

Homogenization of Global Early Instrumental data

Contact Email: elin.lundstad@giub.unibe.ch

Elin Lundstad (PhD-student)

Institute of Geography & Oeschger Centre for Climate Change Research, University of Bern

15.05.2023



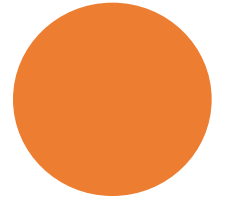
European Research Council
Established by the European Commission

Outline

- HCLIM
- Format
- Geographical view
- Methods
- Some results
- Outlook

u^b

b
**UNIVERSITY
BERN**



Global Early Instrumental Monthly Meteorological Multivariable Database



The screenshot shows the PANGAEA website interface. At the top left is the PANGAEA logo and the text "Data Publisher for Earth & Environmental Science". To the right are navigation links: "SEARCH SUBMIT HELP ABOUT CONTACT". Below this is a citation for "Lundstad, Elin; Brugnara, Yuri; Brönnimann, Stefan (2022): Global Early Instrumental Monthly Meteorological Multivariable Database (HCLIM), PANGAEA, https://doi.org/10.1594/PANGAEA.940724". There are buttons for "RIS Citation", "Bibtex Citation", "Copy Citation", "Facebook", "Twitter", "Show Map", and "Google Earth". A small map shows the distribution of data points across the globe. Below the citation is an abstract starting with "There is a growing need for past weather and climate data to support science and decision-making. This paper describes the compilation and the construction of a global multivariable (air temperature, pressure, precipitation sum, number of precipitation days) monthly instrumental climate database that encompasses a substantial body of the known early instrumental time series. The dataset contains series compiled from existing databases that start before 1890 (though continuing to the present) as well as a large amount of newly rescued data. All series underwent a quality control procedure and subdaily series were processed to monthly mean values. An inventory was compiled, and the collection was deduplicated based on coordinates and mutual correlations. The data are provided in a common format accompanied by the inventory. The collection totals 12452 meteorological records in 118 countries. The data has been merged from 18250 original data files. The data can be used for climate reconstructions and analyses. It is the most comprehensive global monthly climate data set for the preindustrial period."

- **HCLIM**
- Compiled instrumental meteorological data in one database, contains 4 parameters
- New project is to homogenized temperature in the HCLIM database
- Is it possible to use a modern technique?
- How far back in time can we homogenize using modern homogenization methods?
- The first and longest time series in the HCLIM dataset is Paris from 1658. But can it be homogenized?

<https://doi.pangaea.de/10.1594/PANGAEA.940724>

The goal for the dataset (HCLIM)

- Part of PALAEO-RA project (<https://www.palaeo-ra.unibe.ch/>)
- Collection of Early meteorological instrumental data
- More parameters (4)
- The data are quality controlled
- So many duplicates (at most 19 duplicates)
- We find a method for merging and de-duplication of the time series

- *«There is so much data out there: Let's try to collect it»*
- *Now, it is one database! HCLIM*

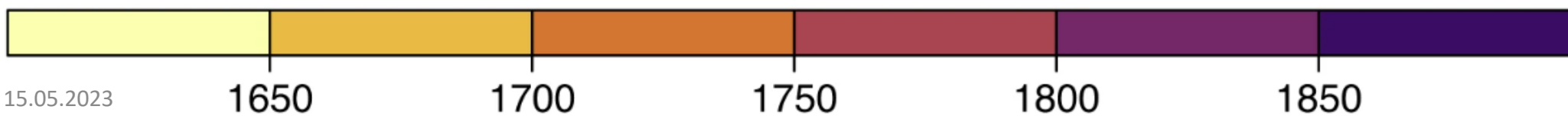
SEF-files

```

SEF      1.0.0
ID       Global-Datasets_Dove-5
Name     Berlin
Lat      52.5
Lon      13.4
Alt      NA
Source   PALAEO-RA
Link     NA
Vbl      ta
Stat     mean
Units    C
Meta     Observer=Brand Gronau. Kirch.Mädler | Obs.times=7 1/2, 12 1/2, 9 1/2. dann 8, 1,11 und tägl. Extr.
Year     Month   Day     Hour     Minute  Period  Value  Meta
1719    1           1           1           1      month   2.8    orig=2.25R
1719    2           2           2           2      month   1.1    orig=0.90R
1719    3           3           3           3      month   5.2    orig=4.15R
1719    4           4           4           4      month   9      orig=7.20R
1719    5           5           5           5      month  15.1   orig=12.08R
1719    6           6           6           6      month  19     orig=15.20R
1719    7           7           7           7      month  21.4   orig=17.15R
1719    8           8           8           8      month  18.8   orig=15.00R
1719    9           9           9           9      month  13.9   orig=11.14R
1719   10          10          10          10     month   9      orig=7.19R
1719   11          11          11          11     month   6.6   orig=5.24R
1719   12          12          12          12     month   0.3   orig=0.22R

```

Temperature (n=3632)



Homogenization methods

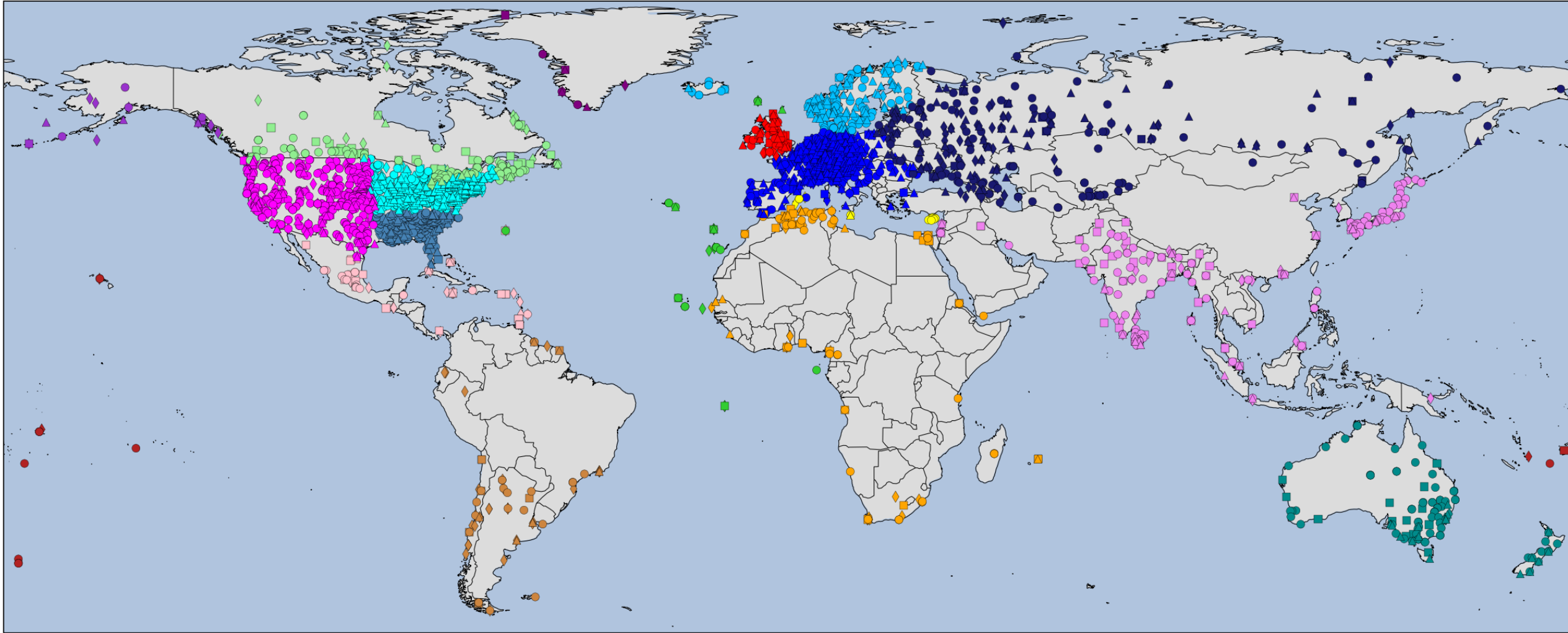
Pha (GHCN)

Climatol (JA
Guijarro)

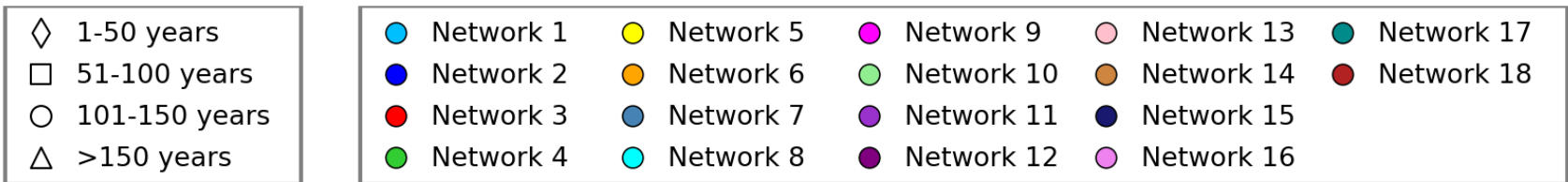
Homer (European
COST-action)
renewed, now
called Bart



