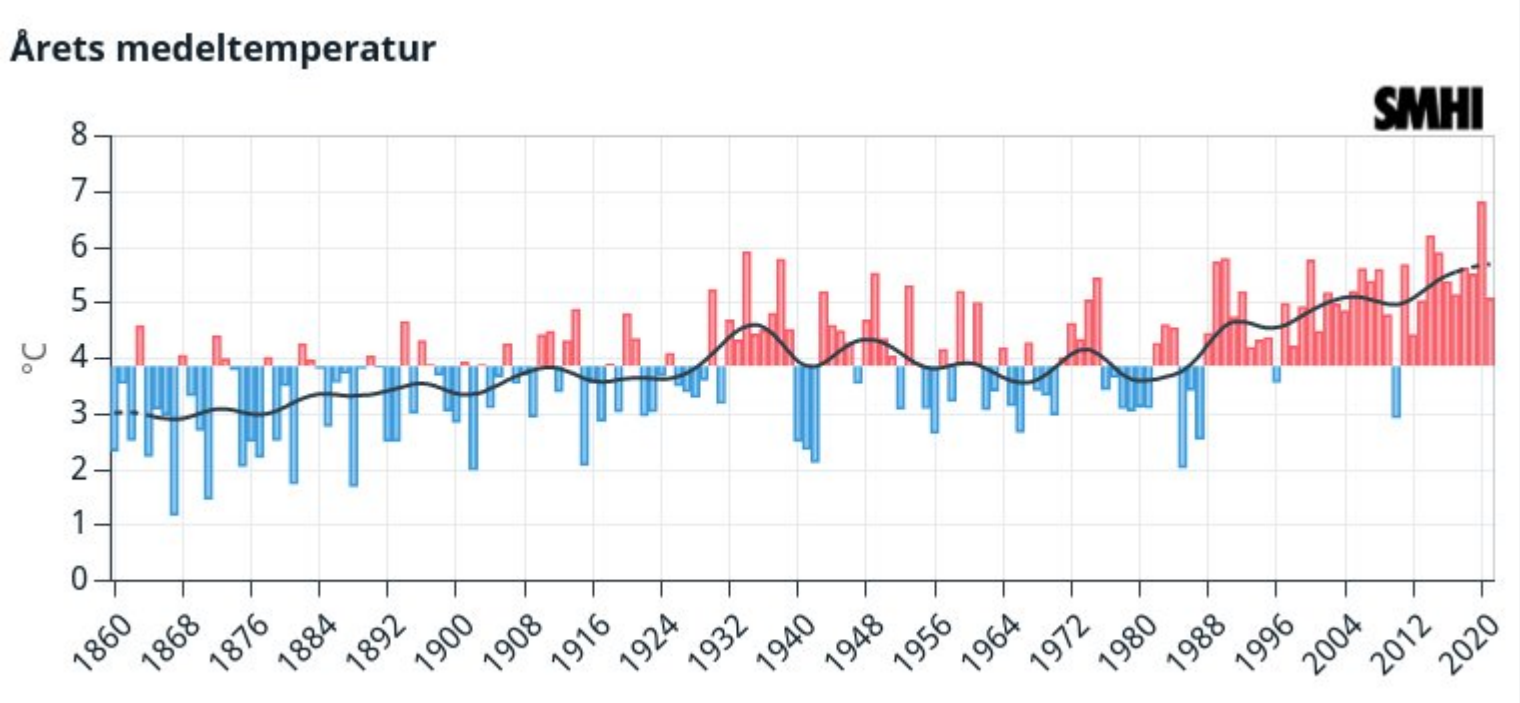
SMHI

EMS 2022, BONN

HOMOGENISATION OF SWEDISH MEAN MONTHLY TEMPERATURE SERIES

Background

Climate indicator, Temperature



SMHI

Motivation

- In 2014 the monthly average temperatures of 35 selected time series (1860–2013) of the Swedish Meteorological and Hydrological Institute (SMHI) data set was homogenised using the SNHT method
- Old data are continuously digitised, erroneous data corrected, and new observations added
- The method is too labour intensive for annual updates
- The data set is too small for regional analysis

SMHI

Objective

A new homogenisation method was to be developed:

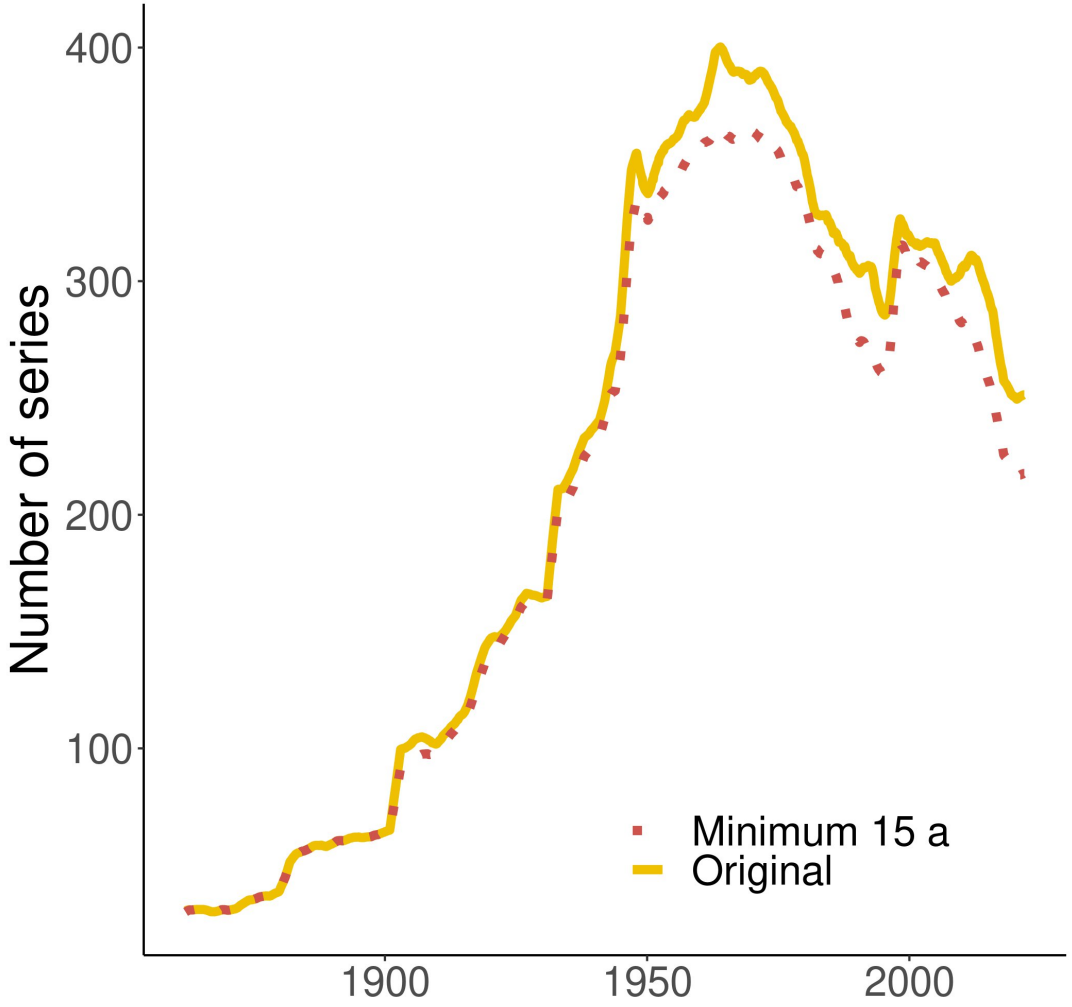
- Annual updates
- Large dataset

Approach

- Merged time series
- Automatic method
- No black box

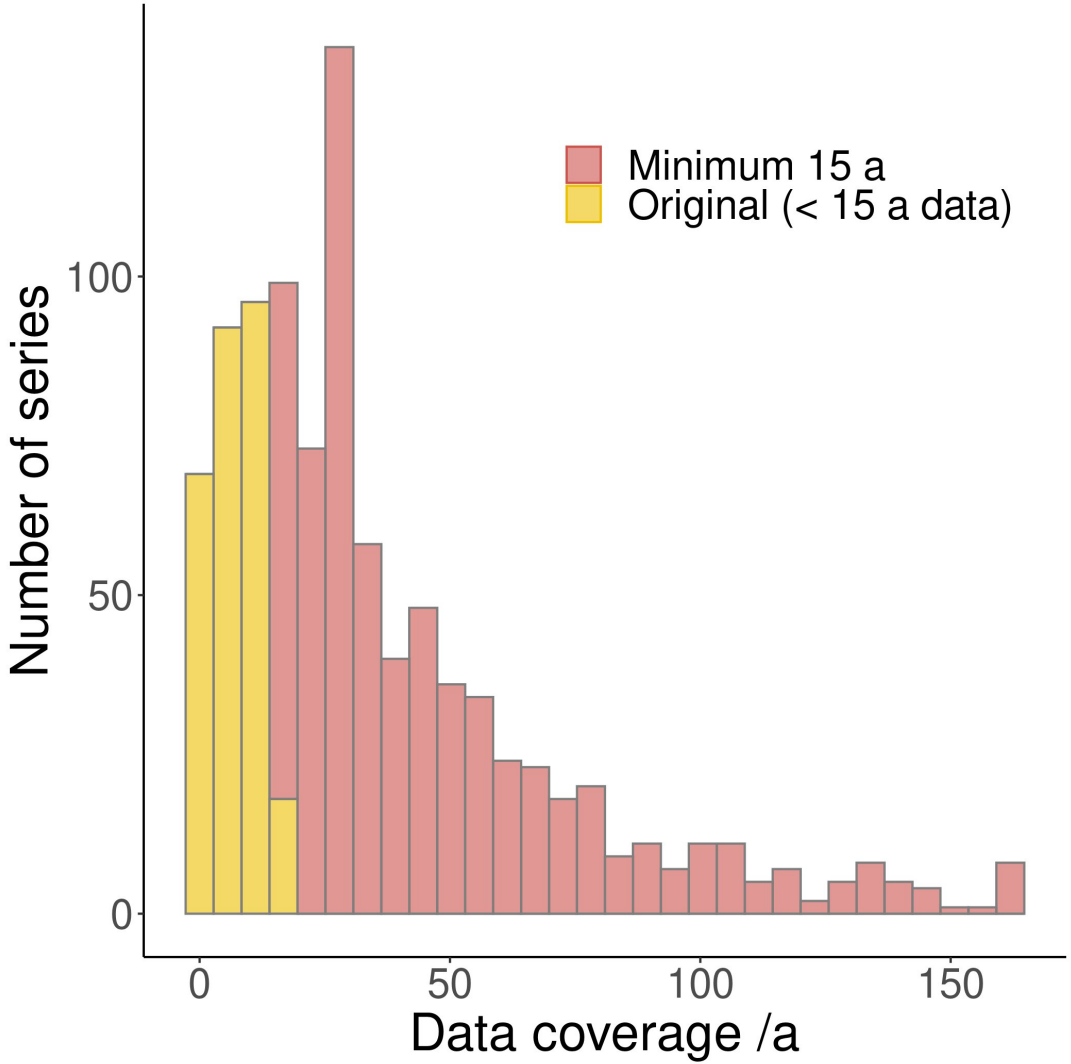
→ HOMER

Description of original data set



SMHI

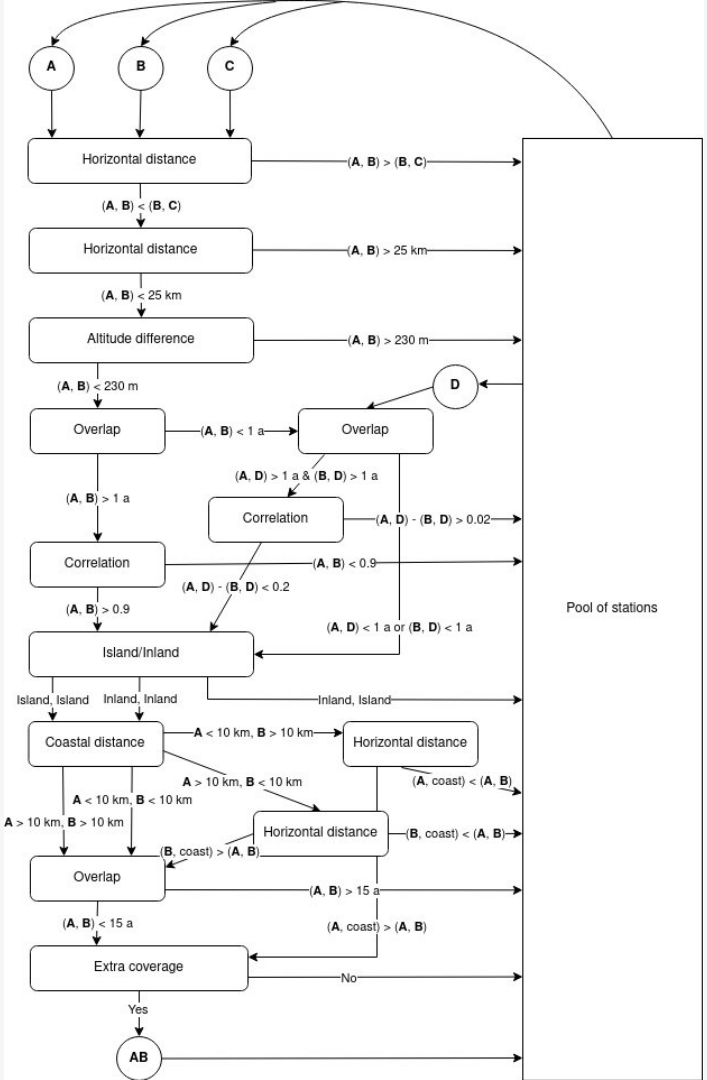
- Number of active stations with digitised monthly average temperature data in SMHI's observational database (yellow solid line)
- More than 400 stations in the early 1960's
- HOMER cannot use time series shorter than 15 years, the red dotted line, show only the stations with at least 15 years of data
- More than 370 stations in the early 1960's



