HOMOGENIZATION OF ECA&D TEMPERATURE DATA-SET

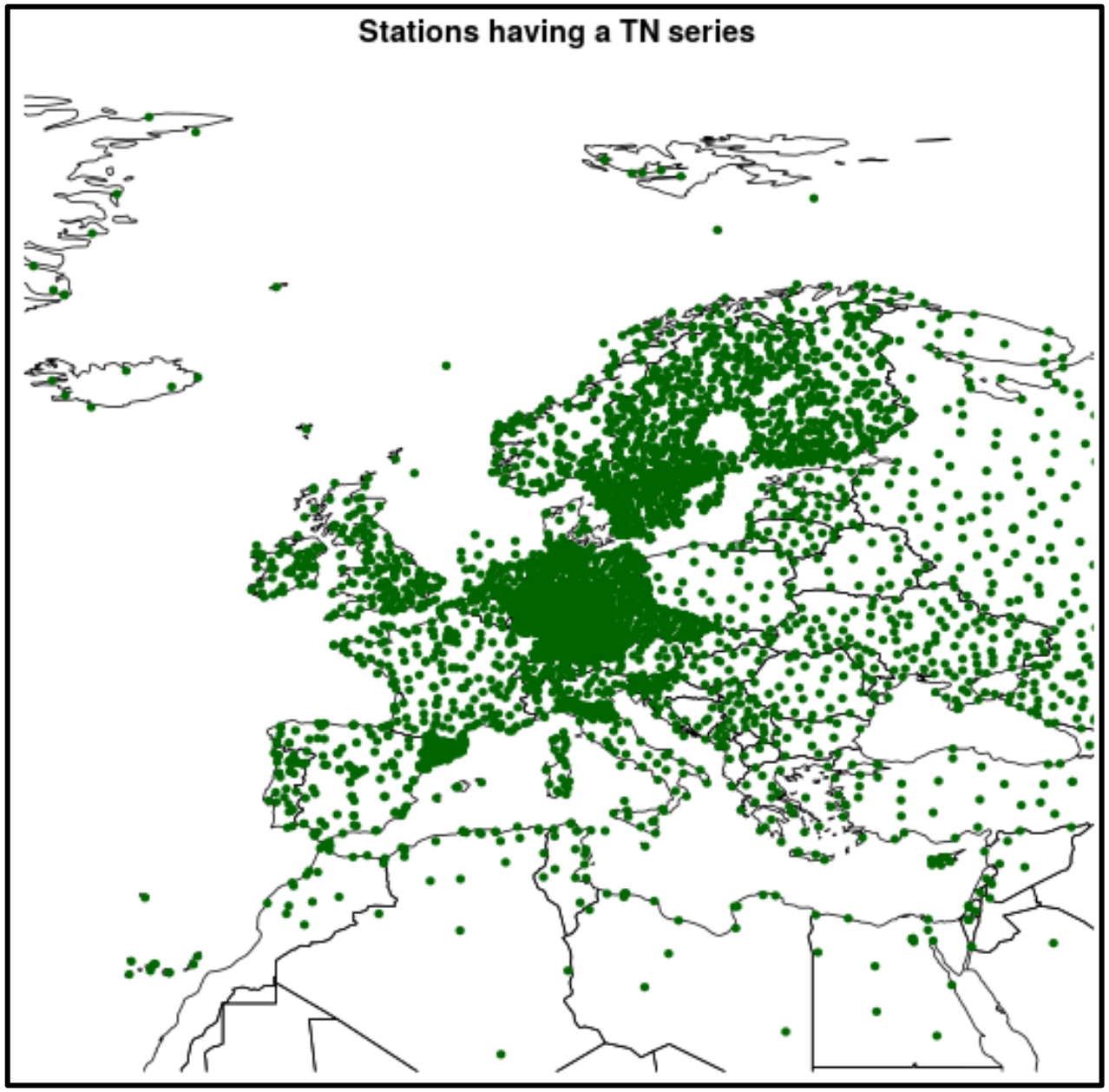
Antonello Squintu[1], Yuri Brugnara[2], Gerard van der Schrier[1], Else van den Besselaar[1]

[1]KNMI, [2]University of Bern

EUSTACE has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no 640171

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

ECA&D DATASET



Collection of ground- based observation all over Europe

Temperature, radiation, precipitation, humidity, pressure, etc. etc.

http://www.ecad.eu/

E-OBS gridded data-set (newly released vers. 14)

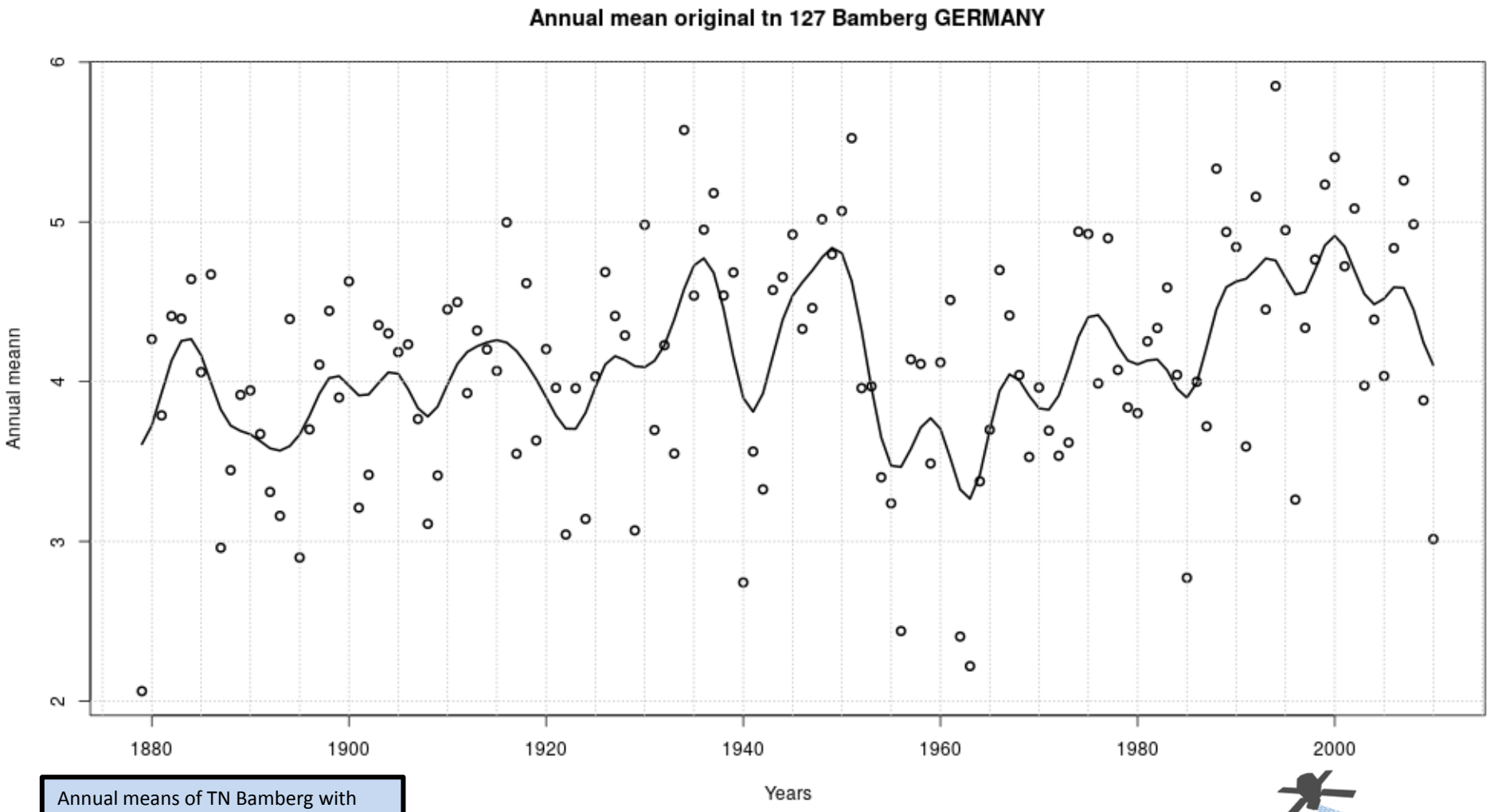
Number of stations and series constantly increases thanks to new participants and updates

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

INHOMOGENEOUS SERIES

Undocumented change points affect calculation of trend indices and gridded data



Annual means of TN Bamberg with running mean.

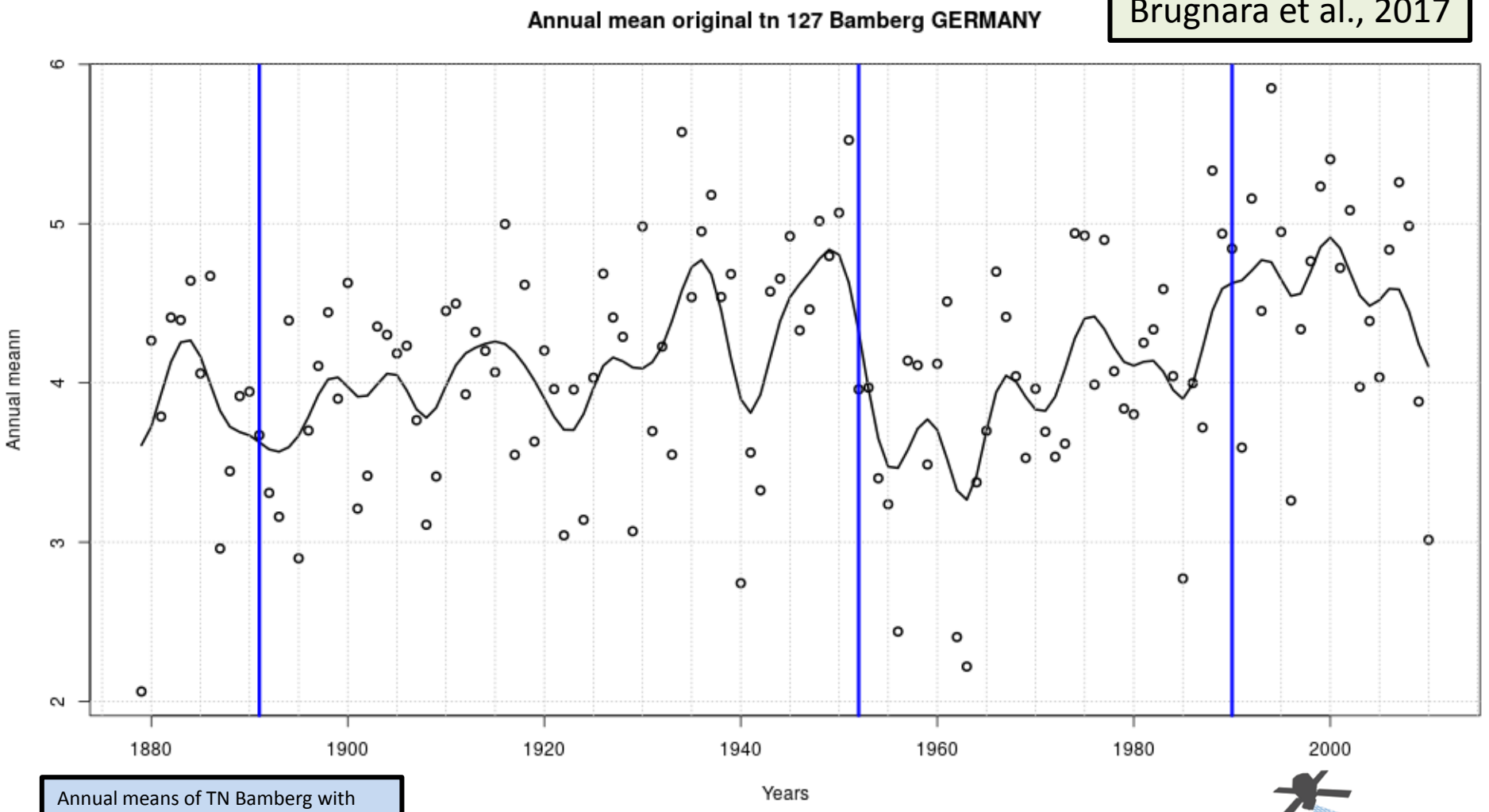
EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

