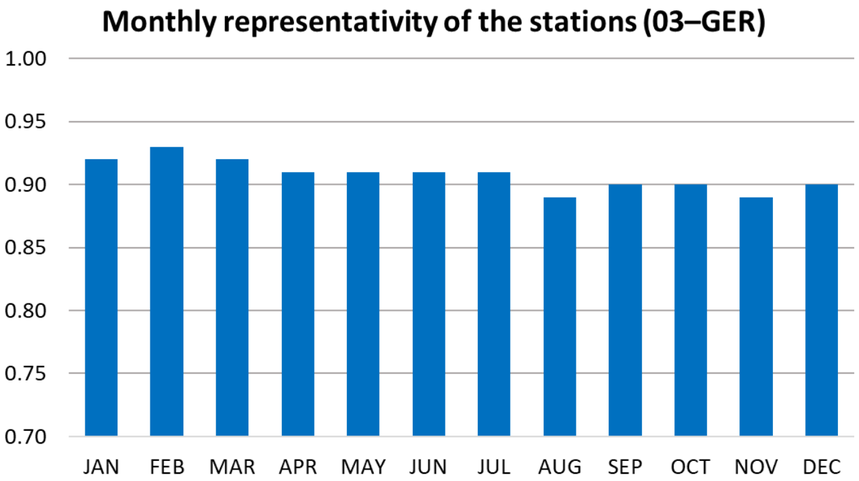


Verification statistics: 03–GER (Germany)

The most important verification statistics at mean temperature:
We used the additive model with a 0.05 significance level,
where the critical value is: 20.91

	Winter	Spring	Summer	Autumn	Year
Test Statistics Before Homogenization	110.01	164.48	246.80	143.28	330.87
Test Statistics After Homogenization	22.36	23.21	21.25	21.90	25.70
Relative Modification of Series	0.09	0.16	0.21	0.15	0.18
Representativity of station network	0.93	0.91	0.91	0.89	0.93





Sub-region: 04–CZE (Czech Republic)

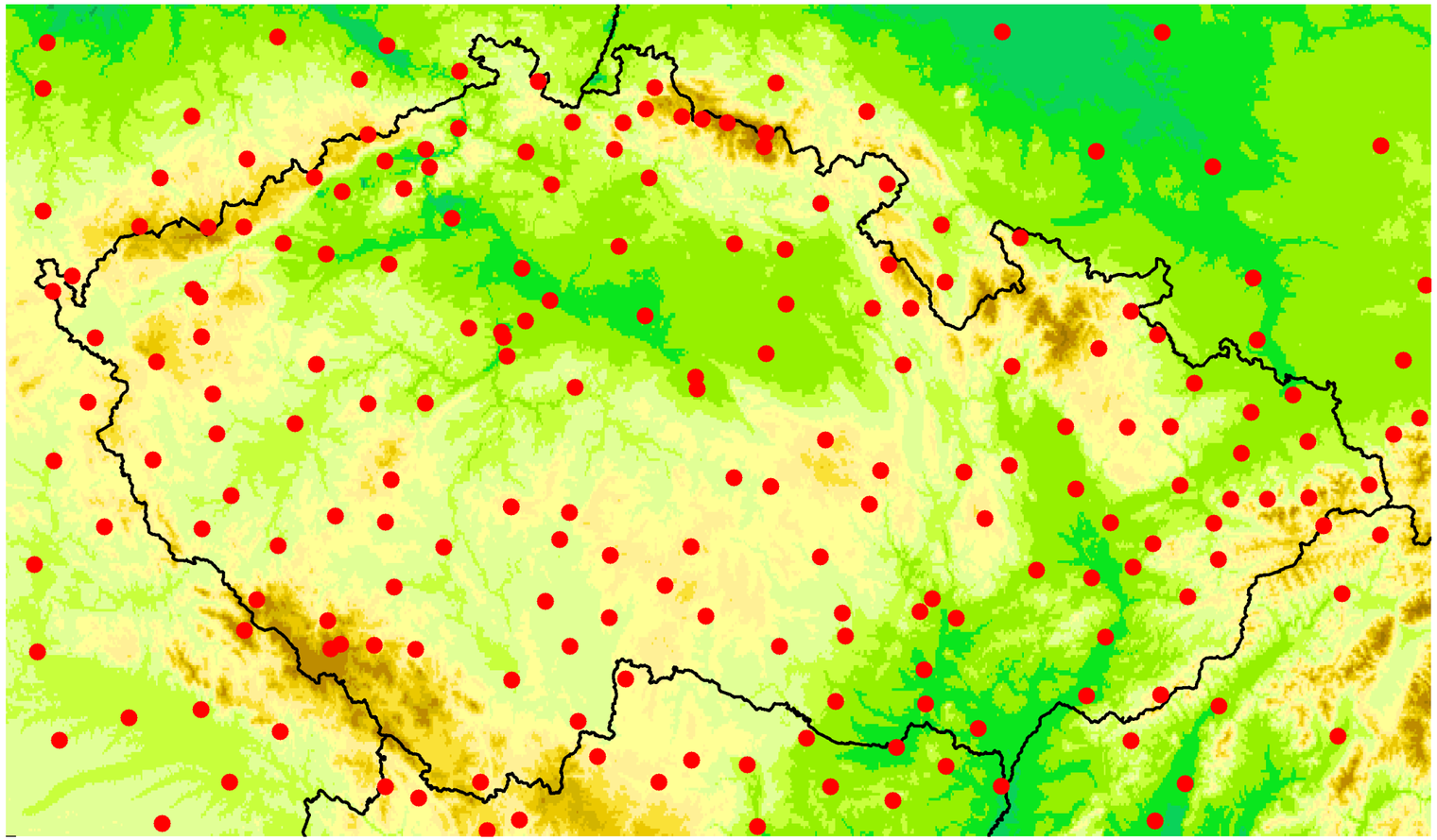
The boundaries of the subregion:

latitude: 48.5° – 51.1°

longitude: 12.0° – 19.0°

number of grids (resolution: 0.1°): 1917

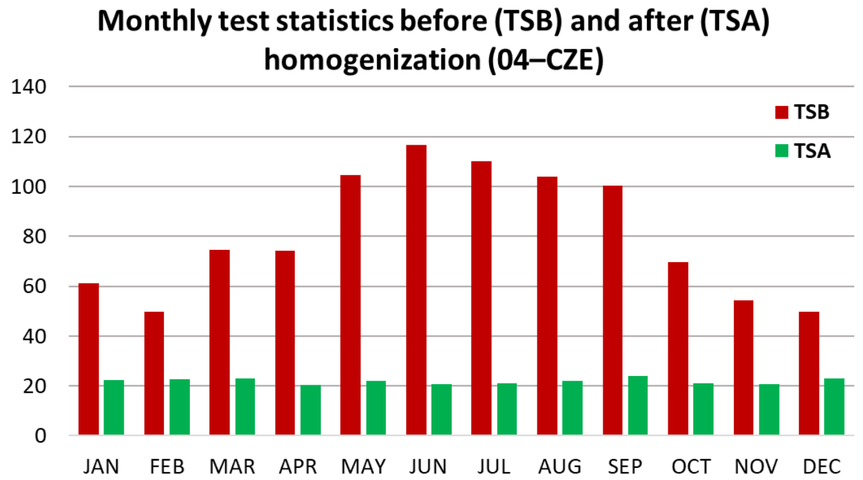
Number of stations: 202

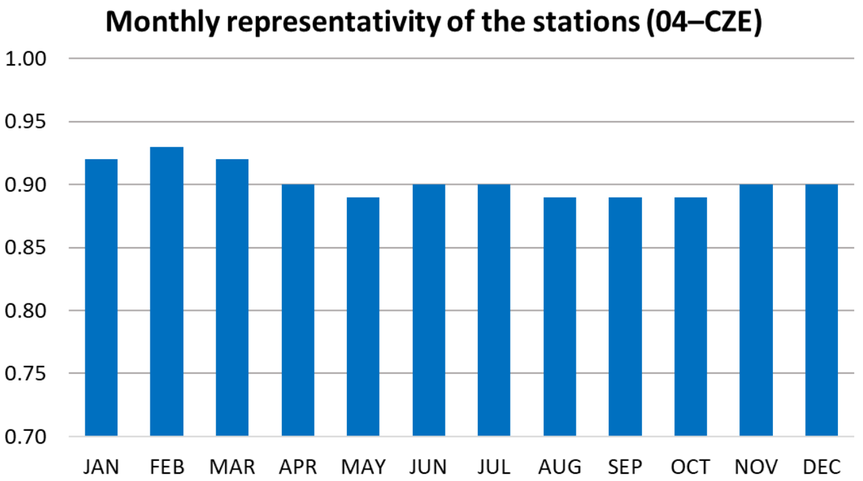


Verification statistics: 04–CZE (Czech Republic)

The most important verification statistics at mean temperature:
We used the additive model with a 0.05 significance level,
where the critical value is: 20.91