18 network



15.05.2023



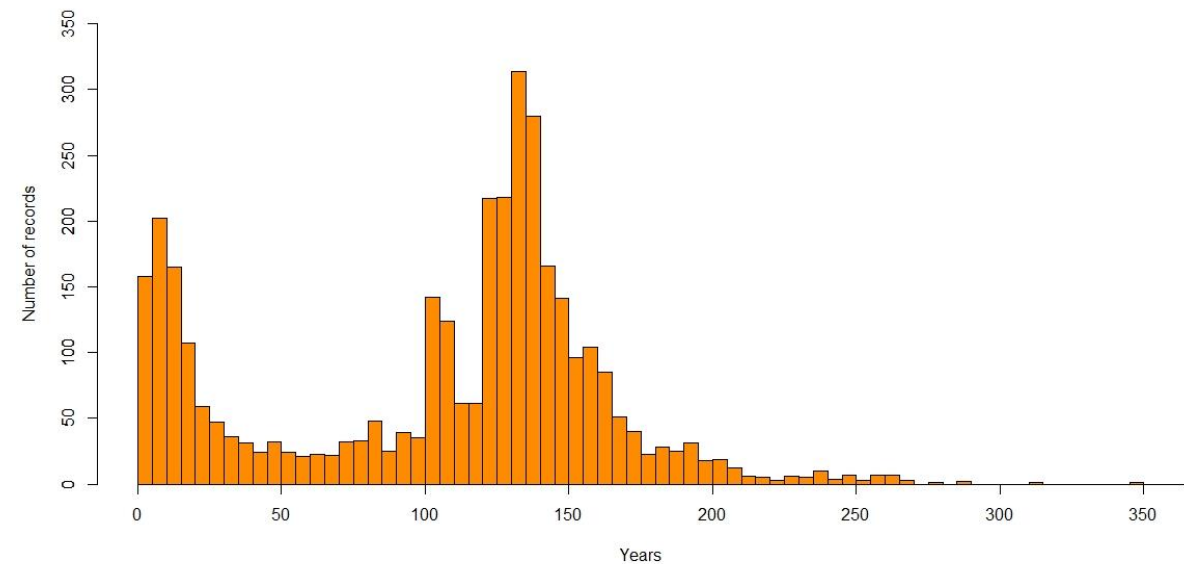
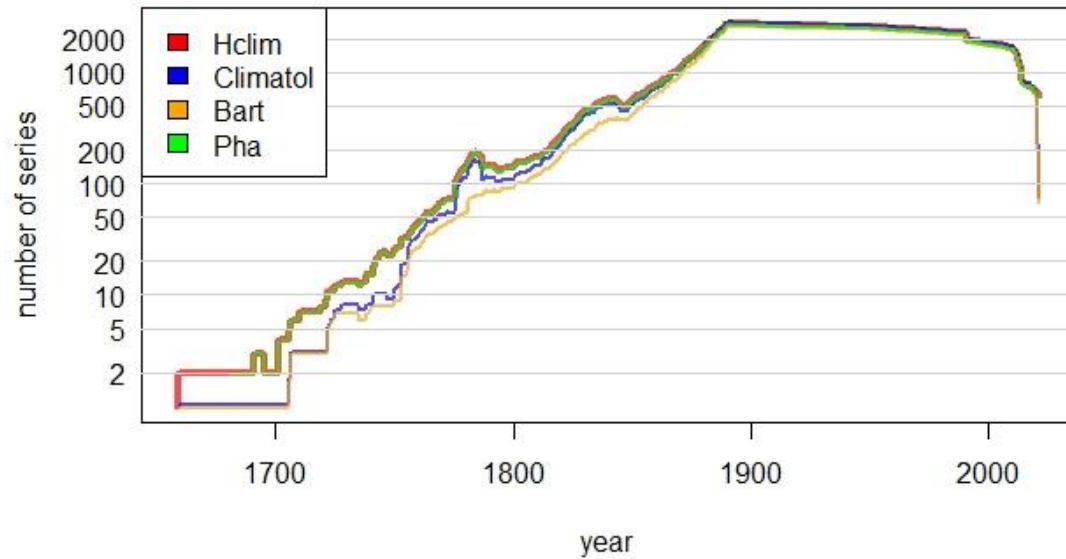
Time table

Network Number	Geographical area	Number of stations	After running Homogenization	Period for homogenization	Period for Hclim
1	Scandinavia	181	181	1722-2020	1722-2021
2	Europe	605	602	1757-2020	1658-2021
3	United Kingdom	88	88	1763-2020	1659-2021
4	Atlantic Ocean	28	28	1856-2013	1749-2013
5	Mediterranean Sea	116	18	1865-2020	1744-2021
6	Africa	90	85	1835-2020	1819-2021
7	USA (South-East)	443	440	1819-2020	1738-2021
8	USA (North-East)	632	628	1780-2020	1743-2021
9	USA (West)	437	435	1843-2020	1837-2021
10	Canada	217	211	1808-2020	1744-2021
11	Alaska	17	9	1899-2020	1828-2021
12	Greenland	12	11	1852-2020	1816-2021
13	Central America	66	55	1855-2013	1764-2021
14	South America	62	62	1851-2020	1825-2021
15	Russia	291	289	1805-2020	1741-2021
16	Asia	214	210	1814-2020	1736-2021
17	Australia	116	109	1851-2020	1788-2021
18	Pacific Ocean	16	16	1873-2019	1867-2021