BREAK DETECTION (BD)

Break detection as agreement between rH-test, CAUME, GAHMDI on a monthly base

Brugnara et al., 2017

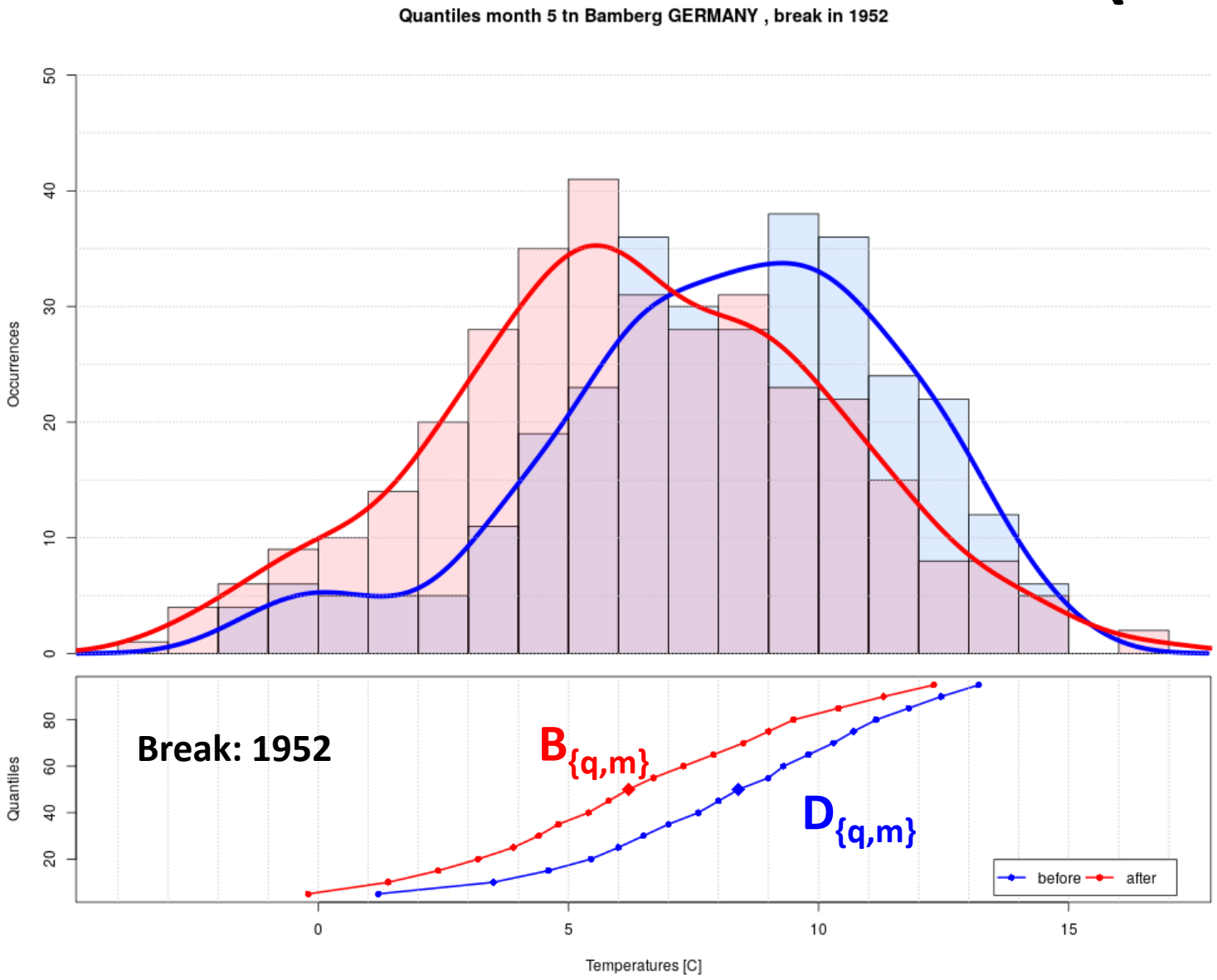


Annual means of TN Bamberg with running mean and change points identified during break detection.

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

QUANTILE MATCHING (QM)



Temperature pdf and quantiles of TN Bamberg (May) 10 years before (blue) and after (red) 1952

Statistical differences in the temperature pdf before and after the break.

If overlap period is present:

$A_{\{q,m\}} = B_{\{q,m\}} - D_{\{q,m\}}$

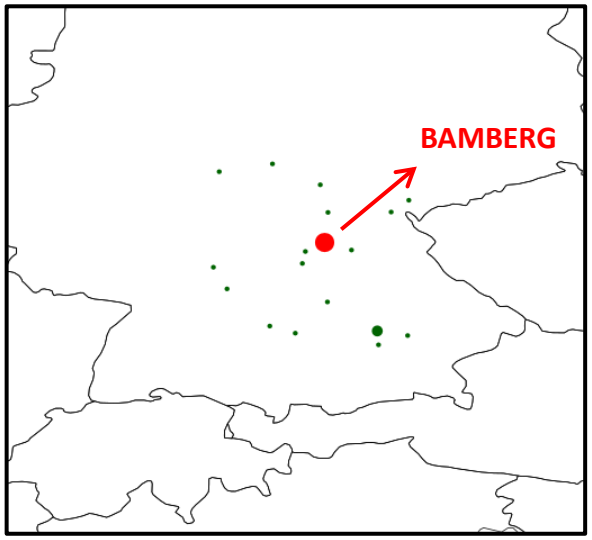
If not, it's necessary to use homogeneous reference series

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

REFERENCE SELECTION

Parallel measurements would make the quantile matching extremely efficient since pdf of the same location should be totally consistent.



In most cases there are no parallel measurements available and a set of reference series is needed.

References are useful to understand how much of the break-jump is due to **artificial intervention** and how much is due to an actual **climatic signal**.

Reference series are selected from the dataset in a 3°x3° box centred on the target series, with no more than 500 m of elevation difference (height/2 for mountain stations).

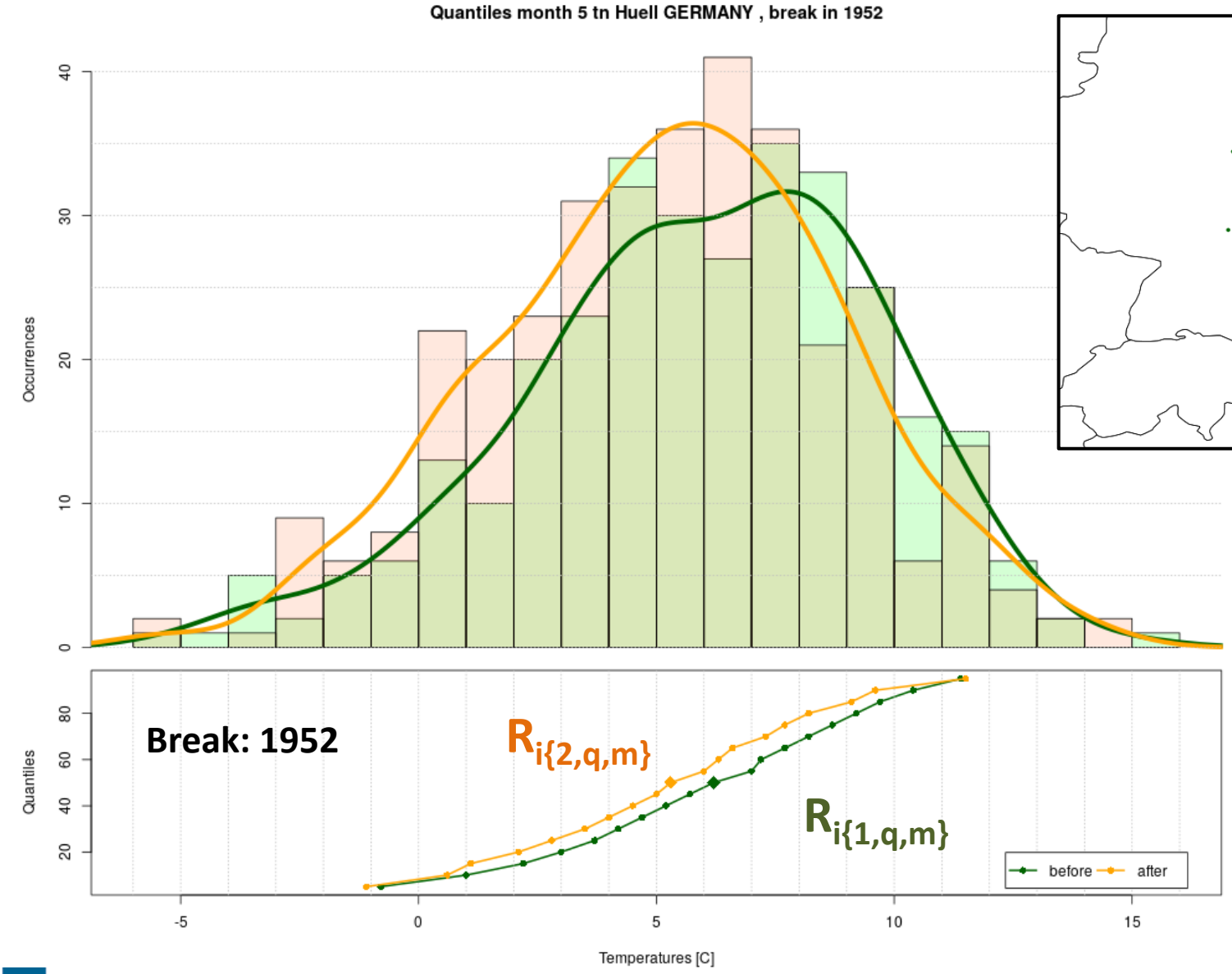
Break detection results are used to split the reference series into homogeneous sub-series since they **only** have to "carry" climatic signal.

