

**11th Seminar for Homogenization and Quality Control in  
Climatological Databases and 6th Interpolation Conference jointly  
organized with the 14th EUMETNET Data Management Workshop**

# **Preliminary results on a new dataset of daily observations from a dense network of weather stations covering the Extended Alpine Region**

**Giulio Bongiovanni, Michael Matiu, Alice Crespi, Anna Napoli,  
Bruno Majone, Dino Zardi**



**PhD SDC**  
SUSTAINABLE DEVELOPMENT  
AND CLIMATE CHANGE

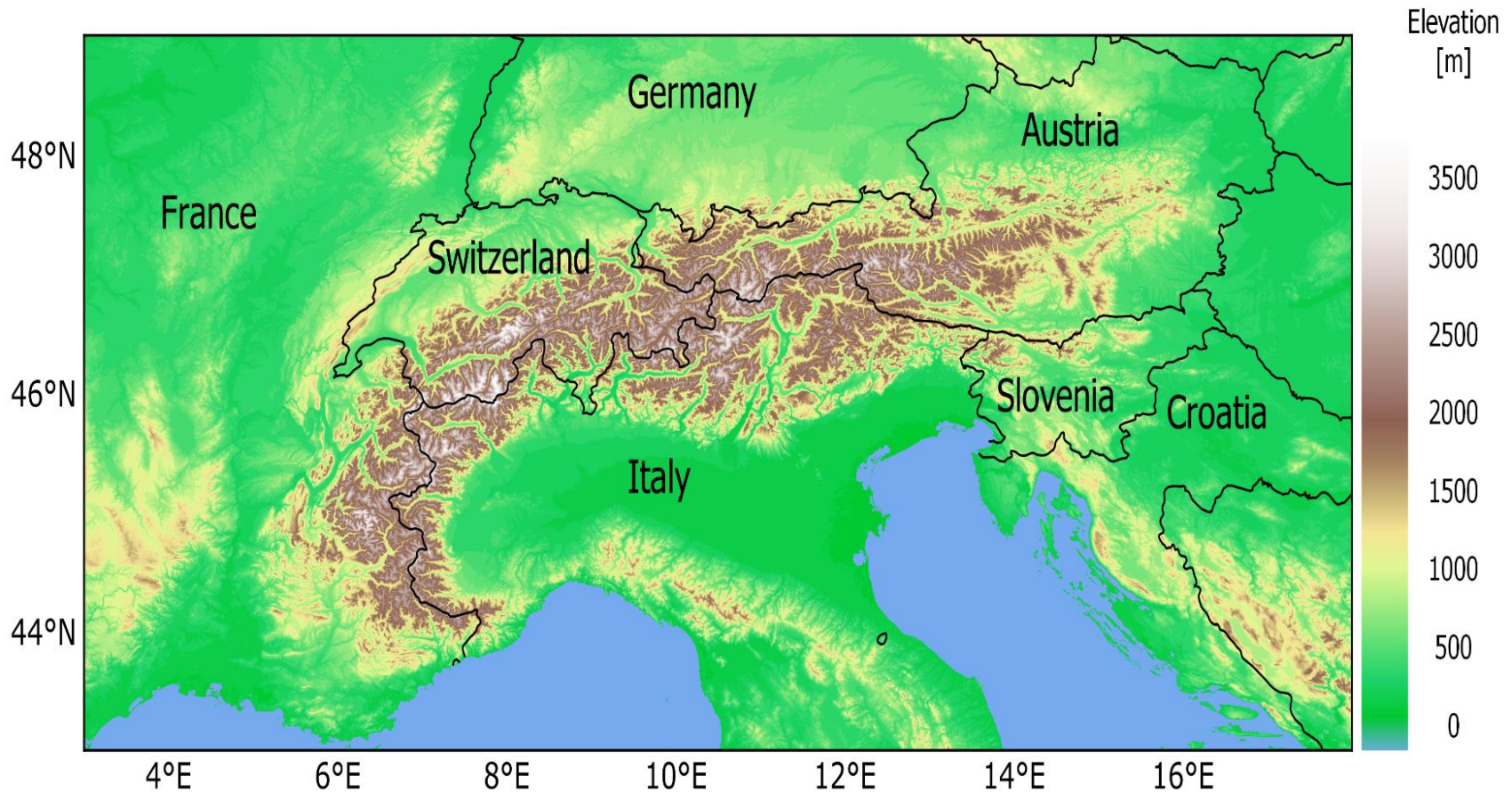


