EU Surface Temperature for All Corners of Earth (EUSTACE): break-detection algorithm for a global daily temperature dataset

Yuri Brugnara[1,2], **Renate Auchmann**[1,2], **Stefan Brönnimann**[1,2]

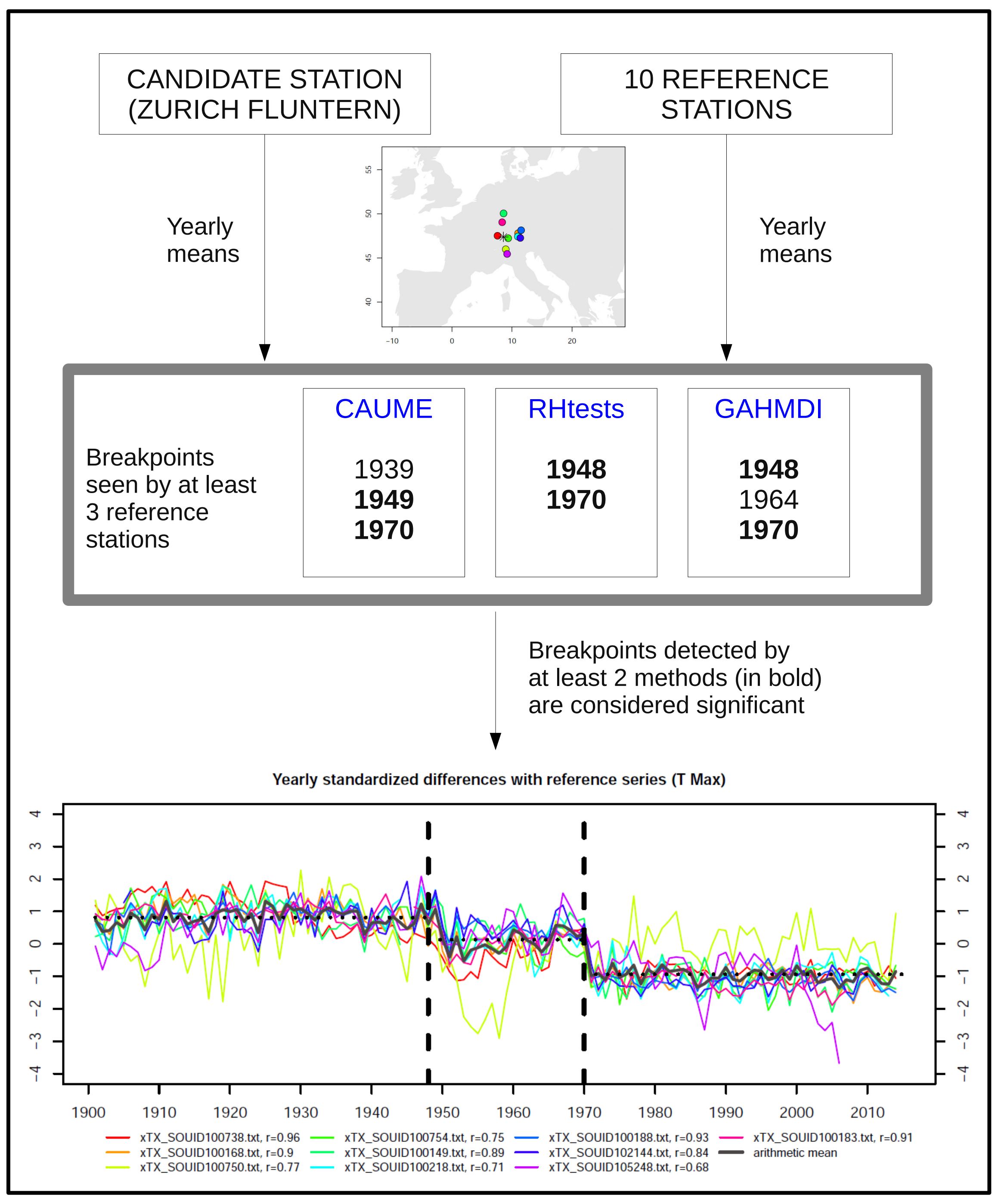
[1]*Oeschger Centre for Climate Change Research – Bern, Switzerland*
[2]*Climatology Group, Institute of Geography, University of Bern – Bern, Switzerland*

Contact: yuri.brugnara@giub.unibe.ch

Abstract

EUSTACE is an EU-funded project which has started in 2015. It will give publicly available daily estimates of surface air temperature since 1850 across the globe for the first time by combining surface and satellite data using novel statistical techniques.

Data quality plays an important role for the success of the project, in particular the assessment of the homogeneity of the available temperature series is crucial. This poster describes the break-detection method that will be applied to a global dataset consisting of tens of thousand of daily maximum and minimum temperature series and analyses its performances in Eurasia.



CANDIDATE STATION (ZURICH FLUNTERN)

10 REFERENCE STATIONS

Yearly means

Yearly means

Breakpoints seen by at least 3 reference stations	CAUME	RHtests	GAHMDI
	1939 **1949** **1970**	**1948** **1970**	**1948** 1964 **1970**

Breakpoints detected by at least 2 methods (in bold) are considered significant

Figure 1 Schematic example of the break-detection algorithm used within EUSTACE.

Break detection

A breakpoint is an inhomogeneity in a climatic time series caused by instantaneous changes in the instrumentation, the observation procedure, and/or the environment of the observatory.

The break detection used within EUSTACE applies a fully automatic algorithm (Fig. 1), adapted from Kuglitsch et al. (2012). After a selection of up to 10 reference series, which satisfy determinate requirements of distance, correlation and data availability, the break detection is performed separately on yearly and semi-yearly means by combining 3 independent detection methods:

- CAUME (Caussinus and Mestre, 2004)
- RHtests (Wang et al., 2007; Wang, 2008)
- GAHMDI (Toreti et al., 2012)

Figure 2 Map of ECA&D stations with at least 30 complete years of daily maximum temperature data.

ECA&D test

The break-detection algorithm was applied on the European Climate Assessment & Dataset (ECA&D, Klein Tank et al., 2002), using the daily maximum temperature series with at least 30 years of data (1'784 stations, Fig. 2).

Only reference stations less than 300 km distant and with a Pearson correlation between yearly first differences of at least 0.6 are used in this test. About 7% of the candidate stations have less than 3 reference stations satisfying these requisites; the break detection was not performed for these stations.

More than 60% of the inhomogeneities for yearly means have an amplitude between 0.1 and 0.4 K, whereas inhomogeneities larger than 1 K are uncommon (Fig. 3). Amplitudes are larger on semi-yearly scale, because inhomogeneities often have a seasonal cycle.

With the tested parameters, the station density is insufficient for a reliable break detection before ca. 1950, resulting in an underestimation of the number of breakpoints (Fig. 4). On ideal conditions (after 1950), a breakpoint is detected on average in 1.7% of the stations every year. A total of 681 series (~38%) were found homogeneous when using only yearly means.

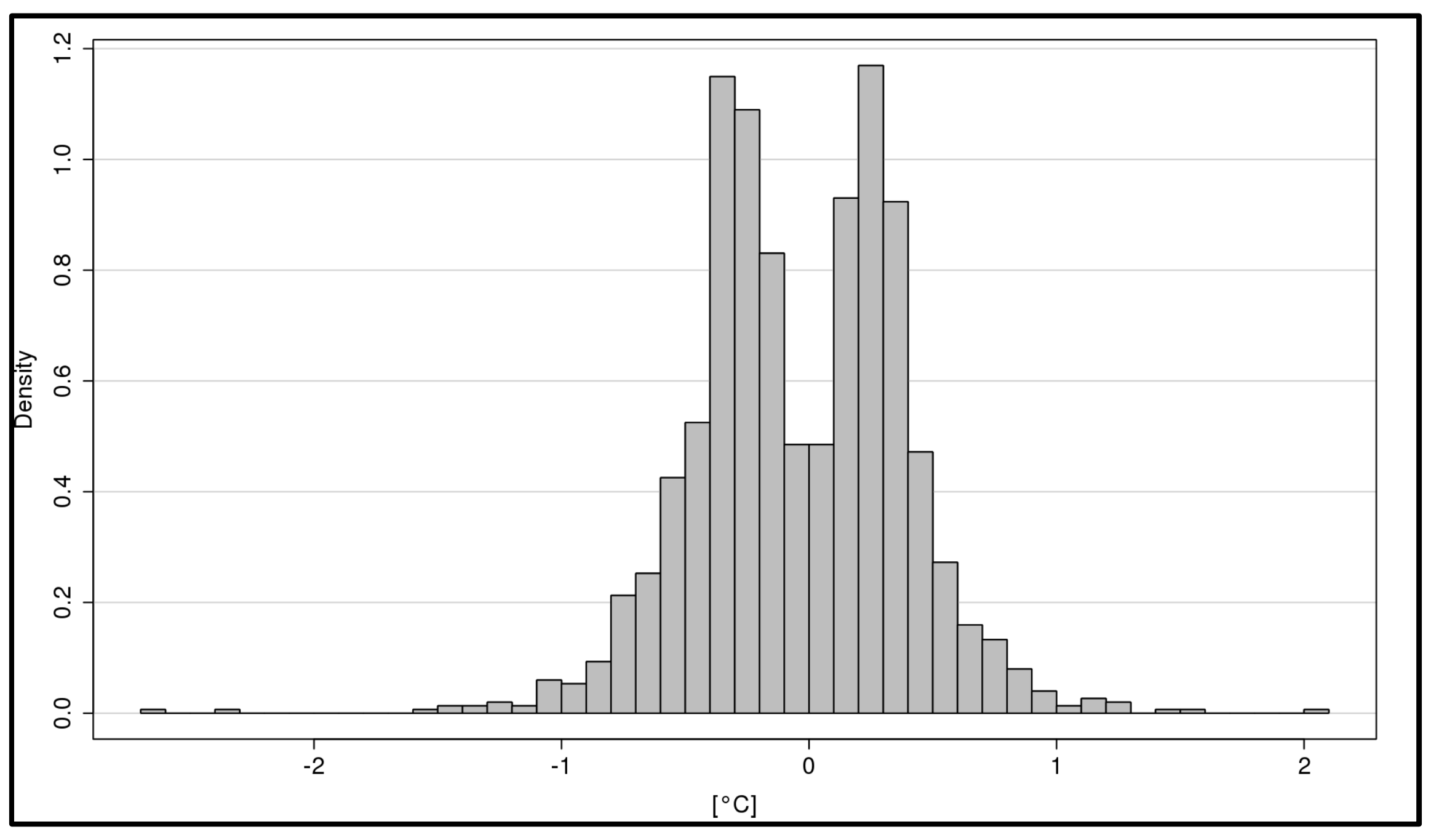


Figure 3 Histogram of amplitudes of the inhomogeneities found in ECA&D yearly means.

Next steps

- The break detection will be applied to a global dataset of quality-controlled daily temperature series, obtained from previous projects (GHCN, ISTI) and from project partners.
- A correction of the inhomogeneities on a daily scale will be attempted for European data.
- Station data will be combined with satellite data to produce a global gridded reconstruction of daily temperature with consistent uncertainty information (EUSTACE Work Package 2, led by the UK Met Office).
- A fraction of temperature series will be diverted and used for validation (EUSTACE Work Package 3, led by the University of Leicester).
- Final data products will be published through established routes to enable access via a large community of users (EUSTACE Work Package 4, led by KNMI).

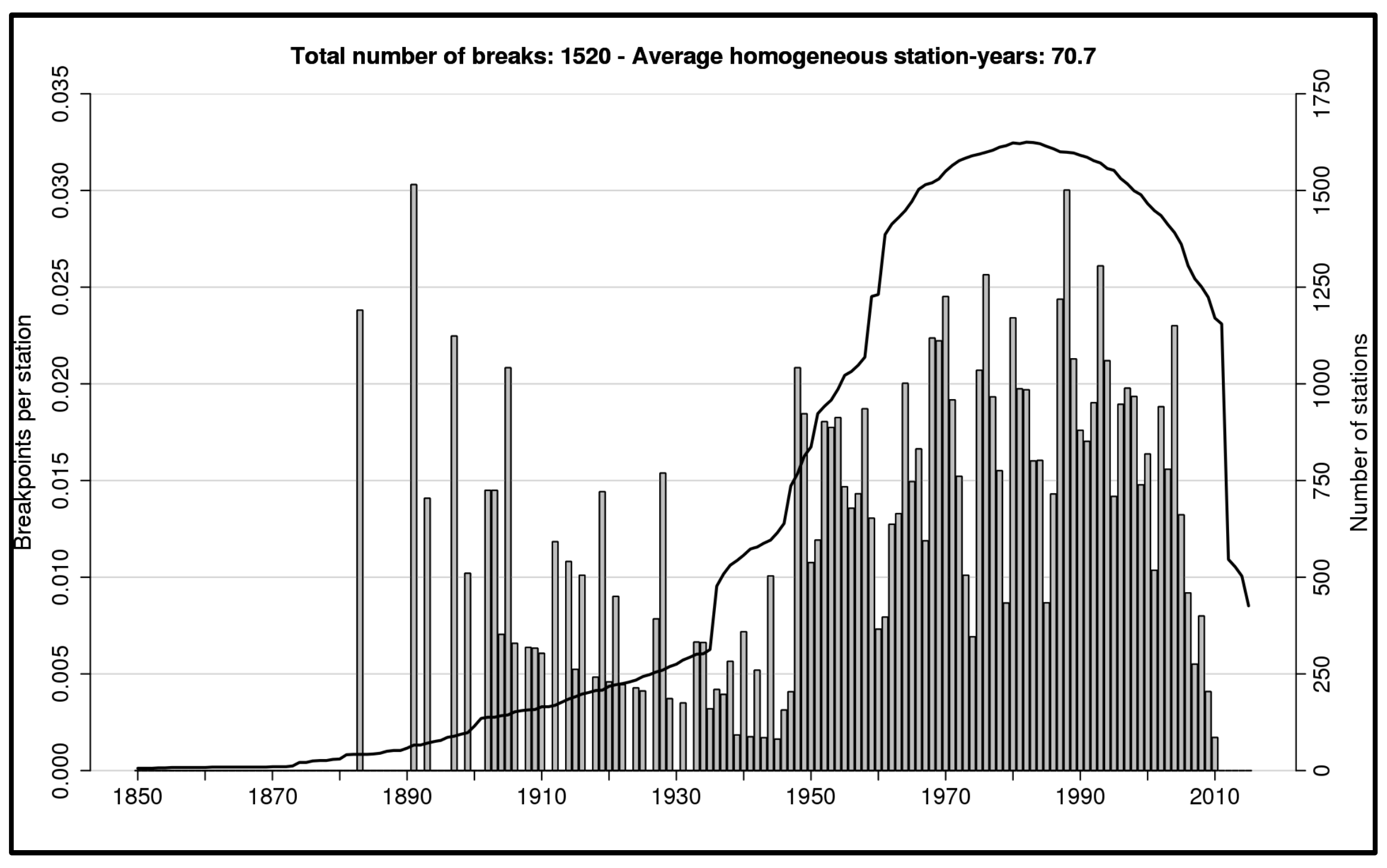


Figure 4 Time series of the number of breakpoints (bars) relative to the number of available stations (line).

Acknowledgements

We thank Andrea Toreti (EC Joint Research Centre) and Yang Feng (Environment Canada) for providing the software for the break-detection methods. ECA&D data was provided by the Royal Netherlands Meteorological Institute (KNMI).

References

H Caussinus and O Mestre. Detection and correction of artificial shifts in climate series. *Journal of the Royal Statistical Society*, 53(3):405–425, 2004.

AMG Klein Tank, JB Wijngaard, GP Können, et al. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology*, 22: 1441–1453, 2002.

FG Kuglitsch, R Auchmann, R Bleisch, S Brönnimann, O Martius and M Stewart. Break detection of annual swiss temperature series. *Journal of Geophysical Research*, 117(D13), 2012.

A Toreti, F G Kuglitsch, E Xoplaki, and J Luterbacher. A novel approach for the detection of inhomogeneities affecting climate time series. *Journal of Applied Meteorology and Climatology*, 51(2):317–326, 2012.

X L Wang. Accounting for autocorrelation in detecting mean shifts in climate data series using the penalized maximal t or f test. *Journal of Applied Meteorology and Climatology*, 47(9):2423–2444, 2008.

X L Wang, Q H Wen, and Y Wu. Penalized maximal t test for detecting undocumented mean change in climate data series. *Journal of Applied Meteorology and Climatology*, 46(6):916–931, 2007.

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

Official website: **www.eustaceproject.eu**

Very early instrumental measurements in Austria

18th century climate data in Austria, still unexploited

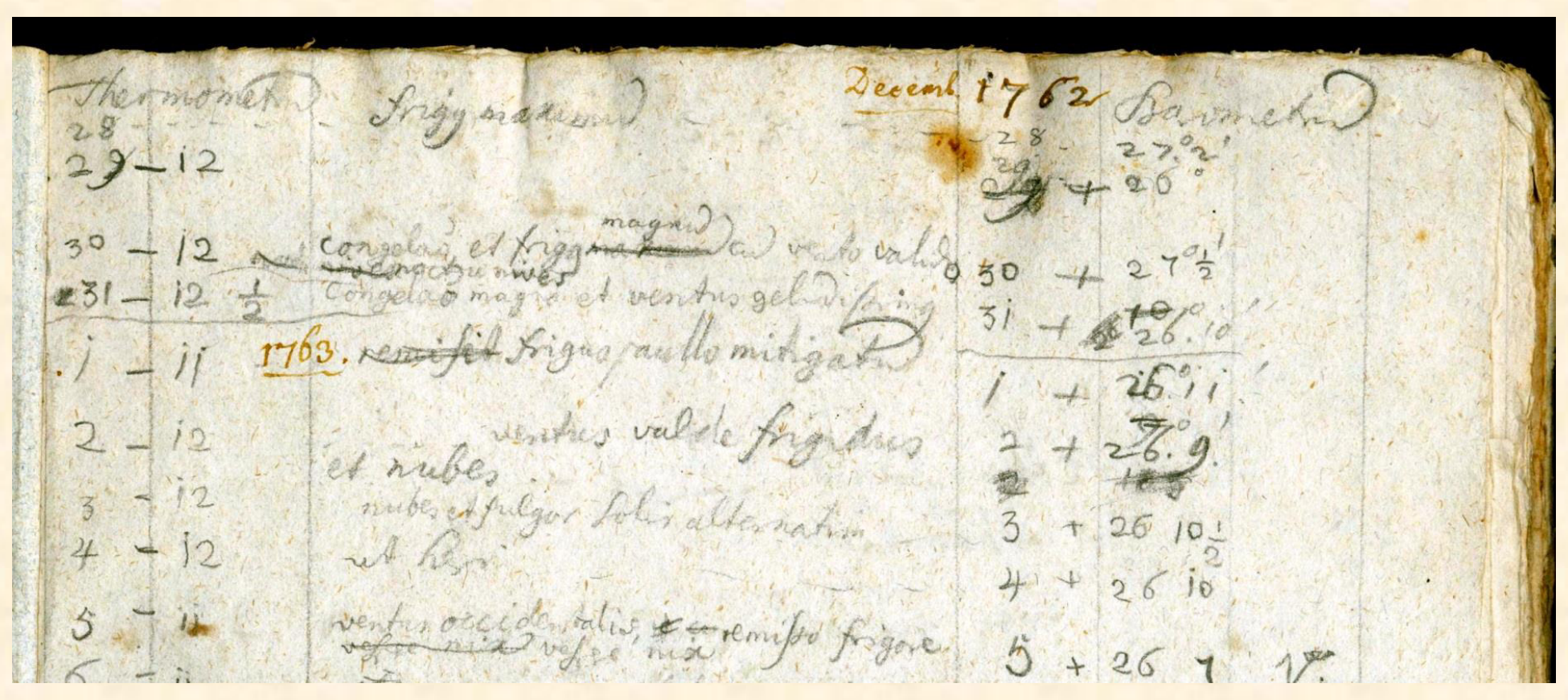
Ingeborg Auer[1)], Barbara Chimani[1)], P. Amand Kraml[2)], Silke Adler[1)]

1 Zentralanstalt für Meteorologie und Geodynamik, Wien
2 Sternwarte Kremsmünster

Instrumental data of the 18th century are very rare worldwide and even in Europe. In Austria, there are three places where measurements started in the second half of the 18th century or even earlier: Kremsmünster, Wien and Innsbruck

KREMSMÜNSTER, 14°08‘ N, 48°03 E, 383 m

observing site without any changes in location since 1762.



Thermomet. frigus [illegible] Decemb. 1762 Barometr.
28 28 27.0 2'
29 — 12 29 + 26 0
30 — 12 congelatio et frigus magnum cum vento valido 30 + 27 0 1/2'
31 — 12 1/2 congelatio magna et ventus gelidus [illegible] 31 + 26.10
1763. frigus paullo mitigatum
1 — 11 1 + 26.11
2 — 12 ventus valde frigidus et nubes 2 + 26.9
3 — 12 nubes et fulgor solis alternatim 3 + 26 10 1/2
4 — 12 [illegible] 4 + 26 10
5 — ventus occidentalis, remisso frigore 5 + 26 7

Father Placidus Fixlmiller performed the first barometer readings (one observation per day) and some short weather descriptions. The first observations are rather unclear but from 1776 onwards until now the station provides high quality data.

Cloudiness and lightning observations started in 1767. Other parameters like precipitation, extreme temperature, air humidity and sunshine duration date back into the 19th century.

Temperature measurements started in June 1763 6 m above the ground in front of a northeast oriented window.

The monthly measurements from 1767 onwards have been homogenized and made available through the HISTALP website. The early biases caused by the unsheltered thermometer has been corrected by using parallel measurements with todays standard installations.

Challenge: Kremsmünster is one of the very few uninterrupted Austrian series on daily time scale. Daily data from 1767 until 1891 still have to be digitized and homogenized until present.