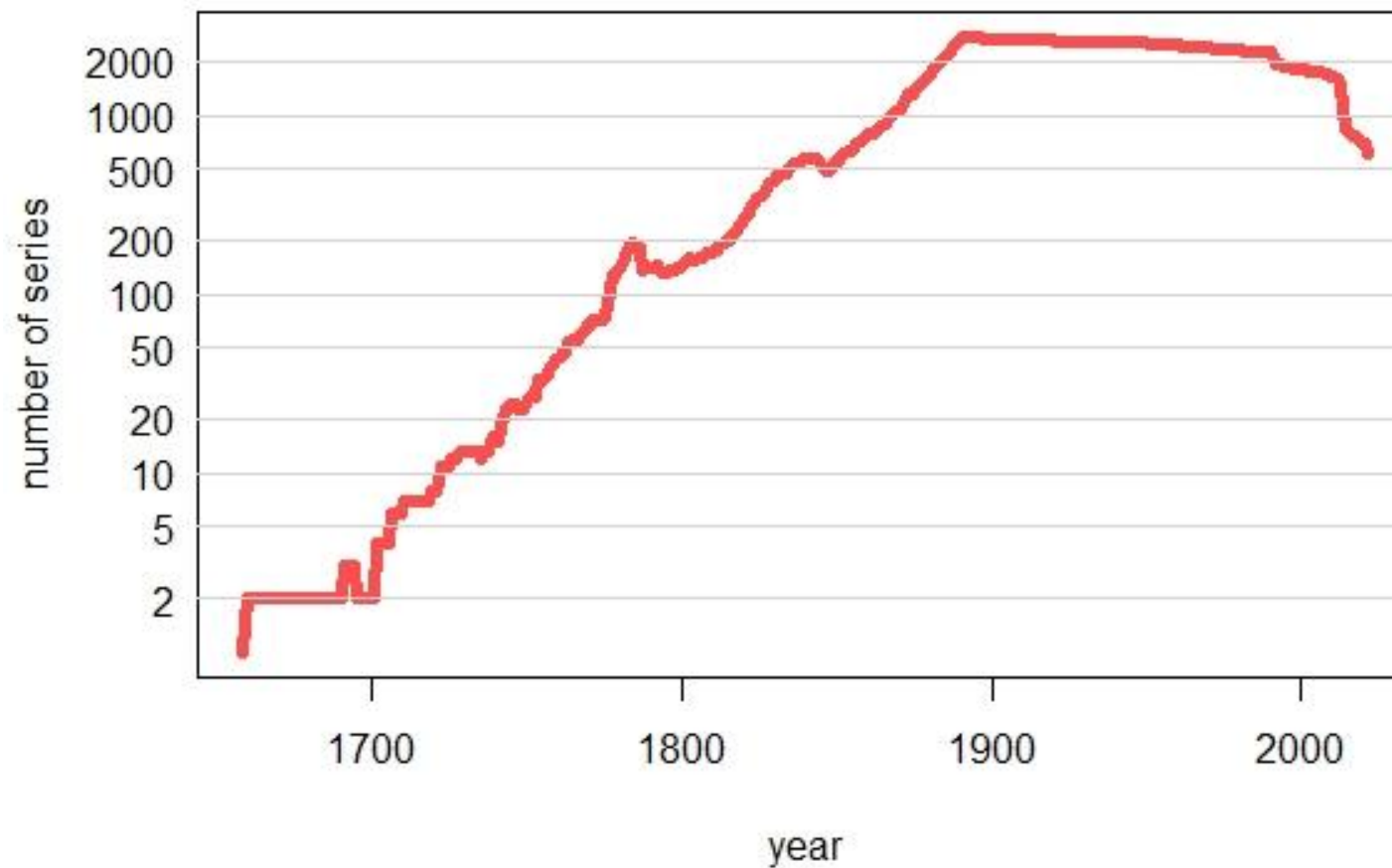
Some numbers from the analyses

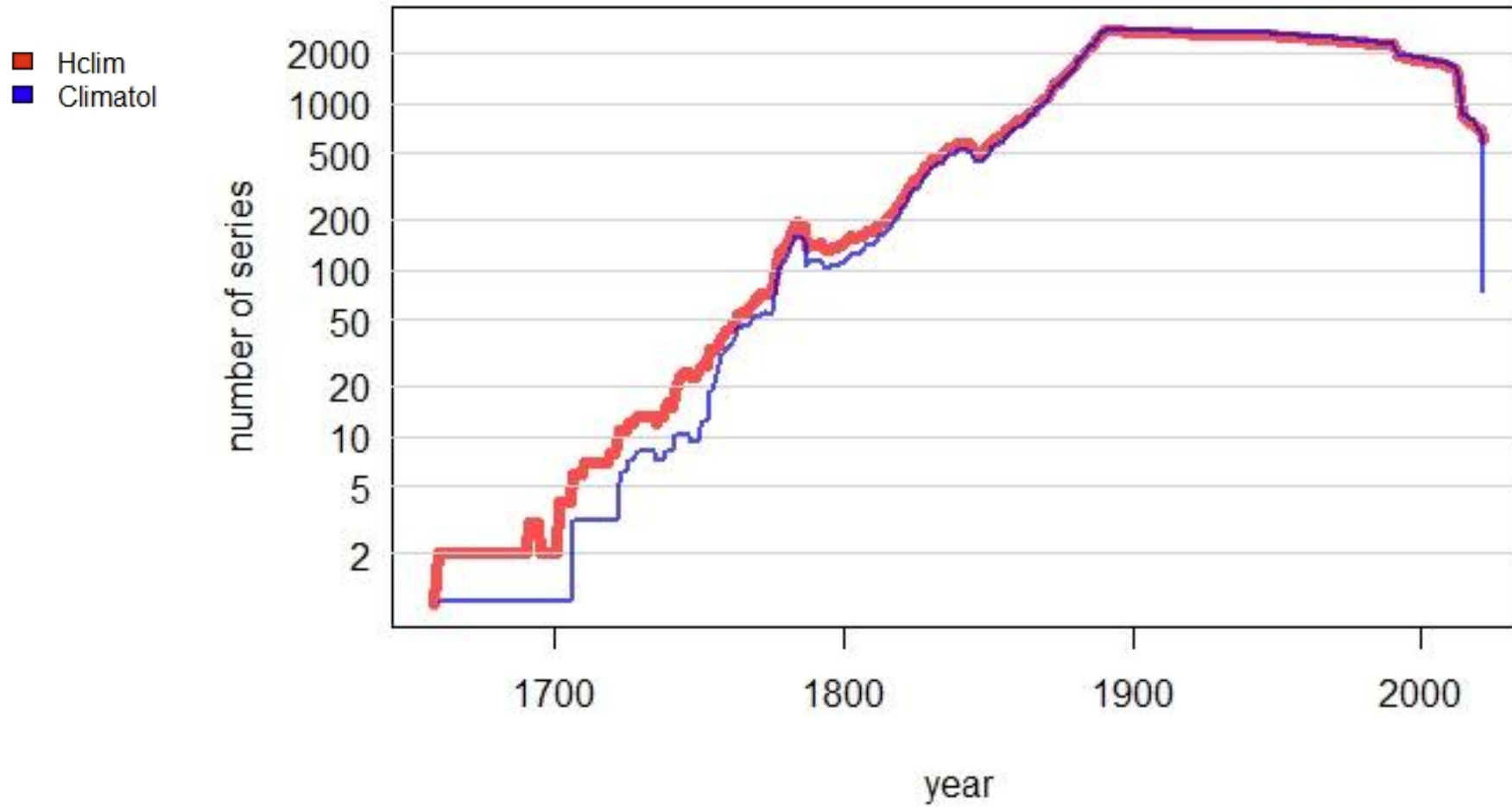
	Climatol	Bart	Pha
Total in	3490	3282	3630
Total kicked out	142	740	2
- Correlation coefficient too low	455		
- Too long distance (defined by the program)	77		
- Record to short		438	
- No neighbours (in distance)		239	
- Outliers as defined by the programme	283		
- Big gaps	1		
- Other reasons			