**IUSS**

Scuola Universitaria Superiore Pavia



# Study Area



**Extended Alpine Region (EAR)**  
**3 – 18° E / 43 – 49° N**

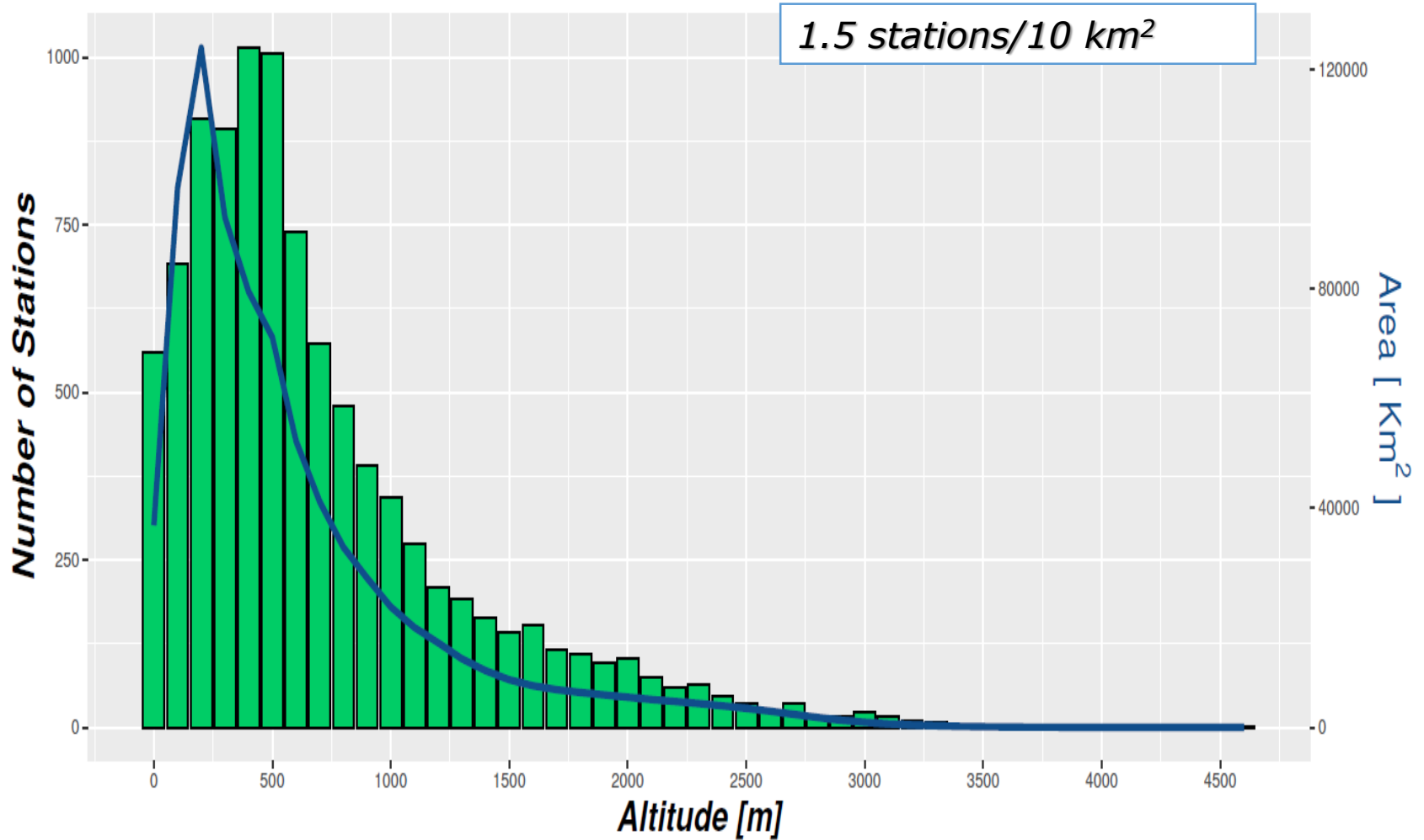
# Dataset Features

- **Higher stations density ( $\sim 10K$ )**
- **High elevations better represented**
- **Multi-variable observations:**
  - **Temperature (T, Tmin, Tmax)**
  - **Precipitation (TP)**
  - **Relative Humidity (RH)**
  - **Snow Depth (HS)**
  - **Wind (WS, WD)**
  - **Radiation (GS)**
  - **Pressure (P)**
- **Daily time frequency ( $\sim 1B$ )**
- **Updated to present days**