SMHI

- The number of series of a certain length (in total 961 series)
- The red bars show the part of the data that can be used by HOMER (in total 686 series)
- The top of the bars represent the total amount
- 173 series have at least 60 years of data (length of two normalperiods)

Merging of time series



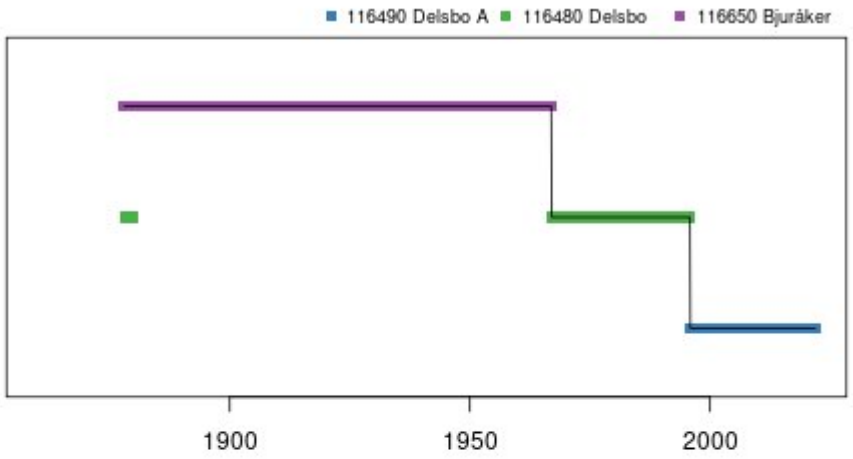
SMHI

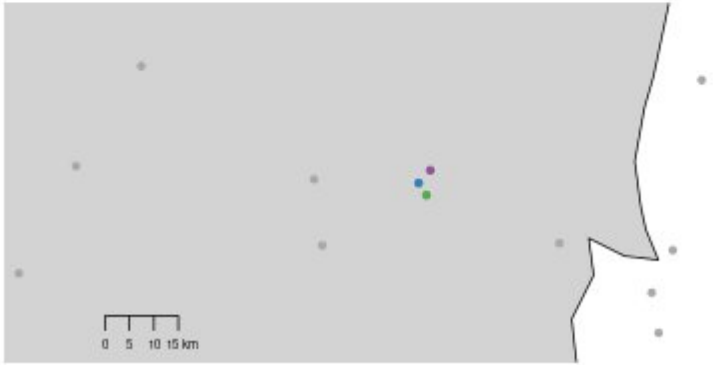
Principles for merging of stations:

i. High data coverage
ii. Proximity
iii. Correlation
iv. Islands/near coastlines/inland
v. Recency
vi. Few shiftpoints
vii. Overlap < 15 years (homogenisation)

Two stages

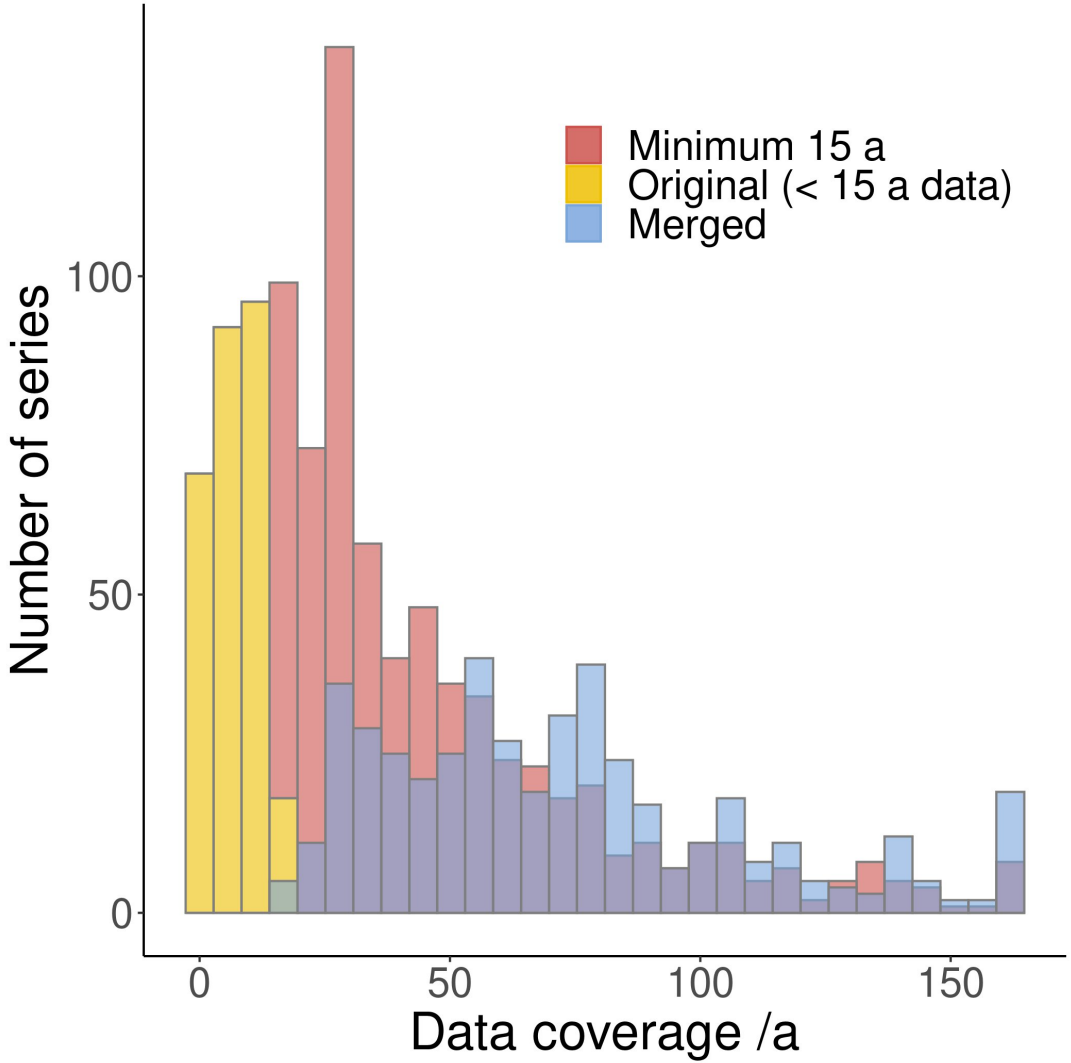
1. Grouping
2. Merging





Example of merged station (Delsbo A):

- Three neighbouring stations
- Delsbo was part of the extensive automation of weather stations in the late 1990's



SMHI

- The number of series of a certain length in the merged data set (blue bars, in total 456 series)
- Merging shifts the distribution towards longer time series
- 256 series have at least 60 years of data (length of two normalperiods), compared with 173 in the original (red bars)

Homogenisation

Method

- The merged network is homogenised with HOMER via the fully automatic script Bart (Joelsson et al. 2022).