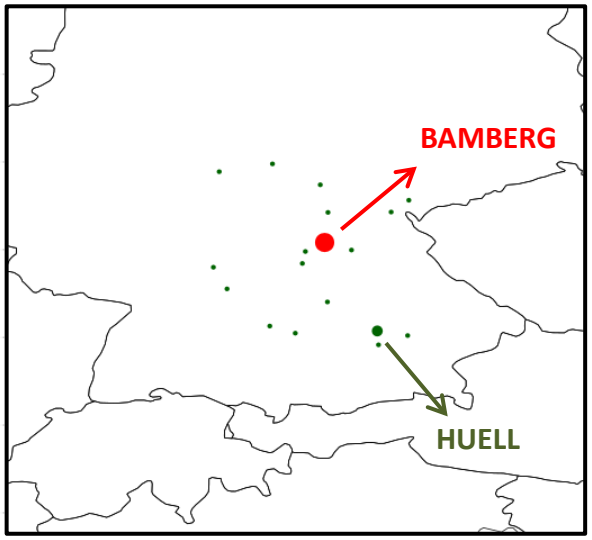
The ref. sub-series must have at least five years of overlap with the target series before the break and at least five years after the break.

Quantile matching is performed if at least three homogeneous reference series with long enough overlapping periods are present.

EUSTACE

HOMOGENEOUS REFERENCE (SUB-)SERIES





For each ref. "*i*":

$$A_{\{i,q,m\}} = =(B_{\{q,m\}}-R_{i\{1,q,m\}})+ -(D_{\{q,m\}}-R_{i\{2,q,m\}})$$

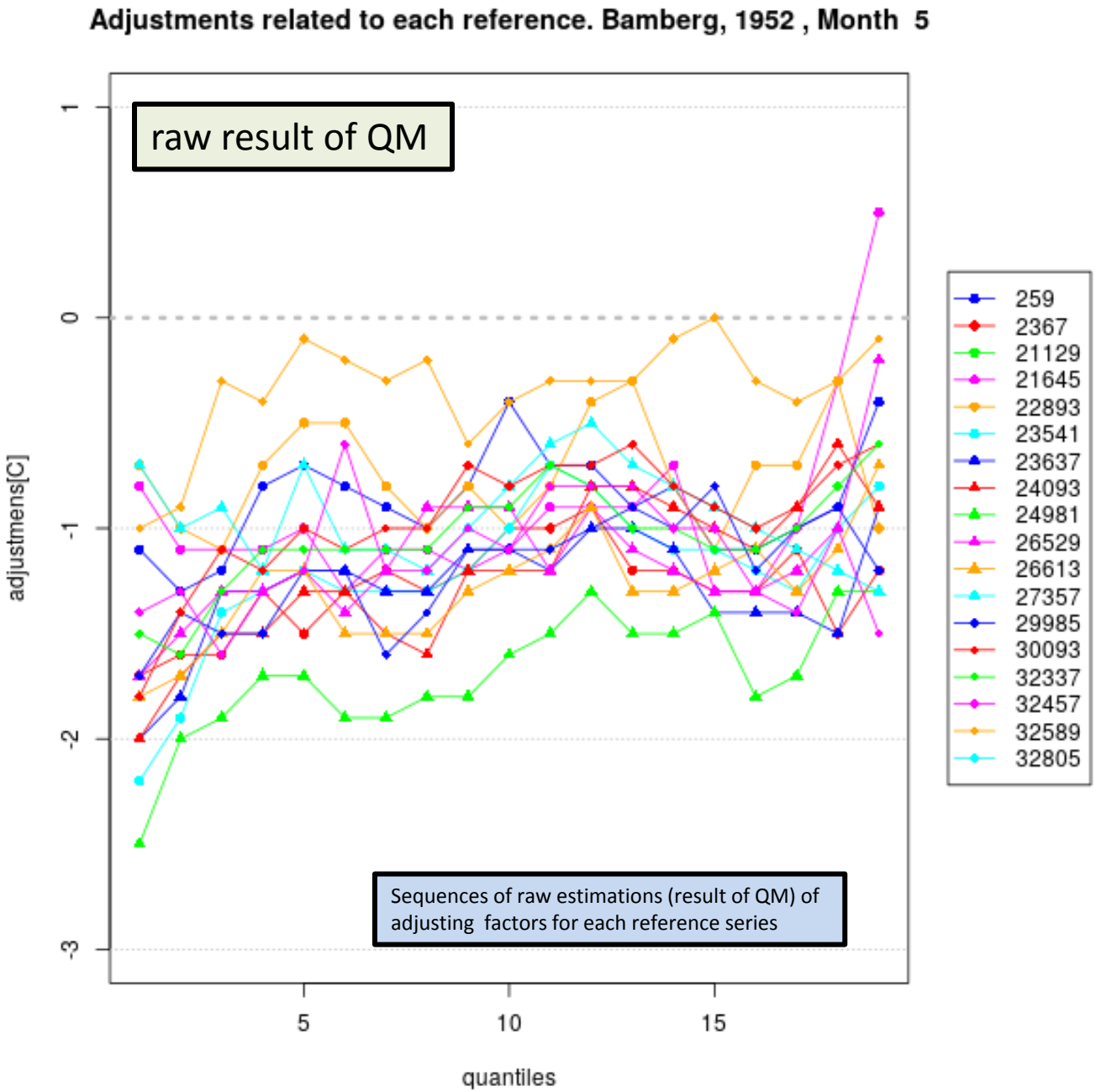
Temperature pdf and quantiles of TN Bamberg (May) 10 years before (blue) and after (red) 1952

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

THE ADJUSTMENTS

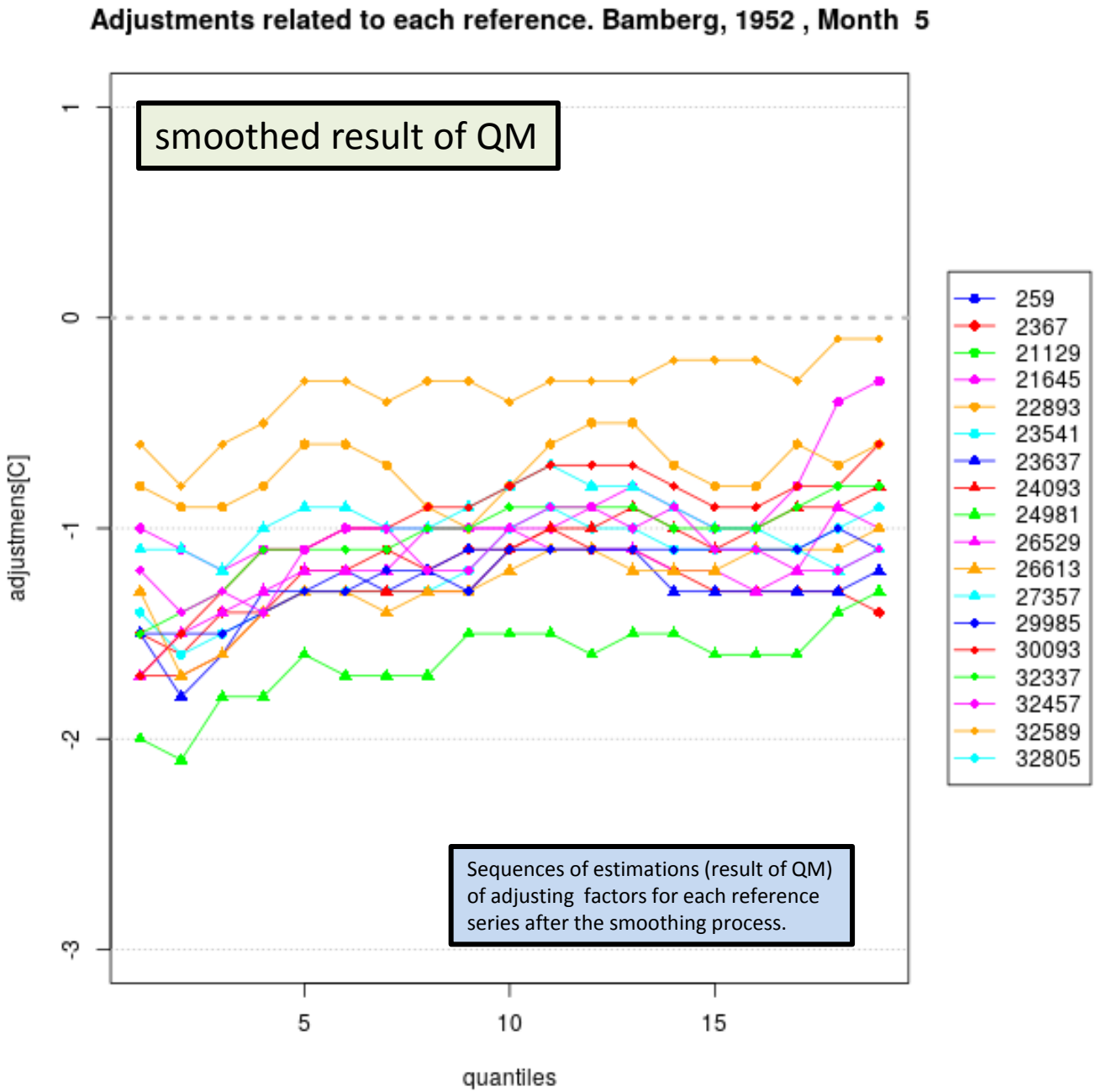


Estimations related to each reference are very noisy and need to be smoothed looking at closest quantiles and adjacent months