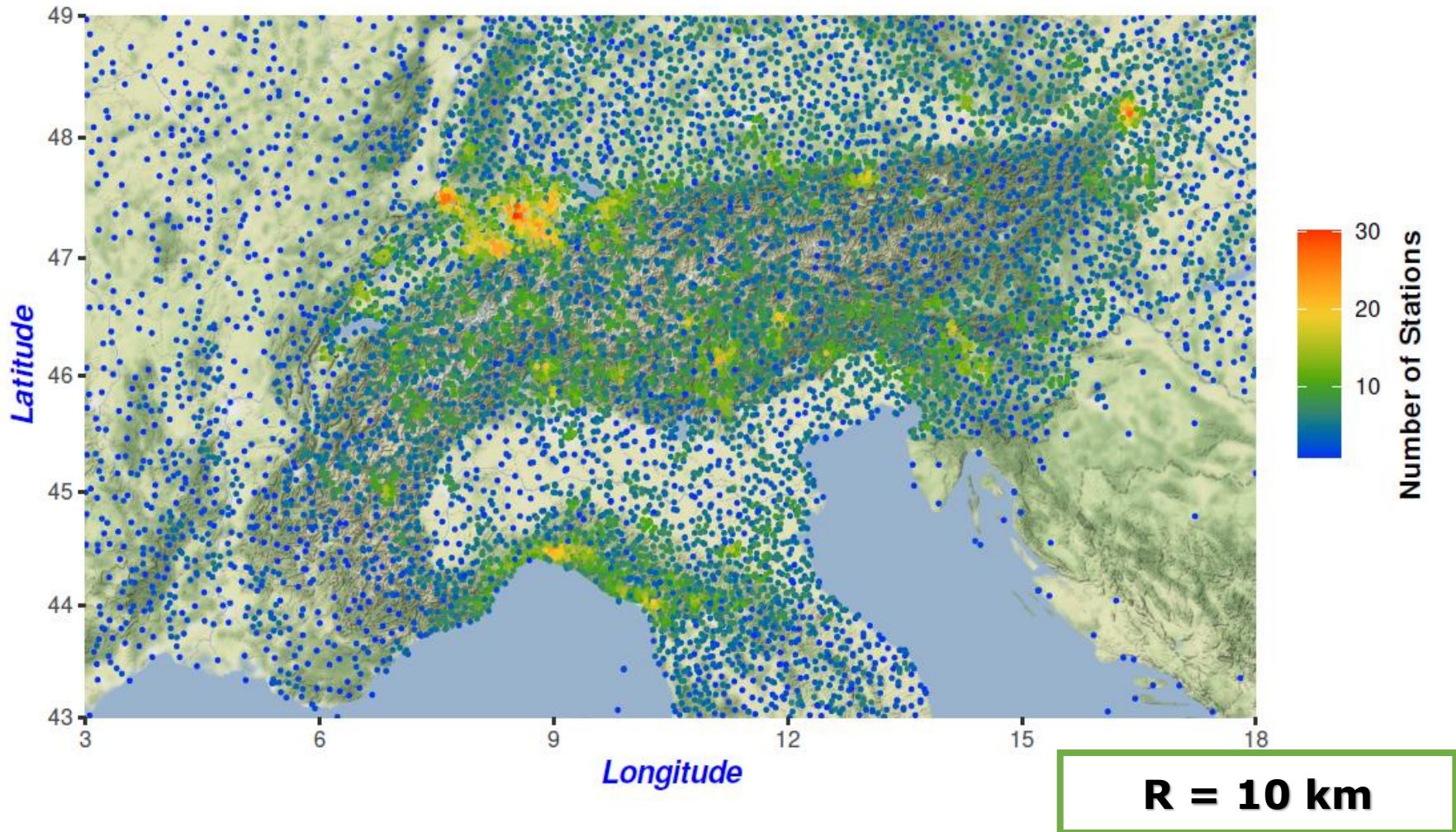
# Data Collection

REGION NAME	PROVIDER	STARTING YEAR	N. STATIONS
<i>Austria</i>	Zentralanstalt für Meteorologie und Geodynamik (ZAMG)	1852	508
	Hydrographische Archivdaten Österreichs (eHYD)	1969	1,193
<i>Croatia</i>	Državni HidroMeteorološki Zavod (DHMZ)	1857	46
<i>Emilia Romagna</i>	Agenzia Regionale per la Prevenzione, l'Ambiente e l'Energia dell'Emilia-Romagna (ARPAE)	1961	540
<i>France</i>	Météo-France	1922	1,124
<i>Friuli Venezia Giulia</i>	Agenzia Regionale per la Protezione dell'Ambiente del Friuli Venezia Giulia (ARPA FVG)	1991	188
<i>Germany</i>	Deutscher WetterDienst (DWD)	1781	1,130
<i>Bosnia Herzegovina</i>	Global Historical Climatology Network (GHCN)	2001	2
<i>Italy</i>	European Climate Assessment & Climate (ECA&D)	1813	75
	Meteo Aeronautica Militare (Meteo AM)	1813	50
	MeteoMont	1980	70
<i>Liguria</i>	Agenzia Regionale per la Protezione dell'Ambiente Ligure (ARPAL)	2002	206
<i>Lombardy</i>	Agenzia Regionale per la Protezione dell'Ambiente della Lombardia (ARPA Lombardia)	1763	480
<i>Marche</i>	Regione Marche	1951	120
<i>Piedmont</i>	Agenzia Regionale per la Protezione Ambientale del Piemonte (ARPA Piemonte)	1913	361
<i>Czech Rep.</i>	Český HydroMeteorologický Ústav (CHMU)	1961	76
<i>Slovakia</i>	Slovenský HydroMeteorologický Ústav (SHMU)	1991	103
<i>Slovenia</i>	Agencija Republike Slovenije za Okolje (ARSO)	1960	470
<i>Sudtirol</i>	Provincia Autonoma di Bolzano	1920	247
<i>Switzerland</i>	MeteoSwiss	1863	1,446
<i>Tuscany</i>	Regione Toscana	1916	393
<i>Trentino</i>	MeteoTrentino	1920	219
	Fondazione Edmund Mach	1983	9
<i>Umbria</i>	Regione Umbria	1916	44
<i>Hungary</i>	Országos Meteorológiai SZolgálat (OMSZ)	1901	206
<i>Aosta Valley</i>	Regione Autonoma Valle d'Aosta	1866	80
<i>Veneto</i>	Agenzia Regionale per la Prevenzione e protezione Ambientale del Veneto (ARPAV)	1956	312
			9,698