Received: 28 December 2020 | Revised: 14 September 2021 | Accepted: 15 September 2021 | Published on: 19 October 2021

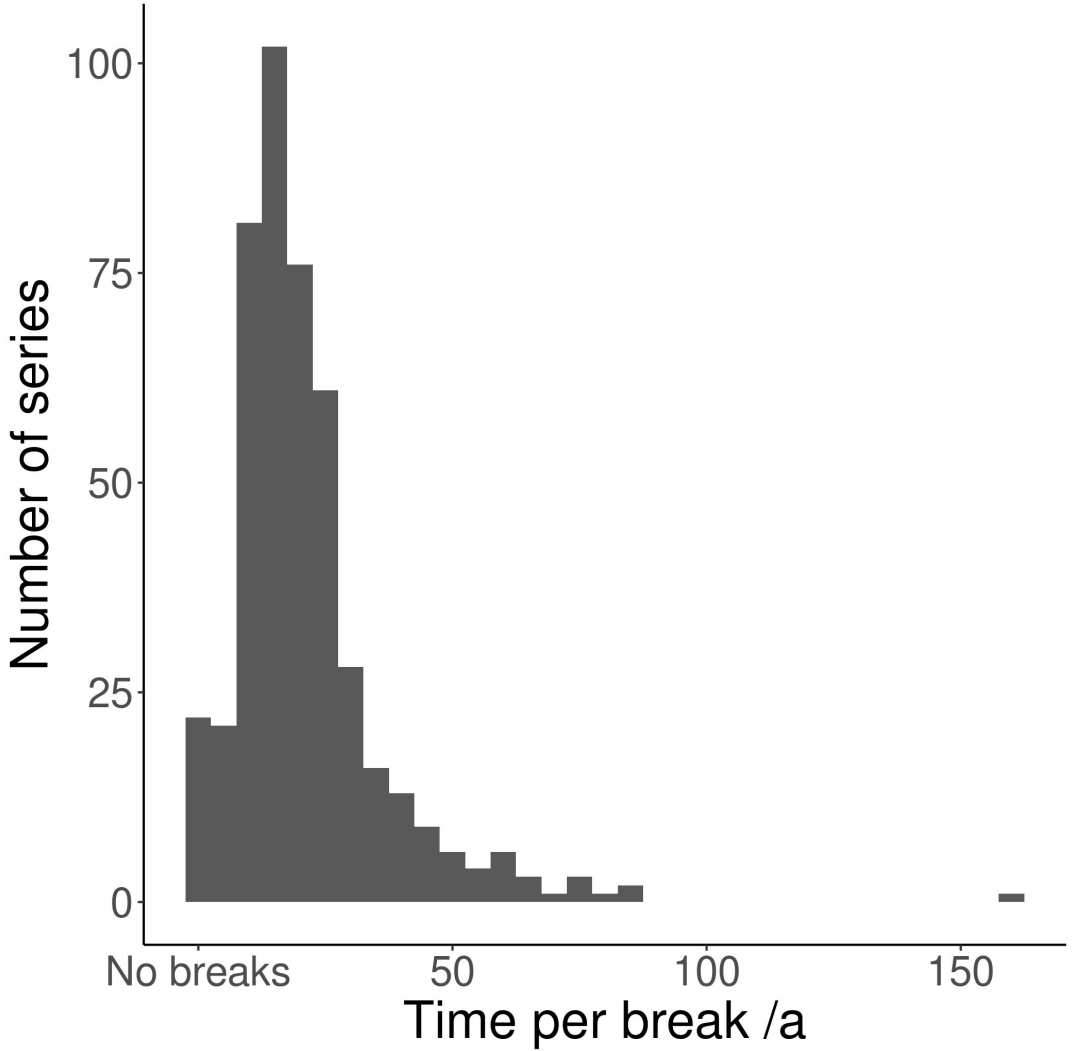
DOI: 10.1002/joc.7394

RESEARCH ARTICLE

International Journal of Climatology | RMetS

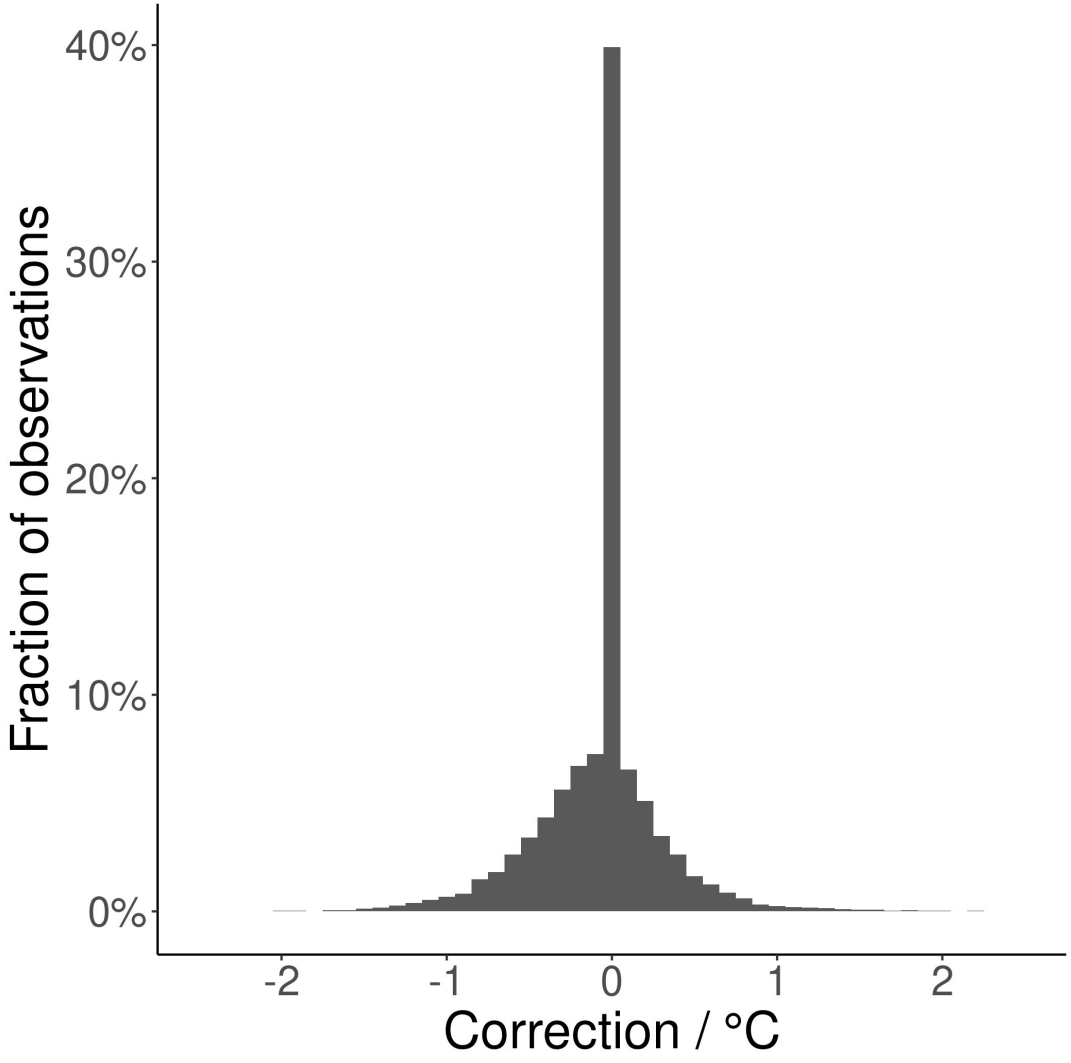
Automation and evaluation of the interactive homogenization tool HOMER