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

THE ADJUSTMENTS



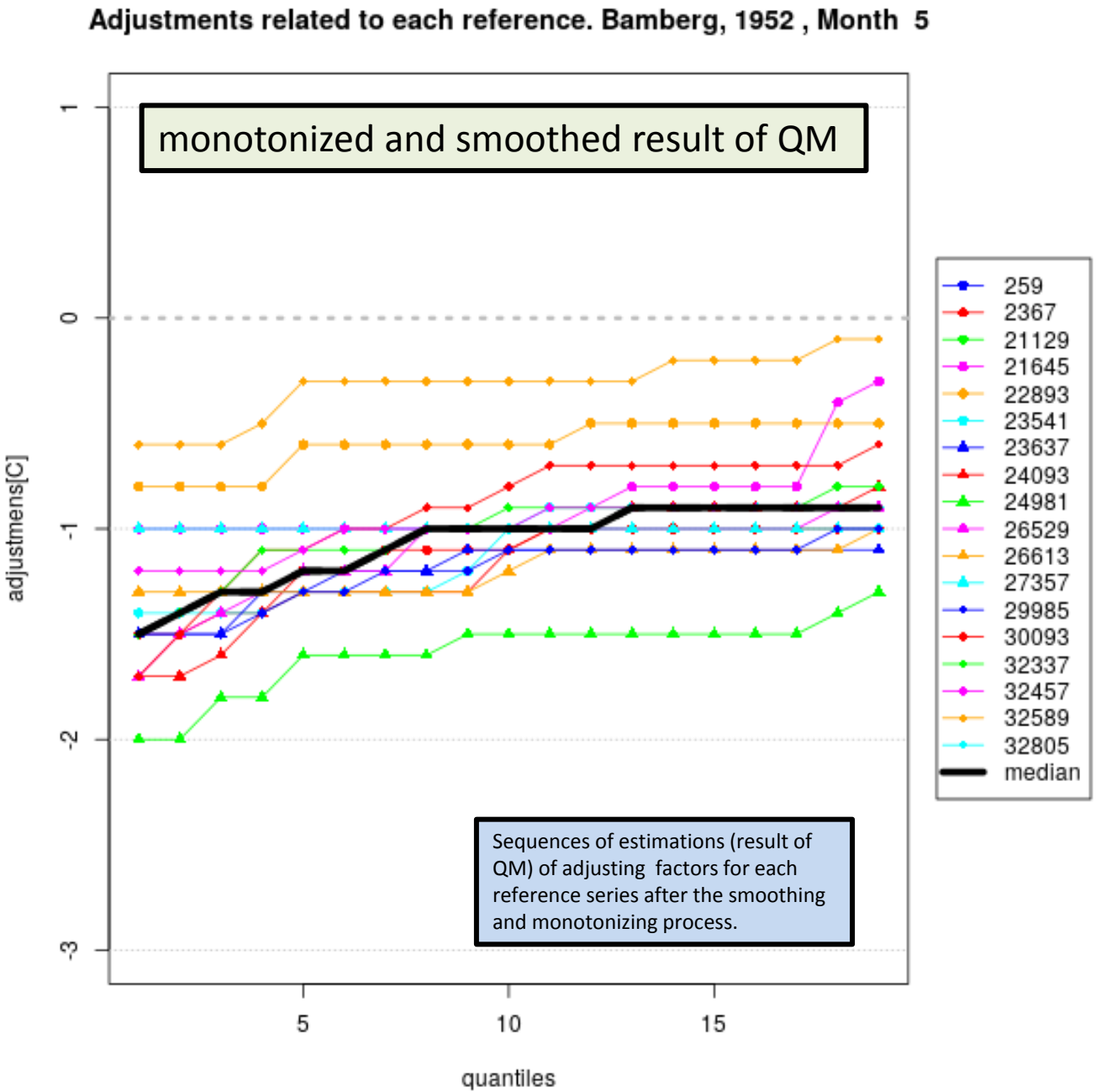
Estimations related to each reference are smoothed looking at close months and quantiles

Each sequence of adjustments has to be monotonous so the data don't change quantile after the correction process.

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

THE ADJUSTMENTS



Estimations related to each reference are smoothed looking at close months and quantiles

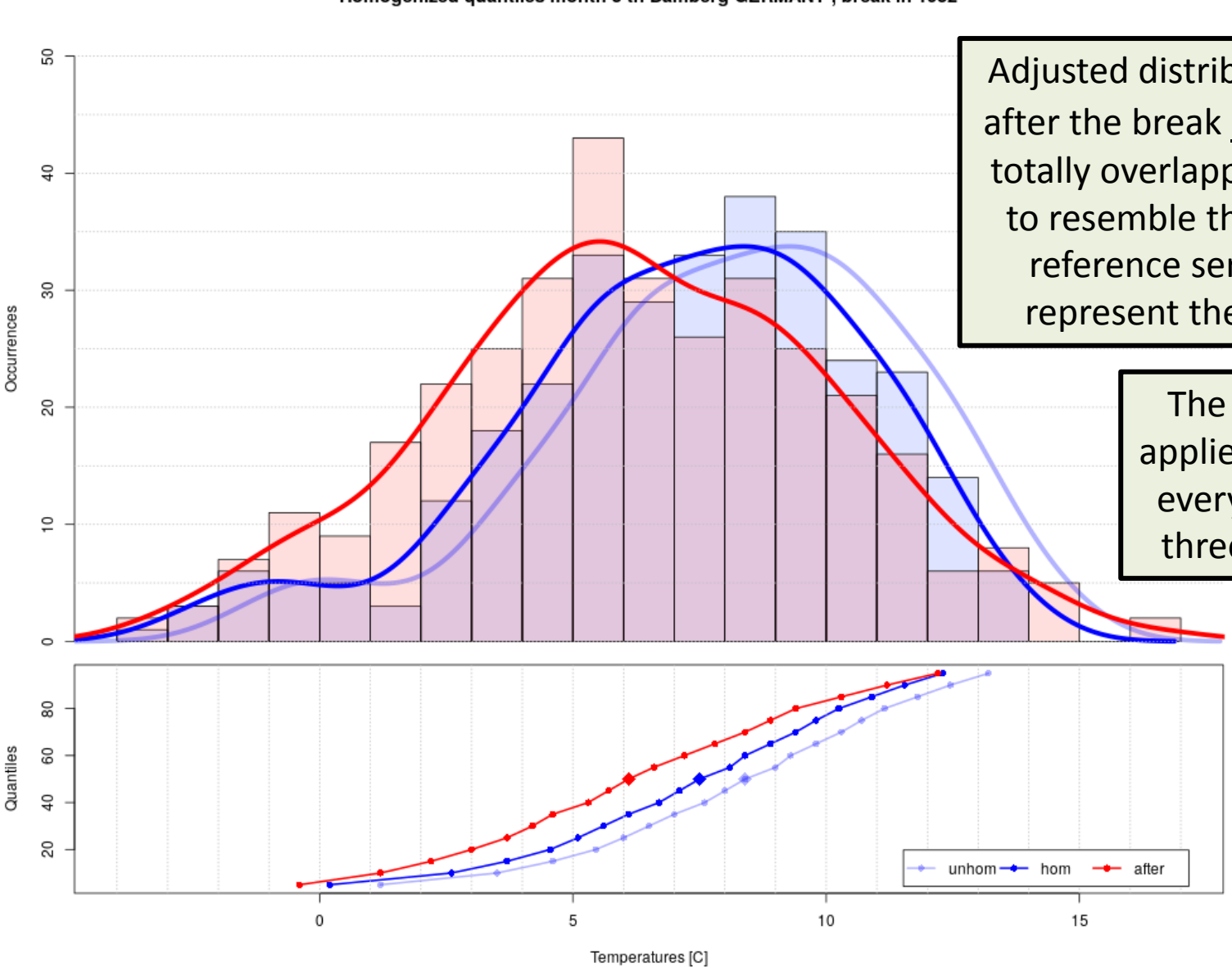
Each sequence of adjustments is then monotonized, moving towards the zero the intervals having different sign of the derivative.

For each month and for each quantile, the median of the extimations related to every references series is calculated.

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

ADJUSTED DISTRIBUTIONS



Adjusted distributions before and after the break **DON'T** have to be totally overlapped, but they have to resemble the features of the reference series, since these represent the climatic signal.

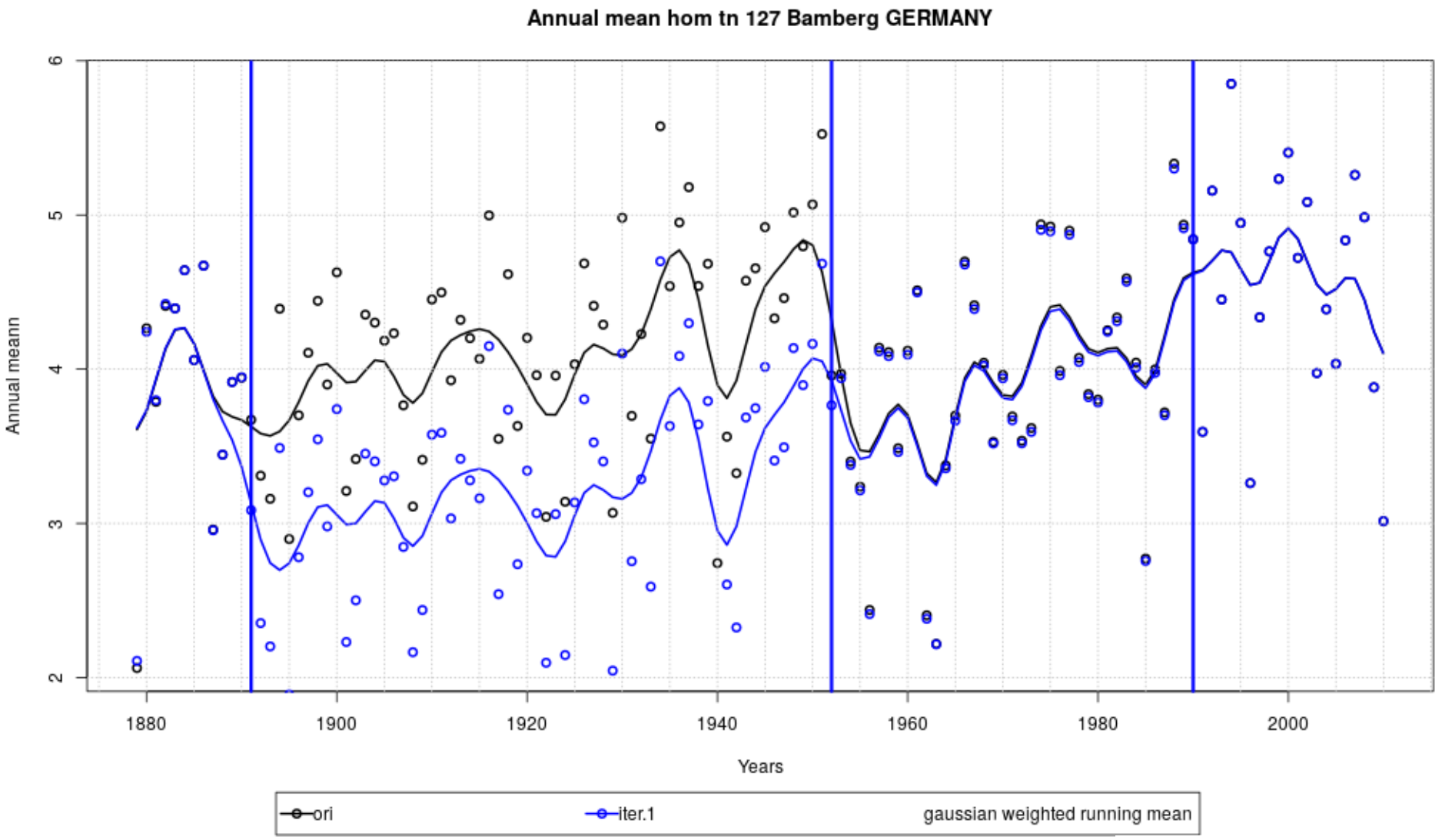
The QM procedure is applied to all months for every breaks if at least three ref. are present.