# Dataset Elevation Distribution



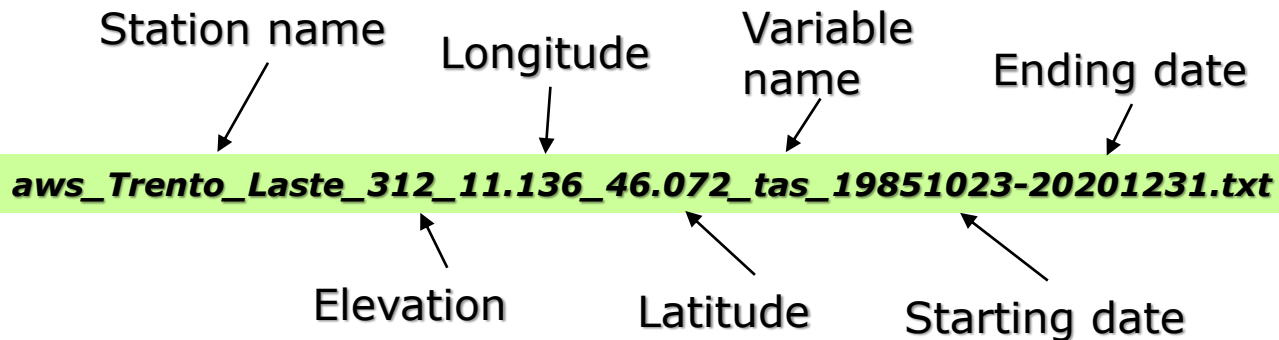
# Dataset Spatial Distribution



# Data Pre-processing

- Data cleaning
- Data storing

	Date	Value
1	1985-10-23	11.1
2	1985-10-24	10.7
3	1985-10-25	7.4
4	1985-10-26	7.0
5	1985-10-27	6.8
6	1985-10-28	6.7
7	1985-10-29	6.1
8	1985-10-30	9.0
9	1985-10-31	10.5
10	1985-11-01	11.0
11	1985-11-02	9.8
12	1985-11-03	7.8
13	1985-11-04	6.5
14	1985-11-05	5.7
15	1985-11-06	9.6
16	1985-11-07	5.7
17	1985-11-08	4.8
18	1985-11-09	5.9
19	1985-11-10	8.3
20	1985-11-11	10.6
21	1985-11-12	5.0
22	1985-11-13	2.9
23	1985-11-14	5.5
24	1985-11-15	1.0
25	1985-11-16	-1.2
26	1985-11-17	-1.2
27	1985-11-18	0.1
28	1985-11-19	2.2
29	1985-11-20	2.1
30	1985-11-21	2.3
31	1985-11-22	2.1
32	1985-11-23	2.0
33	1985-11-24	2.7
34	1985-11-25	0.8
35	1985-11-26	0.3
36	1985-11-27	1.1
37	1985-11-28	-0.4
38	1985-11-29	-0.3
39	1985-11-30	-0.1
40	1985-12-01	2.0
41	1985-12-02	2.0
42		



**Data files example**

# Quality Control Tests

*Cerlini et al. (2020), Durre et al. (2010), Fioravanti et al. (2016)*

## Internal consistency:

- Range check
- Multi-temperature check:
  - $T_{min} < T < T_{max}$
  - $0.01 \text{ } ^\circ\text{C} < T_{min} - T_{max} < 30 \text{ } ^\circ\text{C}$
- Statistical outliers:
  - MAD Test:  $\frac{|x - \text{median}(x)|}{MAD} \leq 3$
  - P95 Test:
    - $TP < 9p_{95} (T \geq 0 \text{ } ^\circ\text{C})$
    - $TP < 5p_{95} (T < 0 \text{ } ^\circ\text{C})$

## Temporal consistency:

- Repeated values (< 5 days) and zeros (< 180 days, TP)
- Temporal step (< 20 °C)

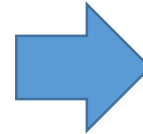
## Manual checking



# Spatial Consistency

## Reference (neighbouring) stations selection:

- Distance:  $D \leq 50$  km
- Elevation difference:  $dh \leq 100$  m
- Correlation:  $r \geq 0.8$
- Time series completeness: valid data  $\geq 80\%$