L. Magnus T. Joelsson | Christophe Sturm | Johan Södling | Erik Engström | Erik Kjellström



SMHI

- Histogram of average time between two breaks (number of years of data / number of detected breaks for each station)
- Median (average) time between breaks is 17 years
- 22 time series are found to be homogeneous
- 37 % of breaks are supported by meta data (or the merging scheme)

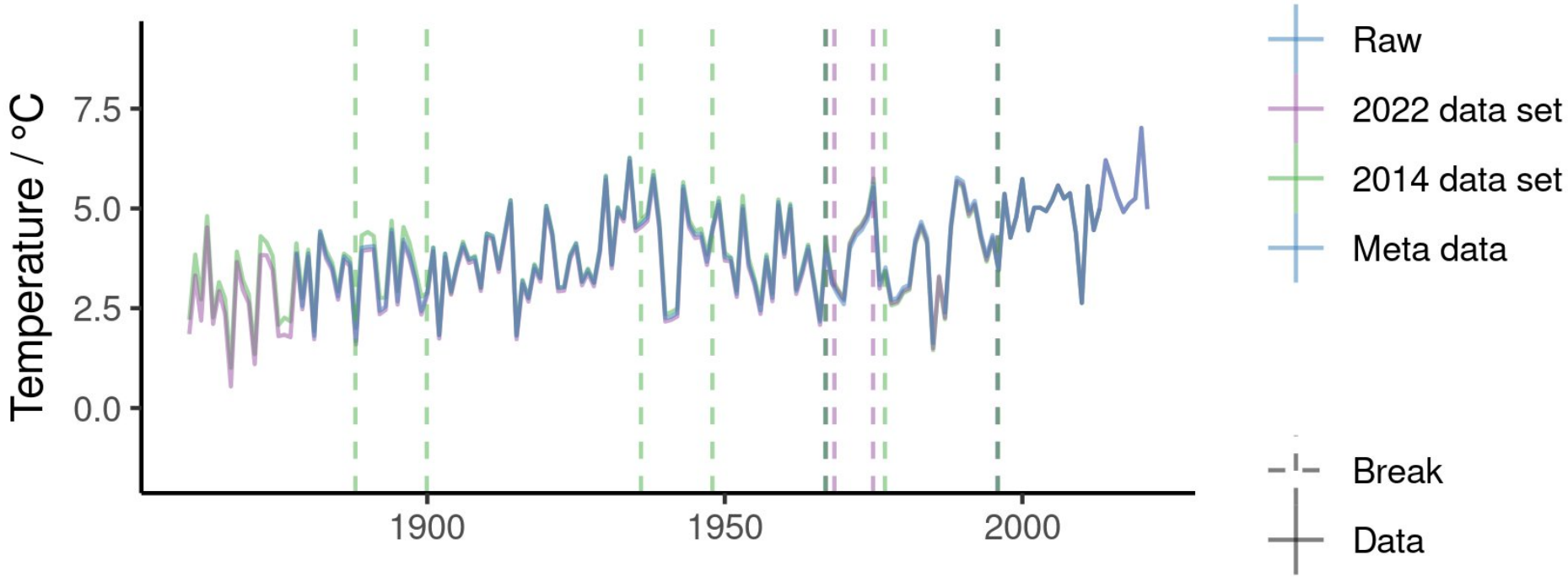


SMHI

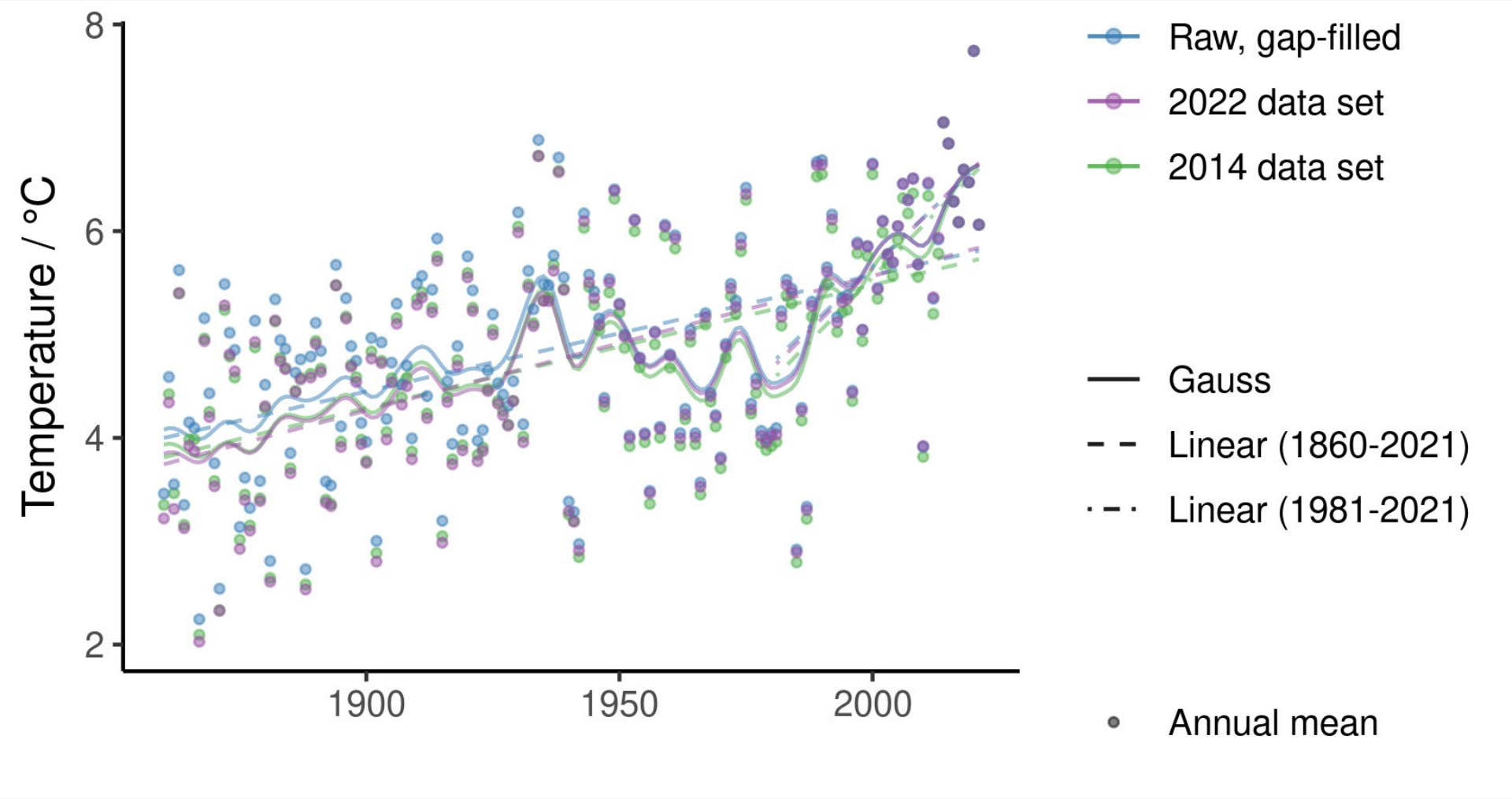
Histogram of corrections:

- 60 % data points corrected
- 42 % corrected by 0.5 ºC or less
- 3 % corrected by 1 ºC or more
- 37 % are negatively corrected
- 24 % are positively corrected

Example of homogenised time series: Delsbo A



Climatological analysis


Temperature / °C
8
6
4
2
1900
1950
2000
Raw, gap-filled
2022 data set
2014 data set
Gauss
Linear (1860-2021)
Linear (1981-2021)
Annual mean

Conclusion and Outlook

SMHI

Conclusion

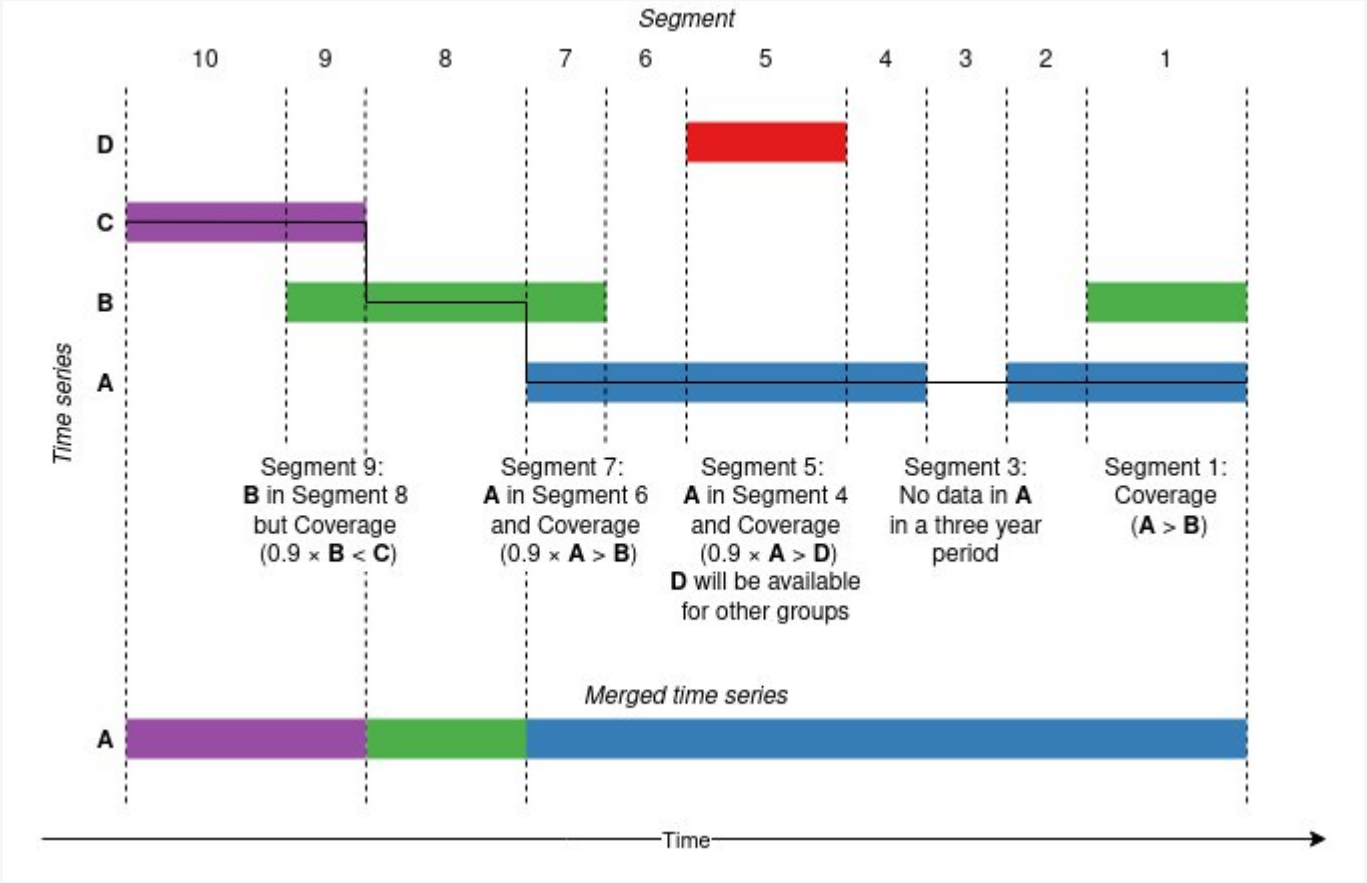
- A monthly mean temperature data set (1860–2021) for 456 Swedish time series is compiled by automatic merging from 836 original time series
- The data set is homogenised by HOMER
- Median time between two breaks is found to be 17 years
- Calculatiuons of temperature trend using raw data underestimates the warming over the entire period, albeit not significantly