Amount and length of the time series

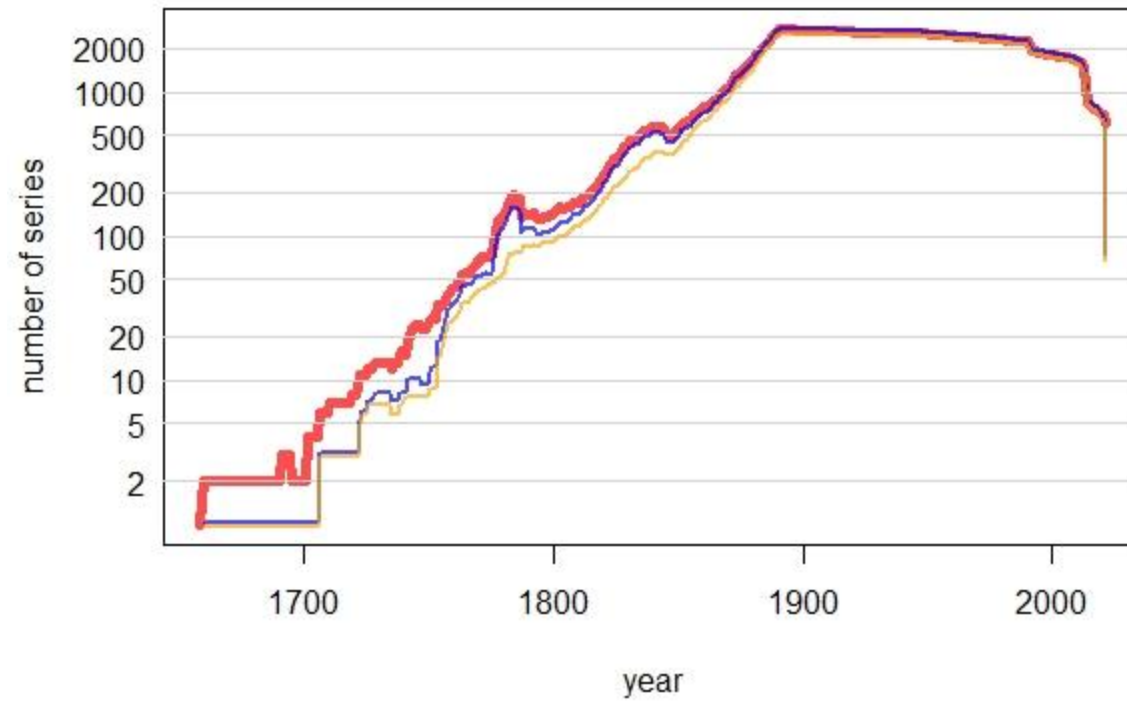


■ Hclim

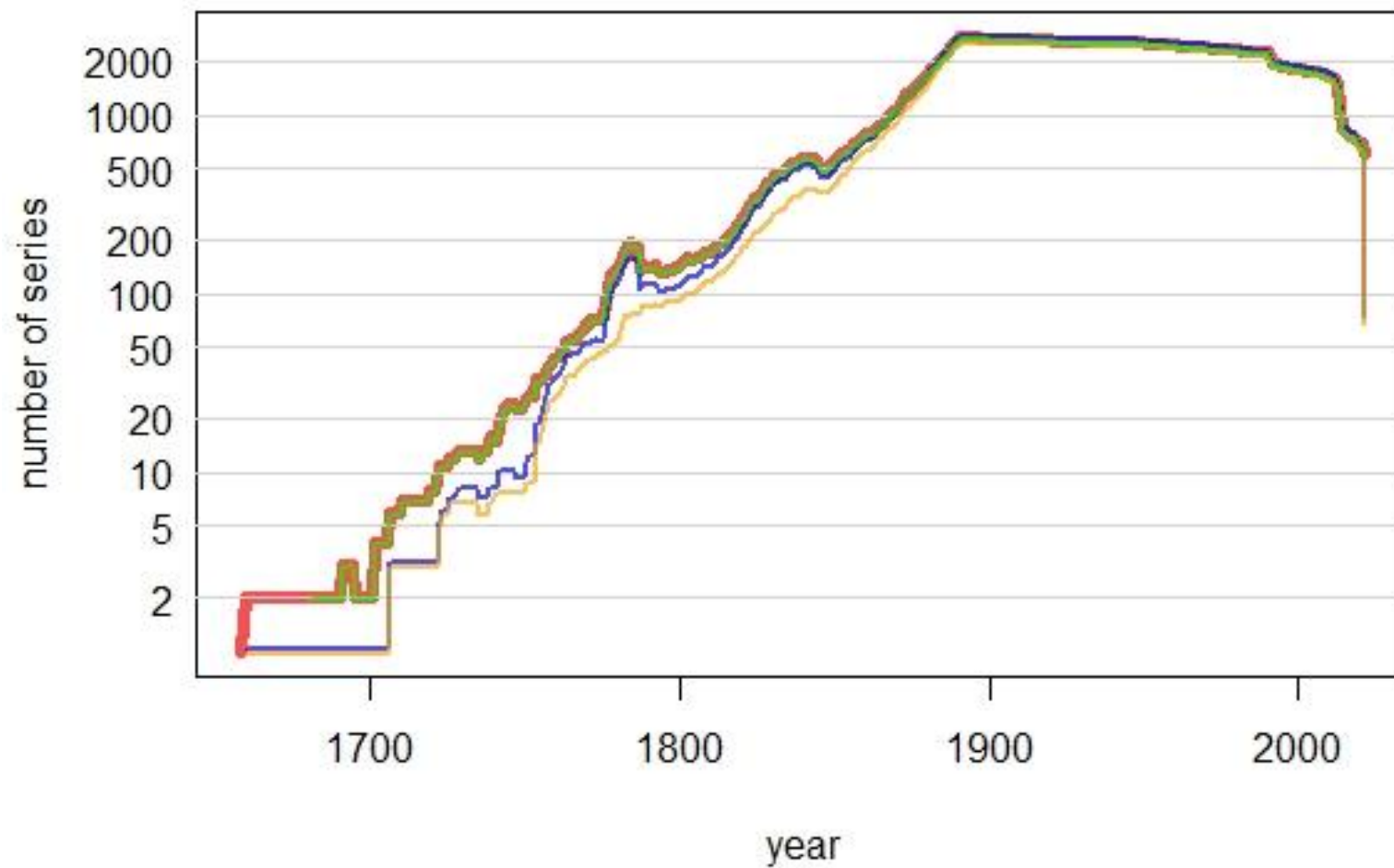




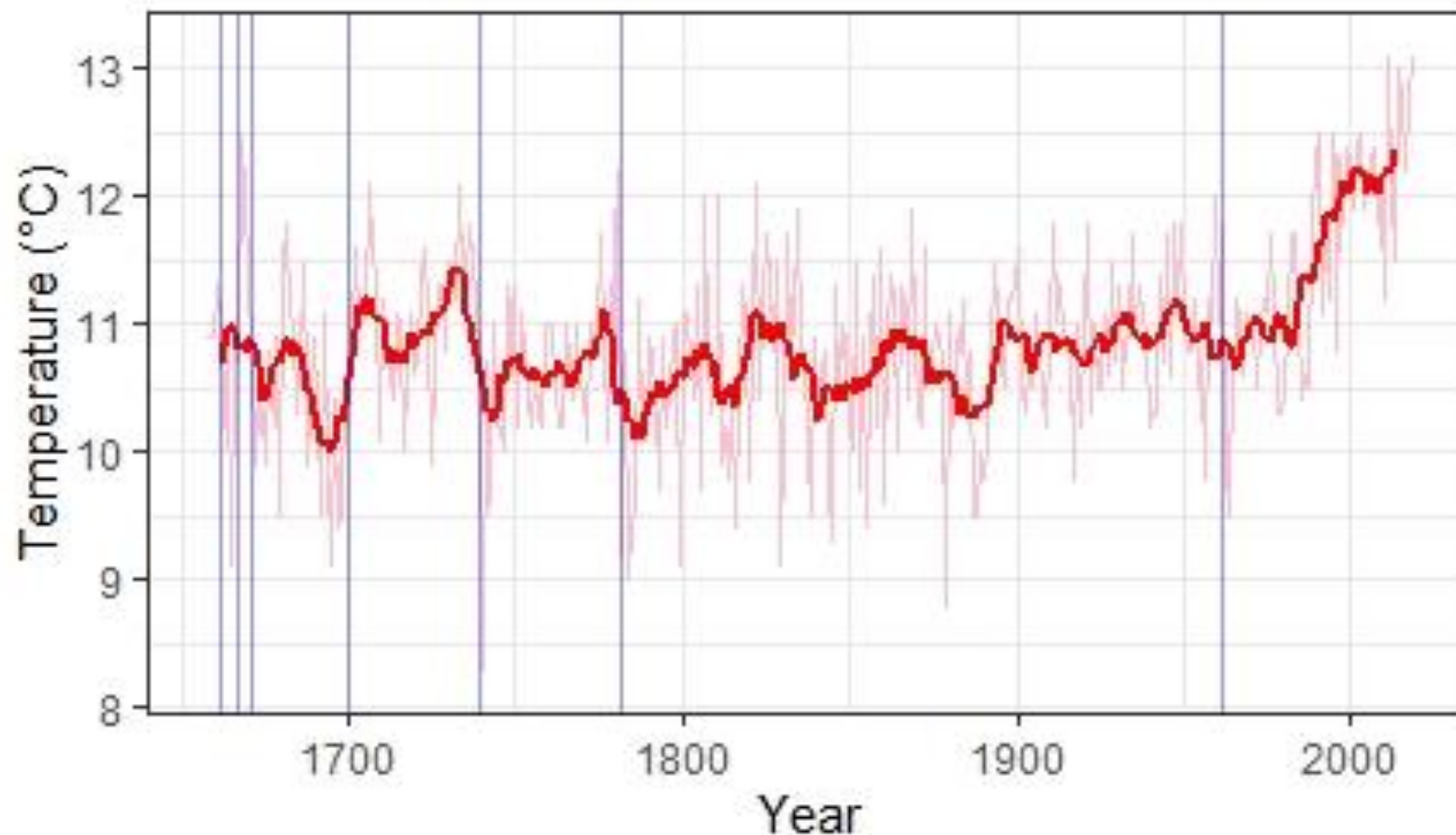
- Hclim
- Climatol
- Bart



- Hclim
- Climatol
- Bart
- Pha

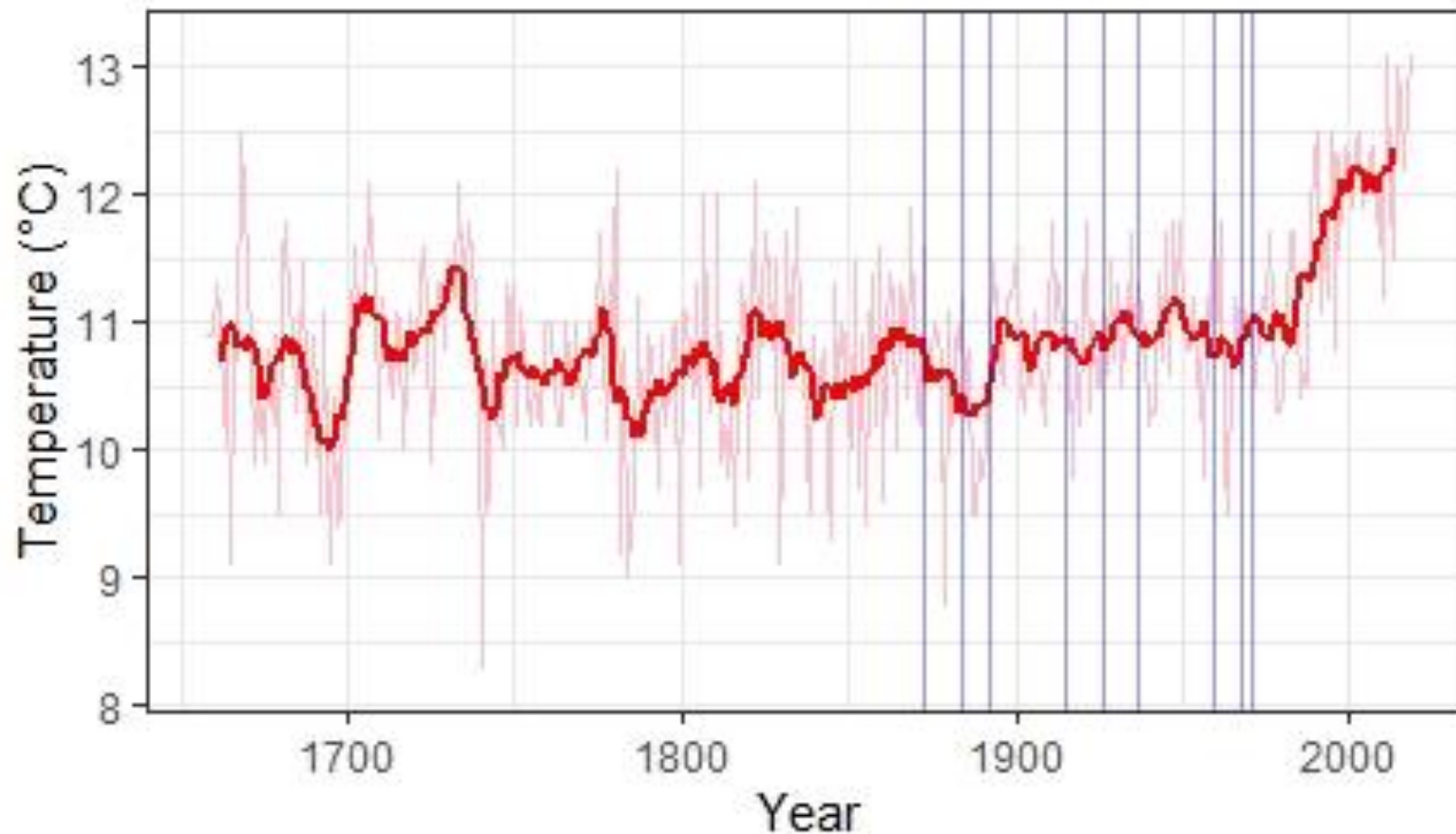


Breakpoint Detection (Hclim)