Hom. temp. pdf and quantiles of TN Bamberg (May) 10 years before (blue) and after (red) 1952. Light blue: unhom. version

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

ADJUSTED SERIES



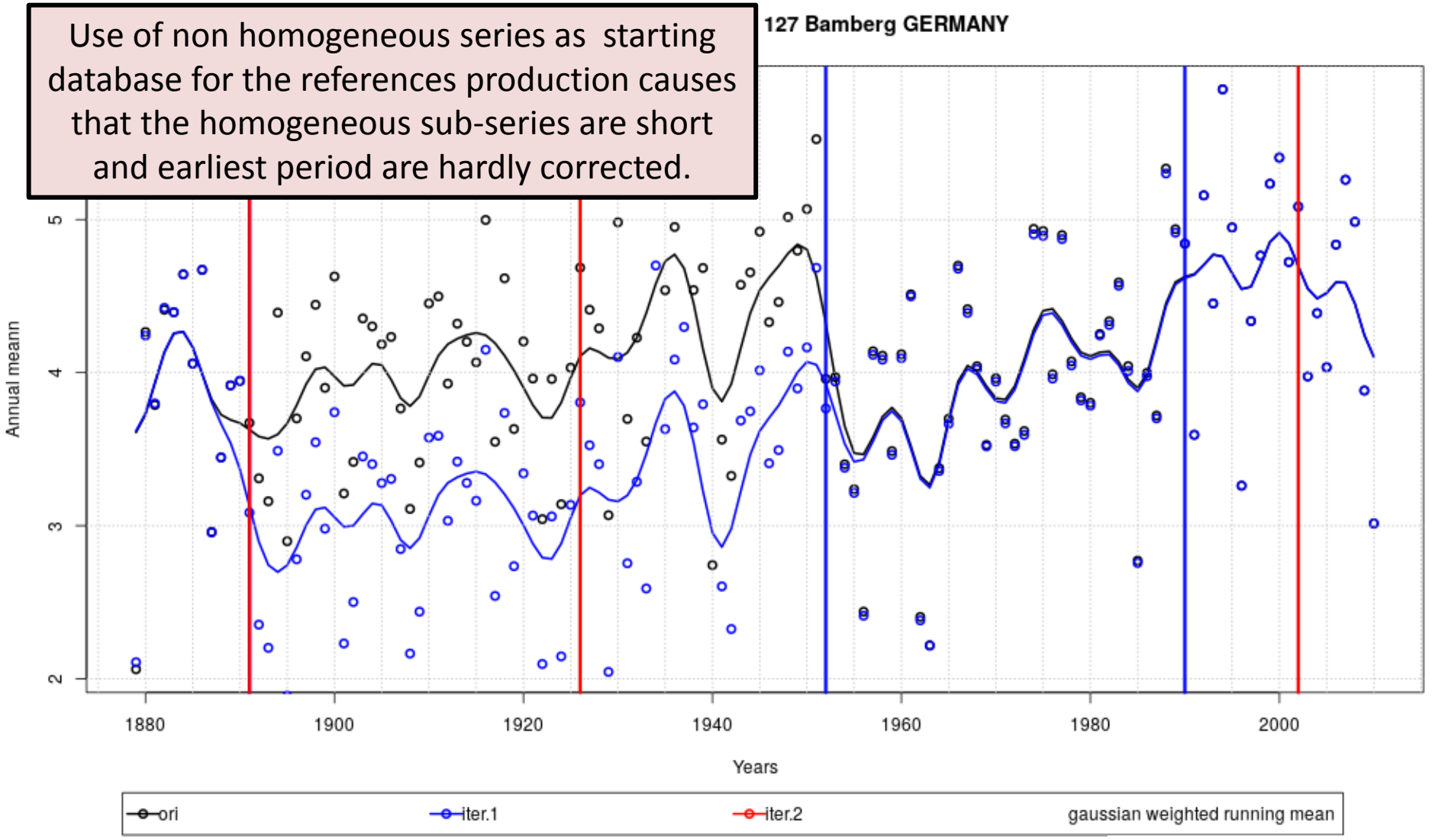
Annual means of TN Bamberg before and after the homogenization with running mean and change points identified during break detection.

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

SECOND ITERATION?

Use of non homogeneous series as starting database for the references production causes that the homogeneous sub-series are short and earliest period are hardly corrected.



Annual means of TN Bamberg before and after the homogenization with running mean and change points identified during the 2 iterations of break detection.

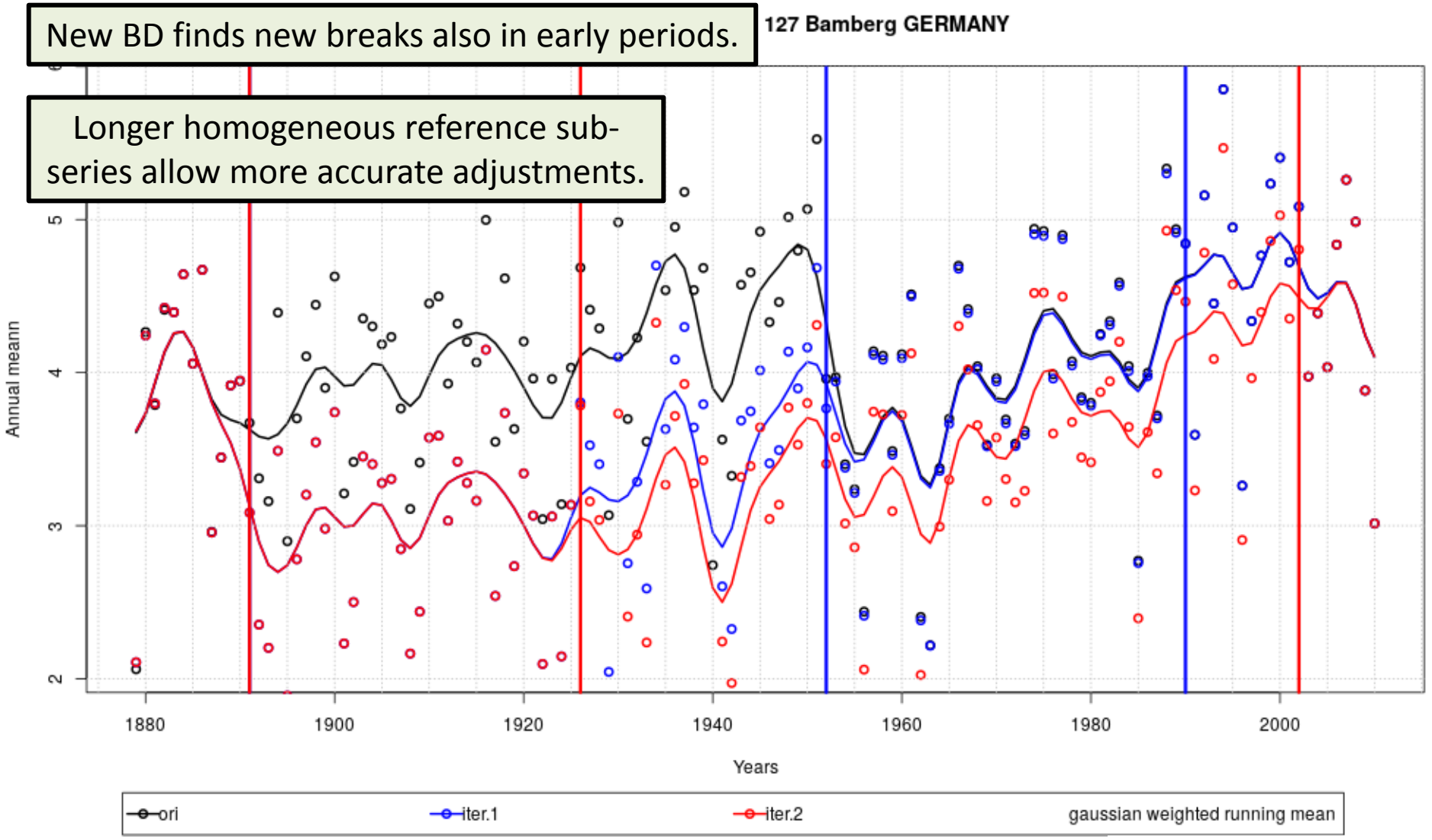
EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

SECOND ITERATION!

New BD finds new breaks also in early periods.

Longer homogeneous reference sub-series allow more accurate adjustments.

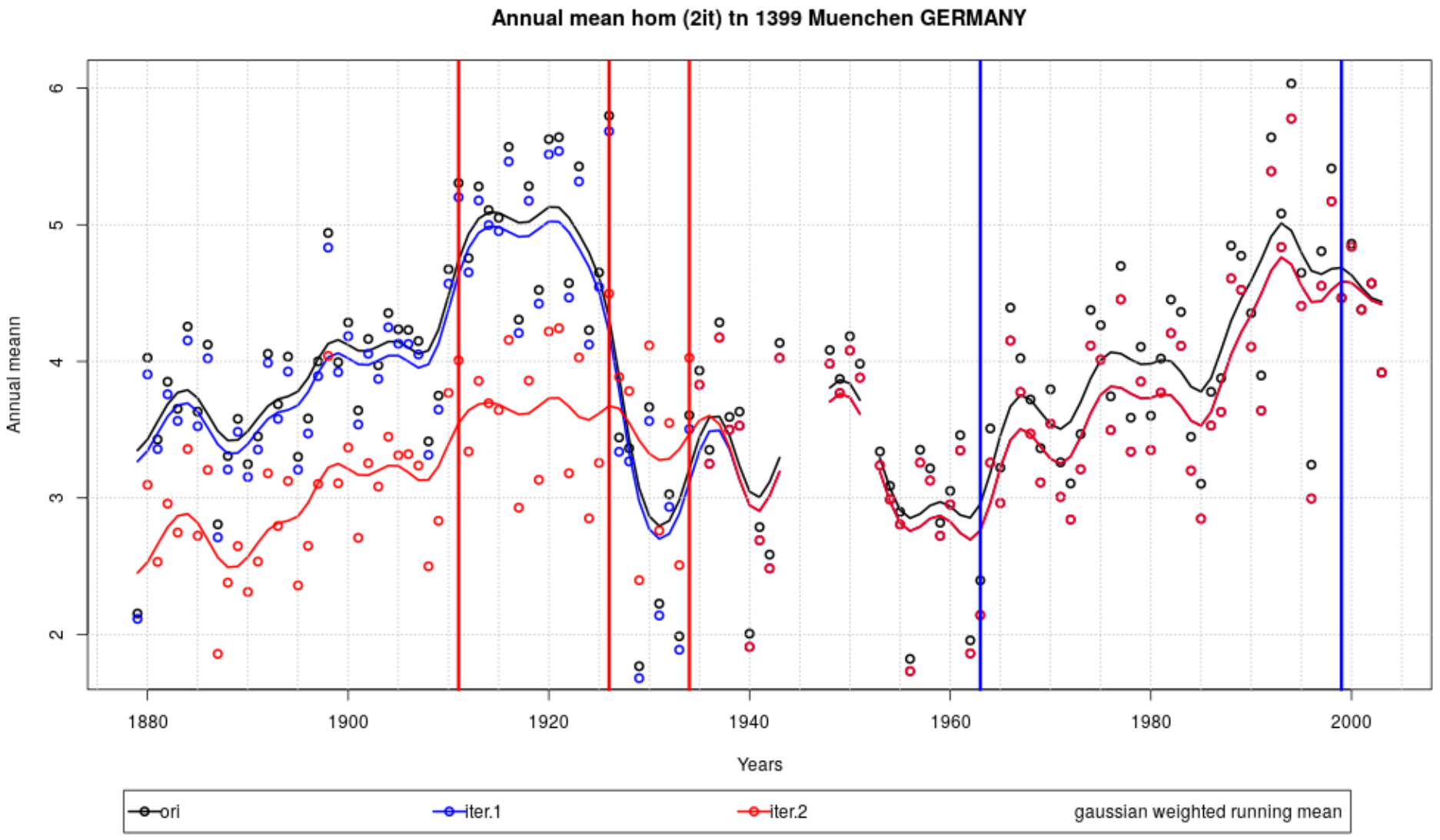


Annual means of TN through the 2 phases of homogenization with running mean and change points identified during the 2 iterations of break detection.