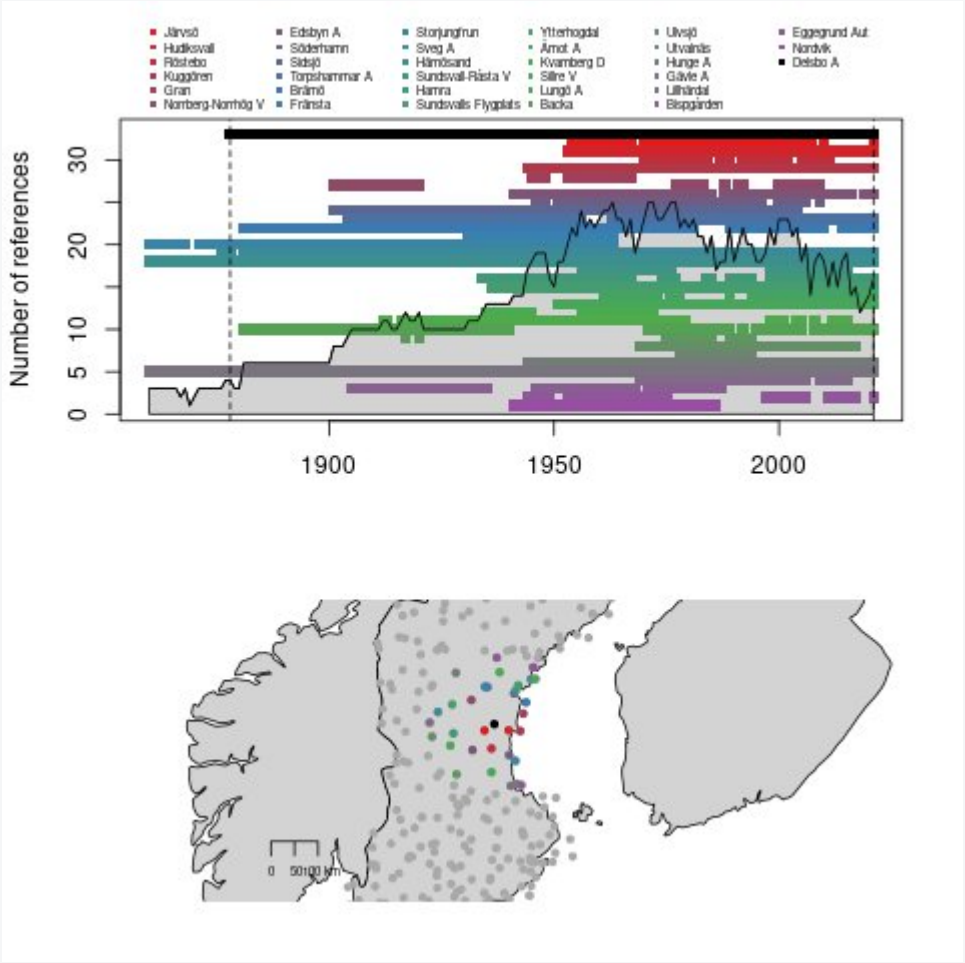
Outlook

- Merging and homogenisation of other variables (e.g. monthly precipitation, monthly mean of daily extreme temperatures
- Further development of merging and homogenisation scripts

Extra

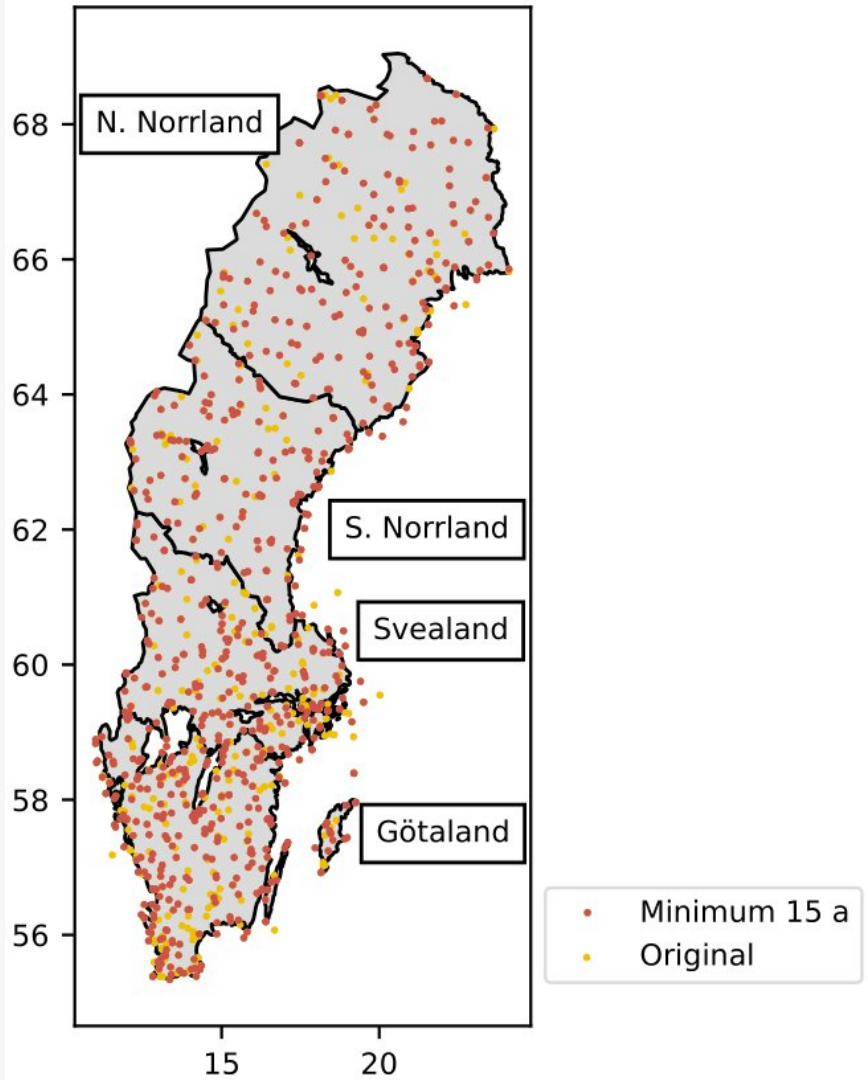
2. Merging


Segment
10
9
8
7
6
5
4
3
2
1
D
C
B
A
Time series
Segment 9:
B in Segment 8
but Coverage
(0.9 × B < C)
Segment 7:
A in Segment 6
and Coverage
(0.9 × A > B)
Segment 5:
A in Segment 4
and Coverage
(0.9 × A > D)
D will be available
for other groups
Segment 3:
No data in A
in a three year
period
Segment 1:
Coverage
(A > B)
Merged time series
A
Time