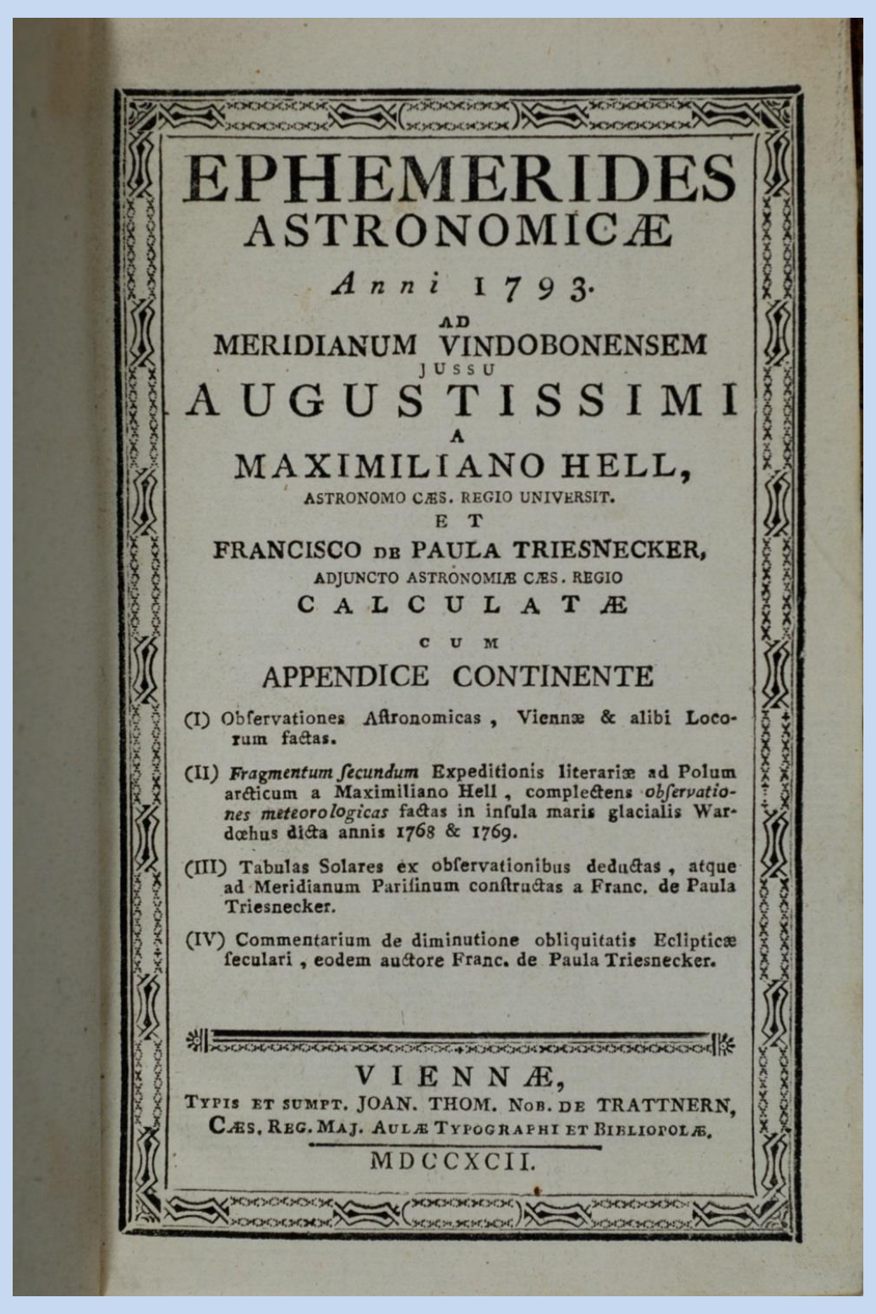
→ A climate time series in daily resolution of nearly 250 years.

WIEN ALTE UNIVERSITÄTSSTERNWARTE, 16 21 E, 48 14 N, 198.5 m

The earliest Viennese measurements from the Jesuit college Vienna (1734 – 1773) have gone lost, likewise the first observations of the Astronomical observatory of the University of Vienna (1762 – 1774). With the help of the present day astronomical observatory we found the Viennese data of October 1768 to June 1769 in the Ephemerides Astronomicae anni 1793. Although there are some gaps in the tables the data seem to be realistic.

site of temperature measurements in 1775

astronomical observatory of the university, 1775 to 1878. Site in the historic centre of Vienna, now the Academy of Sciences



EPHEMERIDES
ASTRONOMICÆ
Anni 1793.
AD
MERIDIANUM VINDOBONENSEM
JUSSU
AUGUSTISSIMI
A
MAXIMILIANO HELL,
ASTRONOMO CÆS. REGIO UNIVERSIT.
ET
FRANCISCO DE PAULA TRIESNECKER,
ADJUNCTO ASTRONOMIÆ CÆS. REGIO
CALCULATÆ
CUM
APPENDICE CONTINENTE

(I) Obſervationes Aſtronomicas, Viennæ & alibi Locorum factas.

(II) *Fragmentum ſecundum* Expeditionis literariæ ad Polum arcticum a Maximiliano Hell, complectens *obſervationes meteorologicas* factas in inſula maris glacialis Wardœhus dicta annis 1768 & 1769.

(III) Tabulas Solares ex obſervationibus deductas, atque ad Meridianum Pariſinum conſtructas a Franc. de Paula Triesnecker.

(IV) Commentarium de diminutione obliquitatis Eclipticæ ſeculari, eodem auctore Franc. de Paula Triesnecker.

VIENNÆ,
TYPIS ET SUMPT. JOAN. THOM. NOB. DE TRATTNERN,
CÆS. REG. MAJ. AULÆ TYPOGRAPHI ET BIBLIOPOLÆ.
MDCCXCII.

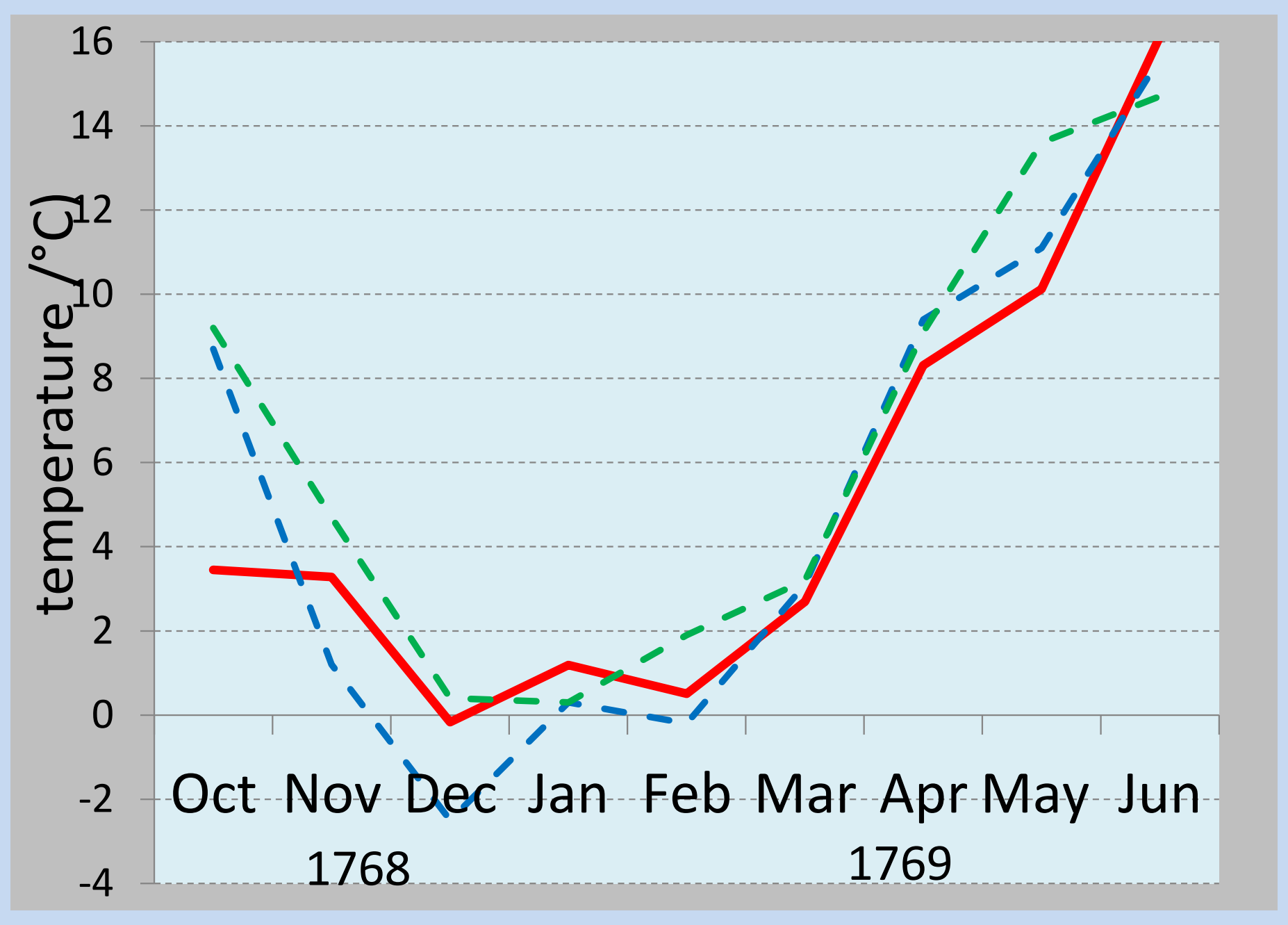
E56

METEOROLOGISCHE
BEOBACHTUNGEN
AN DER
WIENER STERNWARTE
VON 1775 BIS 1855.

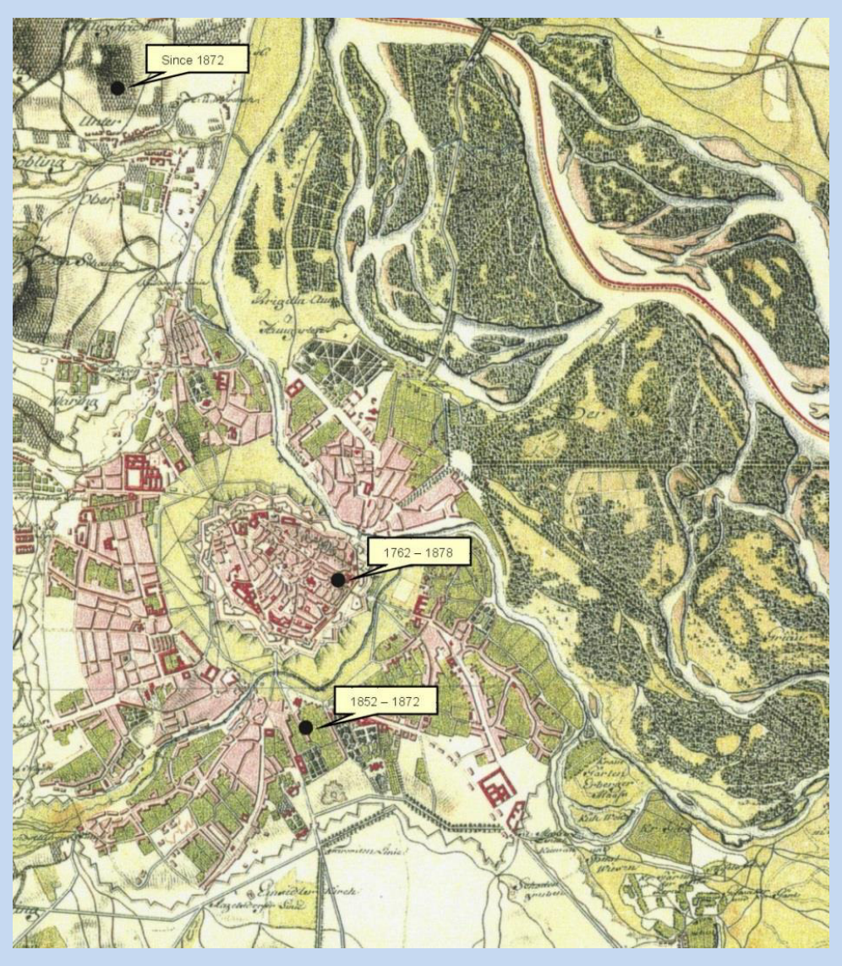
BIBLIOTHEK
der Zentralanstalt für
Meteorologie und Geodynamik
Wien XIX., Hohe Warte 38

ERSTER BAND.
1775—1796.

data after 1775 have been published in a series of „Meteorologische Beobachtungen an der Wiener Sternwarte. They have been quality checked and homogenized. The early instrumental bias has been removed in an second step. The Vienna series since 1775 is part of the HISTALP dataset, http://www.zamg.ac.at/histalp.



Annual course of temperature in Vienna (red solid line) in comparison to Kremsmünster (dashed blue line) and Basel (dashed green line)



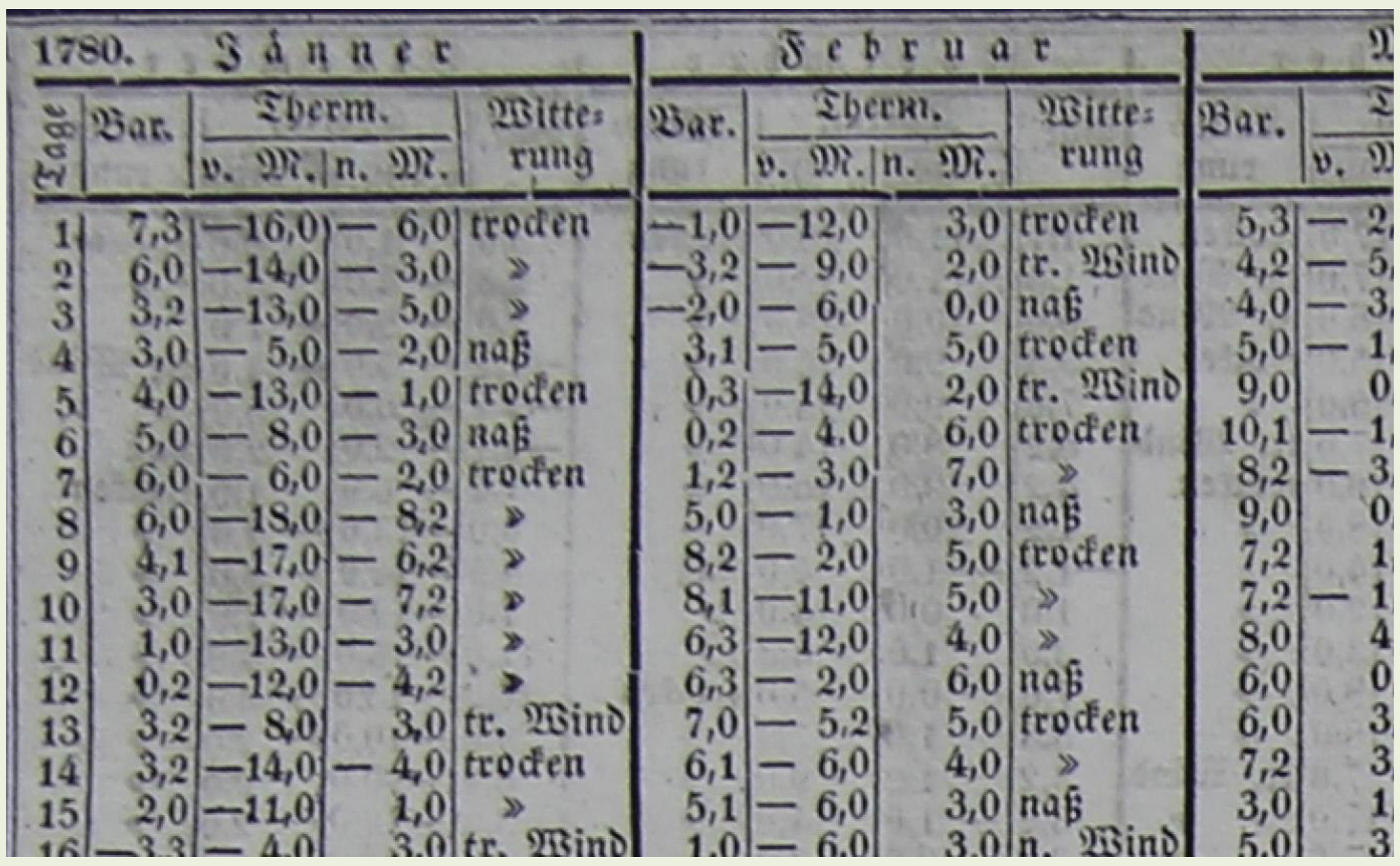
Vienna station site locations (1762 – 1872)

INNSBRUCK SILLGASSE, 11 24 E, 47 16 N, 577 m,

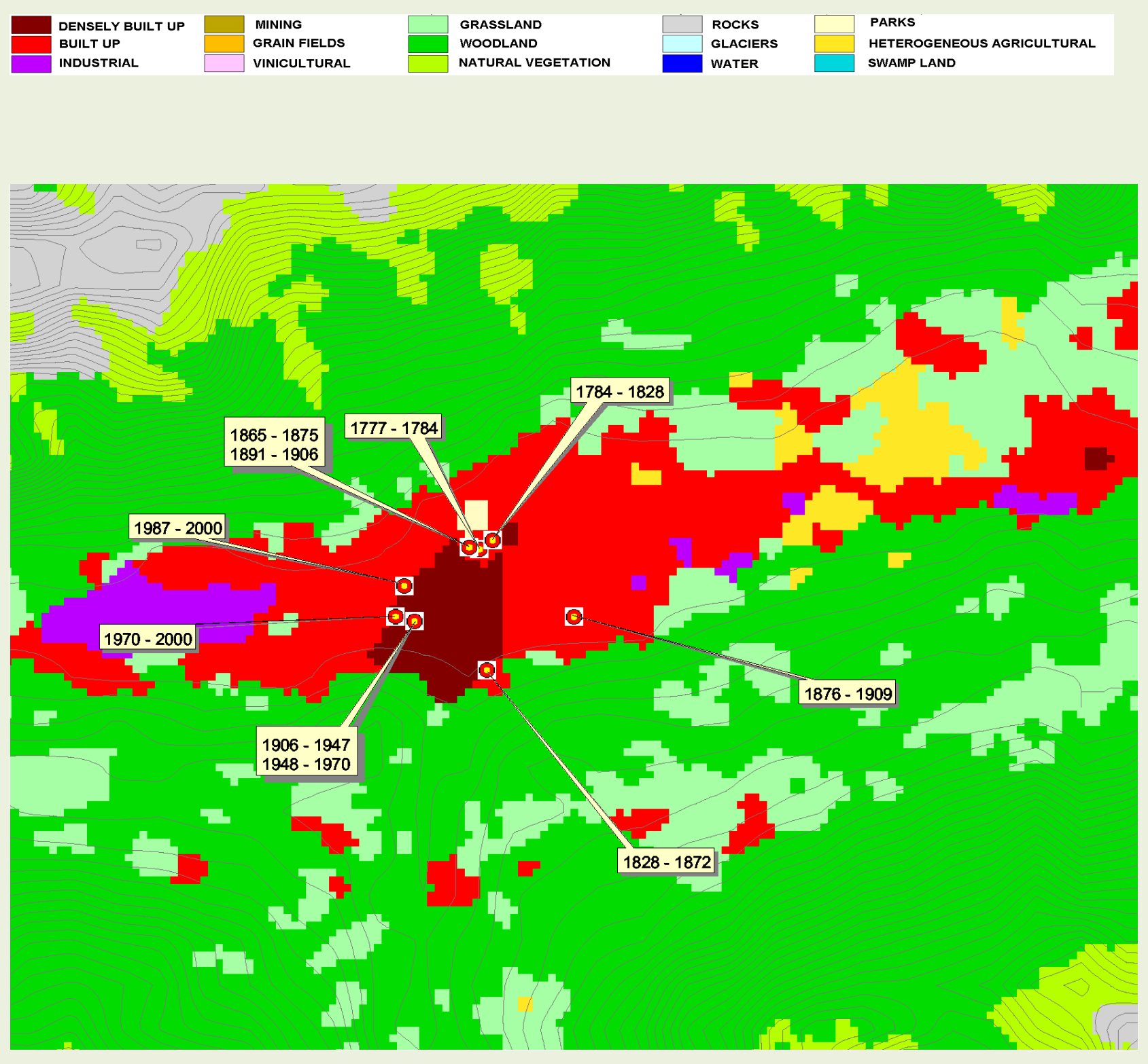
station in the town-centre, however barometer readings without temperature.

The thermometer was situated in a shaded position outside a window in 2nd floor on the NW-wall of the monastery building.

Homogenization was performed on monthly basis and the early bias corrected. Innsbruck is part of the HISTALP database since 1777.



1780.	Jänner				Februar				
Tage	Bar.	Therm.		Witterung	Bar.	Therm.		Witterung	Bar.
		v. M.	n. M.			v. M.	n. M.		
1	7,3	—16,0	— 6,0	trocken	—1,0	—12,0	3,0	trocken	5,3
2	6,0	—14,0	— 3,0	»	—3,2	— 9,0	2,0	tr. Wind	4,2
3	3,2	—13,0	— 5,0	»	—2,0	— 6,0	0,0	naß	4,0
4	3,0	— 5,0	— 2,0	naß	3,1	— 5,0	5,0	trocken	5,0
5	4,0	—13,0	— 1,0	trocken	0,3	—14,0	2,0	tr. Wind	9,0
6	5,0	— 8,0	— 3,0	naß	0,2	— 4.0	6,0	trocken	10,1
7	6,0	— 6,0	— 2,0	trocken	1,2	— 3,0	7,0	»	8,2
8	6,0	—18,0	— 8,2	»	5,0	— 1,0	3,0	naß	9,0
9	4,1	—17,0	— 6,2	»	8,2	— 2,0	5,0	trocken	7,2
10	3,0	—17,0	— 7,2	»	8,1	—11,0	5,0	»	7,2
11	1,0	—13,0	— 3,0	»	6,3	—12,0	4,0	»	8,0
12	0,2	—12,0	— 4,2	»	6,3	— 2,0	6,0	naß	6,0
13	3,2	— 8,0	3,0	tr. Wind	7,0	— 5,2	5,0	trocken	6,0
14	3,2	—14,0	— 4,0	trocken	6,1	— 6,0	4,0	»	7,2
15	2,0	—11,0	1,0	»	5,1	— 6,0	3,0	naß	3,0
16	—3,3	— 4,0	3,0	tr. Wind	1,0	— 6,0	3,0	n. Wind	5,0



site map of observing stations in Innsbruck

contact: ingeborg.auer@zamg.ac.at

Rescue and quality control of European sub-daily data

L. Ashcroft[1], J.R. Coll[1], M. Castellà[1], P. Domonkos[1], E. Aguilar[1], J. Sigró[1] and M. Brunet[1,2]
[1]Centre for Climate Change, University Rovira i Virgili, Av. Remolins 13-15, 43500 Tortosa, Spain
[2]Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK

Questions? Just email lindenclaire.ashcroft@urv.cat or tweet @lindenashcroft.

Introduction

We are rescuing sub-daily European meteorological data to improve regional reanalysis, as part of Uncertainties of Ensembles in Regional Reanalysis (UERRA), an EU/FP7-SPACE-2013 project.

In 18 months we have recovered 5.9 million observations from 99 stations in 13 countries (Figure 1).