- Welch's test

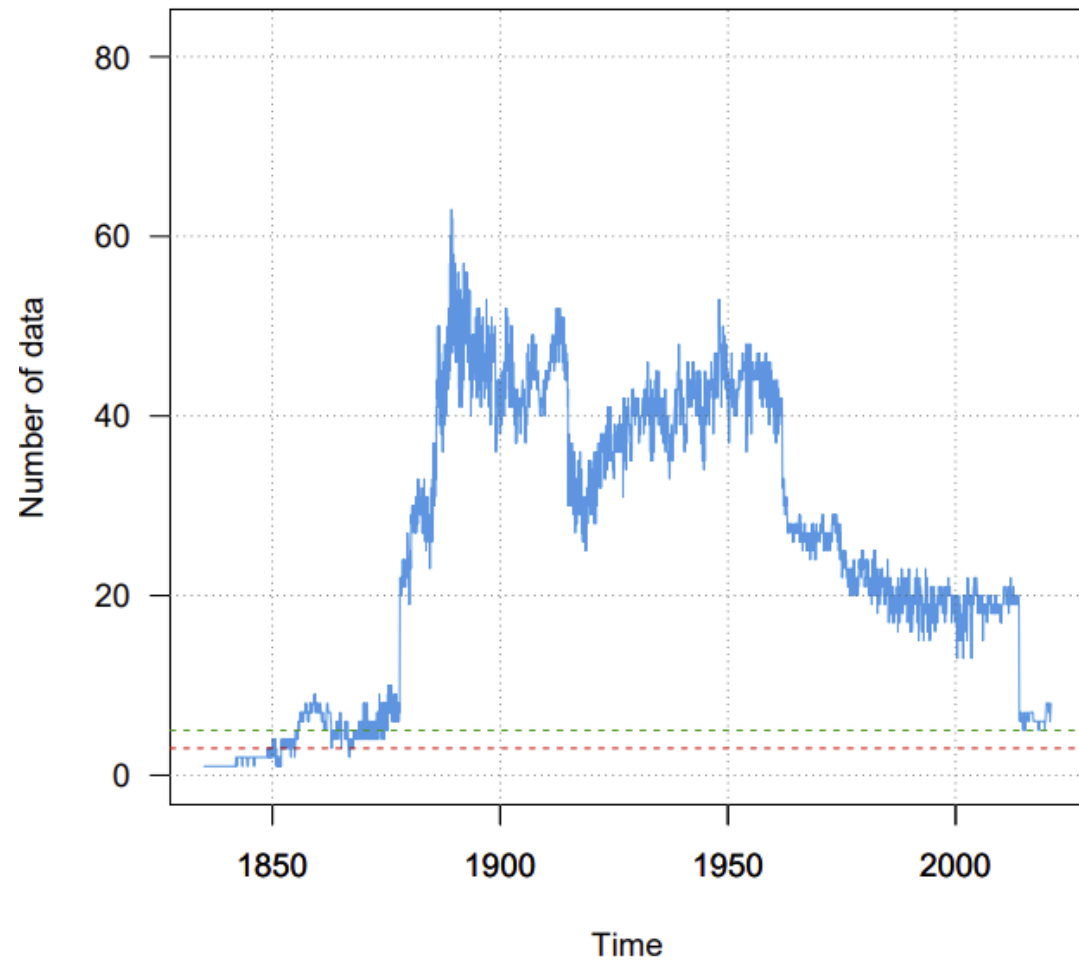
Homogenization(Rousseau)