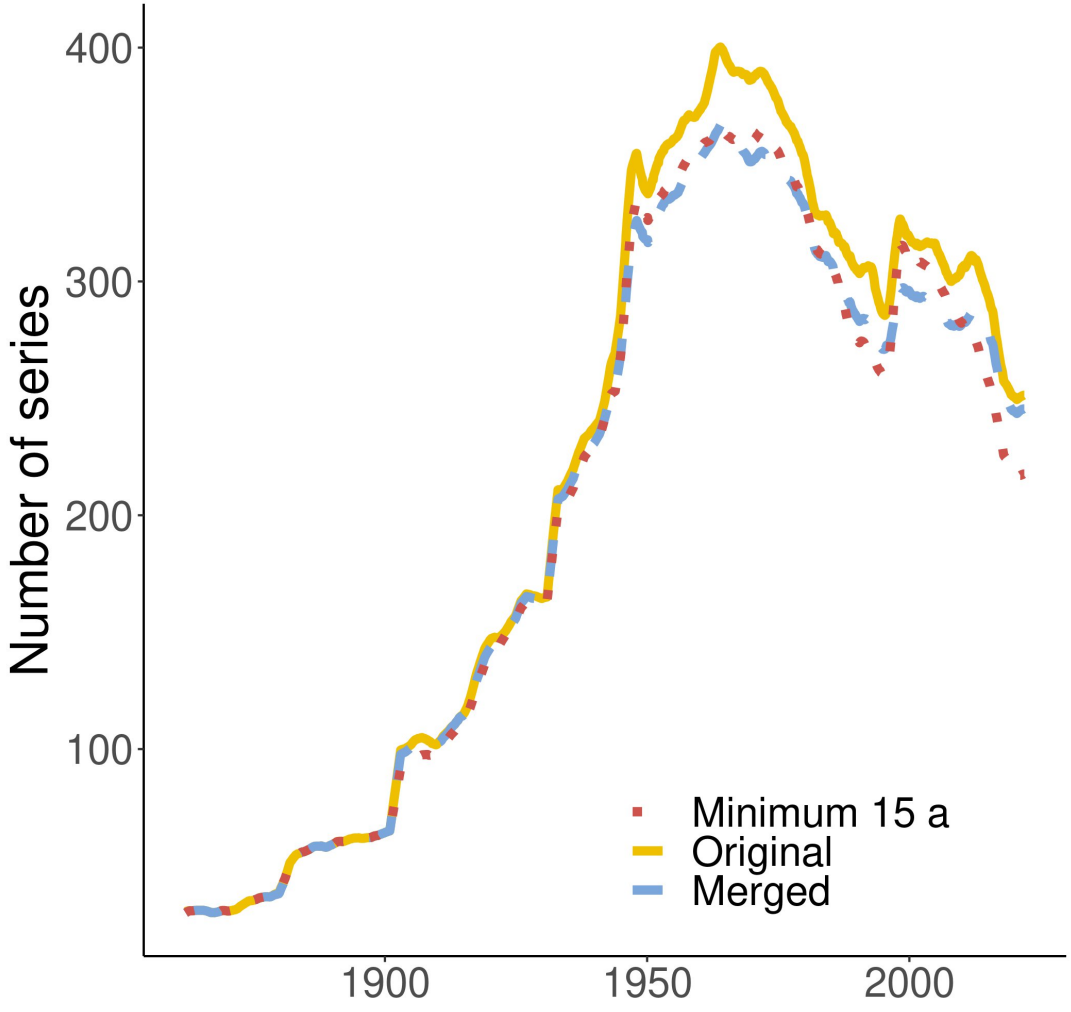
Hybrid reference selection:

- Both correlation and distance taken into account
- At least 4 references for the entire period of the candidate station period
- 32 references (maximum)
- Example of reference network (for candidate station Delsbo A)



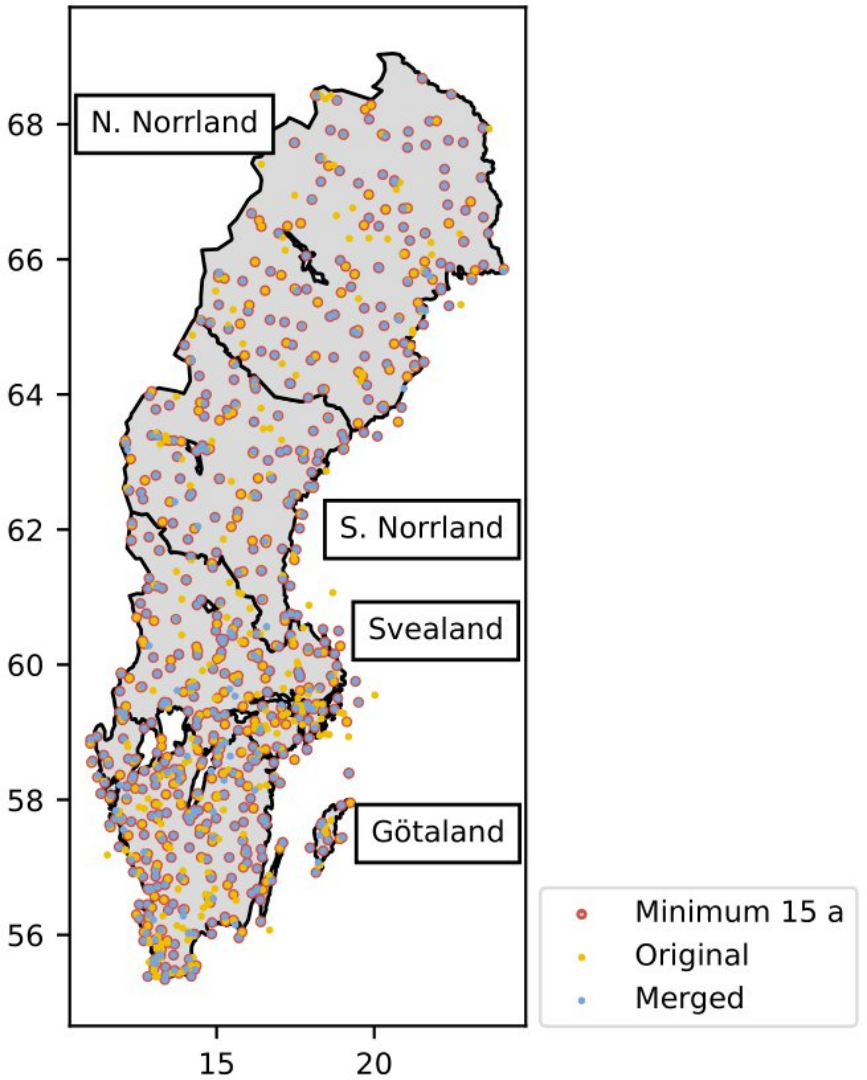
SMHI

- The geographical distribution of all histrorical and active stations with digitised monthly average temperature data (yellow dots)
- Red dots indicate the stations with at least 15 years of data
- The traditional lands of Sweden are included



SMHI

- Number of active stations in the merged data set
- More than 370 stations in the early 1960's
- The merged data set has at least 25 more series than the Minimum 15 a data set from 2017
- The Minimum 15 a data set has at least 20 more series than the merged data set in the late 1990's



SMHI

- Geographical distribution of merged stations (blue points)
- Original data set (yellow + blue points) is included for comparison
- Stations with (originally) at least 15 years of data (red circles)
- The merged data comapre well with the Minimum 15 a data set