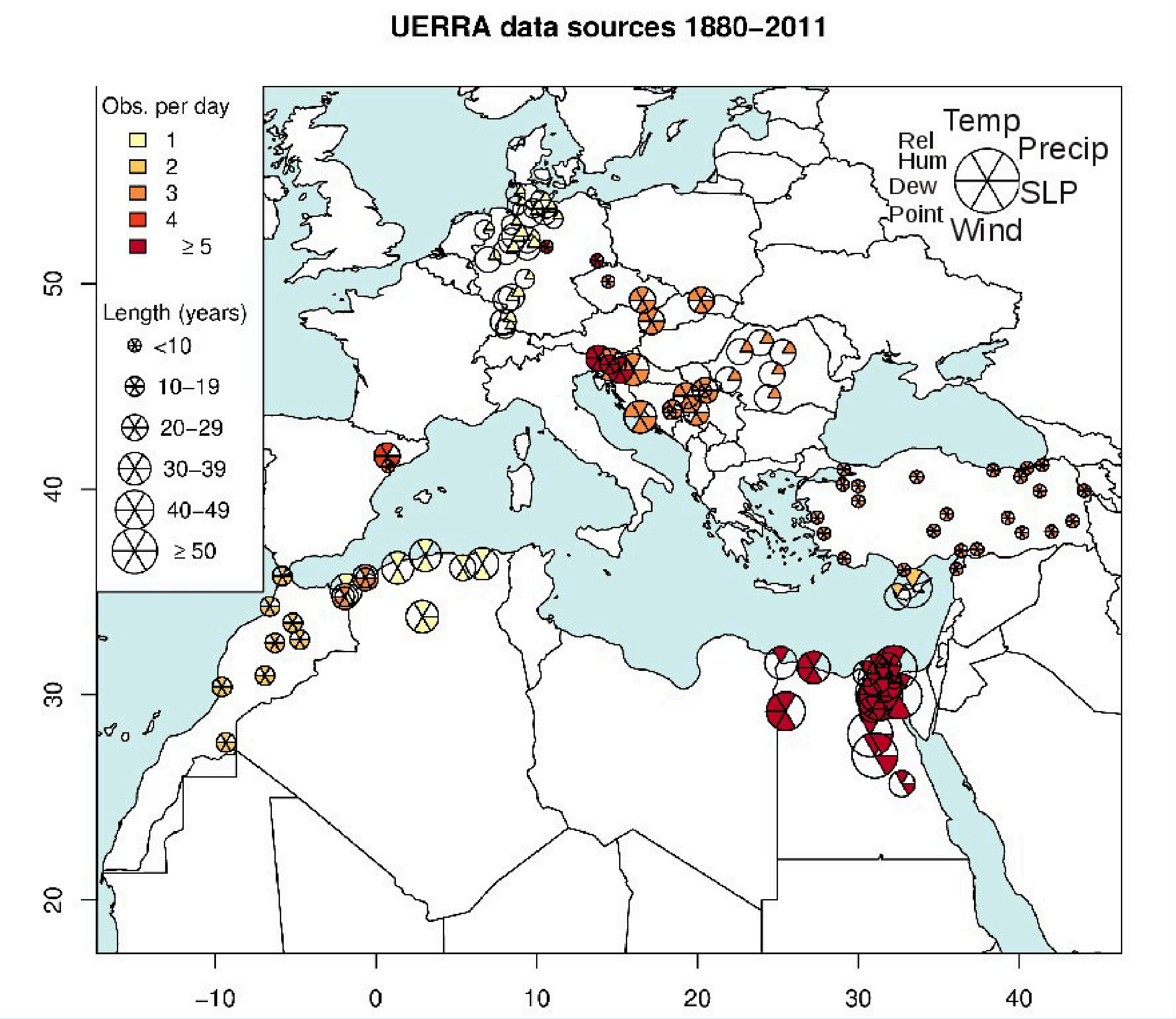
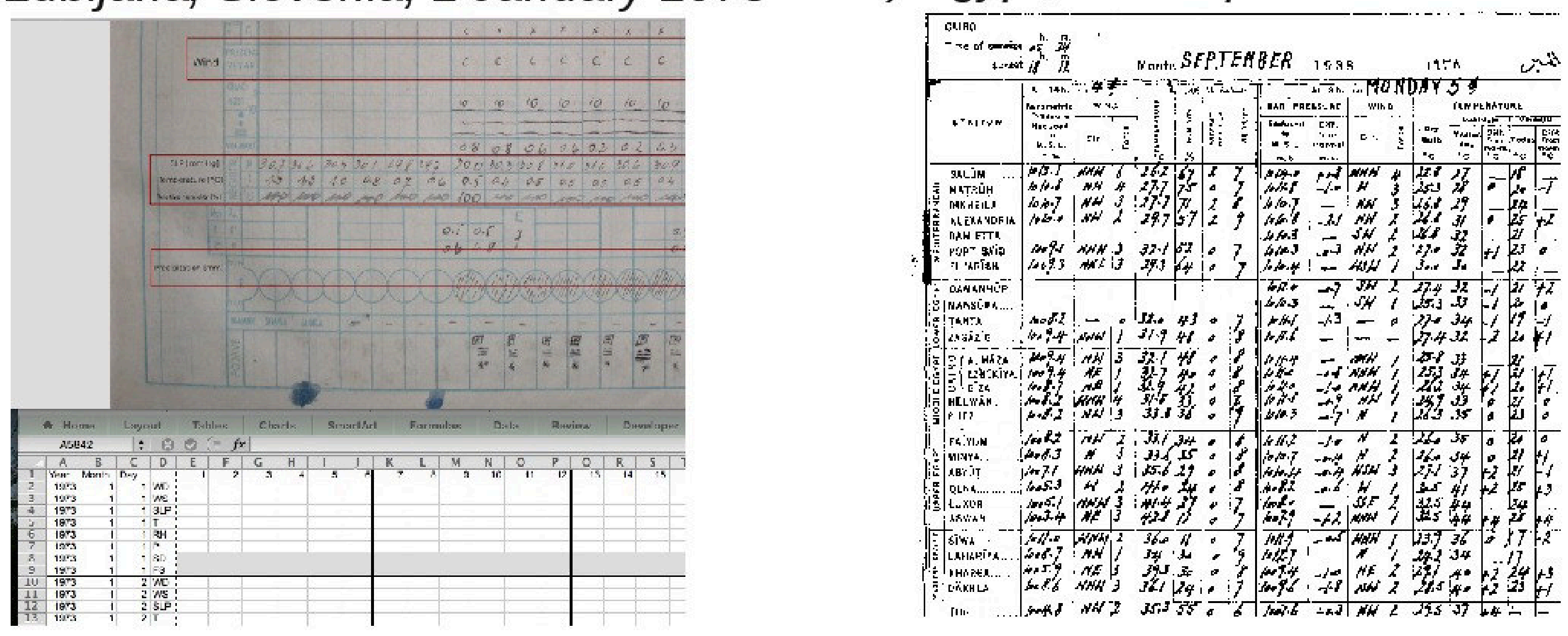


Figure 1. Stations with sub-daily data digitised by UERRA.

What have we learned so far about digitising subdaily data?

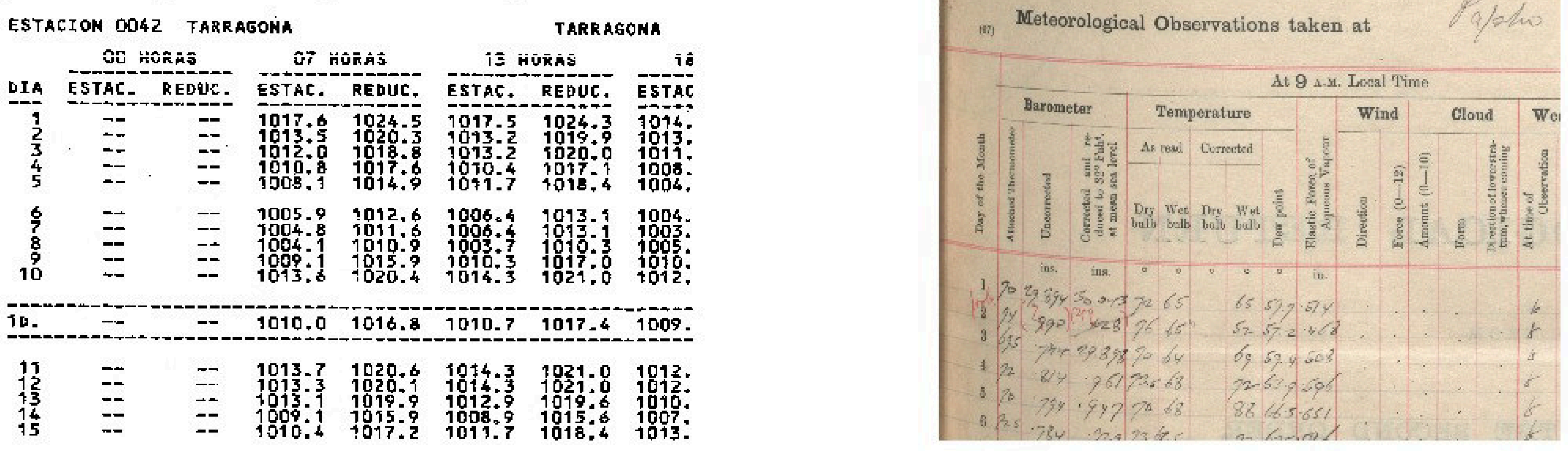
1. All data sources require different preparation: templates are important

- Digitisers need clear instructions about which variables to digitise.
- Spreadsheet templates that match the data source are helpful and reduce errors.
- This is particularly important with sub-daily data, which can be difficult to read (e.g. Figure 2).



a) Lubljana, Slovenia, 1 January 1973

b) Egypt, 4–5 September 1938



c) Tarragona, Spain, July 1977.

d) Paphos, Cyprus, July 1901

Figure 2. Examples of data sources. An example template is shown in 2a).

2. Timely feedback is important to reduce subtle errors and biases

- We conduct systematic visual cross-checking when new sources are digitised.
- Three days of every 4 months for 25% of years are checked, looking for common mistakes.
- Feedback is given to the digitisers to improve future work. Common errors include:
  - \> Consistent misreading of confusing values (e.g. a 3 as 8)
  - \> Skipped dates, leading to out-of-order data
  - \> Incorrect rows or columns digitised (these last two problems are minimised with the use of templates)

These errors are hard to detect using automatic methods.

Table 1: Number of errors found using systematic visual cross-checking.

Number of systematically checked values	Total number of errors found	Gross errors	Subtle and systematic errors	Source errors
20312	418 (2.05%)	8	339	69

3. Automatic quality control is important for gross errors and outliers

- Our programs check for **14 different error types**:

Date order
Pattern repetition
Record breakers
Repeated values
Frequency biases
Strange distributions
Climatic outliers
Jumps and spikes
Logical failures
Intervariable comparison
Calculated vs observed values
Strange scattering
Bivariate distribution outliers
Unit changes

For example, recovered data from Tarragona, Spain:

- 0700, 1300, 1800 for 1977–1984, 6 variables = 52,020 values.
- Less than 3% of observations were flagged as errors.
- 62% of flagged observations were corrected as typos or removed as clear source errors.

More errors in data digitised without a template.

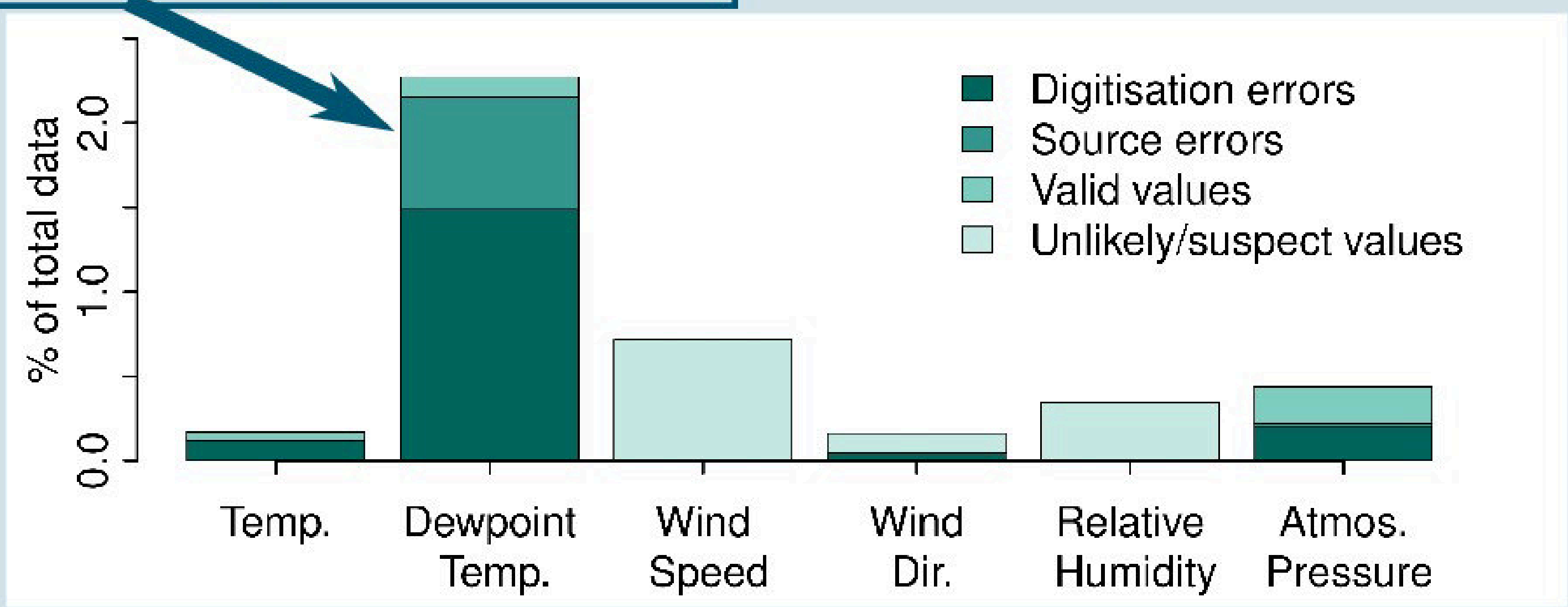


Figure 3. Percentage of total Tarragona subdaily data flagged by the automatic quality control procedure.

Many thanks to the partner National Meteorological Services for providing data for digitisation. We are also indebted to the digitisation team: I. López, D. Azuara, J. Paradinas, A. Domènech, E. Pla, E. Romero, A. Robert, J. Tarragó, C. Bonfill, G. Vandellòs, R. Guerra, A. Àvila, A. Balart, M. Martin, E. Cebrián, P. Sabaté and A. Boqué.

Enhanced climate resilience through user engagement and data rescue in Tanzania

Martyn Sunter, Met Office

Fig 1: Location of the four main sites used during the project

Challenge

Tanzania is highly susceptible to climate variability and change. Despite having rich water and land resources, there is low productivity and persistent poverty. Agriculture is central to Tanzania's economy, with approximately 80% of livelihoods dependent on it. Agriculture production suffers around $200m in annual losses, mostly due to the weather, and climate change is likely to amplify existing pressures. Many communities are highly vulnerable, such as small scale farmers who depend on cereal crops for survival.

Solution

Timely, relevant and useful weather and climate services are essential to manage the challenges Tanzania faces. The Met Office has a longstanding relationship with the Tanzania Meteorological Agency. With support from the UK Department for International Development through UKaid, the Met Office led a project to help build improved climate services for Tanzania. The project focussed on two key elements within TMA's services; user engagement and data rescue, both building towards national implementation of the Global Framework for Climate Services (fig 2). A Met Office team worked closely with their counterparts at TMA to build capacity through knowledge transfer.

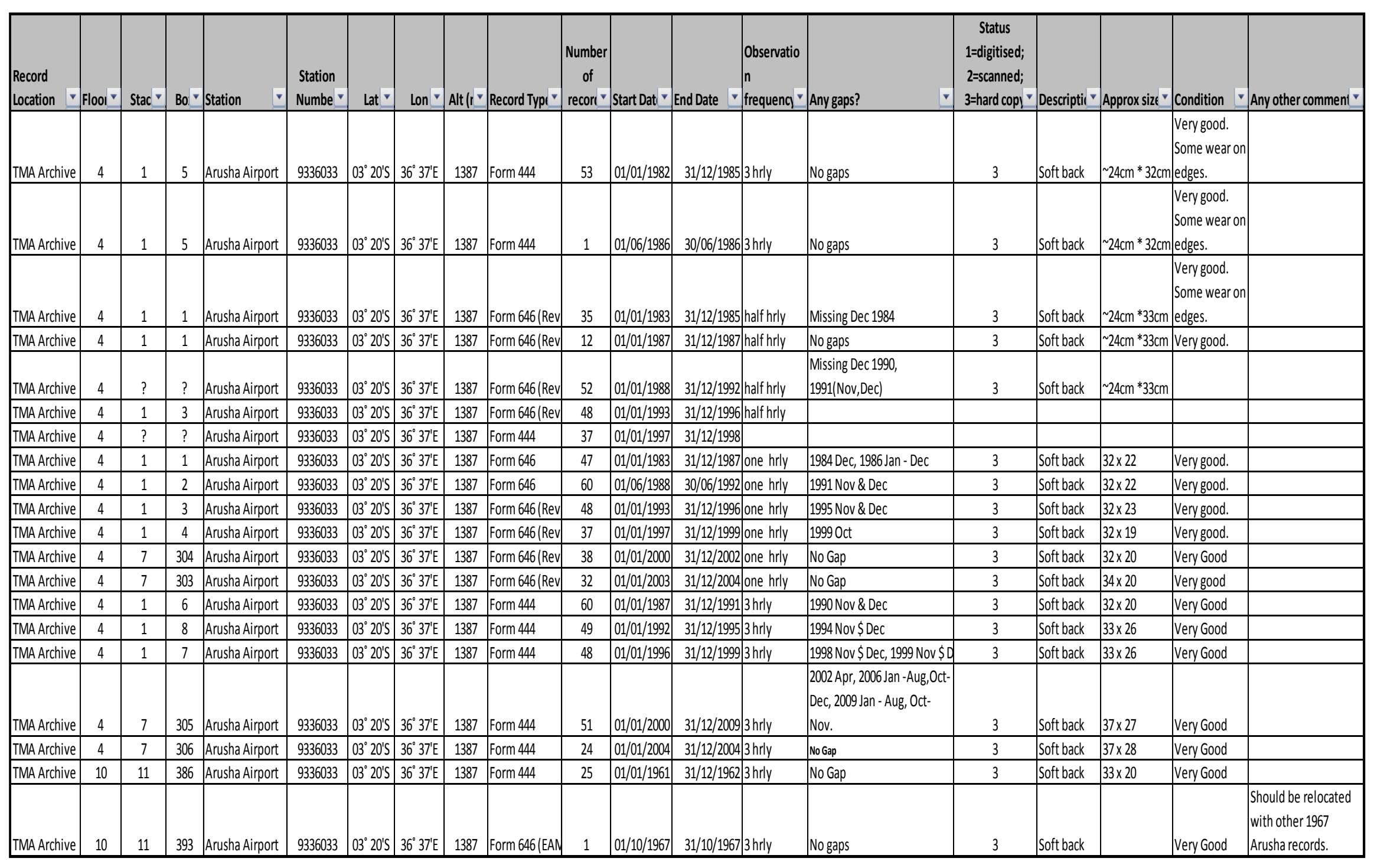
Project objectives

Under this project, the Tanzania Meteorological Agency (TMA) had three broad objectives:

- To introduce enhanced working practices for user engagement and data management activities at the TMA resulting in improved service delivery and increased climate awareness of users.
- To understand user requirements for the highest priority climate vulnerable sectors in Tanzania to enable delivery of relevant climate services.
- To digitise the high priority paper-based observation archive and use it to develop climate services.

Creating the inventory

All boxes in the TMA climate archive were relabelled to provide a consistent approach. The contents of boxes were then examined and the details of each box were noted in an inventory spreadsheet (fig 3). This meant that gaps between the paper records and the TMA climate database could be identified.



Record Location	Floo	Stac	Bo	Station	Station Numbe	Lat	Lon	Alt (r	Record Typ	Number of recor	Start Dat	End Date	Observation frequenc	Any gaps?	Status 1=digitised; 2=scanned; 3=hard copy	Descripti	Approx size	Condition	Any other comment
TMA Archive	4	1	5	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	53	01/01/1982	31/12/1985	3 hrly	No gaps	3	Soft back	~24cm * 32cm	Very good. Some wear on edges.	
TMA Archive	4	1	5	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	1	01/06/1986	30/06/1986	3 hrly	No gaps	3	Soft back	~24cm * 32cm	Very good. Some wear on edges.	
TMA Archive	4	1	1	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	35	01/01/1983	31/12/1985	half hrly	Missing Dec 1984	3	Soft back	~24cm *33cm	Very good. Some wear on edges.	
TMA Archive	4	1	1	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	12	01/01/1987	31/12/1987	half hrly	No gaps	3	Soft back	~24cm *33cm	Very good.	
TMA Archive	4	?	?	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	52	01/01/1988	31/12/1992	half hrly	Missing Dec 1990, 1991(Nov,Dec)	3	Soft back	~24cm *33cm		
TMA Archive	4	1	3	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	48	01/01/1993	31/12/1996	half hrly						
TMA Archive	4	?	?	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	37	01/01/1997	31/12/1998							
TMA Archive	4	1	1	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646	47	01/01/1983	31/12/1987	one hrly	1984 Dec, 1986 Jan - Dec	3	Soft back	32 x 22	Very good.	
TMA Archive	4	1	2	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646	60	01/06/1988	30/06/1992	one hrly	1991 Nov & Dec	3	Soft back	32 x 22	Very good.	
TMA Archive	4	1	3	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	48	01/01/1993	31/12/1996	one hrly	1995 Nov & Dec	3	Soft back	32 x 23	Very good.	
TMA Archive	4	1	4	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	37	01/01/1997	31/12/1999	one hrly	1999 Oct	3	Soft back	32 x 19	Very good.	
TMA Archive	4	7	304	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	38	01/01/2000	31/12/2002	one hrly	No Gap	3	Soft back	32 x 20	Very Good	
TMA Archive	4	7	303	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (Rev	32	01/01/2003	31/12/2004	one hrly	No Gap	3	Soft back	34 x 20	Very good	
TMA Archive	4	1	6	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	60	01/01/1987	31/12/1991	3 hrly	1990 Nov & Dec	3	Soft back	32 x 20	Very Good	
TMA Archive	4	1	8	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	49	01/01/1992	31/12/1995	3 hrly	1994 Nov $ Dec	3	Soft back	33 x 26	Very Good	
TMA Archive	4	1	7	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	48	01/01/1996	31/12/1999	3 hrly	1998 Nov $ Dec, 1999 Nov $ D	3	Soft back	33 x 26	Very Good	
TMA Archive	4	7	305	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	51	01/01/2000	31/12/2009	3 hrly	2002 Apr, 2006 Jan -Aug,Oct-Dec, 2009 Jan - Aug, Oct-Nov.	3	Soft back	37 x 27	Very Good	
TMA Archive	4	7	306	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	24	01/01/2004	31/12/2004	3 hrly	No Gap	3	Soft back	37 x 28	Very Good	
TMA Archive	10	11	386	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 444	25	01/01/1961	31/12/1962	3 hrly	No Gap	3	Soft back	33 x 20	Very Good	
TMA Archive	10	11	393	Arusha Airport	9336033	03° 20'S	36° 37'E	1387	Form 646 (EAM	1	01/10/1967	31/10/1967	3 hrly	No gaps	3	Soft back		Very Good	Should be relocated with other 1967 Arusha records.

Fig 3: Extract from inventory spreadsheet

Digitisation trial

A digitisation trial was carried out to compare the accuracy of an automated data capture technique (intelligent character recognition) with the manual keying by a TMA member of staff. A full year of maximum and minimum temperature data for Dar es Salaam from 1953 was used. The cumulative sum of the number of errors in each month is shown in figure 6. The higher number of errors for minimum temperature produced by the member of staff was due to them getting out of step by one day over a 16 day period in January, but they were more accurate than the automated technique for the maximum temperature. The accuracy obtained by the automated technique was 97.1% and from the TMA member of staff was 96.2% (or 98.4% ignoring the 'wrong day' error). The automated technique occasionally had problems distinguishing particular characters, occasionally mixing up 7's and 2's, for example, but additional QC checks implemented at the data capture stage would reduce these. The paper documents used were hand written, but of high quality, with each measurement contained within a box in a table, which made it easier for the automated technique to interpret than other paper records that were scanned during the project.