	Winter	Spring	Summer	Autumn	Year
Test Statistics Before Homogenization	88.57	122.46	162.47	107.92	228.14
Test Statistics After Homogenization	23.62	22.23	20.66	21.69	25.78
Relative Modification of Series	0.08	0.16	0.20	0.15	0.17
Representativity of station network	0.93	0.90	0.90	0.89	0.92





Sub-region: 05–POL (Poland)

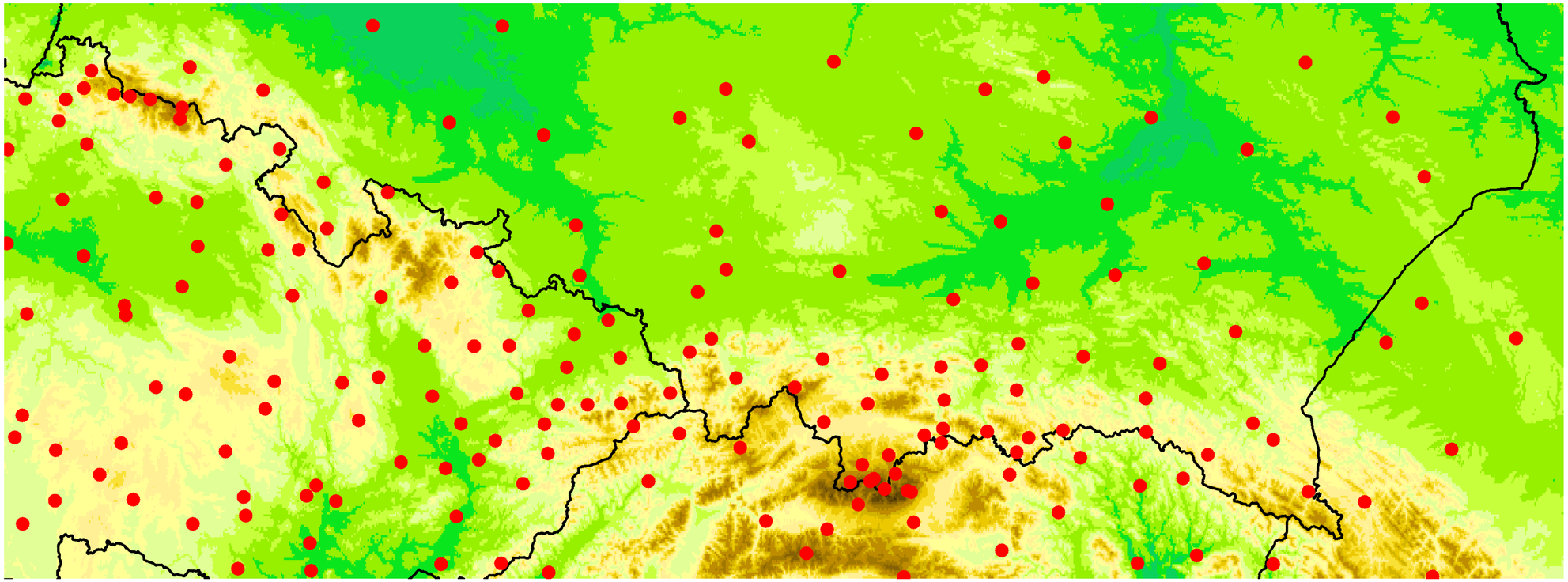
The boundaries of the subregion:

latitude: 48.9° – 51.1°

longitude: 14.7° – 24.2°

number of grids (resolution: 0.1°): 2208

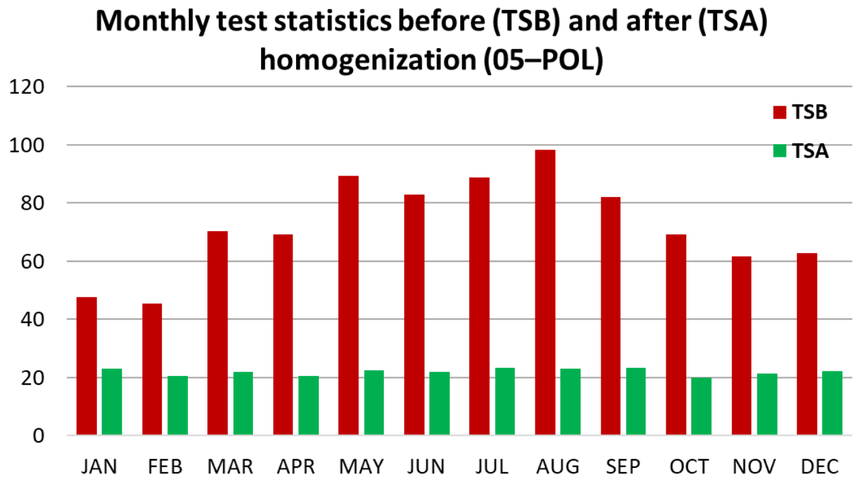
Number of stations: 182

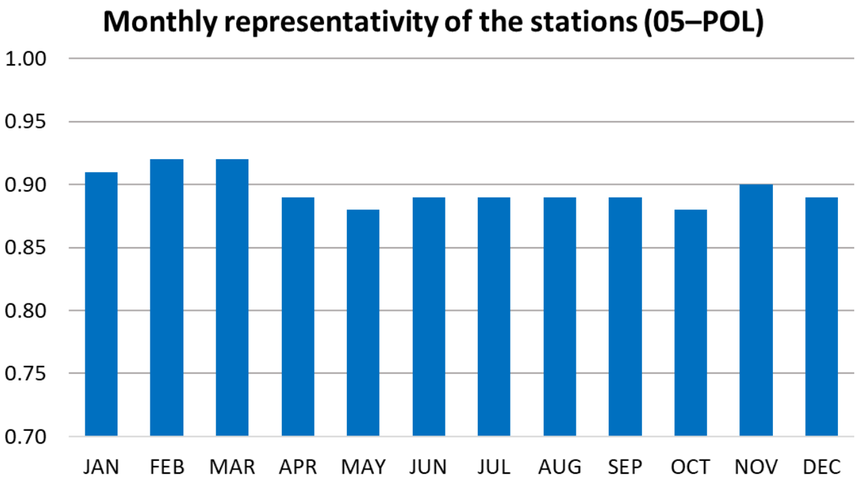


Verification statistics: 05–POL (Poland)

The most important verification statistics at mean temperature:
We used the additive model with a 0.05 significance level,
where the critical value is: 20.91

	Winter	Spring	Summer	Autumn	Year
Test Statistics Before Homogenization	85.72	106.38	125.62	104.71	200.43
Test Statistics After Homogenization	23.23	20.24	20.94	20.60	23.11
Relative Modification of Series	0.09	0.16	0.18	0.17	0.18
Representativity of station network	0.92	0.90	0.90	0.88	0.92





Errors

The data series may contain different types of errors.