**$3 \leq N. \text{ Stations} \leq 10$**

## Tests:

- Isolated wet/dry reports (TP)
- Corroboration test
- Anomaly-based reconstruction test

# Isolated wet/dry reports

Isotta et al. 2014

(Distance <  $D_{th}$ )

Oct-Apr:  $c=2.7$ ,  $D_{th} = 15$  Km  
May-Sep:  $c=3.2$ ,  $D_{th} = 20$  Km

$D_{min}$ : nearest station distance

## Isolated wet reports

- Station test:  $tp > \left(0.3 + c \frac{D_{min}}{15000}\right) \rightarrow \llcorner \text{wet} \llcorner$
- Reference stations:  $tp < 0.3$  mm  $\rightarrow$   
**station test =  $\llcorner \text{isolated wet} \llcorner$**

## Isolated dry reports

- Station test:  $tp < 0.3$  mm  $\rightarrow \llcorner \text{dry} \llcorner$
- Reference stations:  $tp > \left(0.8 + c \frac{D_{min}}{15000}\right) \rightarrow$   
**station test =  $\llcorner \text{isolated dry} \llcorner$**

# Spatial Consistency Tests

Durre et al. 2010, Curci et al. 2020

$X_{anom}$  -> moving window daily climatologic anomalies

- window length: 15 (temperature) or 30 (precipitation) days

$X_{test}$  -> tested time series

$X_{ref}$  -> reference time series

## Corroboration test

- $|x_{test,anom} - x_{ref,anom}| \leq (10^{\circ}\text{C}; 50 \text{ mm})$  for each  $x_{ref}$
- If no  $x_{ref}$  corroborate tested time series -> flagged day

# Spatial Consistency Tests

Matiu et al. 2021

## Anomaly-based reconstruction test

$x'$  -> tested «weighted gap-reconstructed» time series:

- weight =  $cf * w$

cf: difference (temperature) or ratio (precipitation) between time series

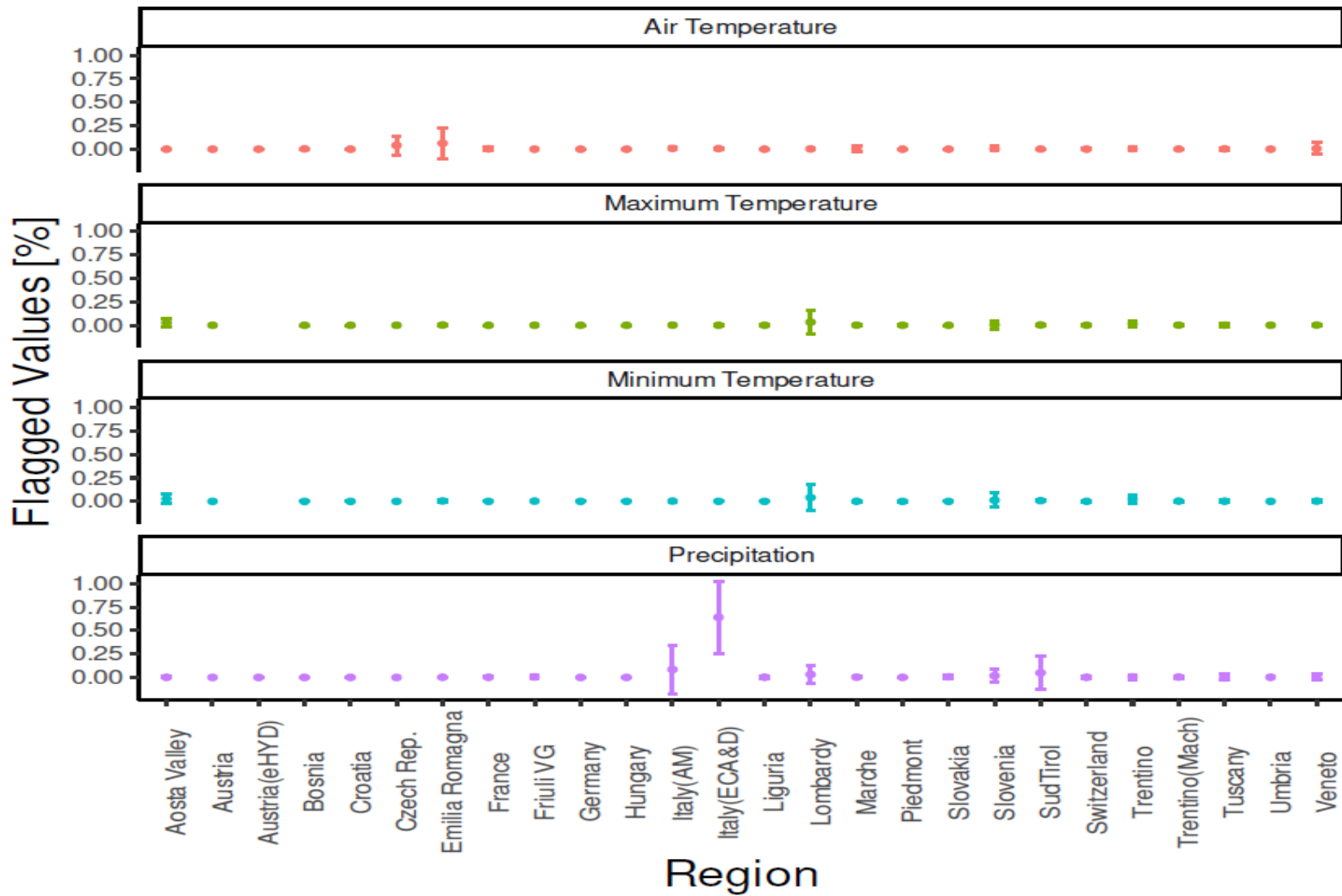
$$w = e^{-\frac{1-r^2}{\tau_c^2 \log 2}}$$