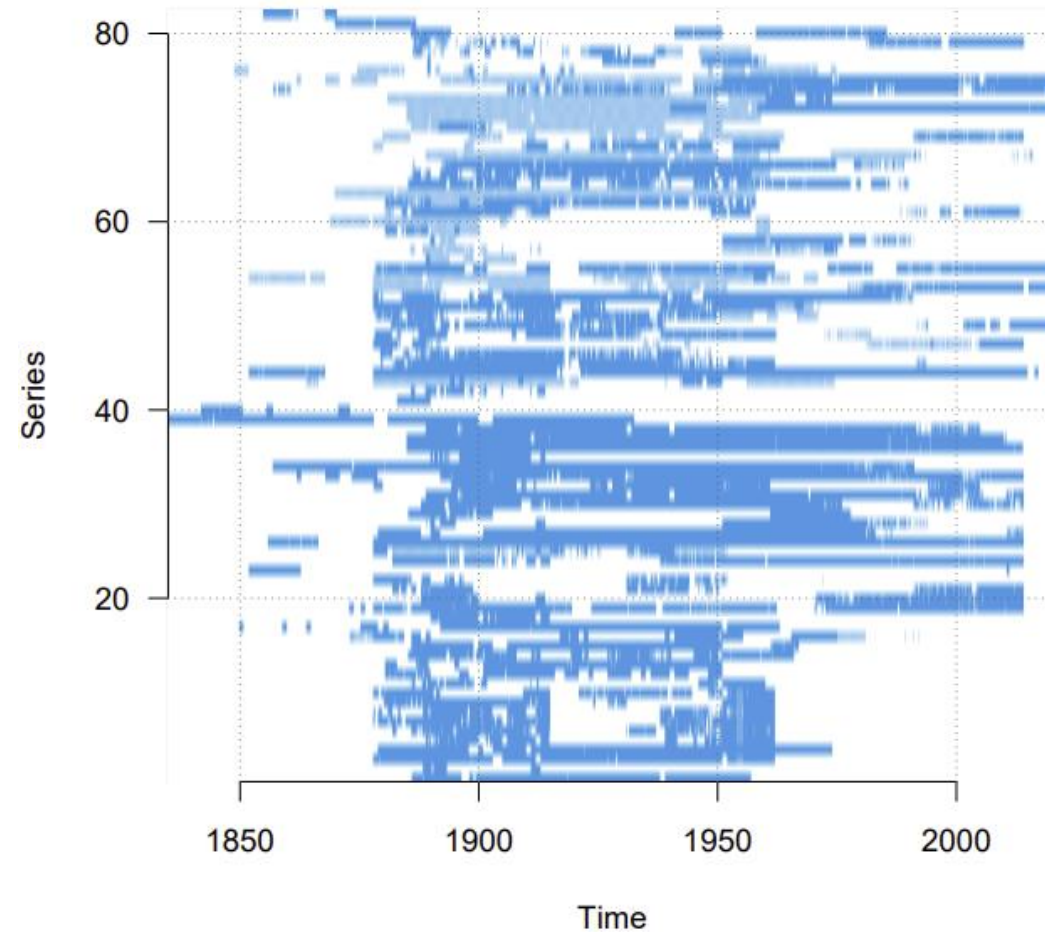
Corrections

Data - Africa

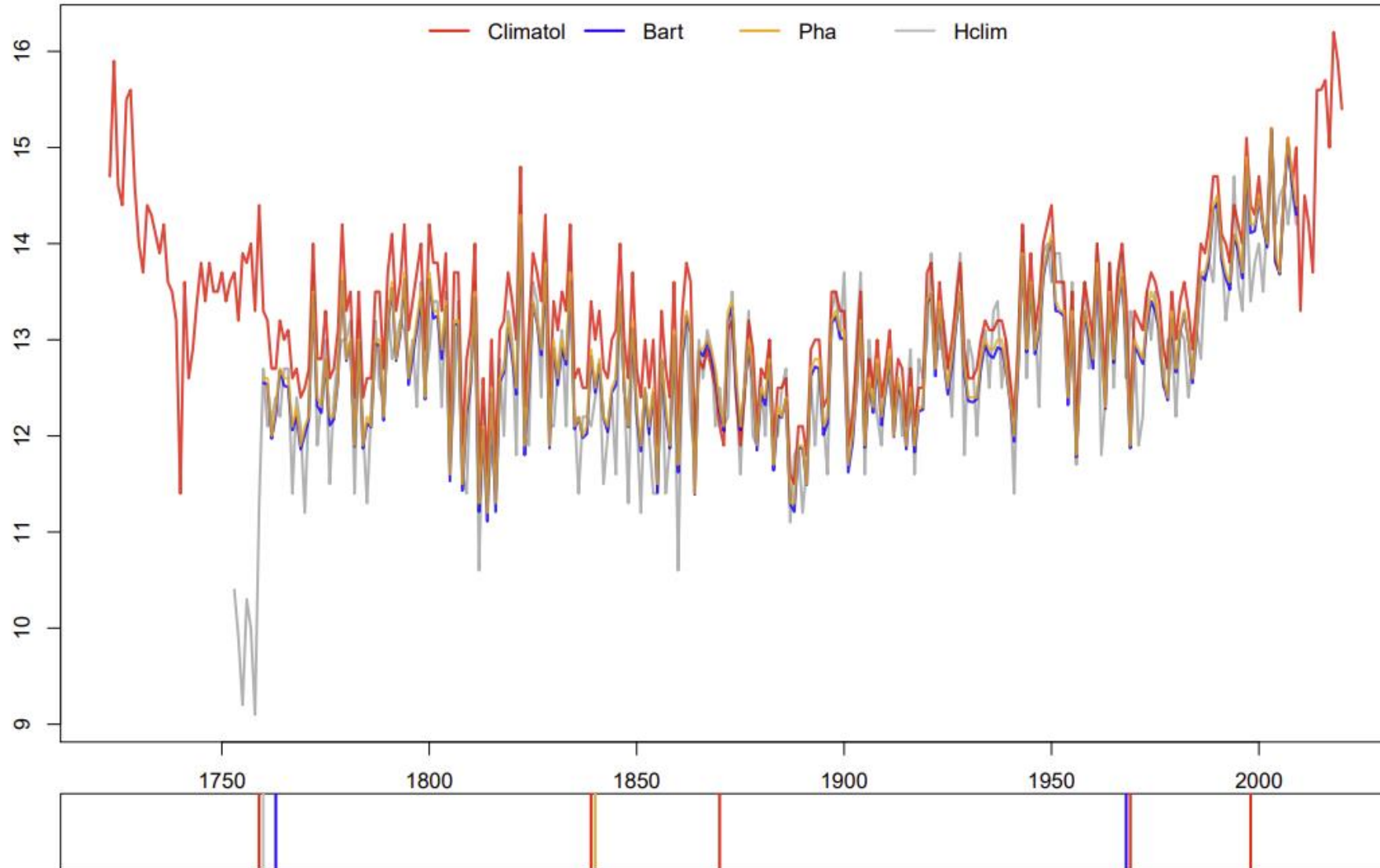
Number of ta data in all stations



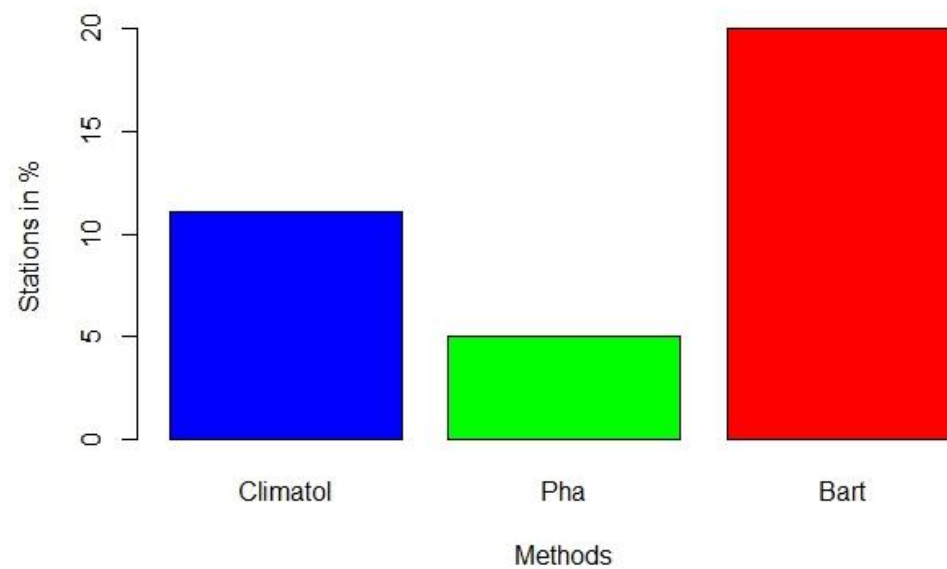
ta data availability



Annually homogenized temperature in Turin

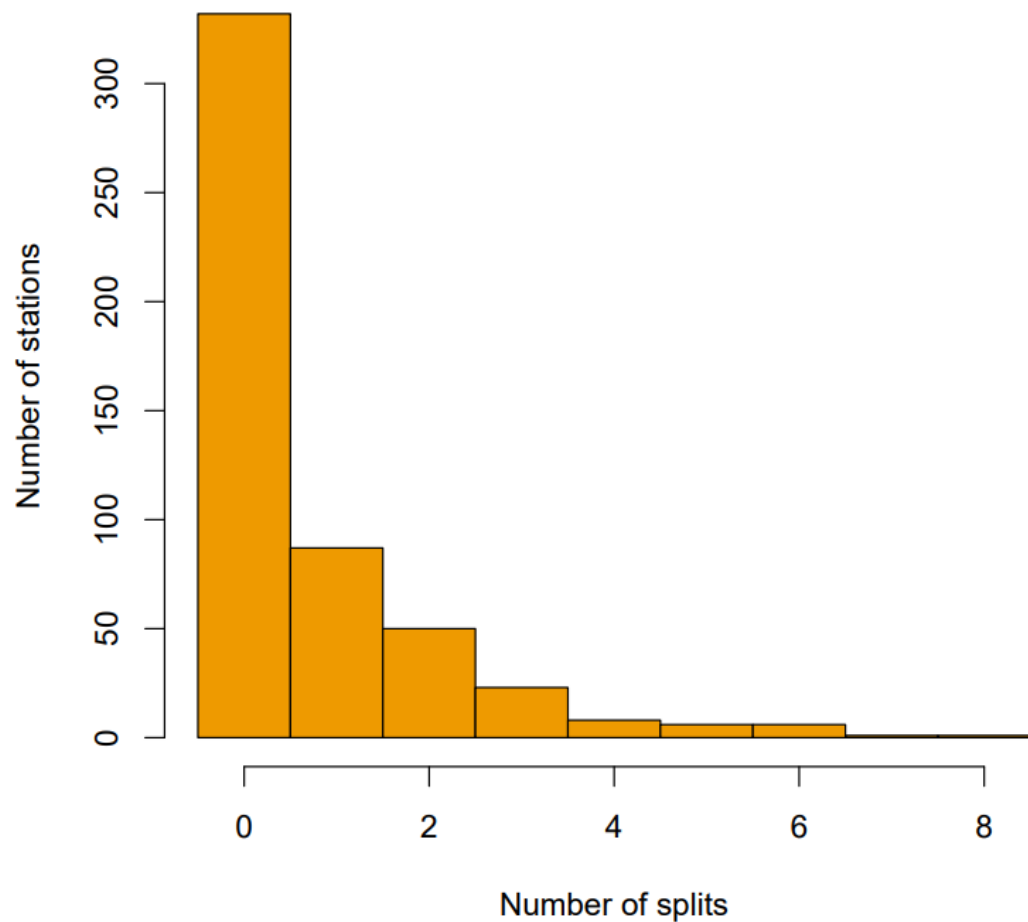


Not possible to homogenize?