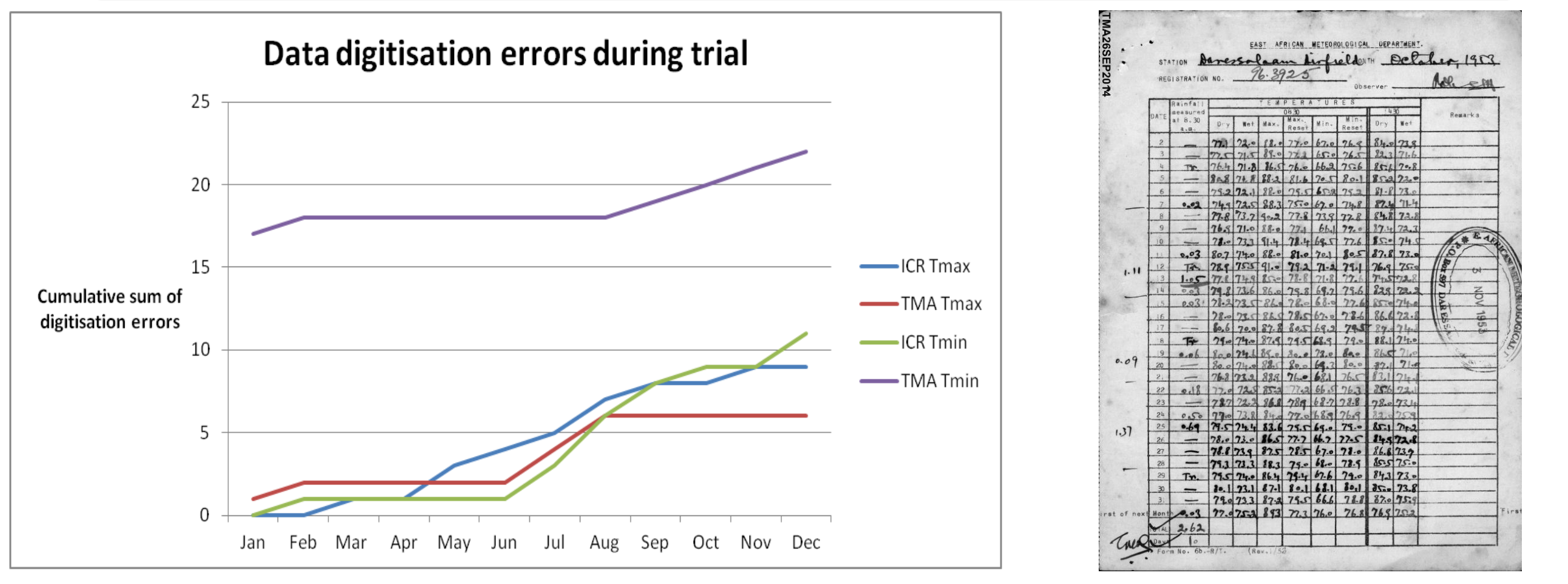


Fig 6: Data digitisation errors during trial and example of paper record from 1953 used.

Strategy and processes

A data rescue strategy was developed and documented, along with an inventory process and scanning and digitisation process. This followed the approach as described in WMO/TD No.1210 "Guidelines on Climate Data Rescue" (fig 4).

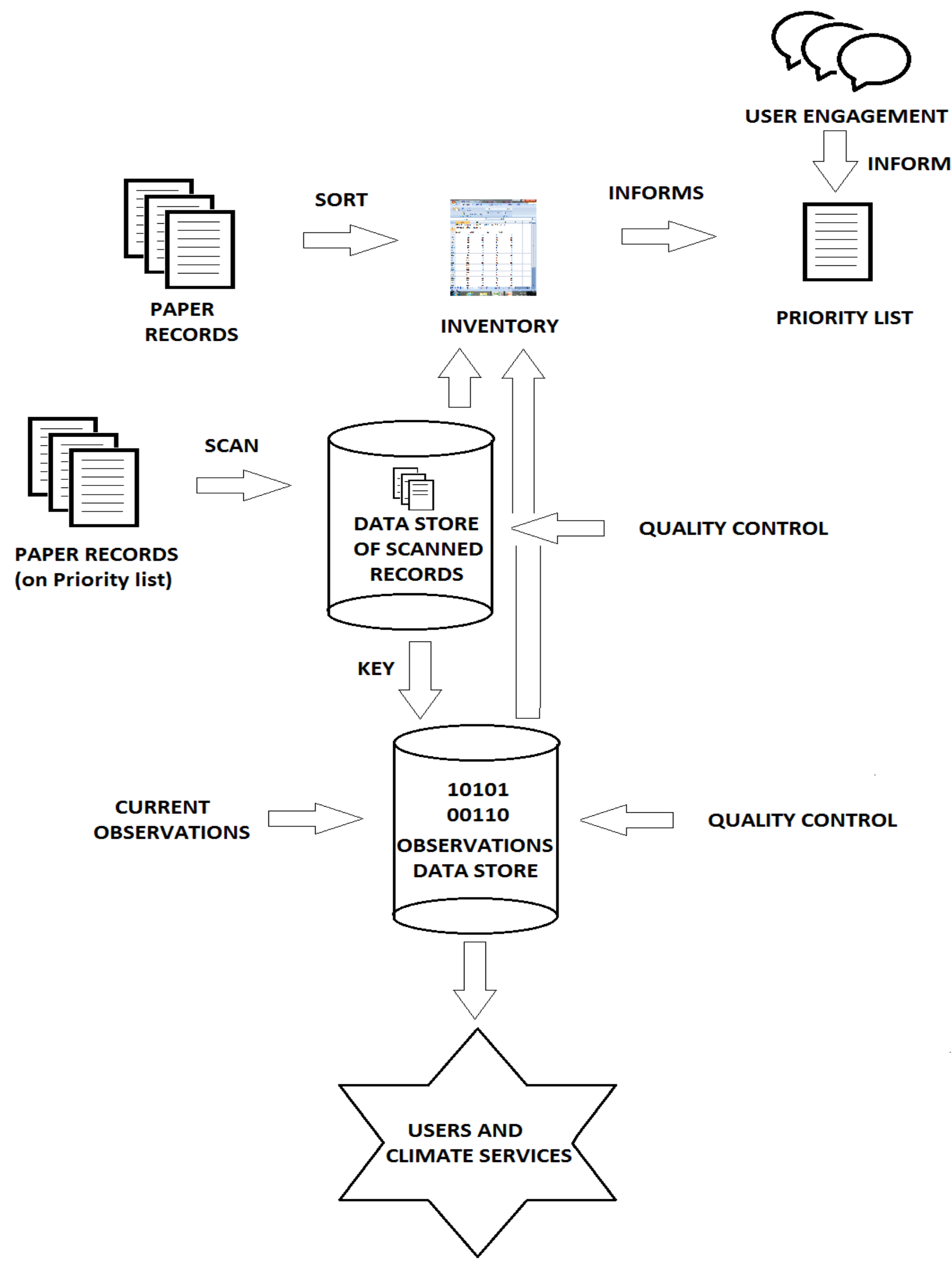


Fig 4: Schematic of TMA Data Rescue Approach based on WMO/TD No.1210 "Guidelines on Climate Data Rescue".

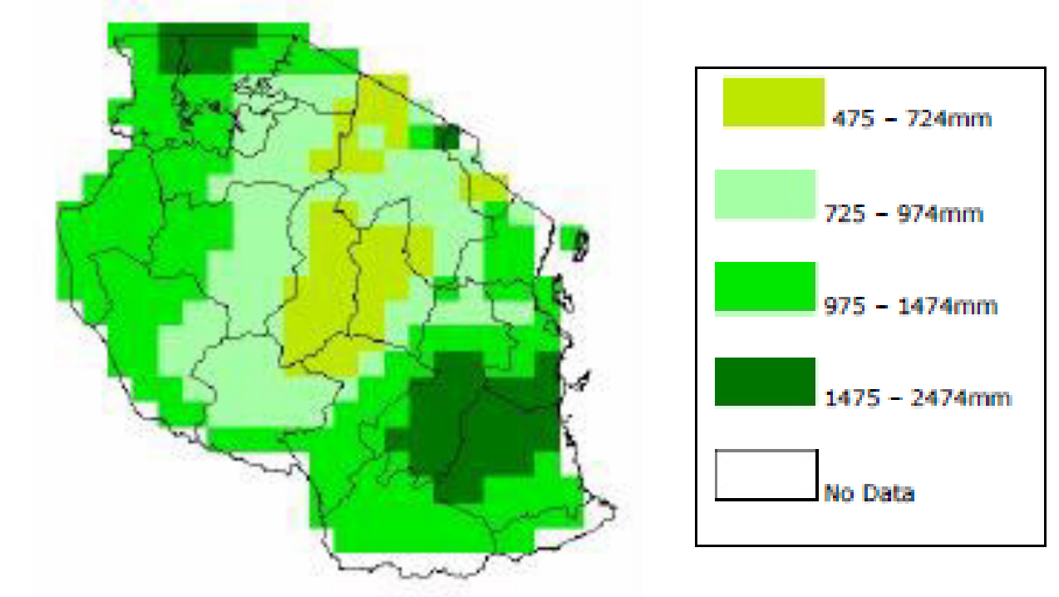


Fig 5: Average annual rainfall in Tanzania
Source FAO (2010)

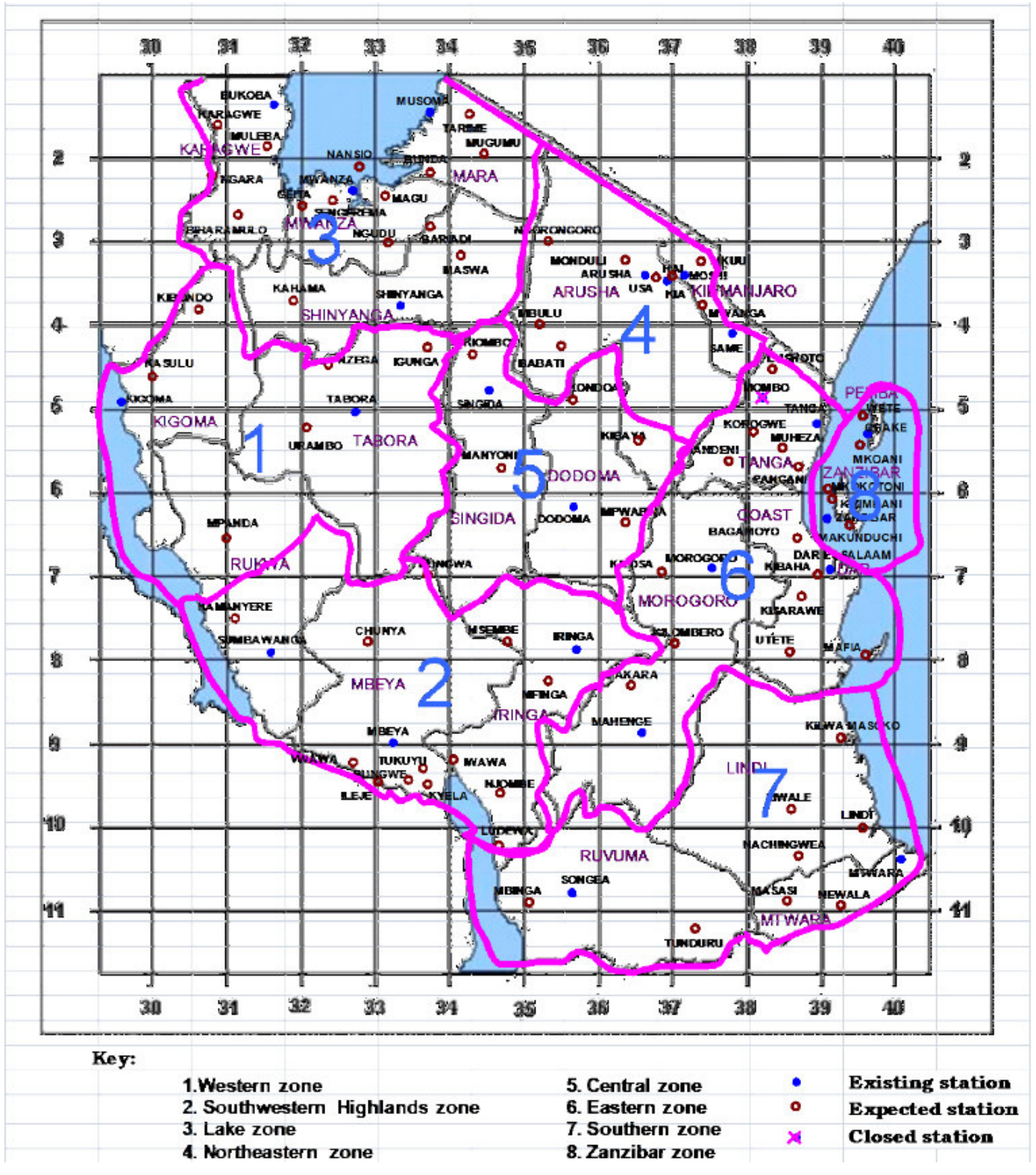


Fig 7: Tanzania Meteorological Agency – Zonal Areas

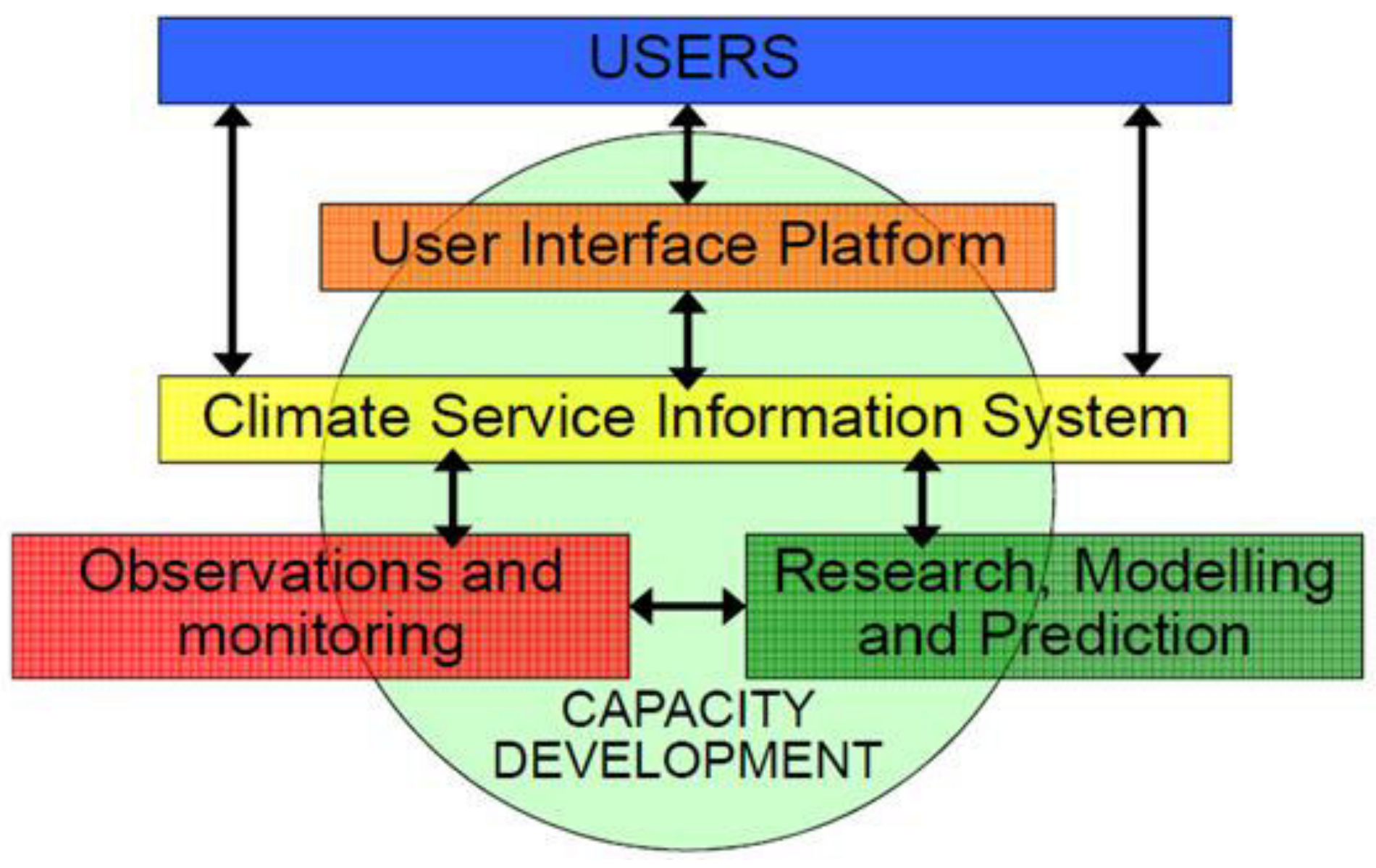


Fig 2: The Global Framework for Climate Services is based around five components identified as being necessary for producing and delivering effective climate services.

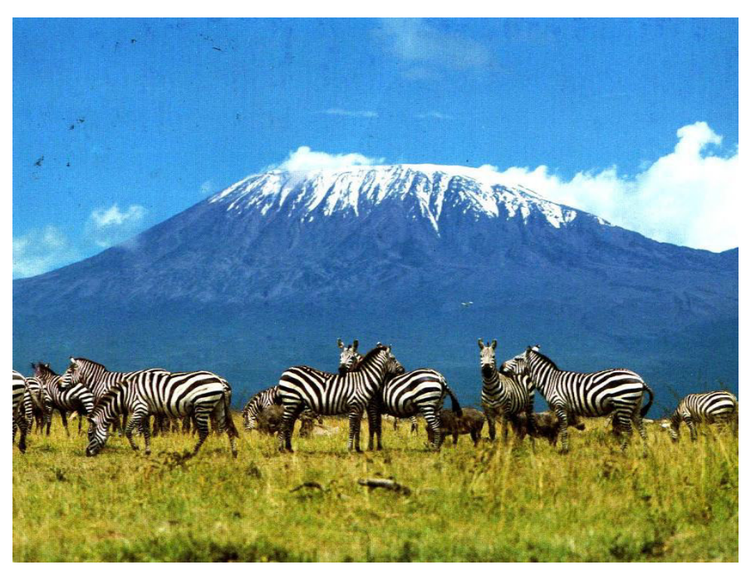
Scanning the historical paper records

Four synoptic stations were selected for the project: Dar es Salaam, Zanzibar, Arusha and Mtwara (fig 1). All historical observations of daily rainfall, maximum temperature and minimum temperature that were not already digitised were scanned. The original documents varied in size and condition and a planetary scanner was required for most. A number of large ledgers (some of which had data from the 19th Century), and were deteriorating in the archive, were also scanned to rescue these historical data.

Results and Future Work

The synoptic weather stations of Dar es Salaam, Zanzibar, Arusha and Mtwara provided one station in four of TMA's zonal areas (fig 7). For these stations the gaps in the data were identified by comparing the archive inventory produced during the project with the data already stored within TMA's climate database. For the four stations, this represented 44 years of data for each of maximum and minimum temperature and 34 years of rainfall data (about 44,500 observations). All the paper records, which were on different types of form, were scanned. These documents were found to be not suitable for automated data capture techniques due to their quality and legibility so data from these sites had to be keyed manually.

A follow on project is planned to take forward the user engagement and data digitisation work and the development of climate services. The data rescue work will build on the strategy and processes developed during this initial project with the aim of developing this capacity as a business as usual operation within the TMA. The aim is to complete the inventory of the vast archive of meteorological data, digitise more of the priority data as identified by the user engagement activities and incorporate quality management processes to release the potential of TMA's climate data. This will mean that climate services based on a reliable archive of data can be developed in the future.



> *"The project supports TMA's mission to provide quality, reliable and cost effective meteorological services meeting stakeholders' expectations thereby contributing to the protection of life and property, environmental and national poverty eradication goals."*
>
> **Dr Agnes Kijazi, Director General, Tanzania Meteorological Agency**

Supported by:

UKaid
from the British people

Met Office FitzRoy Road, Exeter, Devon, EX1 3PB United Kingdom
Tel: +44 1392 885680 Fax: +44 1392 885681
Email: name.name@metoffice.gov.uk

Historical climate data rescue in the Slovak Hydrometeorological Institute

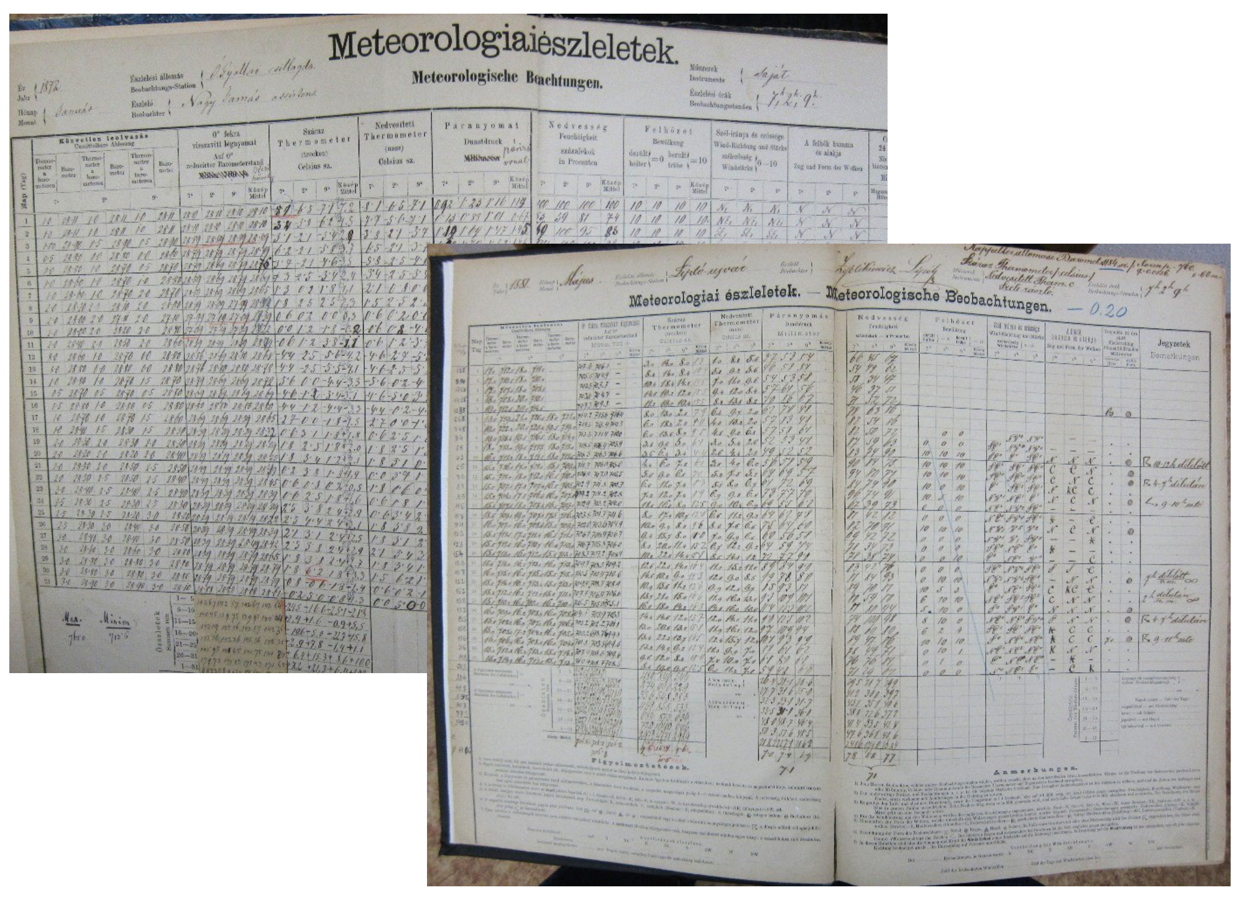
Šťastný, P., Mikulová, K.

pavel.stastny@shmu.sk; katarina.mikulova@shmu.sk

Slovak Hydrometeorological Institute,
Department of Climatology, Jeséniova 17, Bratislava

Introduction

Systematic meteorological observations in our area started in the mid-19th century, after the establishment of the Central Institute for Meteorology and Geomagnetism in Vienna in 1851. This Institute has formed the first state monitoring network of meteorological stations in the former Austro-Hungary. On the territory of Slovakia number of 8 stations was in operation in the year 1860. Number of stations gradually grew and in 1914 it was about 45 meteorological stations of higher degree in our territory. Between the two world wars over 100 stations was observed. Later, till today the number of stations has not changed significantly. Valuable observations and measurements are in stations where their position was changed only marginally and their observations were without significant breaks. Such a station is Hurbanovo (former Stará Ďala, Ógyalla in Hungarian), which was so far observing continuously and its position remained in the same place.