$$X = \overline{x'_{test} - x'_{test,anom} + x_{ref,anom}}$$

*tested anomaly-reconstructed time series*

- $|(X - x_{test})| \leq (10^\circ\text{C}; 50 \text{ mm})$
- $\frac{|(X - x_{test})|}{x_{test}} \leq 0.5$  (**only TP**)

# Quality Control Results



# Homogenization

## Single Stations (Monthly) Homogenization

### Reference stations selection:

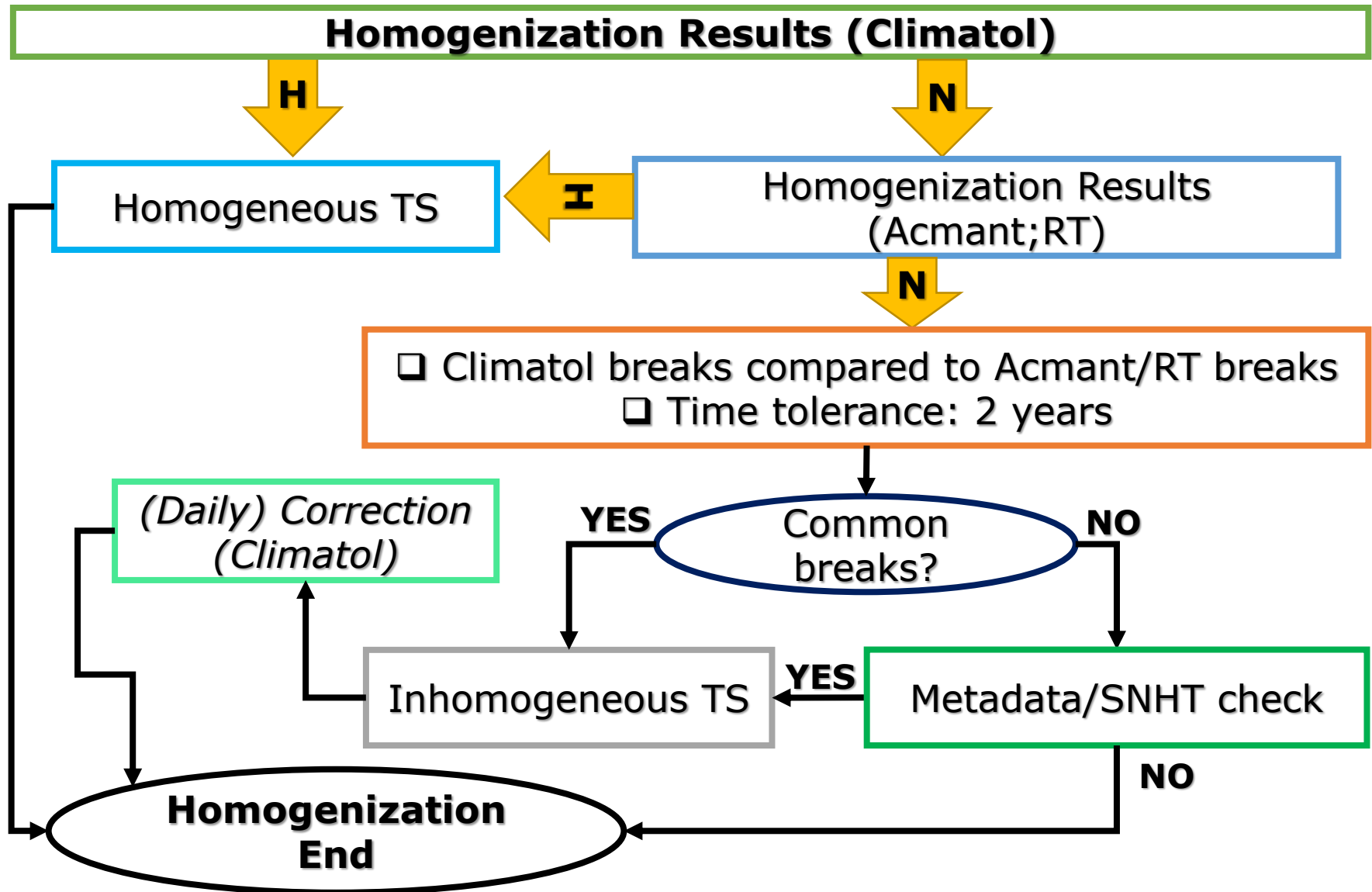
1. Distance < 100 Km
2. Correlation > 0.9
3. Height difference < 300 m
4. Number of references: depending time series completeness (Valid data > 70%, 4÷25 stations)
5. Time series length > 20 years