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

SECOND ITERATION



Annual means of TN through the 2 phases of homogenization with running mean and change points identified during the 2 iterations of break detection.

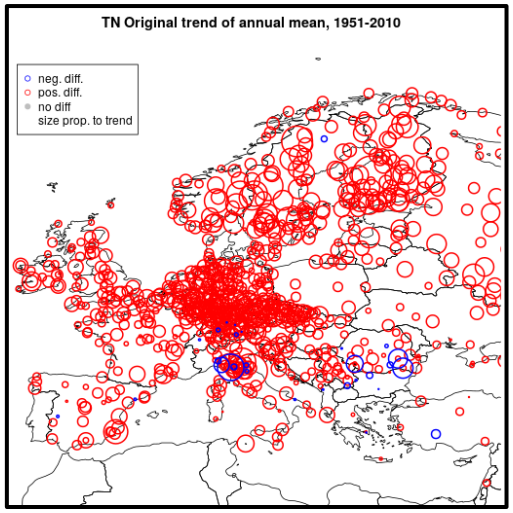
EUSTACE

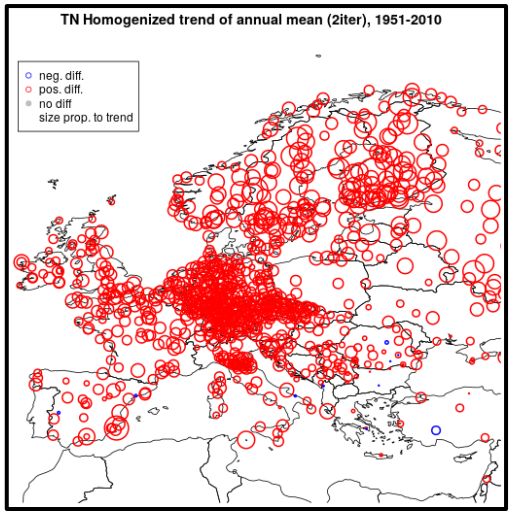
Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

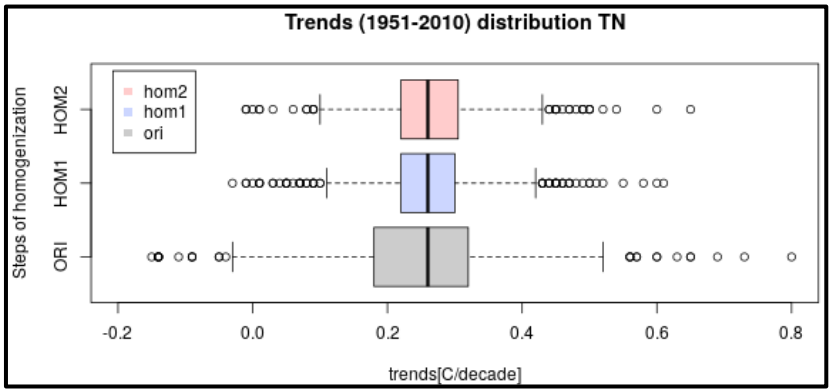
EFFECTS ON TRENDS

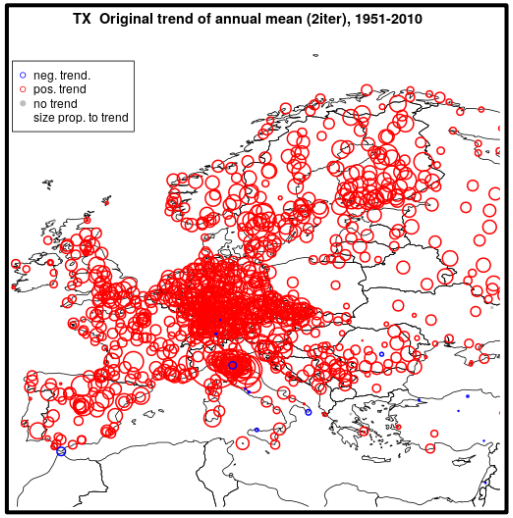
BEFORE HOM

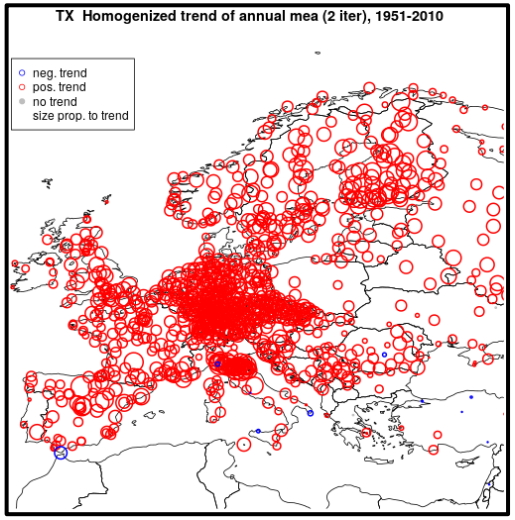
AFTER HOM

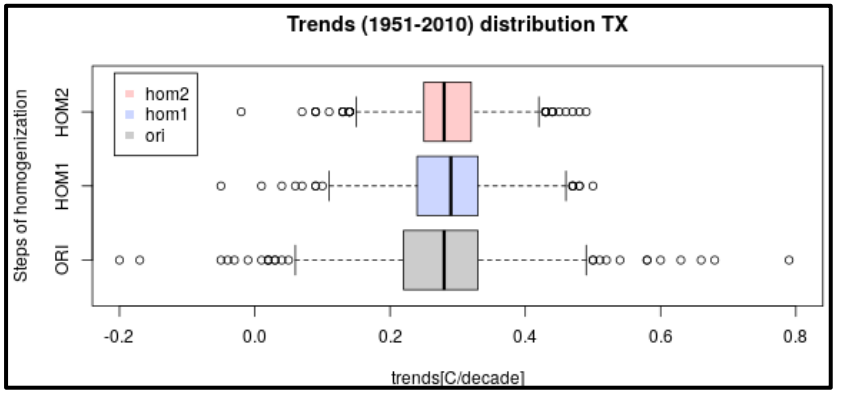










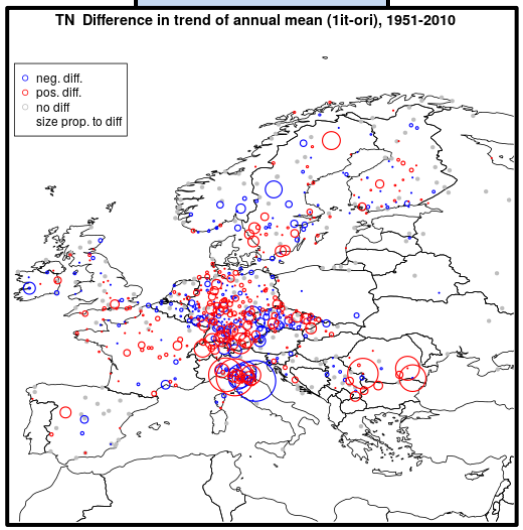


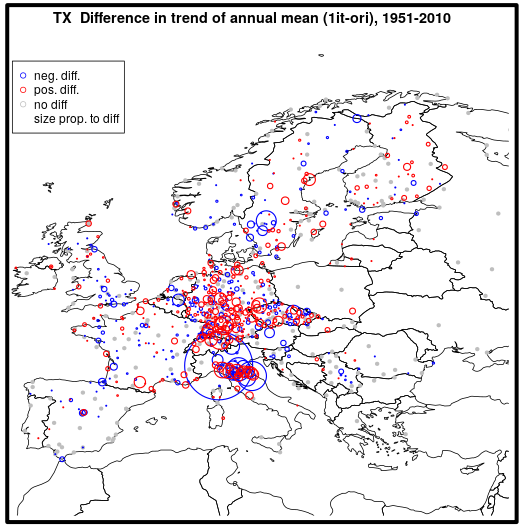
EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