Meteorologiai észleletek. Meteorologische Beobachtungen.

Fig. 1. Meteorological observation records from Hurbanovo (Ógyalla) and Liptovský Hrádok (Liptó-Ujvár)

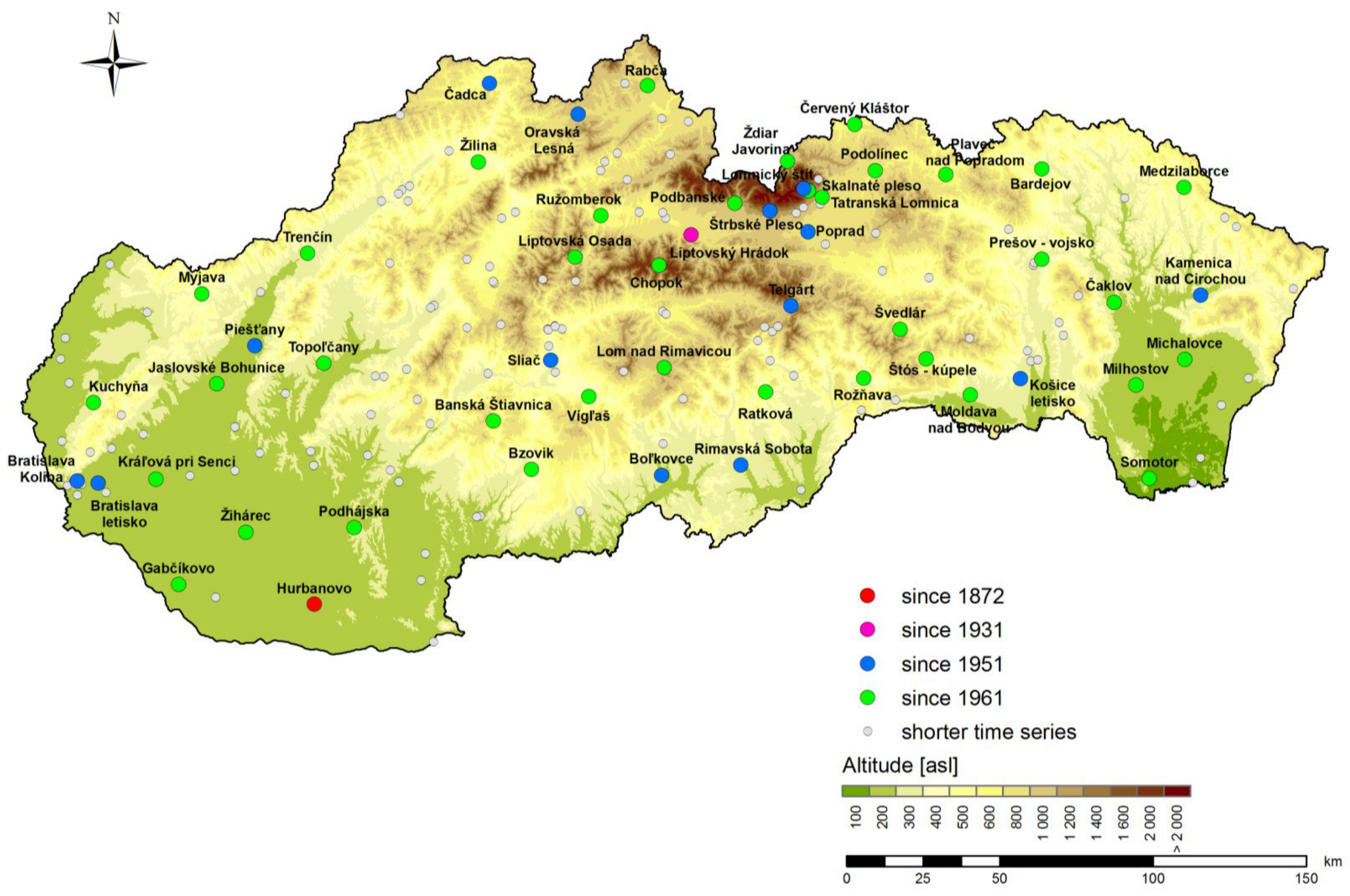


Fig. 2. Map of meteorological stations in Slovakia with data in electronical database

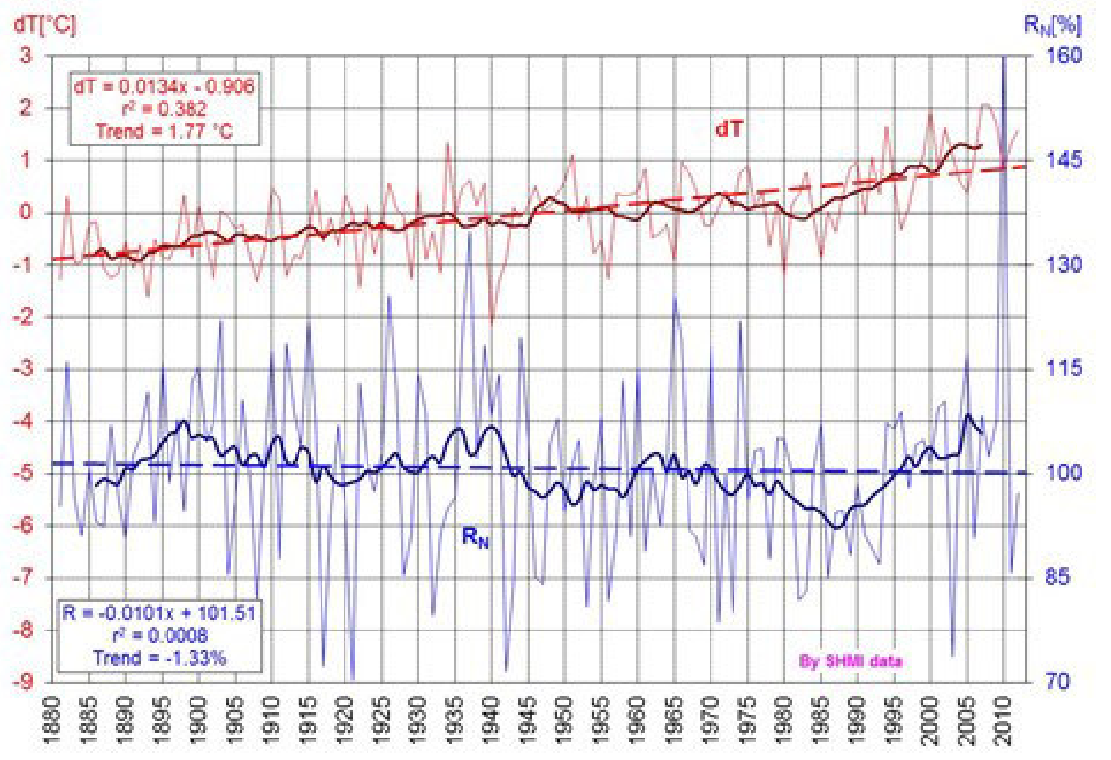
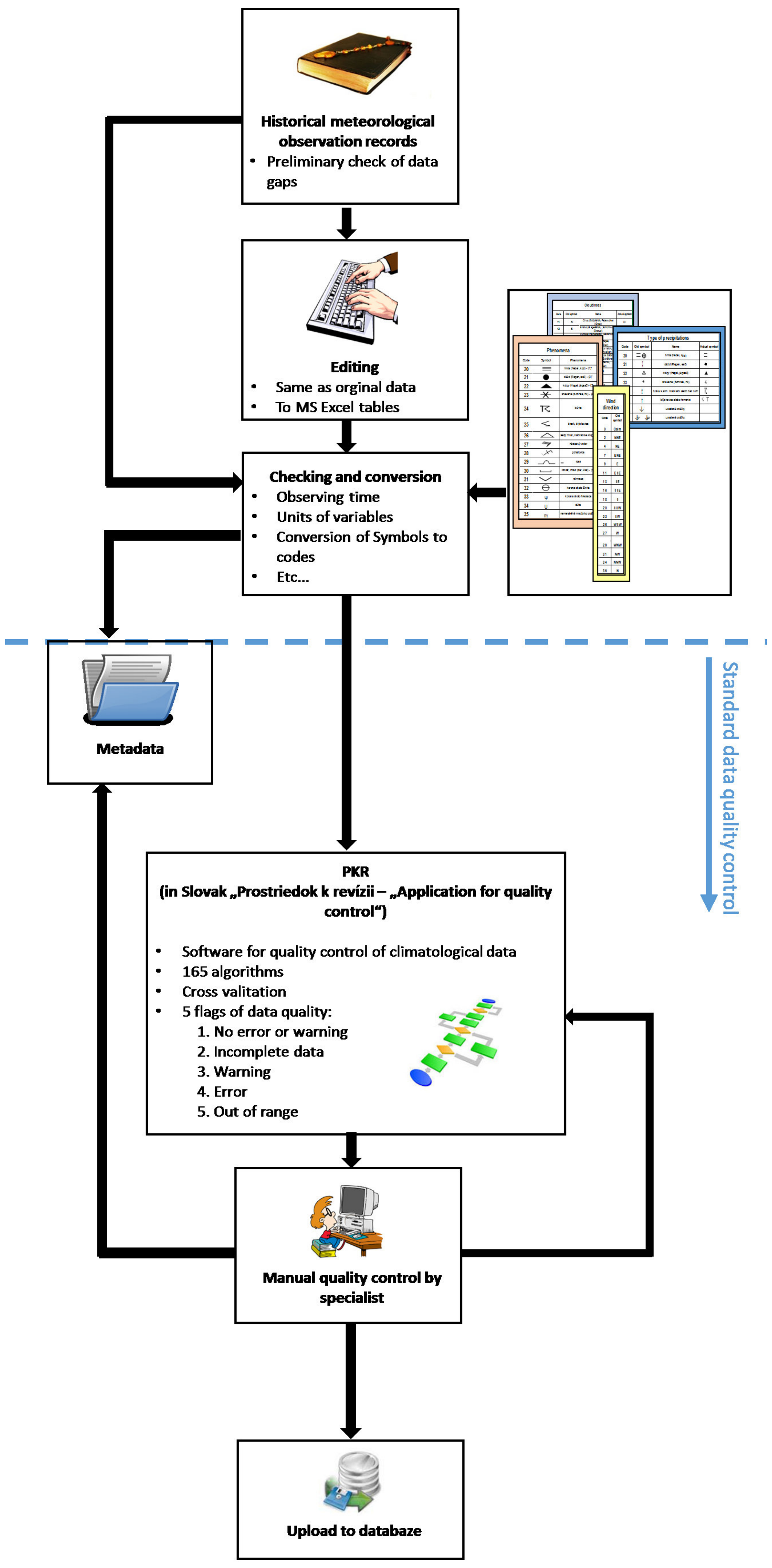


Fig. 3. Mean annual air temperature trend (dT) and atmospheric precipitation totals trend (RN) in Slovakia

Methods

Paper reports, including historical meteorological observation records from stations are placed in the archive of SHMI (Fig. 1). The electronic database of SHMI contains digitized climatological data consistently since 1961 (Fig.2). Later on the portion of data from 14 stations since 1951 was added into the database. Next effort of SHMI was to extend electronic data of daily observations as far into the past and data from Liptovský Hrádok were digitized since 1931. The time series of daily data from Hurbanovo station was extended in electronic database to the year 1872. We are currently preparing data from Liptovský Hrádok station since 1881. Hurbanovo and Liptovský Hrádok stations are often used for trend analyses for the past climate (Fig. 3). Some photos from stations surrounding are in Fig. 4.

As part of historical data rescue activities we have prepared a technological tool consisted from digitization, quality control and archiving of climatological data in the database system KMIS (Fig. 5). For each period monthly weather records vary in the number measurements and observations elements, in the order of recorded data, in some variables units and observation time. Therefore we created forms for each period in MS Excel reports, identical to paper forms for easier editing. In the next step the edited data are transformed into the current shape of meteorological data forms for semi-automatic quality control data processing. A separate chapter of data rescue effort are meteorological phenomena, which were systematically began to record in the 30-ies of the last century. The occurrence and duration of some phenomena can be identified using other meteorological elements (eg. snowing). Some observers reported dangerous phenomena especially such as data on the occurrence of storms, hail or in the notes.



Historical meteorological observation records
- Preliminary check of data gaps

Editing
- Same as orginal data
- To MS Excel tables

Checking and conversion
- Observing time
- Units of variables
- Conversion of Symbols to codes
- Etc...

Metadata

Standard data quality control

PKR
(in Slovak „Prostriedok k revízii – „Application for quality control“)

- Software for quality control of climatological data
- 165 algorithms
- Cross valitation
- 5 flags of data quality:
  1. No error or warning
  2. Incomplete data
  3. Warning
  4. Error
  5. Out of range

Manual quality control by specialist

Upload to databaze

Fig. 4. Meteorological stations in Hurbanovo and Liptovský Hrádok

Fig. 5. Process flow diagram of meteorological data rescue used in SHMI

Conclusions

The data rescue effort in SHMI will be continued in the near future. Our electronic database will be enriched with adding data from suitable stations with historical records, their number is estimated at about 15 stations.

Acknowledgment: *This contribution was supported by Slovak Research and Development Agency (SRDA) in the frame of the agreement No. APVV-0303-11.*

Data Rescue for precipitation station network in Slovak Republic

Oliver Bochníček[1] – Pavol Faško[1] – Marek Švec[1] – Zuzana Paľušová[1] – Ladislav Markovič[1]

(1) Slovak hydrometeorological institute Bratislava, Slovakia

ABSTRACT

Transparency of archive catalogues presents very important task for the data saving. It helps to the further activities e.g. digitalization and homogenization. For the time being visualization of time series continuation in precipitation stations (approximately 1250 stations) is under way in Slovak Republic since the beginning of observation (meteorological stations gradually began to operate during the second half of the 19th century in Slovakia). Visualization is joined with the activities like verification and accessibility of the data mentioned in the archive catalogue, station localization according to the historical annual books, conversion of coordinates into x-JTSK, y-JTSK and hydrological catchment assignment. Clustering of precipitation stations at the specific hydrological catchment in the map and visualization of the data duration (line graph) will lead to the effective assignment of corresponding precipitation stations for the prolongation of time series. This process should be followed by the process of turn or trend detection and homogenization. The risks and problems at verification of records from archive catalogues, their digitalization, repairs and the way of visualization will be seen in poster. During the searching process of the historical and often short time series, we realized the importance of mainly those stations, located in the middle and higher altitudes. They might be used as replacement for up to now quoted fictive points used at the construction of precipitation maps. Supplementing and enhancing the time series of individual stations will enable to follow changes in precipitation totals during the certain period as well as area totals for individual catchments in various time periods appreciated mainly by hydrologists and agro-climatologists.

Fig. 1 **Time Series in alphabetical order**

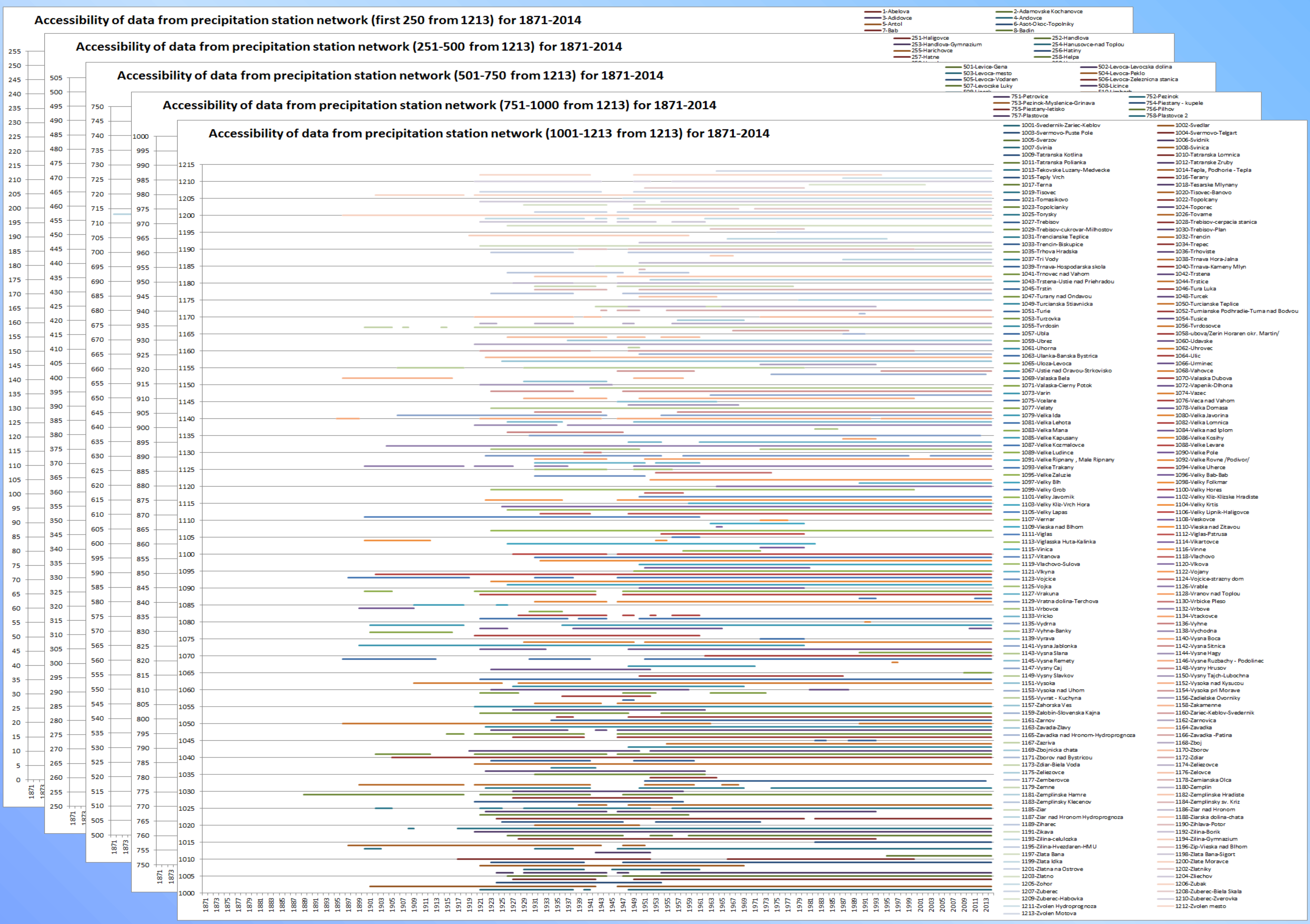
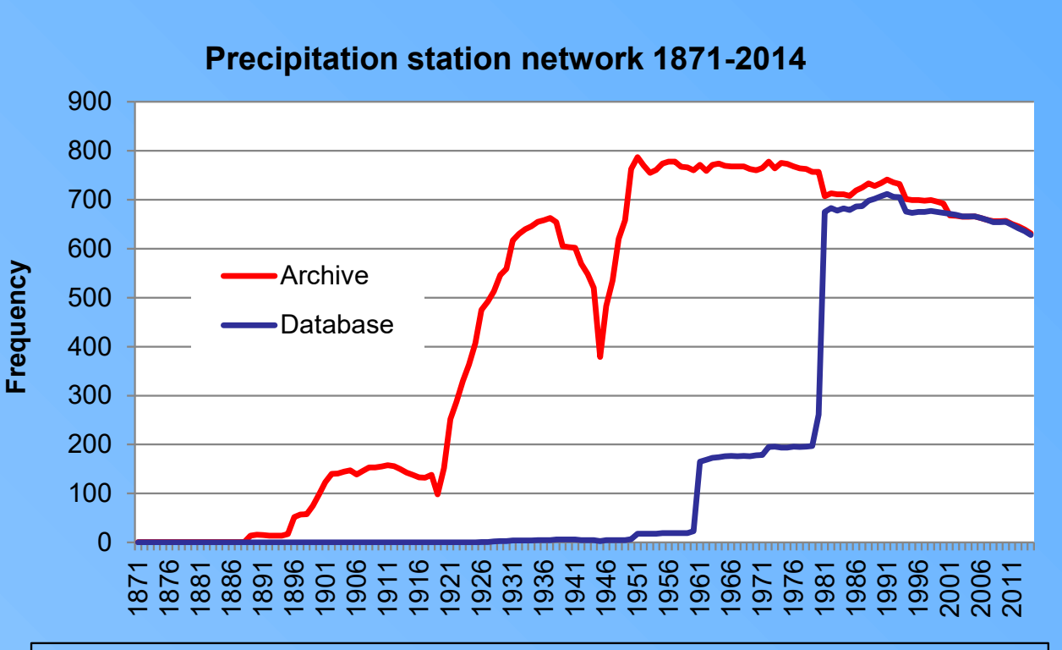


Fig. 2 **Station Frequency**



Basic steps and facts during the process of **Data Rescue** of precipitation station network in conditions of Slovakia:

1. Archive catalogue, which consists precipitation station name, period of observation (from-to), station indicative (if exists) and the place (drawer) in archive. Stations were in alphabetical order, with many errors and duplicities coming from the fact that many stations have changed location during the history of observation so they got two or three names with changeable order in the list. These duplicities have been eliminated after first round of station visualization in tabular form.
2. Stations have been sorted in alphabetical order (to better find if indicative does not exist) and each precipitation station has adopted code number. Line graphs have been drawn (Fig. 1). Because of excel sheet does not enable to display more than 250 time series (lines), 1213 precipitation stations have been divided into five line graphs sorted by station name. This is useful for knowing the length of the time series of concrete precipitation station.
3. Number of precipitation time series since 1871 can be seen in Fig. 2, archived (red line) and digitalized (blue line). It means that almost 60% of precipitation data are not in database.
4. In coincidence with the annual books the latitude, longitude as well as altitude have been added for the exact period and precipitation station. Geographic coordinates have been transformed to JTSK coordinates for GIS visualization.
5. Precipitation data are useful for many sectors especially for hydrology. That's why we use hydrological catchments (Fig. 3). Precipitation stations have been placed into the catchments of Slovakia (Fig. 4). To be clear concrete catchment should be chosen (example Fig. 5 4-23-01 and later on other partial sub catchments of 04-23).
6. Overlapping periods of observations at precipitation stations in chosen 04-23-01 catchment are shown (Fig. 6).
7. Data processing according to standard procedures and rules (e.g. MISH, MASH, AnClim etc.) (Fig. 7).
8. Special (future) task: Mountain stations (often in rare location) not yet digitalized have been separated from the list (Fig. 8). Their time series have various length lying in various altitudes and catchments. According to the digitalized data stored in database and not digitalized data observing at the same time mountain stations may be supplemented and homogenized.
9. Most of the stations not in database but important for mountain precipitation investigation lies in the altitude 500-2000 m a. s. l. (Fig. 9) with the peek of occurrence in altitude interval 701-800 m. Benefit consists of better identification of precipitation and snow conditions in higher altitudes for space data processing dealing with river drain and flow from concrete catchment as well as prediction of avalanches alert.