**CLIMATOL**

**ACMANT (v4.4)**

**RH TEST (RT)**

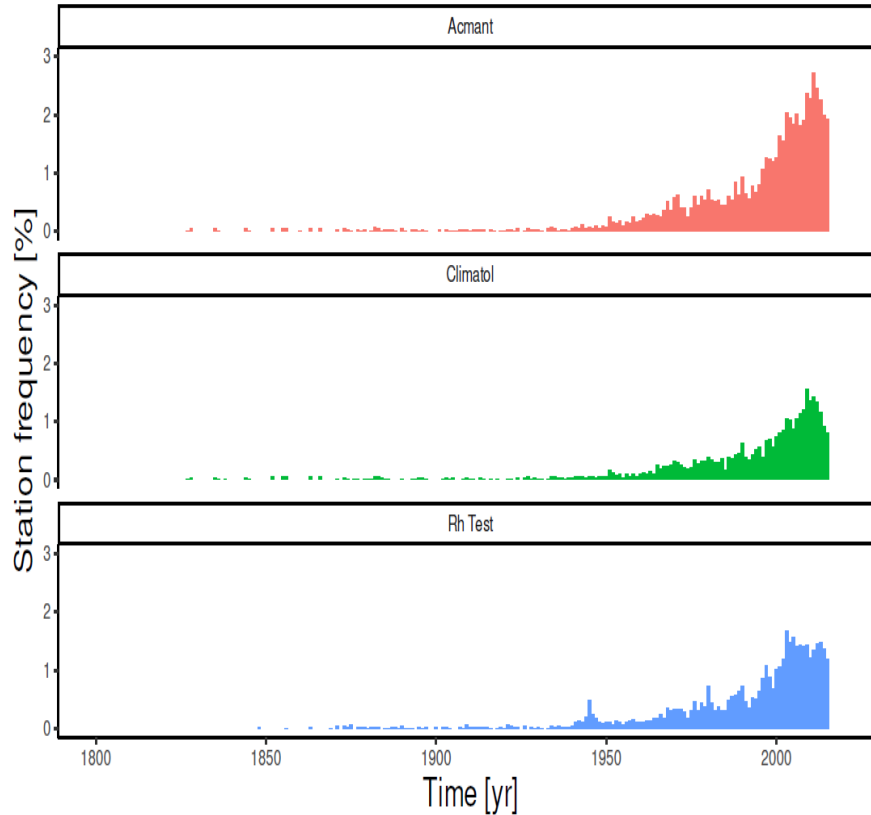
# Homogenization Workflow



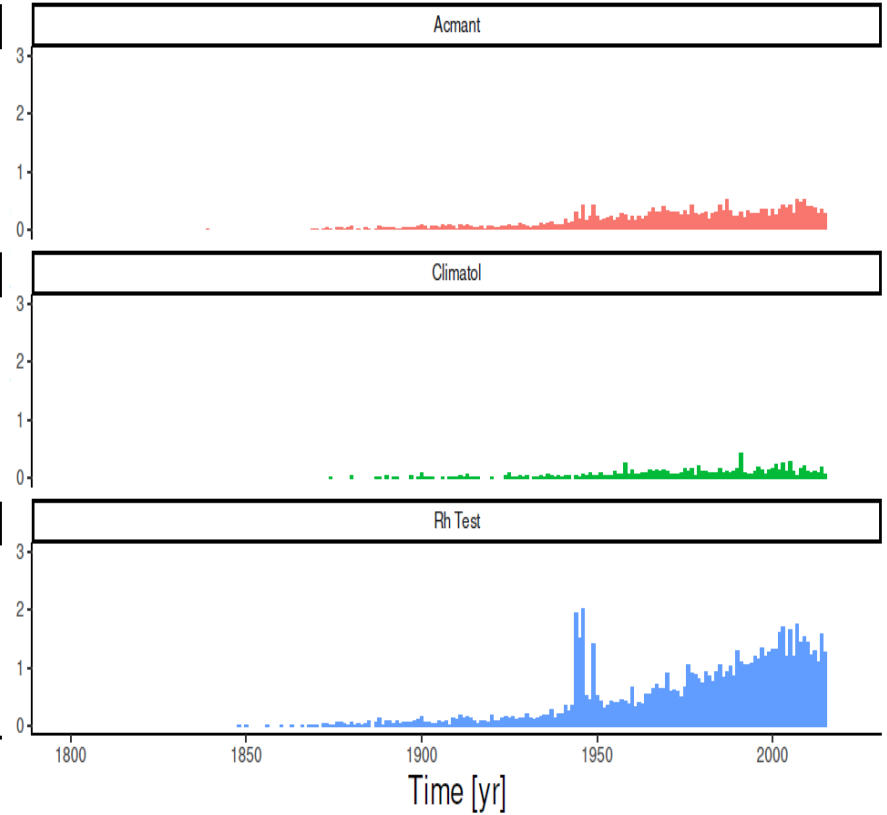
# Homogenization Results

## TIME DISTRIBUTION OF BREAKS

*Air Temperature*



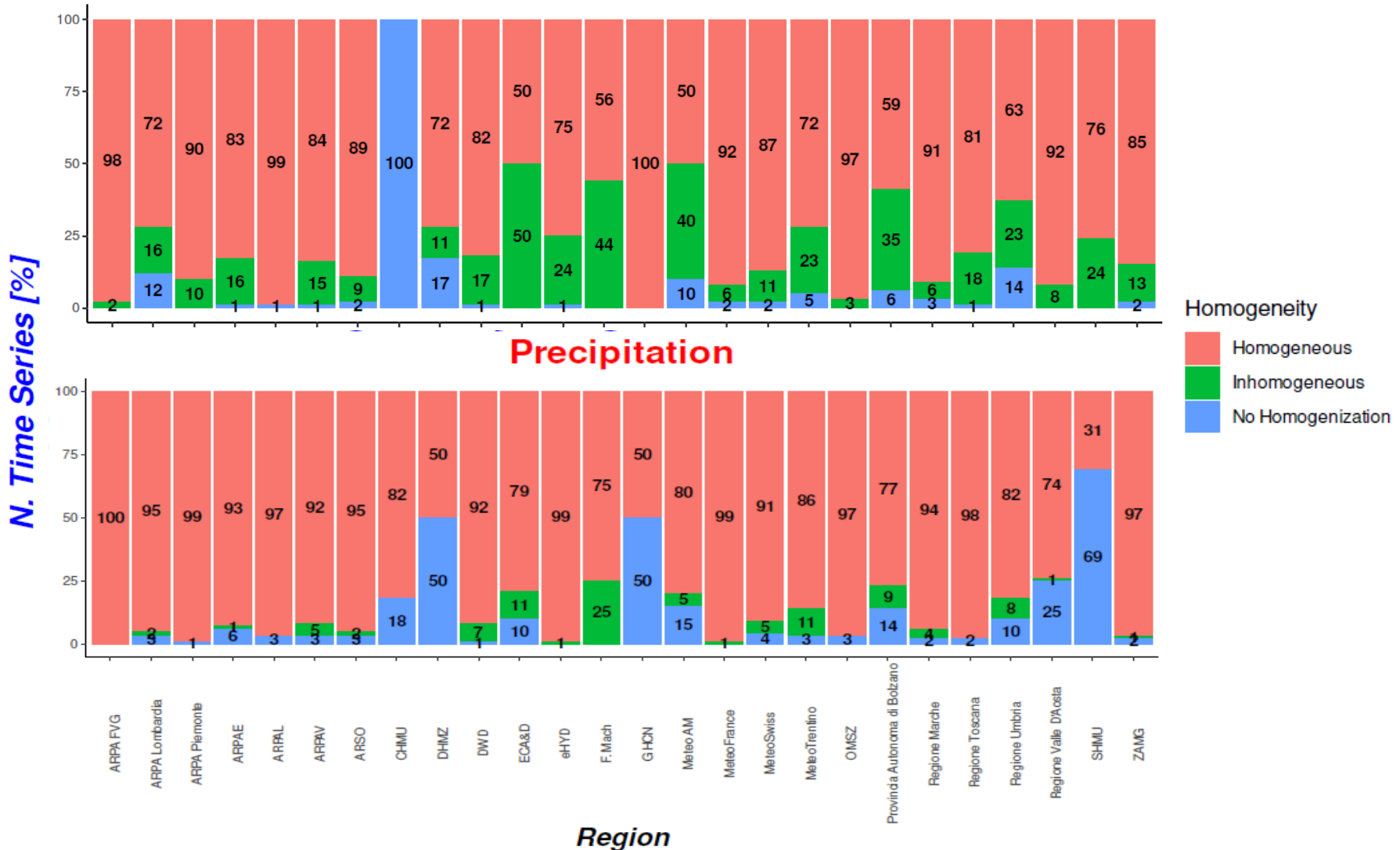
*Precipitation*





# Homogenization Results

## Air Temperature



# Future Developments

- *Validation of homogenization results*
- *Climatological analysis*
- *Application of quality control and homogenization procedures to other variables*
- *Spatial interpolation*

THANKS FOR THE ATTENTION