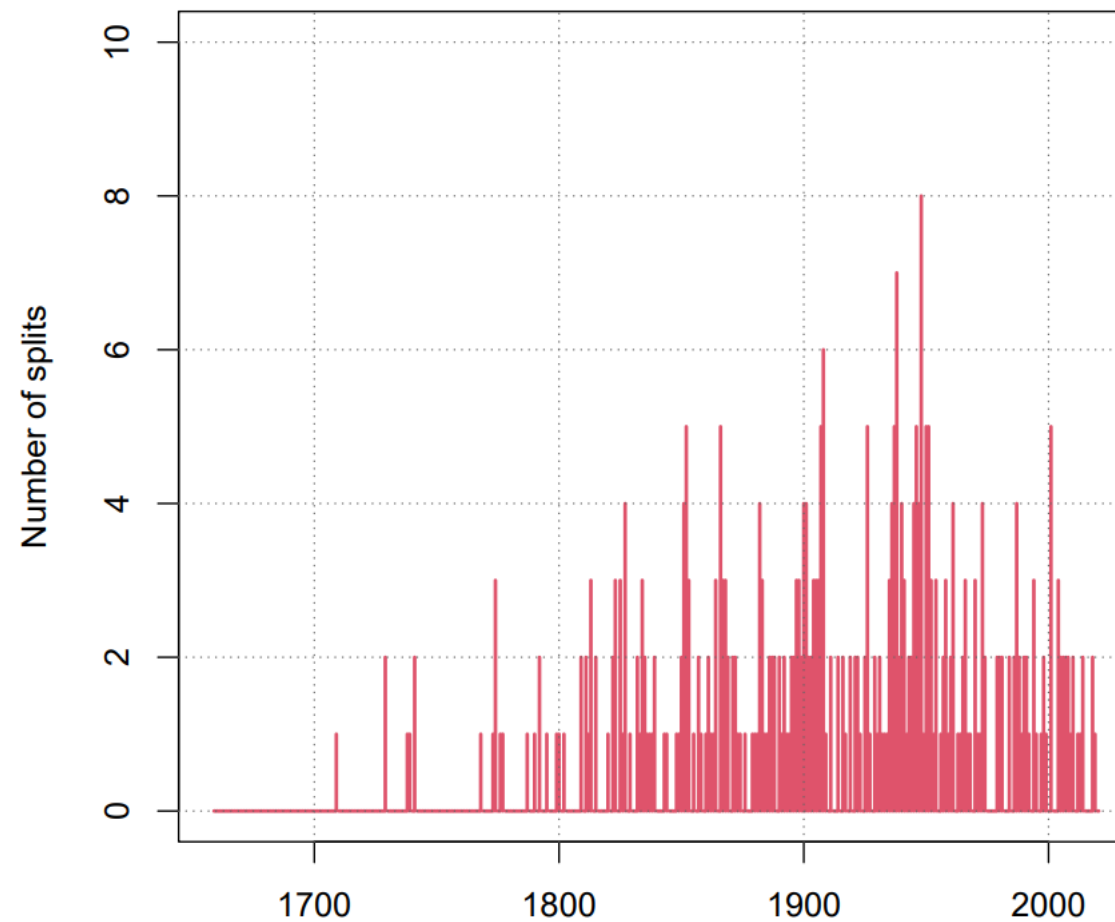


Homogenization - Europe

Number of splits per station

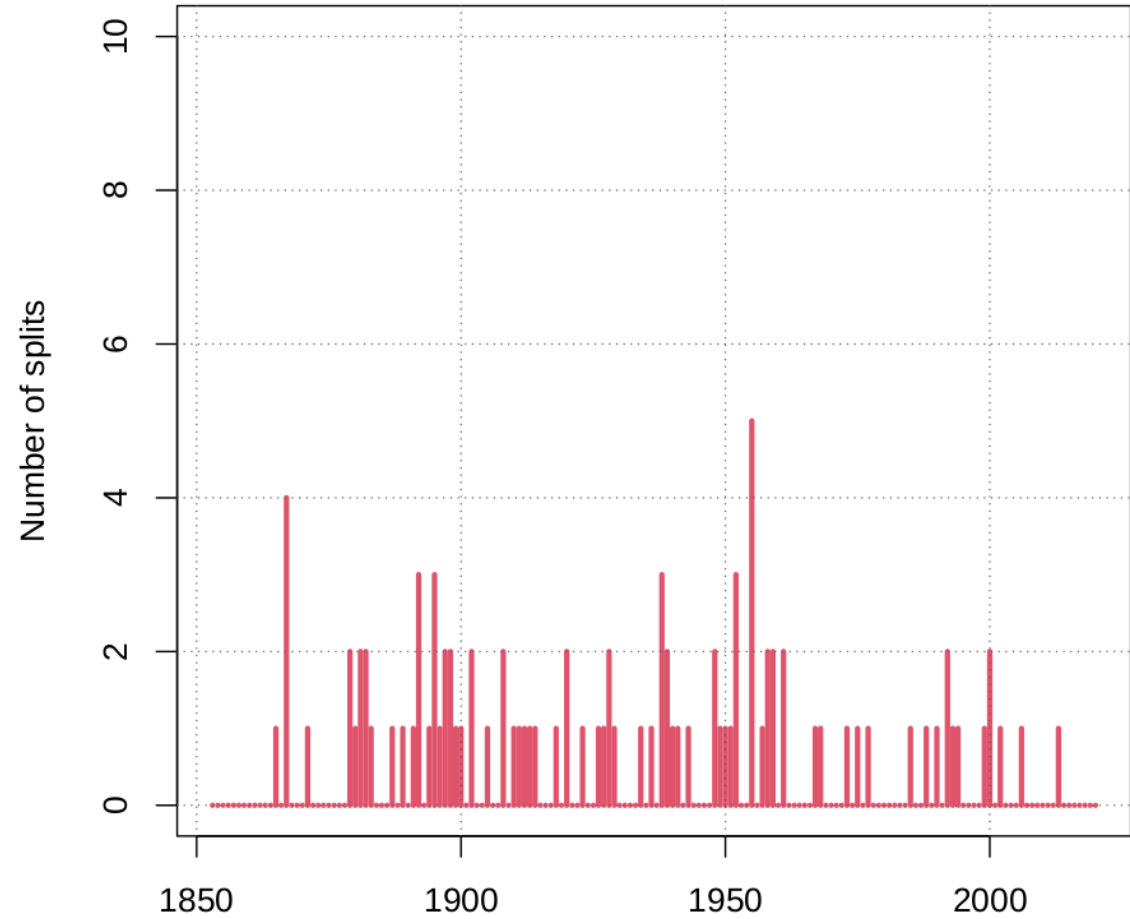


Number of splits per year

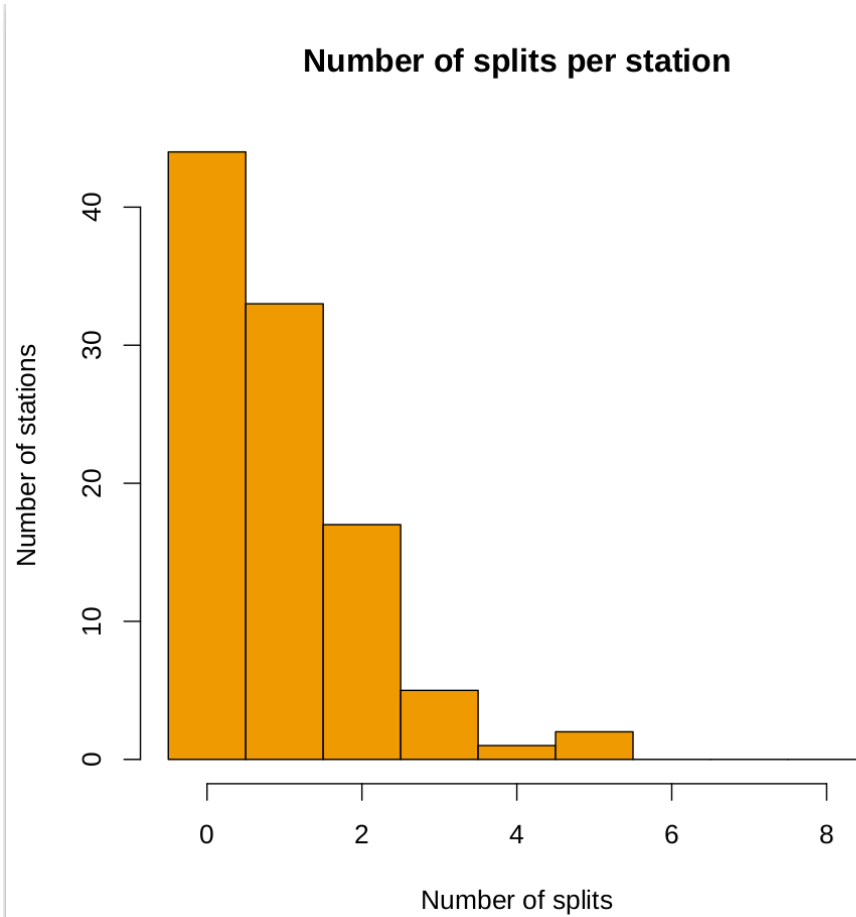


Homogenization - Australia

Number of splits per year

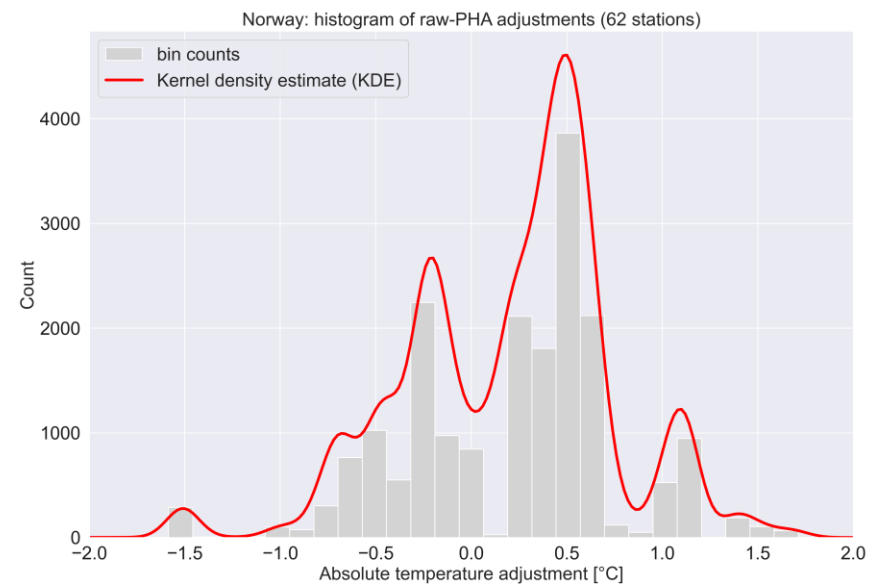
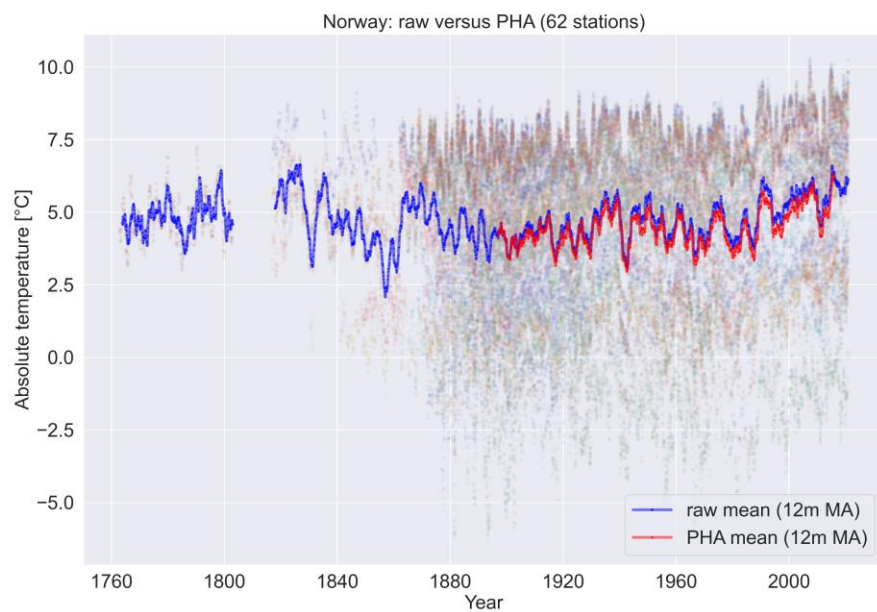
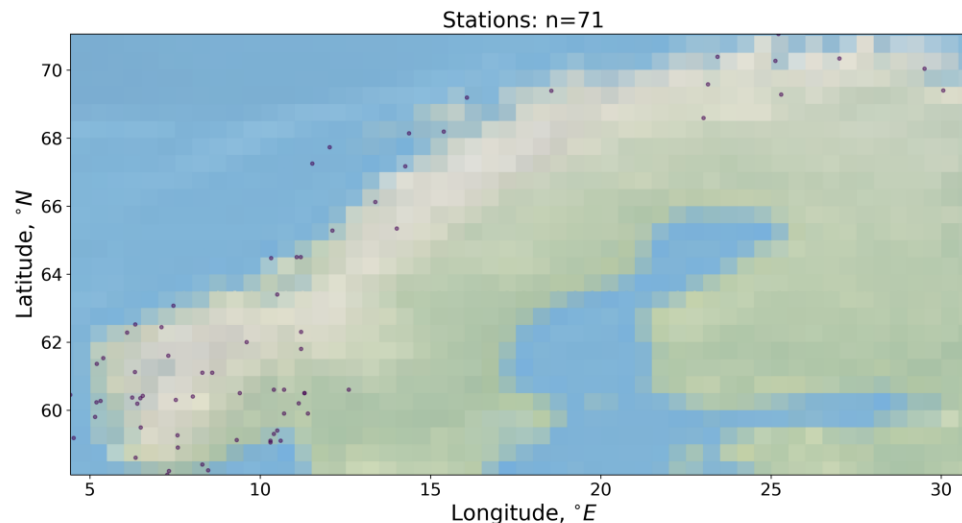
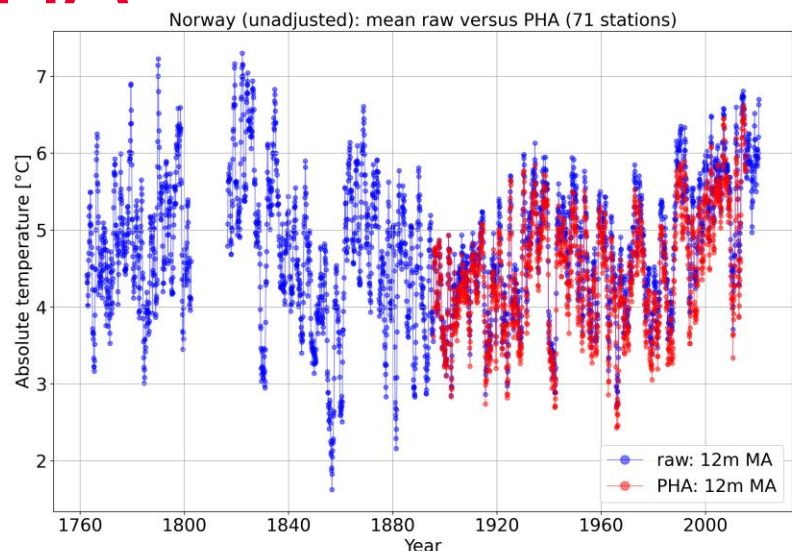