Fig. 3 **Catchments**

Fig. 4 **Precipitation station network 1871-2014 in catchments**

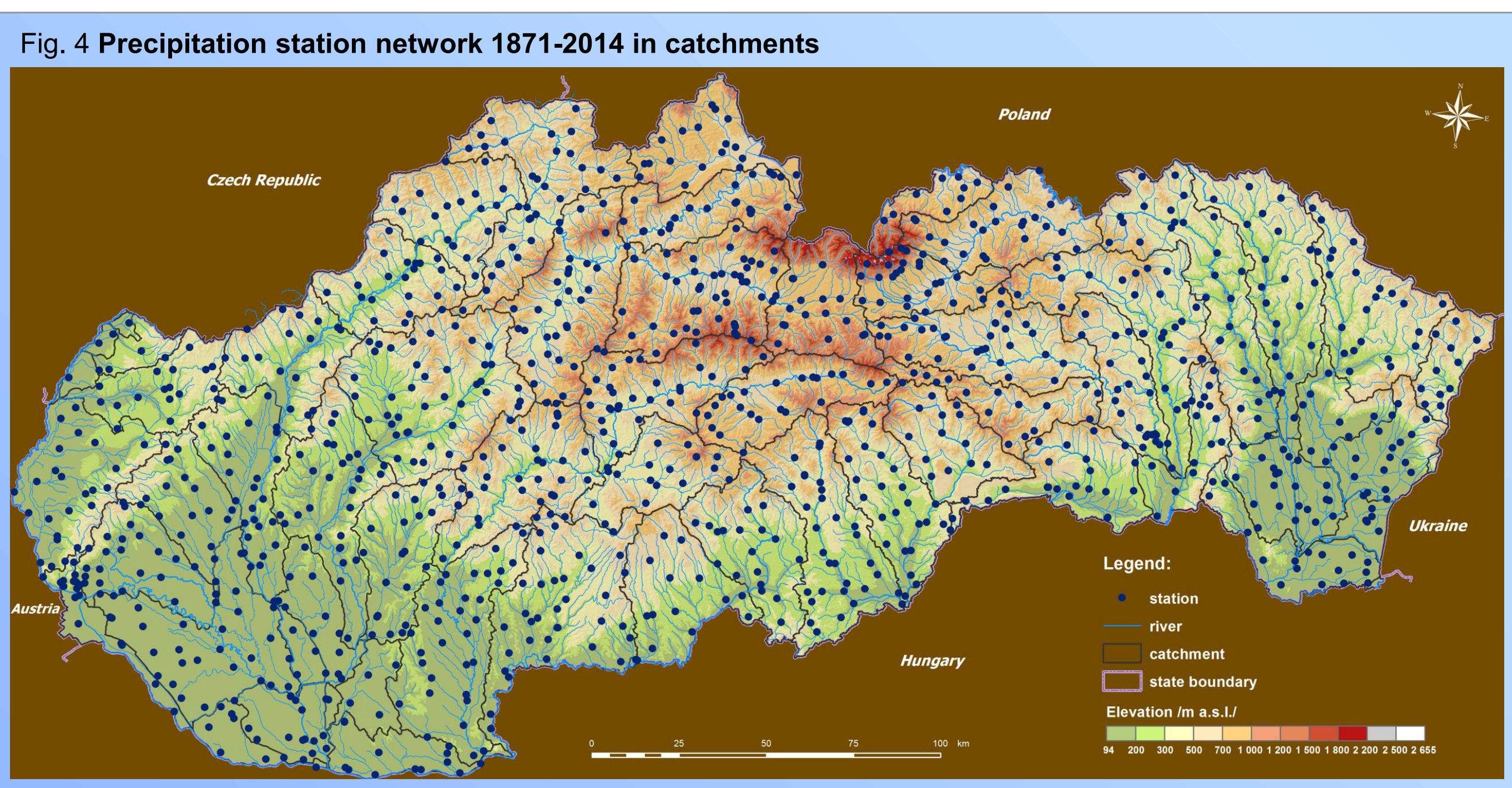


Fig. 5 **Hron River catchments** 4-23-01 (4-23-02, 4-23-03, 4-23-04, 4-23-05)

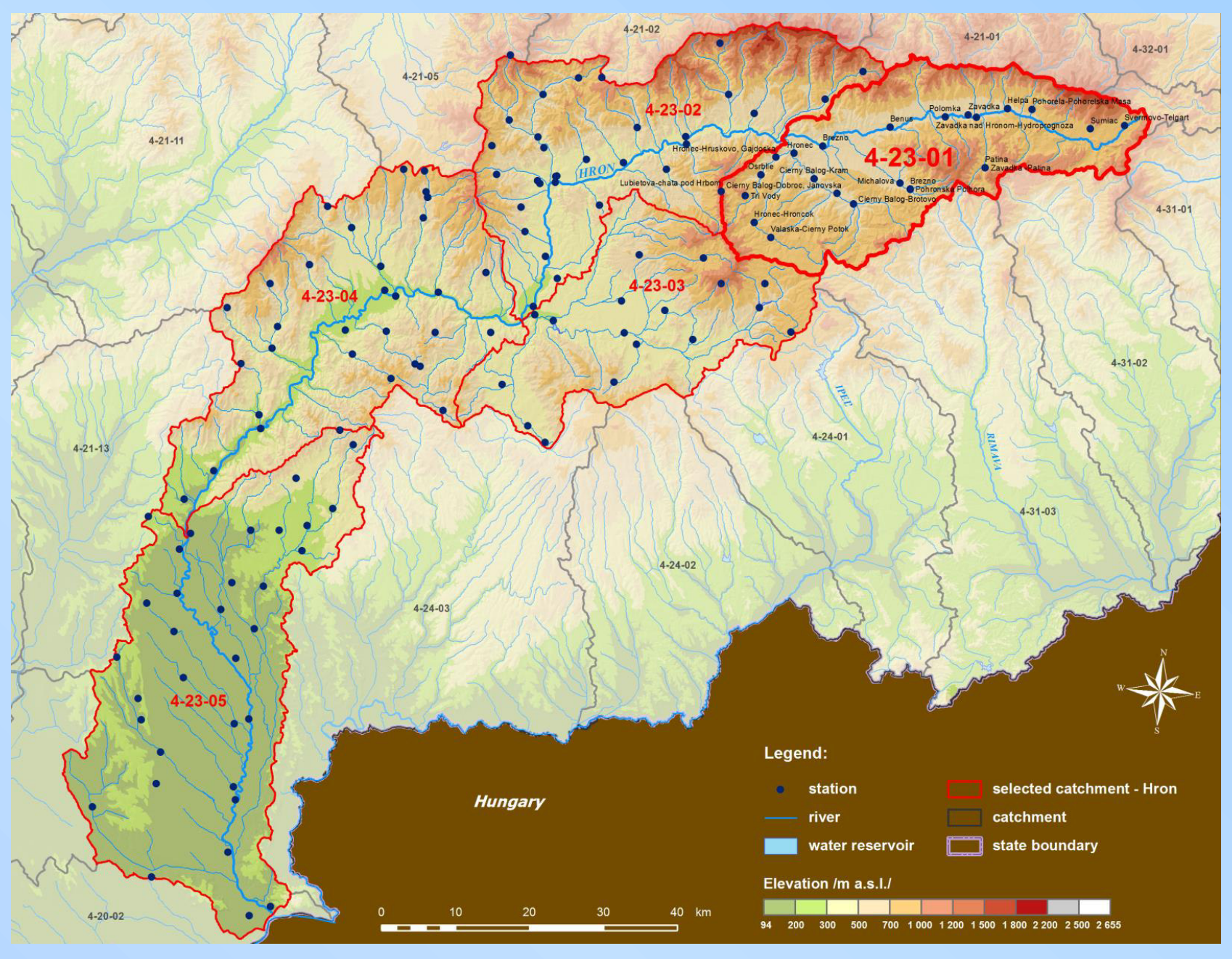


Fig. 6 **Precipitation station network** (04-30-03 catchment) **in time series**

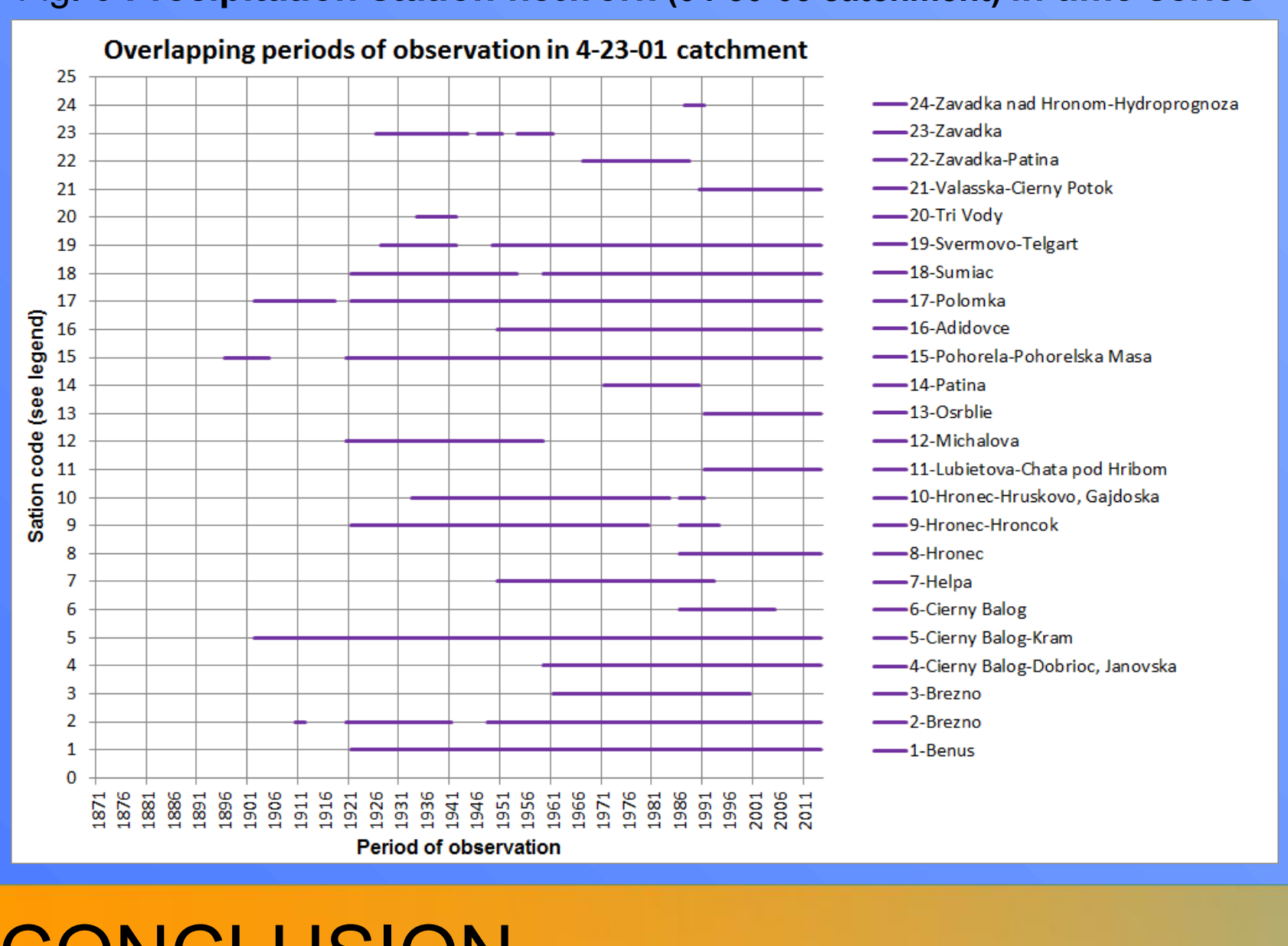
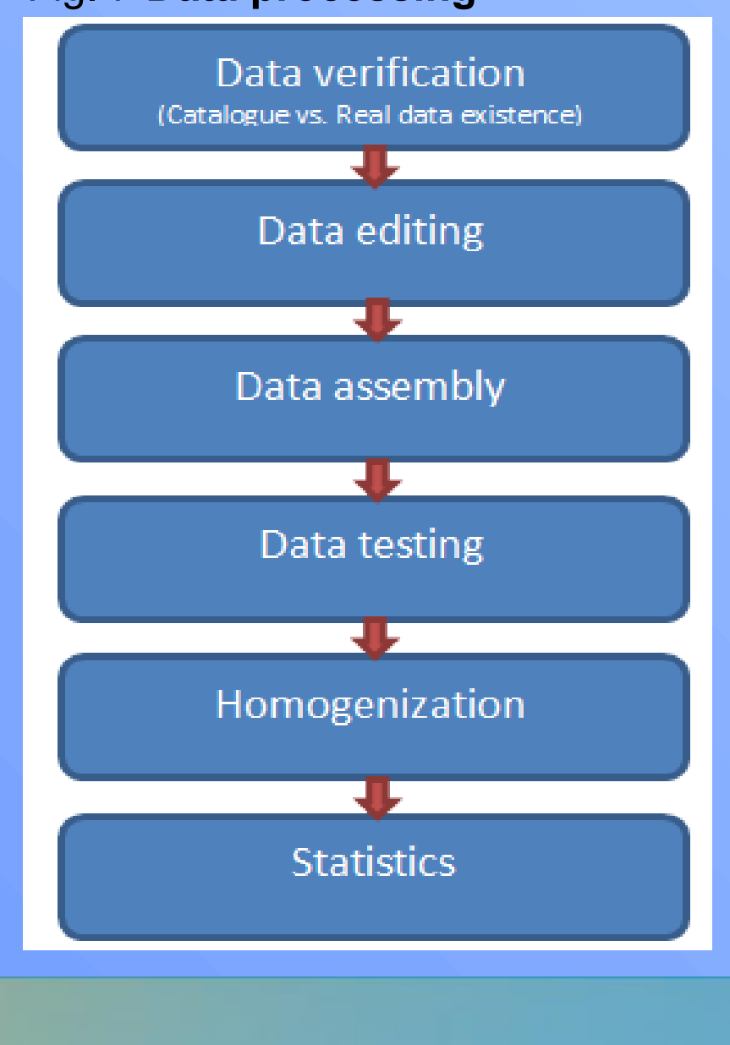


Fig. 7 **Data processing**



Data verification
(Catalogue vs. Real data existence)

Data editing

Data assembly

Data testing

Homogenization

Statistics

Fig. 8 **Mountain stations (not digitalized) – placement in catchment**

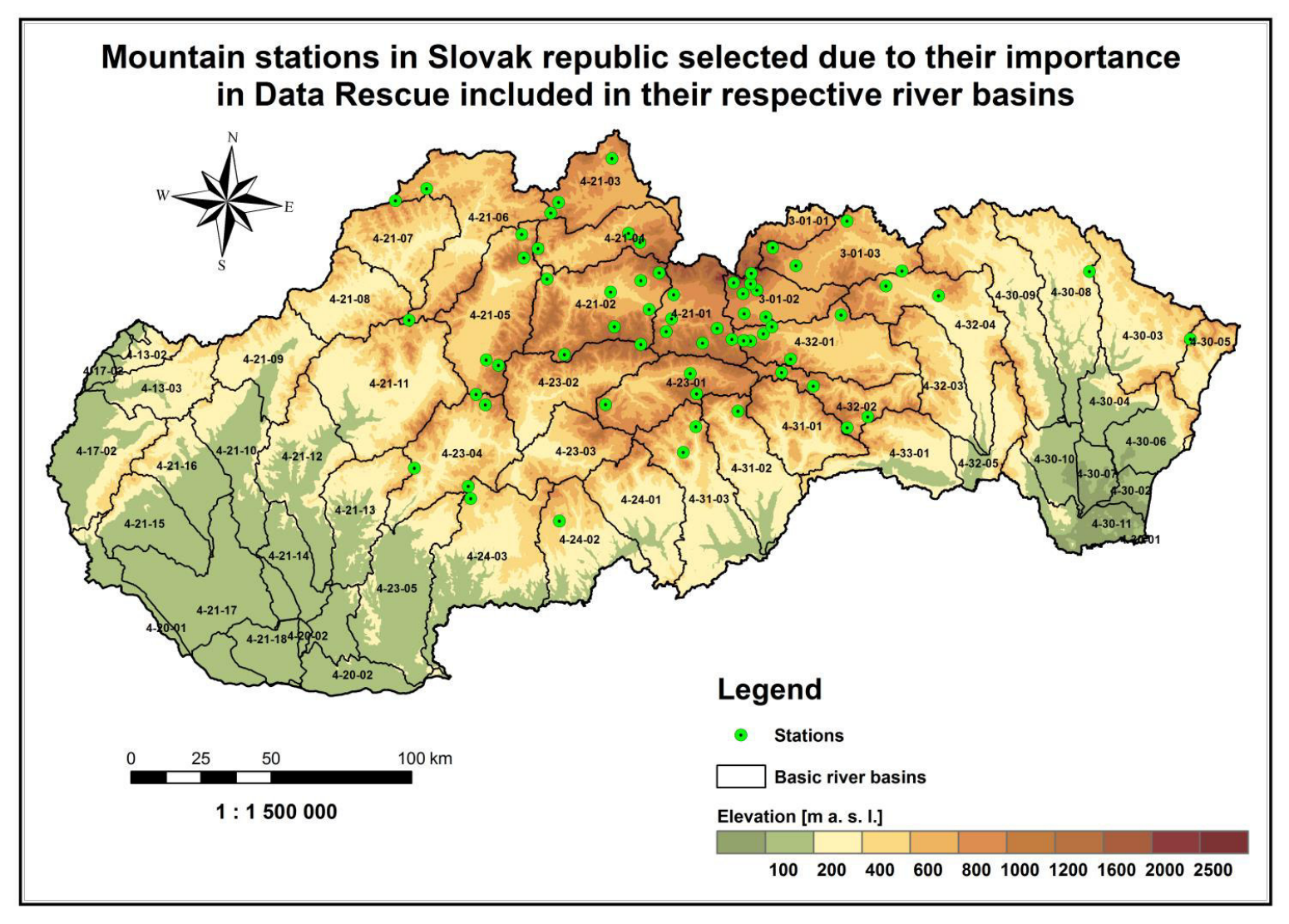
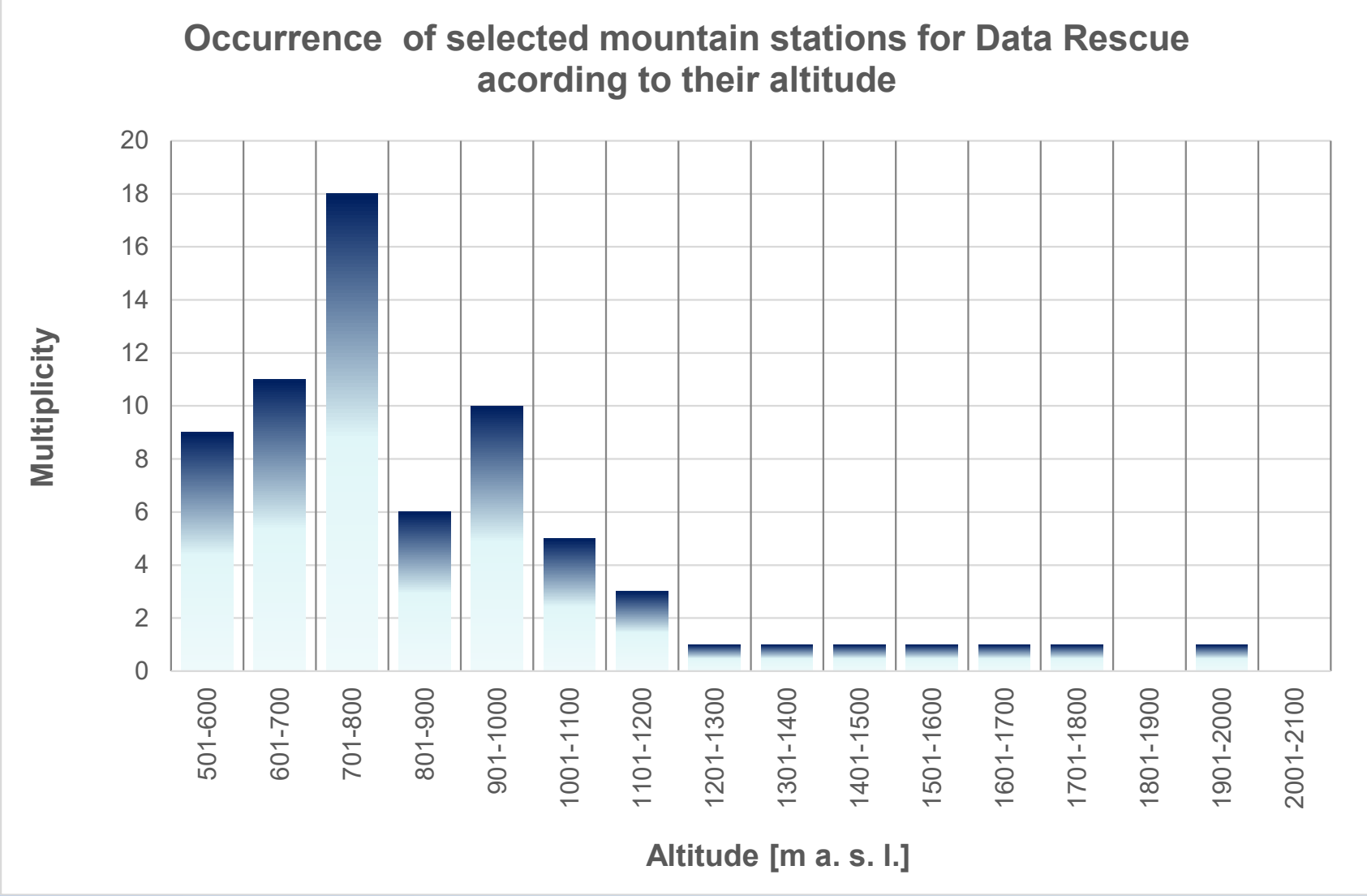


Fig. 9 **Mountain stations (not digitalized) – altitude occurrence**



CONCLUSION

Archive data represent national heritage. Data Rescue is the ongoing process of preserving all data at risk of being lost due to deterioration of the medium and digitizing current and past data into computer compatible form for easy access. The establishment of the data rescue team plays important role and its activities have a high priority within WMO Programs that's why national activities have been encouraged to start. In poster the basic steps during the data rescue process have been mentioned. Many problems dealing with station names consisting of more words, or having been written down twice or more times in the catalogue under the identical time of observation concerning place and position of the station in annual books had to be solved before assigning code for line graph (Fi. 1) describing the time series. Most of the precipitation data have not been digitalized yet (Fig. 2). Professional staff is shrinking (retiring) and students cannot substitute this kind of work. Daily or monthly data have to be decided for digitalizing. Project for data rescue should be supplied and managed internationally. Precipitation data dropping in certain catchment (e.g. there is depicted the catchment of the Hron River) play significant role for the assessment of river flow rate and runoff downstream (especially for hydrology Fig. 3, 4 and 5), snow loading and water snow cover, preparation of standards and determination of areas from the point of avalanches alert as well as tourism. These knowledges lead us to the future solution of mountain stations operated often by foresters (Fig. 8 and 9) and situated in very interesting and significant altitudes in such a way as to include the data from these stations into the future data processing. Homogenization of precipitation data of individual catchments will result in possibility to compare precipitation relations in the respective season in connection with hydrological relations of the same season and catchment for surface as well as ground water.