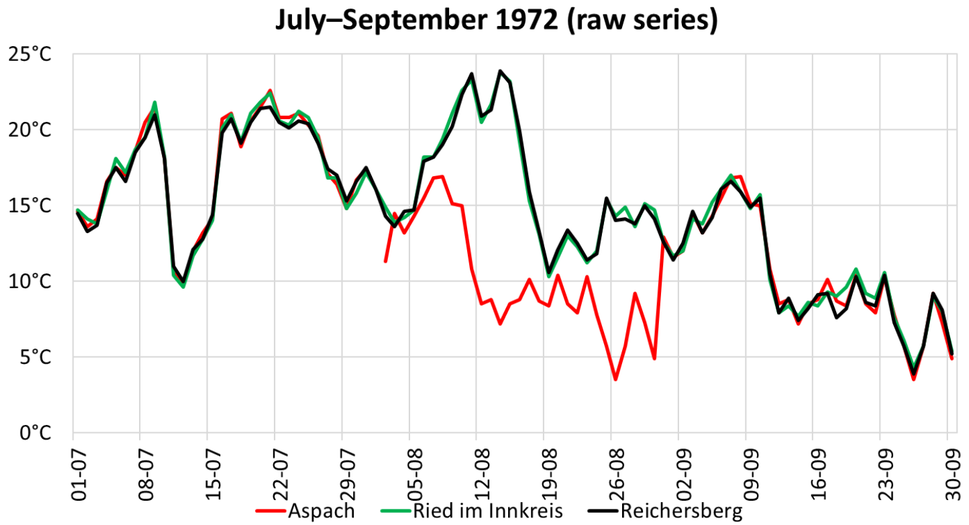
The homogenization of daily data also includes automatic quality control (QC) procedure.

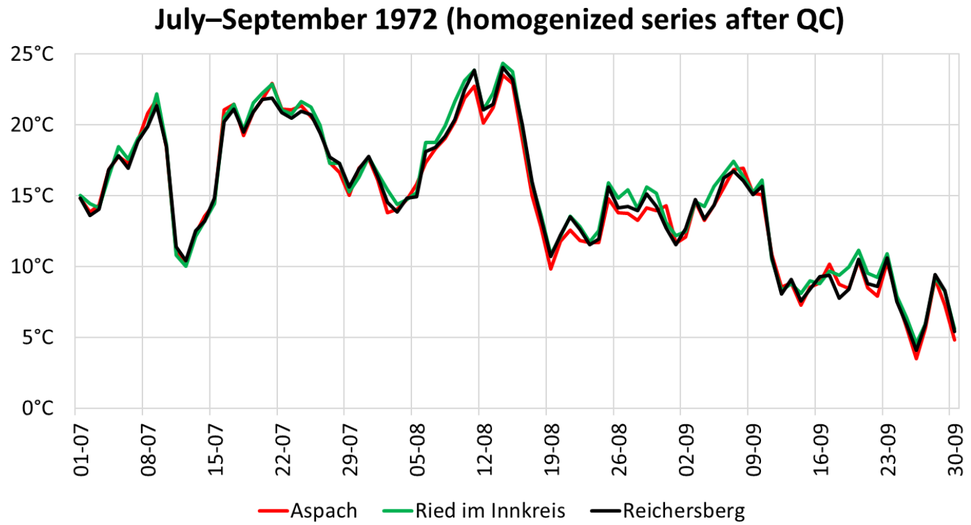
August 1972 – Aspach (AUT)

Repeated data: the September's data are also in August.

MASH detected the errors and corrected the values very well.

			161	162
1972	8	2	0	-0.99
1972	8	4	0	0
1972	8	6	0	-0.34
1972	8	8	0	-0.69
1972	8	9	0	-3.64
1972	8	10	0	-5.41
1972	8	11	0	-10.48
1972	8	12	0	-10.2
1972	8	13	0	-10.98
1972	8	14	0	-14.89
1972	8	15	0	-13.01
1972	8	16	0	-8.78
1972	8	17	0	-3.51
1972	8	18	0	-2.67
1972	8	19	0	0
1972	8	20	0	0
1972	8	21	0	-2.65
1972	8	22	0	-2.51
1972	8	23	0	0
1972	8	24	0	-2.48
1972	8	25	0	-7.66
1972	8	26	0	-8.89
1972	8	27	0	-6.66
1972	8	28	0	-2.65
1972	8	29	0	-5.41
1972	8	30	0	-7.65





Thank you for your attention!

Interreg
Danube Region

Co-funded by
the European Union

Danube-ADAPT