Number of splits per station

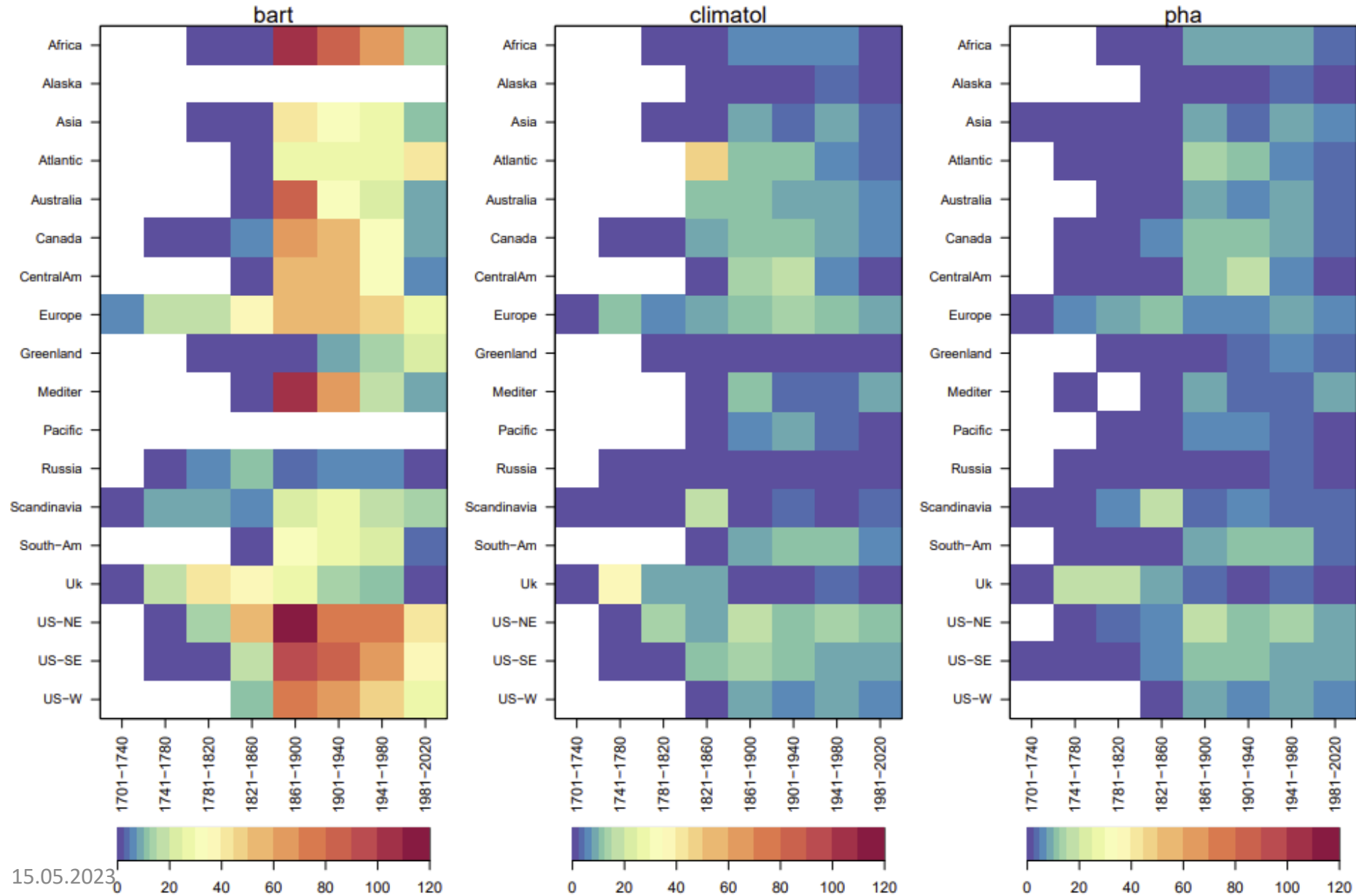


PHA

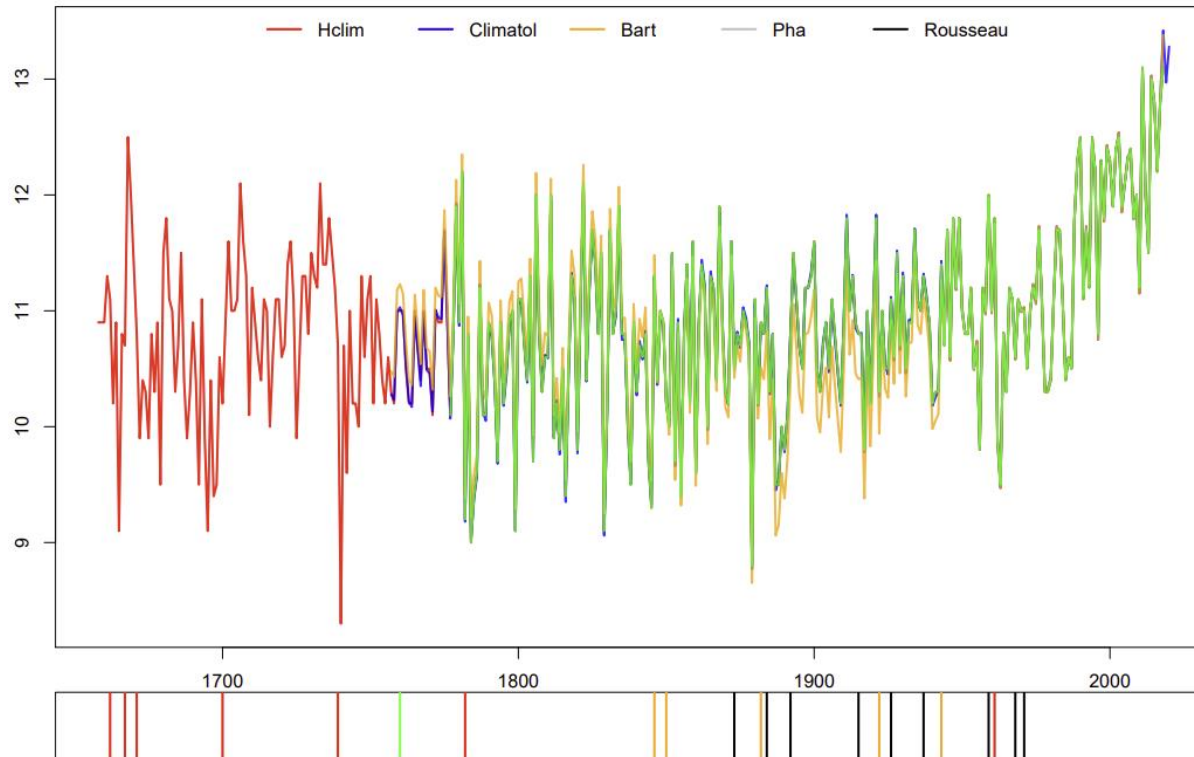
u^b



Comparing number of breaks in a heat map



Paris



- **HOM HCLIM**
- Is it possible to use modern techniques?
- How far back in time can we homogenize using modern homogenization methods?
- The first and longest time series in the HCLIM dataset is Paris from 1658. But can it be homogenized?
- The answer is Yes
- from 1757
- And the oldest station that can be homogenized in this project is Uppsala (1722)

A circular inset image on the left side of the slide shows a close-up of a spiral-bound notebook. The pages are lined and have some numbers (16, 17, 18, 30) printed on them. The spiral binding is visible on the left edge.

Outlook

- Finishing the homogenization (next paper)
- Which method are best for the early instrumental data?

- Extreme Events?
- My goal now:
- Paper + Dissertation report and the Defense during the autumn 23



PANGAEA.

Data Publisher for Earth & Environmental Science

<https://doi.pangaea.de/10.1594/PANGAEA.940724>

scientific **data**

[Explore content](#) ▾ [About the journal](#) ▾ [Publish with us](#) ▾

[nature](#) > [scientific data](#) > [data descriptors](#) > article

Data Descriptor | [Open Access](#) | [Published: 19 January 2023](#)

The global historical climate database HCLIM

[Elin Lundstad](#) , [Yuri Brugnara](#), [Duncan Pappert](#), [Jérôme Kopp](#), [Eric Samakinwa](#), [André Hürzeler](#), [Axel Andersson](#), [Barbara Chimani](#), [Richard Cornes](#), [Gaston Demarée](#), [Janusz Filipiak](#), [Lydia Gates](#), [Gemma L. Ives](#), [Julie M. Jones](#), [Sylvie Jourdain](#), [Andrea Kiss](#), [Sharon E. Nicholson](#), [Rajmund Przybylak](#), [Philip Jones](#), [Daniel Rousseau](#), [Birger Tinz](#), [Fernando S. Rodrigo](#), [Stefan Grab](#), [Fernando Domínguez-Castro](#), ... [Stefan Brönnimann](#) [+ Show authors](#)

Scientific Data **10**, Article number: 44 (2023) | [Cite this article](#)



European Research Council

Established by the European Commission

Thanks for the attention!