Acknowledgment:

This work was a part of the projects: "Development of technology of spatial data processing of the climate system" ITMS 26220220102, and "Applied research of methods for determining the climatic and hydrologic design parameters", ITMS 26220220132 funded from the EU Structural Funds
The authors thank to the Join Research Centre of the EC for supporting the project "CLIMATE OF THE CARPATHIAN REGION", by force of some digitalization data arose.

REFERENCES

http://www.wmo.int/pages/prog/hwrp/datarescue.php


EUMETNET
European Meteorological Services Network

www.zamg.ac.at

ZAMG
Zentralanstalt für Meteorologie und Geodynamik

Progress of European Data Rescue Activities – developments during the last two years

Ingeborg Auer, Barbara Chimani, Silke Adler (ZAMG), Michael Begert (MeteoSwiss), Anna Frey (FMI), Jose A. Guijarro (AEMET), Monika Lakatos (OMSZ), Hermann Mächel (DWD), Marc Prohom (Meteocat) , Dubravka Rasol (DHMZ), Miroslav Řepka (CHMI), Weine Josefsson (SMHI)

Introduction

Long-term datasets are of great importance for climate research. They allow describing past climate variability highly resolved in space and time, are important for re-analyses and model evaluation. Especially early instrumental series are the connecting link to the paleoclimatic community. In 2011 the EUMETNET data rescue portal has been launched presenting European data rescue activities. So far 23 EUMETNET members plus the Catalonian and the Georgian meteorological Services take part in this activity. The poster wants to summarize the progress achieved during the last two years.

Progress

Sweden, Switzerland, Spain, Finland, Hungary, Croatia, Germany and Czech Republic reported about their progress during the last two years. For details please have a look at https://www.zamg.ac.at/dare/activities/data-inventory

Sweden

SMHI has a digitizing project running for many years. Most Swedish observations are in digital form since 1961 and can also be found in the Meteorological database called MORA, including most variables and for every three hours. Quite a lot stations are available as monthly values for temperature and precipitation since the start of observations 1859/1860.

A substantial amount of digitized data prior to 1961 is not yet available in MORA, but a project has started to put these data into MORA.

Switzerland

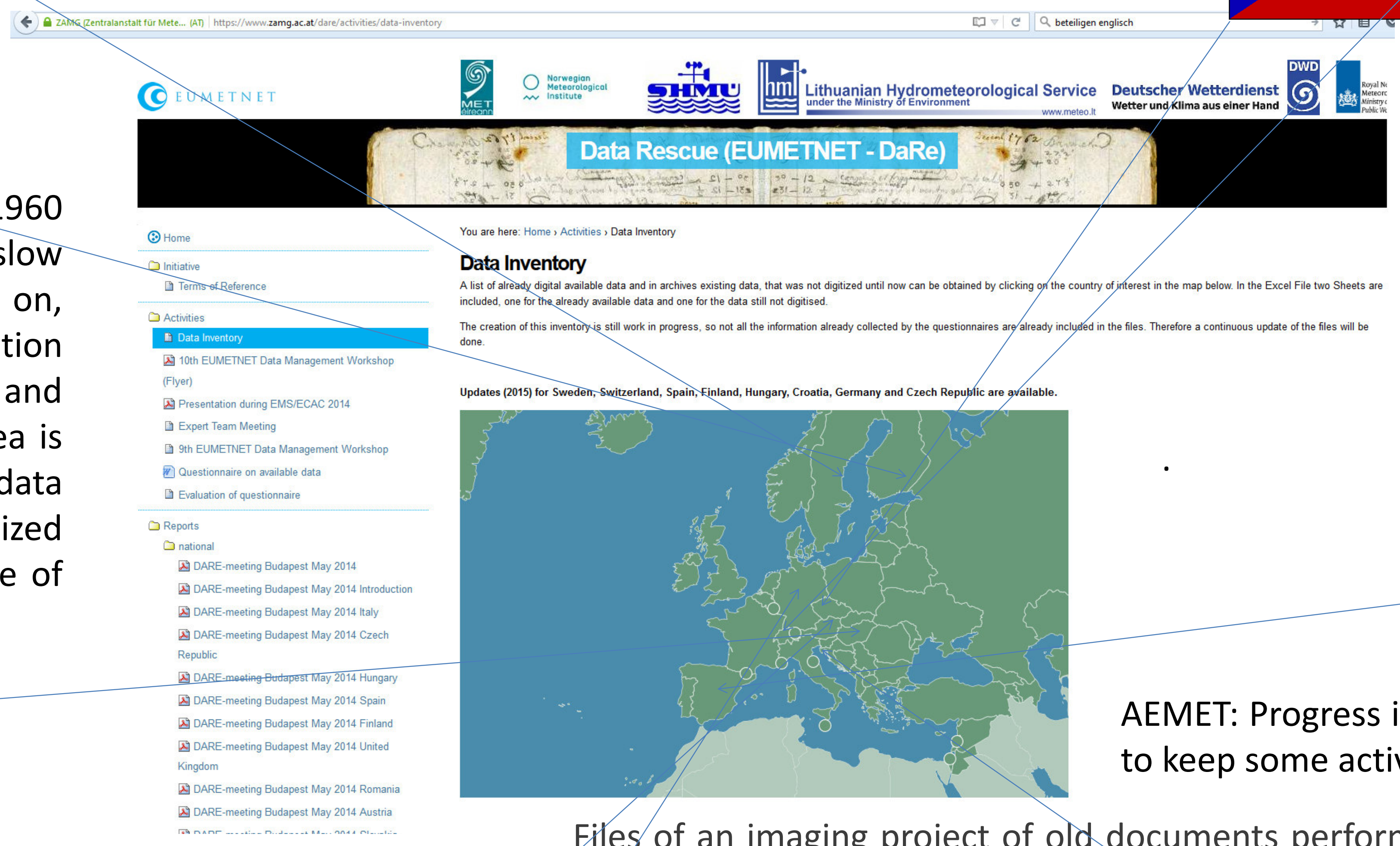
MeteoSwiss has run several digitizing projects within the last decades. The latest one called DigiHom aimed to digitize all temperature (Tmean, Tmin, Tmax), precipitation and sunshine duration series of the Swiss National Basic Climatological Network (Swiss NBCN) back to 1864. The Swiss NBCN combines the most important climate stations of Switzerland and consists of 29 climate and 46 additional precipitation stations. Based on an internal requirement analysis with respect to completeness of available digitized data and homogenization purposes a set of stations/variables were defined to be digitized next. An internal project was started where two people digitize the values beside their operational work. MeteoSwiss will continue to digitize station data step by step.

Finland

Most Finnish observations are in digital form since 1960 and are available in operational database. Some slow progress for the older datasets has been going on, basically concerning digitizing of historical precipitation forms. It is planned to start a more controlled and organized process concerning historical data. The idea is to have organized documentation about the types of data to be digitized. The rescue of the data will be prioritized based mainly on the customer needs and big picture of the importance (defined by experts opinions).

Czech Republic

Almost all precipitaton and temperature data of very long series (with some exceptions) have been digitized and quality control has started. Hundreds of photos of archived historical reports (books) have been made in Brozany central archive. Many books and reports were found that have not been digitized yet there (particularly stations with shorter series). All regional offices had some money for digitization this year, but grant is finished now. This will cause problems in the future.



Data Inventory

A list of already digital available data and in archives existing data, that was not digitized until now can be obtained by clicking on the country of interest in the map below. In the Excel File two Sheets are included, one for the already available data and one for the data still not digitised.

The creation of this inventory is still work in progress, so not all the information already collected by the questionnaires are already included in the files. Therefore a continuous update of the files will be done.

Updates (2015) for Sweden, Switzerland, Spain, Finland, Hungary, Croatia, Germany and Czech Republic are available.

Hungary

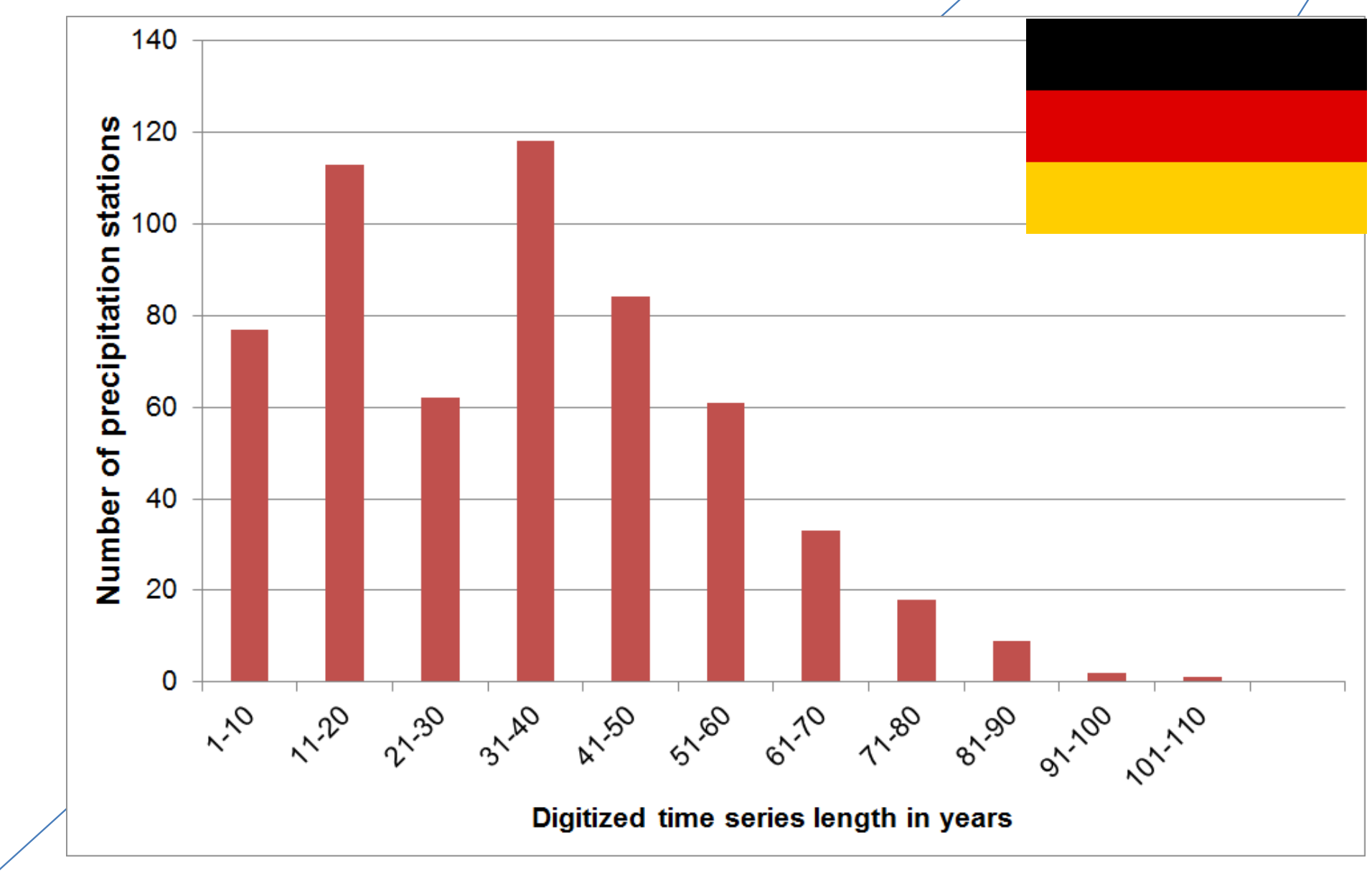
The progress of the digitization depends on the availability of resources. Records of several climate stations were digitized in the project "Complex risk management system for agriculture" at the Hungarian Meteorological Service in 2014. The observers contributed also to this activity beyond their operational tasks. Hungarian Meteorological Service will continue data rescue step by step on regular basis.

Spain

AEMET: Progress is limited to individual initiatives of the available staff to keep some activities alive:

Files of an imaging project of old documents performed in the nineties have been incorporated to the national climatological data bank; digitization of old daily data (focusing on the longest series), old phenological data and old radiosonde printed outputs; paper strips pluviograph records of old observatories are being digitized in cooperation with the Madrid Polytechnic University.
An integral Data Rescue project is being devised in AEMET to coordinate and stimulate a more consistent approach to the recovery of the oldest data.

Additional meteorological data have been digitized at the Meteorological Service of Catalonia (Autonomous Government of Catalonia) including metadata. A project to recover new data for late 19th century from historical press sources is on the way.



Germany

The project KLIDADIGI ended in 2013, but the digitization work at Deutscher Wetterdienst will be **continued** with reduced staff. During the last ten years about 580 daily precipitation series (with 21890 station years, or 2500 station years during 2013-2015) were digitized at DWD. The progress of digitizing climate stations (3 observations per day of up to 9 variables) during the years 2013-2015 was about 670 station years (or 19 series). During the last ten years the amount of digitized climate series was 3530 station ears or 67 series.

Croatia

Croatia has started collaboration with the Croatian State Archives which should give support in data rescue. The scanning process should start by the beginning of 2016.
Croatia digitized precipitation data: 36 stations, most of them beginning around 1950 (the oldest from 1921), altogether 1028 years of data.

Austria

Due to staff shortage data rescue is progressing rather slowly. Five stations (Wien Schönbrunn, Krippenstein, Reichenau/Rax, Fuchsenbigl and Traisen) could be extended back in time. ZAMG will continue the digitization of data step by step.

contact: ingeborg.auer@zamg.ac.at

STRUCTURE AND PRACTICAL APPLICATIONS OF METEOROLOGICAL METADATA

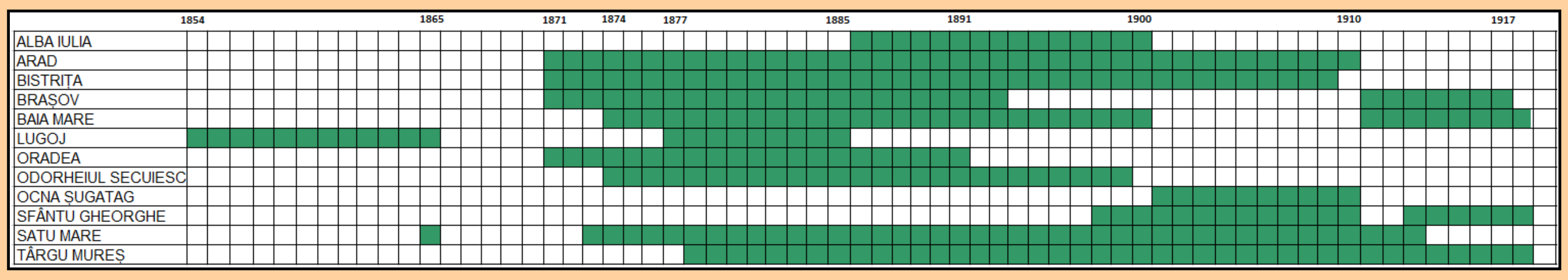
Ancuta Manea, Elena Toma, Marius-Victor Birsan, George Tudorache, Alexandru Dumitrescu

Meteo Romania (National Meteorological Administration), Sos. Bucuresti-Ploiesti 97, Bucharest, Romania

> "The mind is not a vessel to be filled, but a fire to be kindled." — Plutarch
>
> "The important thing is to not stop questioning. Curiosity has its own reason for existence. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend a little of this mystery each day." —"Old Man's Advice to Youth: 'Never Lose a Holy Curiosity.'" LIFE Magazine (2 May 1955) p. 64" — Albert Einstein

I. Data recovery in 2015

As a result of data and metadata continuous recovery action that is taking place in 2015 within the Romanian National Meteorological Administration, we can present the following information about scanning documents with old meteorological data recorded in Romania, in many cases prior to the establishment of the National Meteorological Institute in Romania (1884), for some of the meteorological stations that are functioning today:



In some cases, the recovered data present some gaps.

After the scanning procedure, data are digitized and included in the National Meteorological Database.

II. Metadata recovery in 2015

The Romanian Meteorological Network comprises today a number of 158 meteorological stations.

Meteorological station

OMU Peak

SYNOP index: 15280

Settled on 8th September 1927

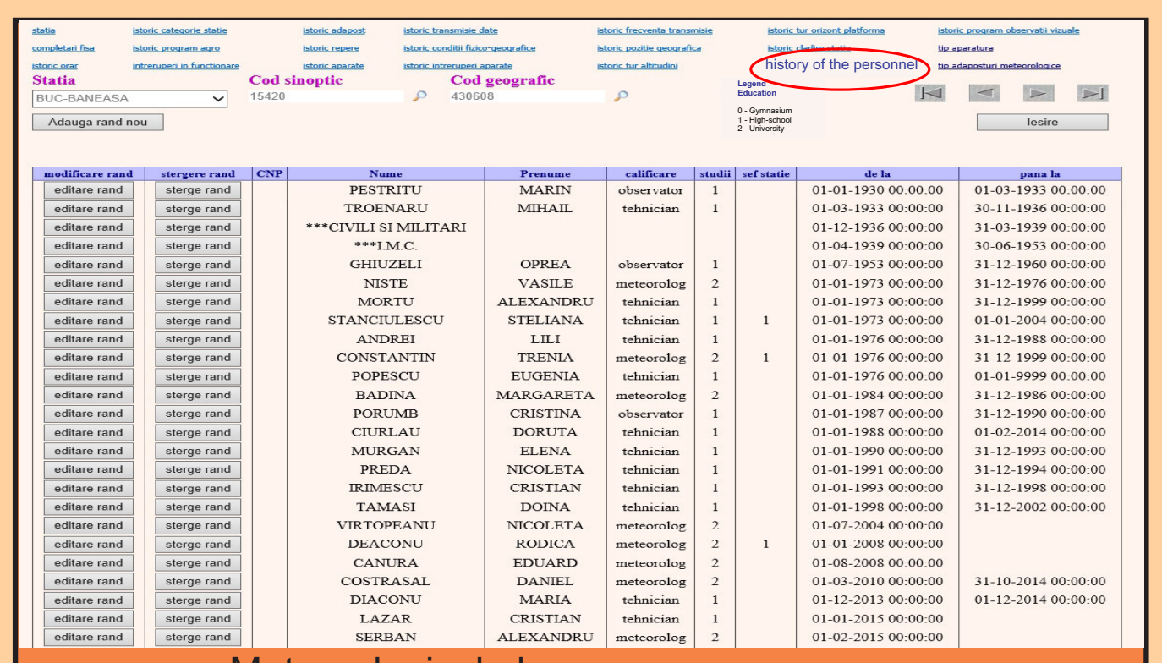
Peak mountain station

For all these stations we have historical references. In 2015, 24 historical reference sets were updated. Each historical set is organized as follows by chapters:

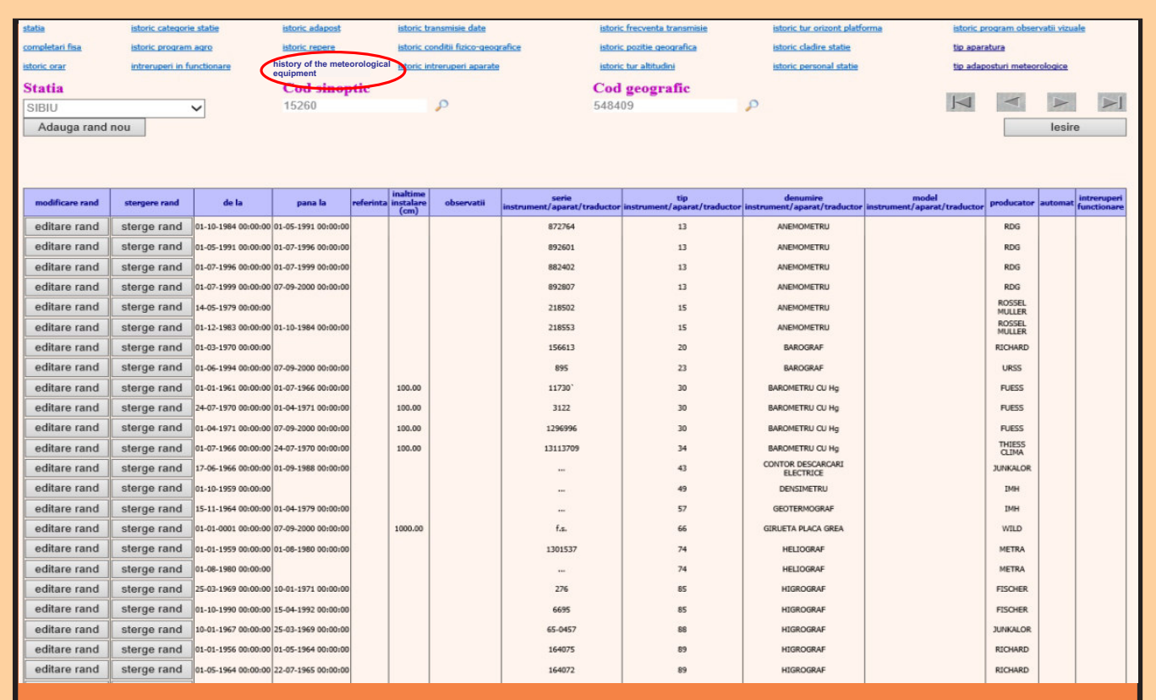
- History of the settlement of the meteorological station
- Access to the station.
- Description of the geographical features of the surroundings and the human activity
- History of the meteorological platform
- History of the visibility landmarks
- Administrative situation
- History of the personnel
- History of the meteorological equipment
- Present-day photos
- Historical photos
- Satelite images with the surrounding area
- Historical documents where the meteorological station is mentioned (scans)