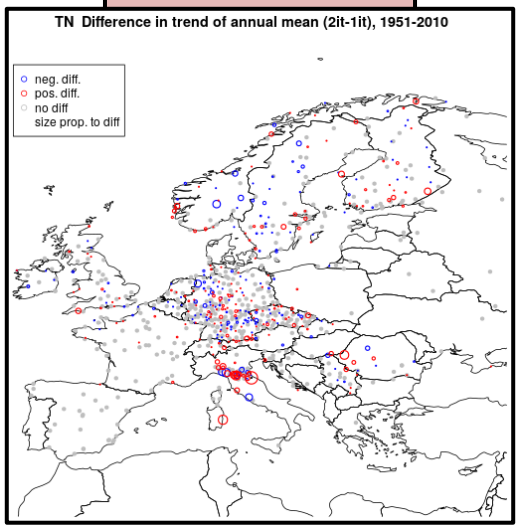
DIFFERENCES IN TRENDS

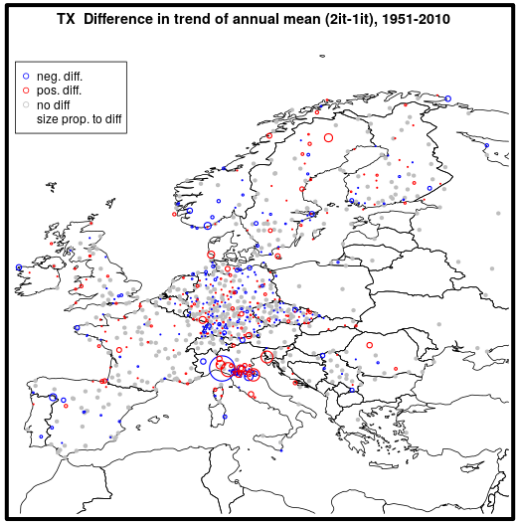
HOM1-ORI



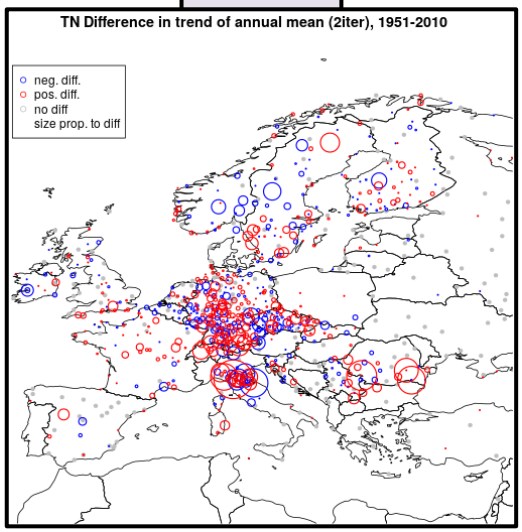


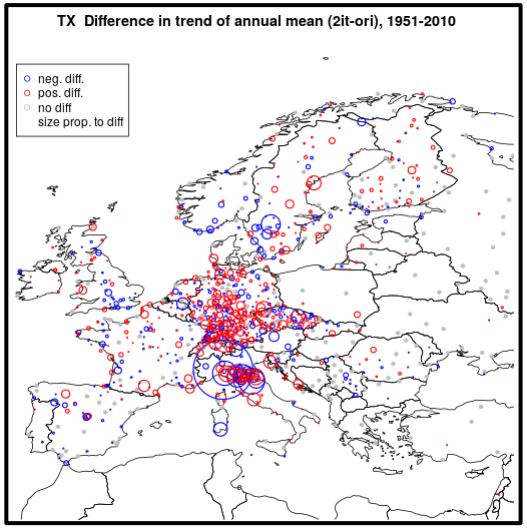
HOM2-HOM1





TOTAL





EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

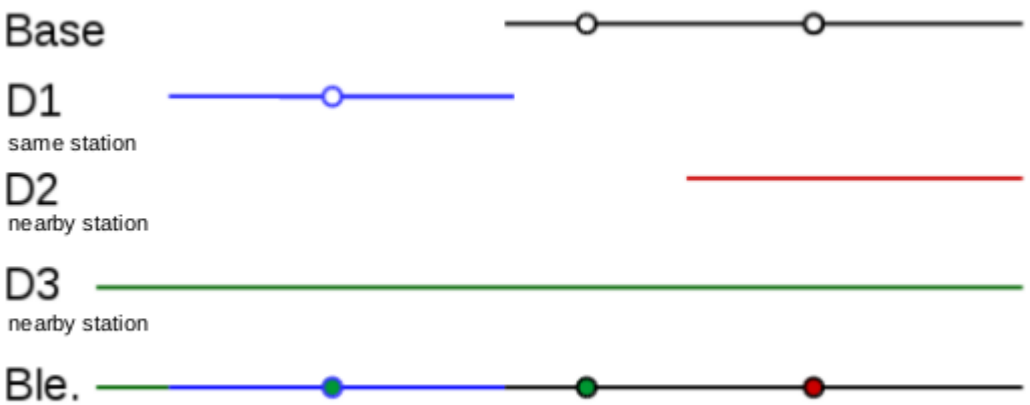
BLENDING OF SERIES

Hom. series from same station and from stations at a maximal distance of 12.5 km and height difference of no more than 100 meters are blended.

Very useful in case of dismissed or moved stations (for example from the centre to the airport).

Latest ending series from target station is selected as “base”, the other ones as “donating”

Blended series are longer and eventual gaps are filled with data from donating series.

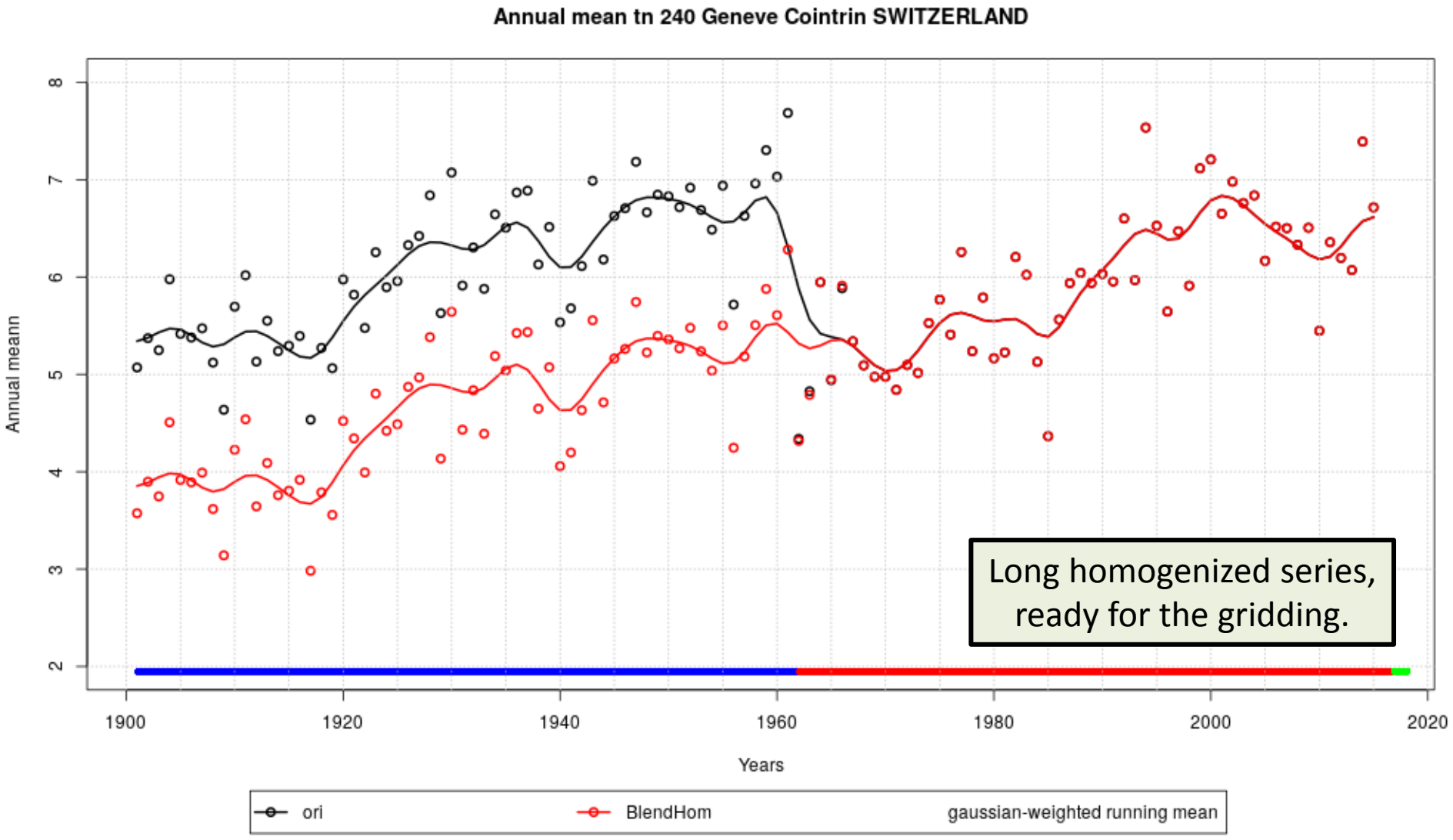


Though, blending procedure introduces new inhomogeneities that need to be adjusted with an “ad-hoc” homogenization process

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

HOMBLENDED SERIES



Annual means of TN Geneva/Geneva Airport before and after the blending homogenization with running mean. Donating series are identified by the colors below: Geneva(blue), Geneva Airport (red), synoptical extension to current time (green)

EUSTACE

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

CONCLUSIONS

For a reliable calculation of trends, gridded data-sets and more products homogeneous series are indispensible.

Thanks to a process of Break Detection and Quantile Matching Homogenization a homogenized data-set is produced.

High number of portions of the series are not adjusted due to lack of long enough reference homogeneous sub-series.

Clear improvement is obtained on homogeneity of the series and on spatial consistency of trends.

A 2nd iteration of BD and QM is done, using the outcomes of 1st it. as input.

Dismission and replacement of stations causes series to be short. For the best gridded data and trend calculation, long series are required.

Series from neighbour stations are blended, caring of induced inhomogeneities that are adjusted with a QM procedure

NEXT STEPS TO UNDERTAKE:

- test the influence of reference locations around the target
- put constraint on adjustments not to “flip” the distribution
- develop a validation system based on manually homogenized series

EUSTACE

Science & Technology Facilities Council

University of Leicester

Met Office Hadley Centre

University of Reading

UNIVERSITY OF BATH

u^b

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

THE UNIVERSITY OF EDINBURGH

UNIVERSITÄT BERN
OESCHGER CENTRE
CLIMATE CHANGE RESEARCH

DMI

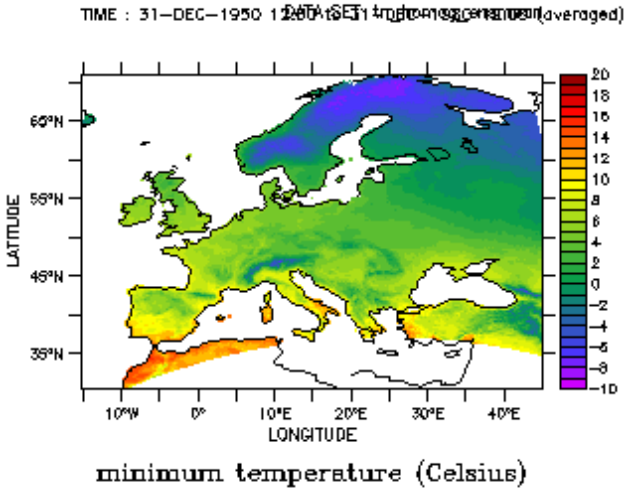
THANKS FOR YOUR ATTENTION

QUESTIONS?
VRAGEN?
DOMANDE?

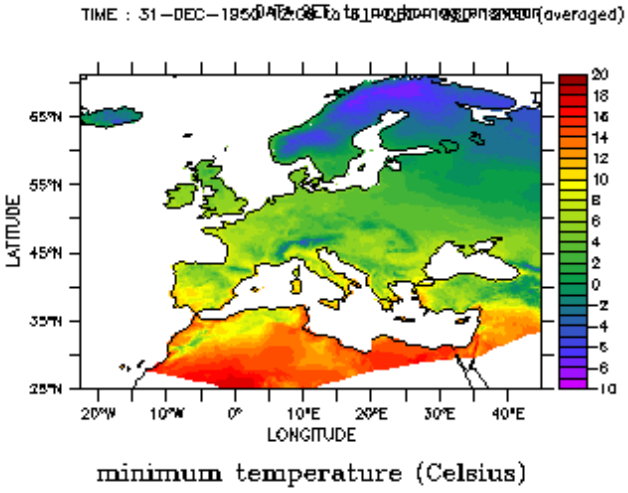
EUSTACE has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no 640171

EUSTACE

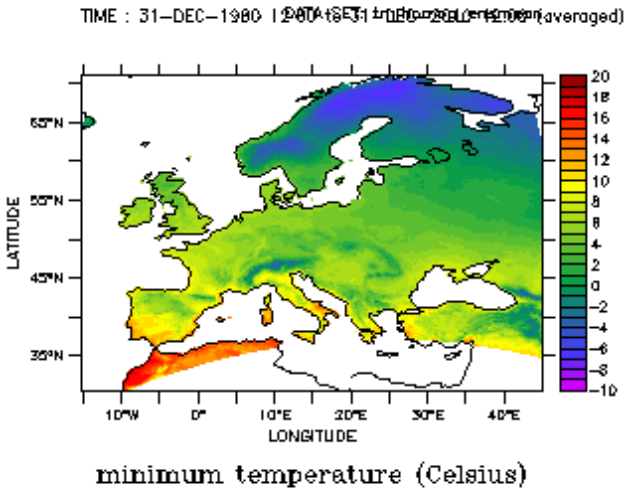
GRIDDED DATASET - TN



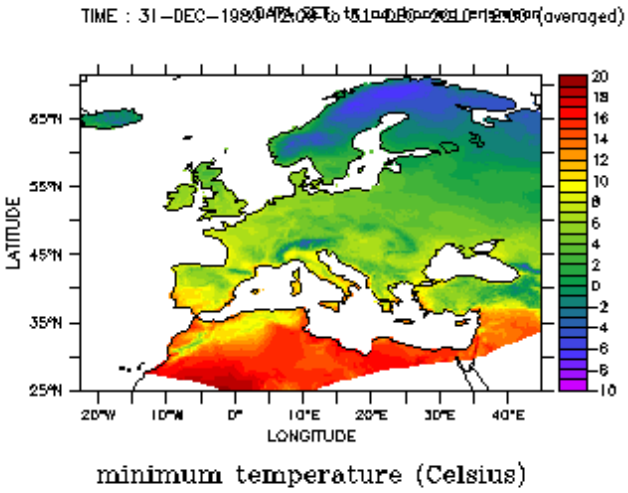
minimum temperature (Celsius)



minimum temperature (Celsius)



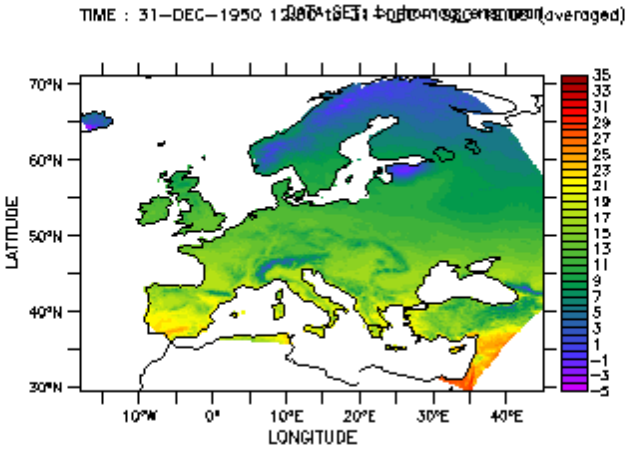
minimum temperature (Celsius)



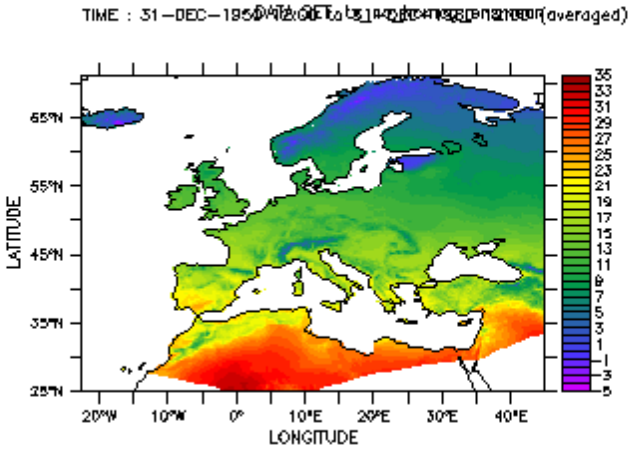
minimum temperature (Celsius)

EUSTACE

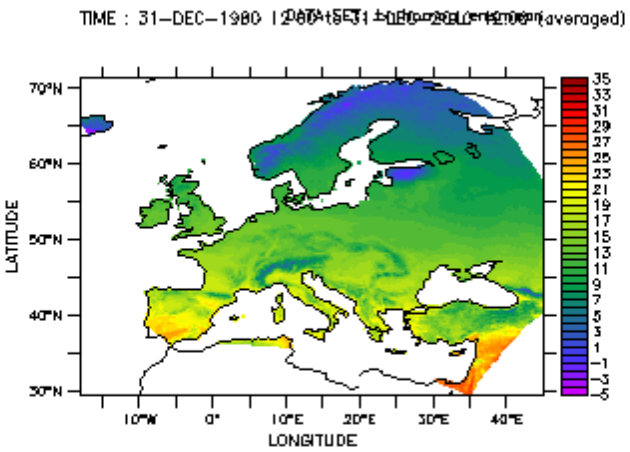
GRIDDED DATASET - TX



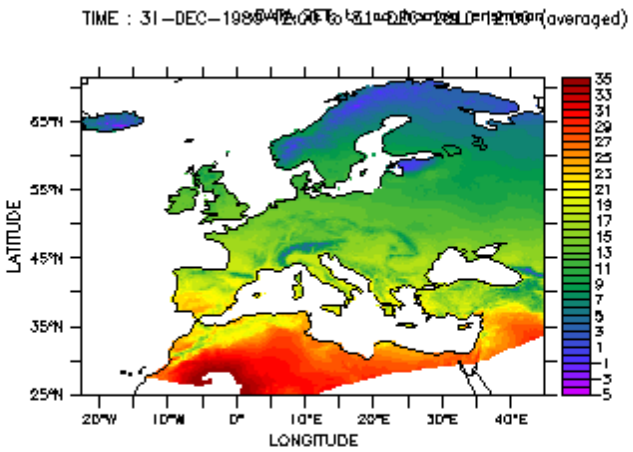
maximum temperature (Celsius)



maximum temperature (Celsius)



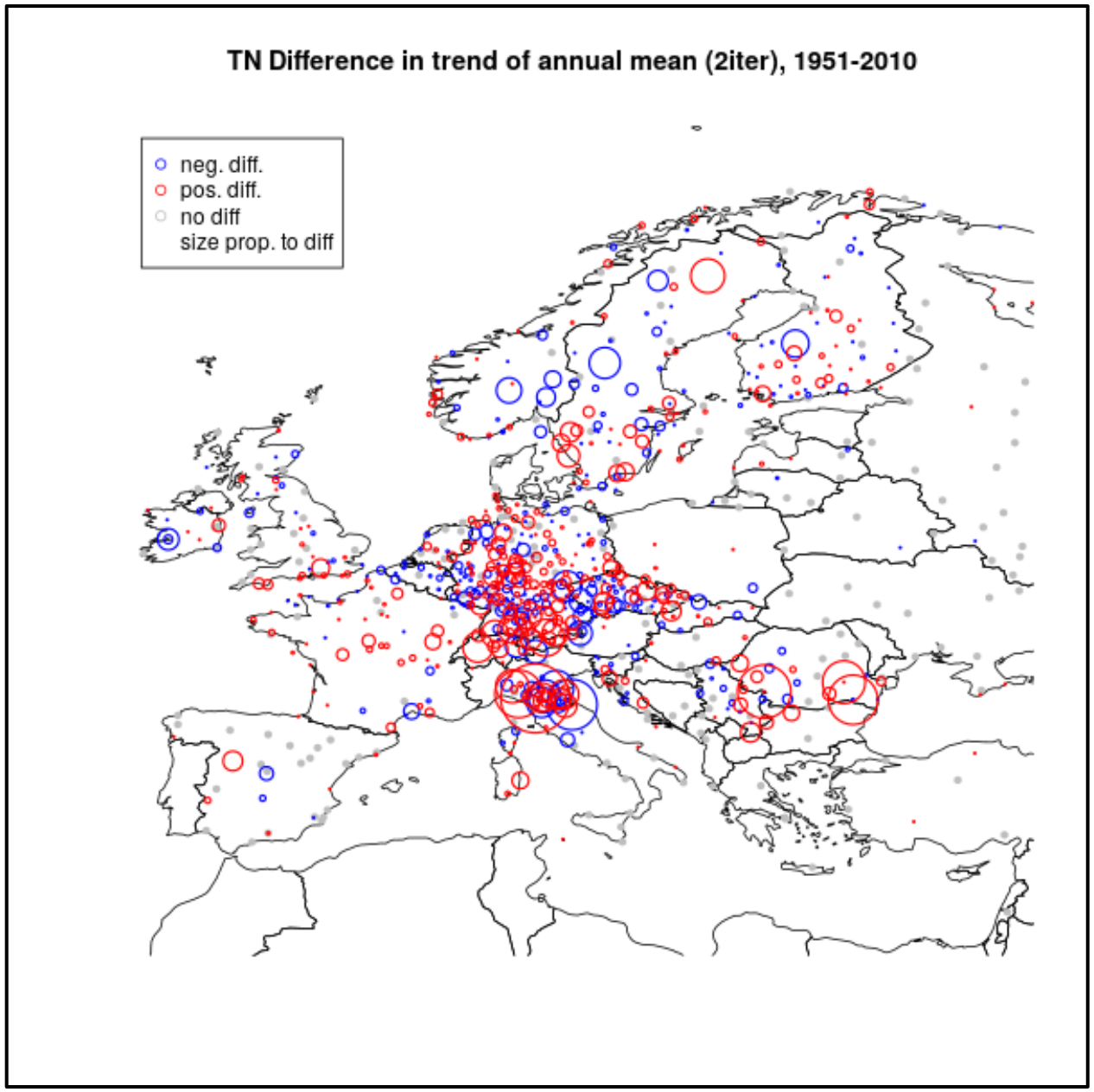
maximum temperature (Celsius)



maximum temperature (Celsius)

EUSTACE

DIFFERENCE IN TRENDS



Difference signs in the corrections.
The aim is not to have warmer trends, therefore corrections on trends can also be negative

EUSTACE