III. The management of the meteorological metadata – national level

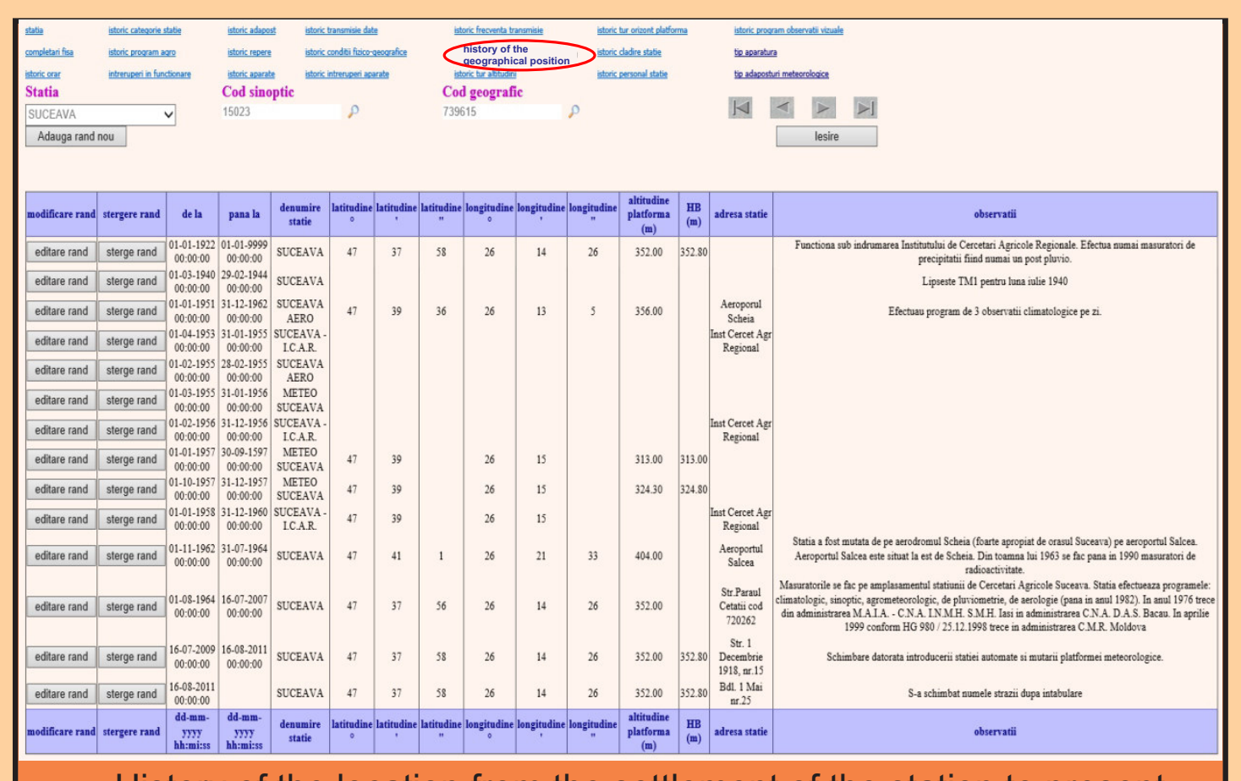
III.1. Meteorological station metadata



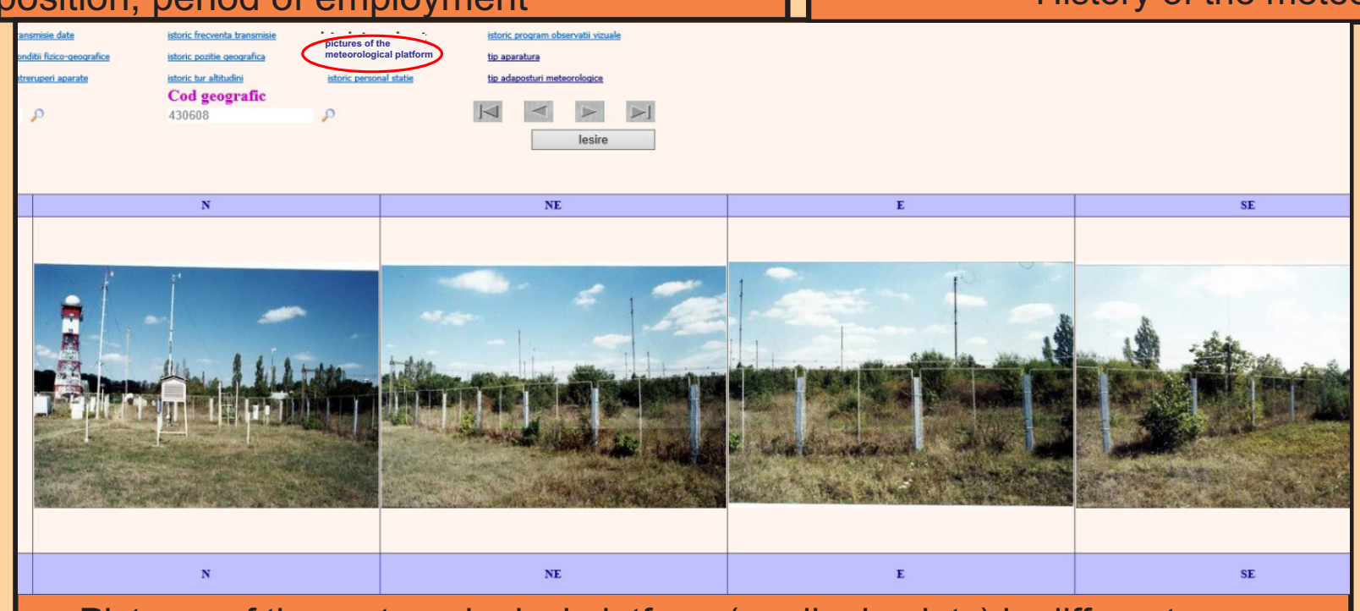
Meteorological observers – name, education, position, period of employment



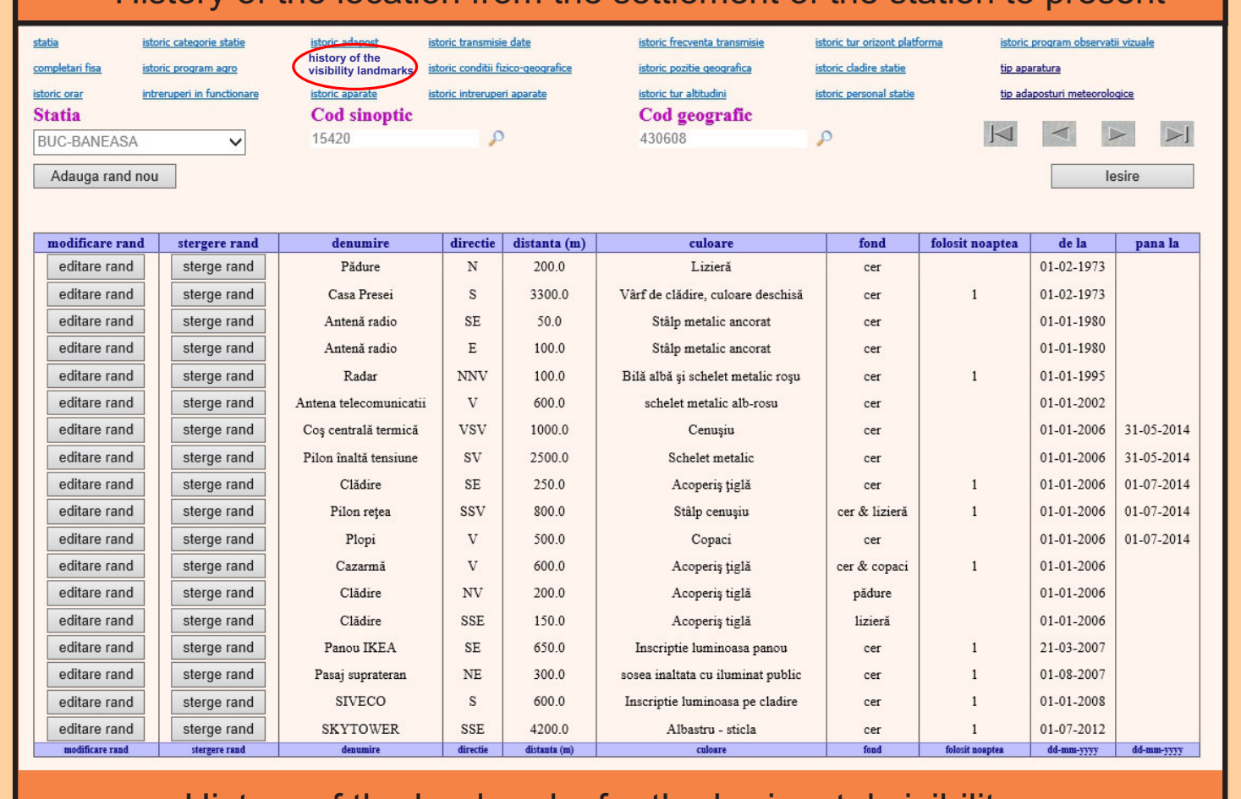
History of the meteorological equipment



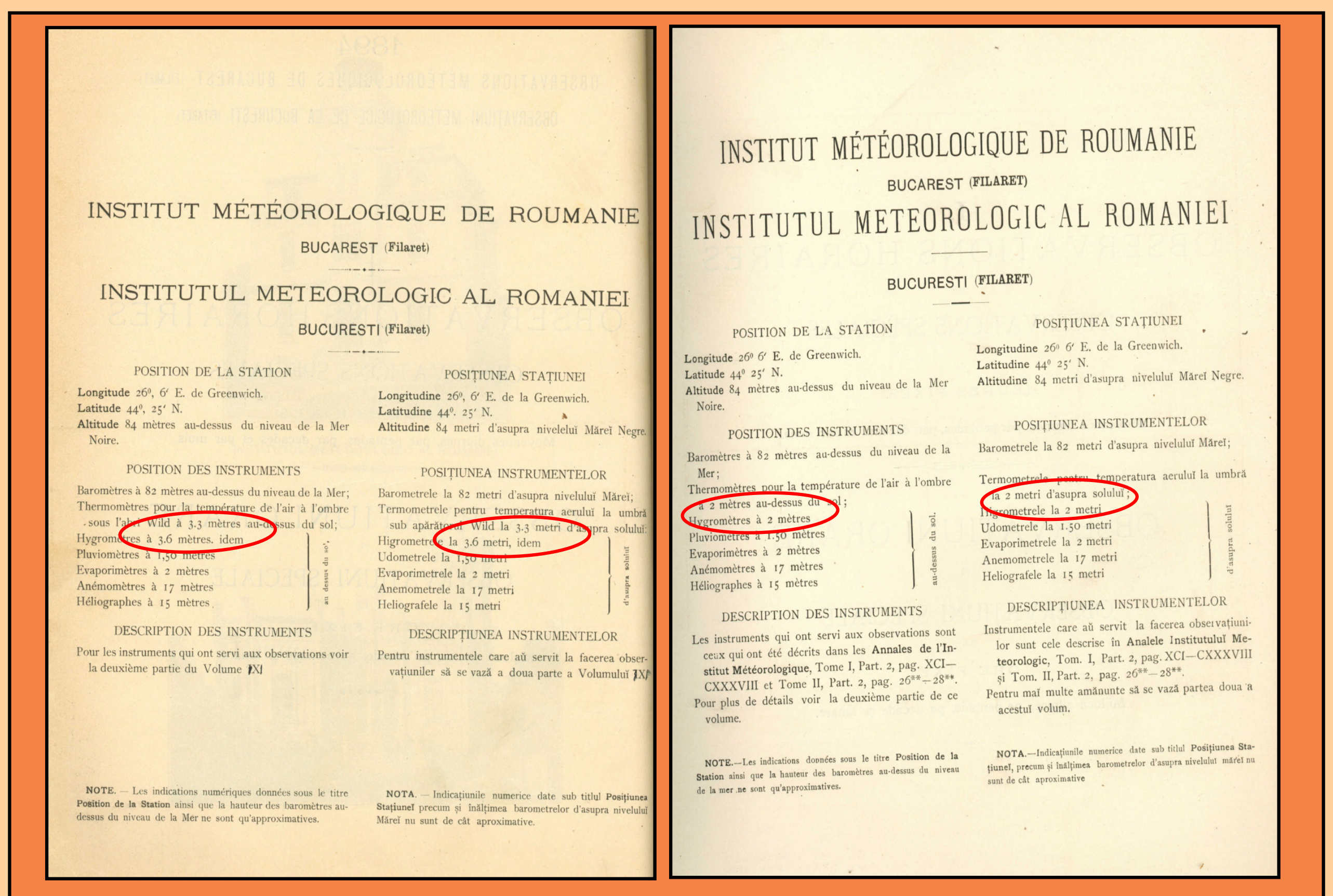
History of the location from the settlement of the station to present



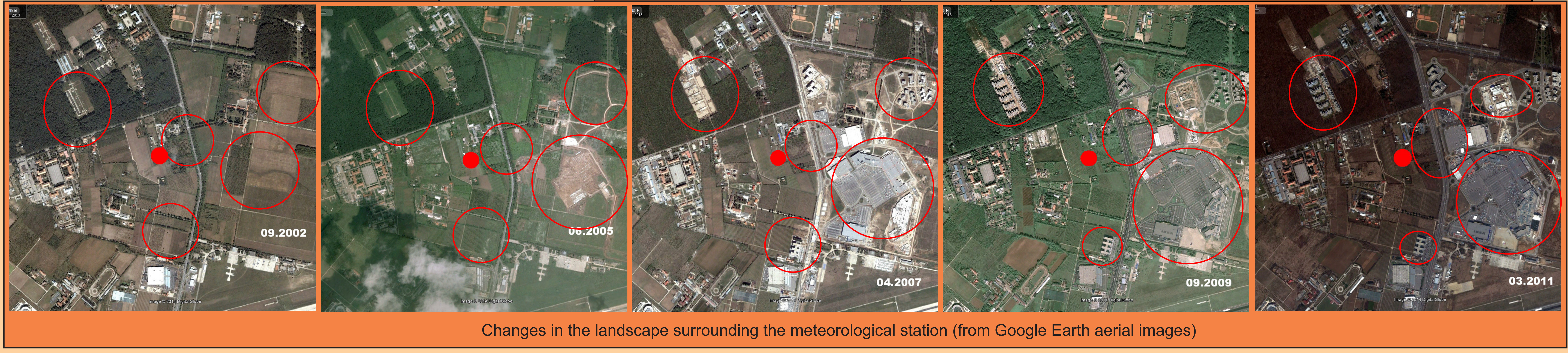
Pictures of the meteorological platform (cardinal points) in different years



History of the landmarks for the horizontal visibility

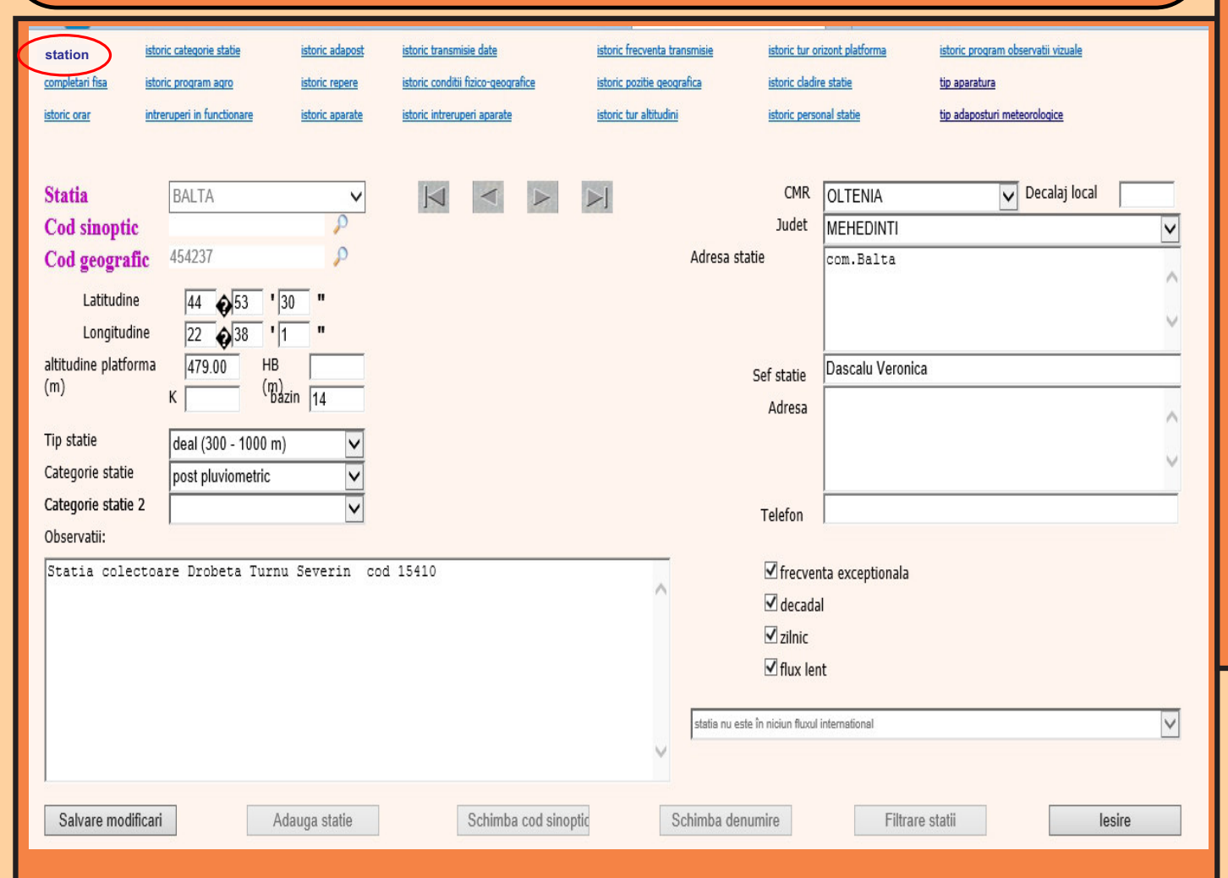


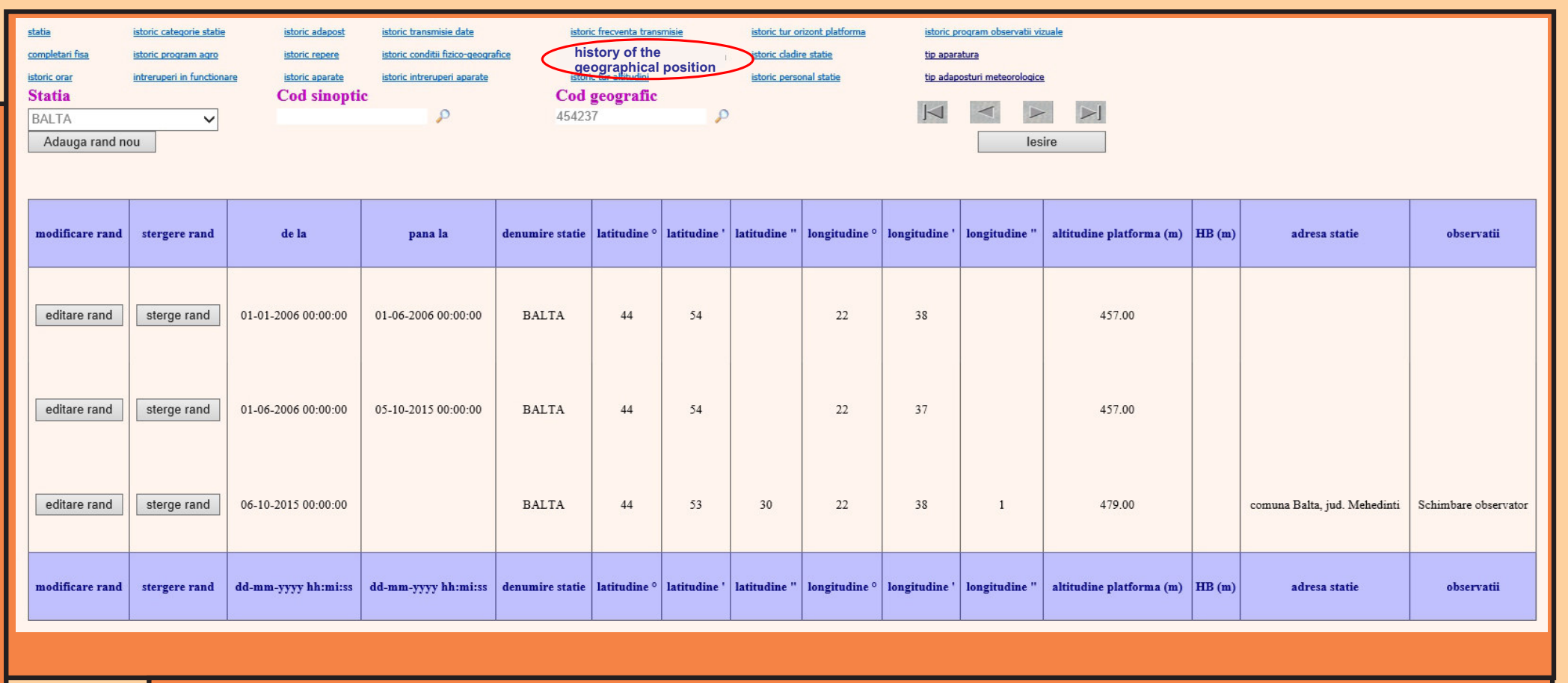
Annales de L'Institut Météorologique de Roumanie pour l'année (1889 Tome V, 1894 Tome X)

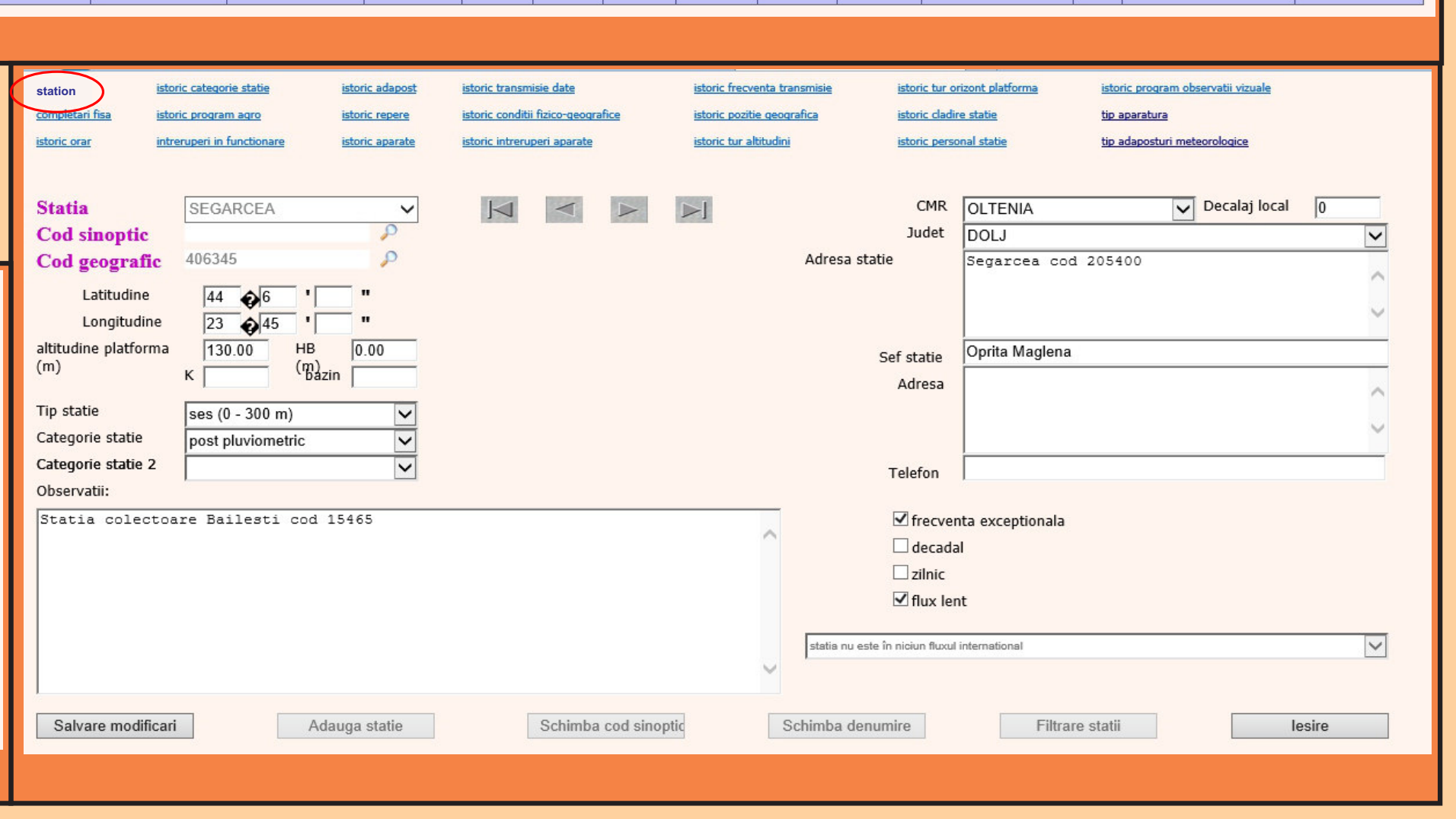


Changes in the landscape surrounding the meteorological station (from Google Earth aerial images)

III.2. Rain measuring points metadata







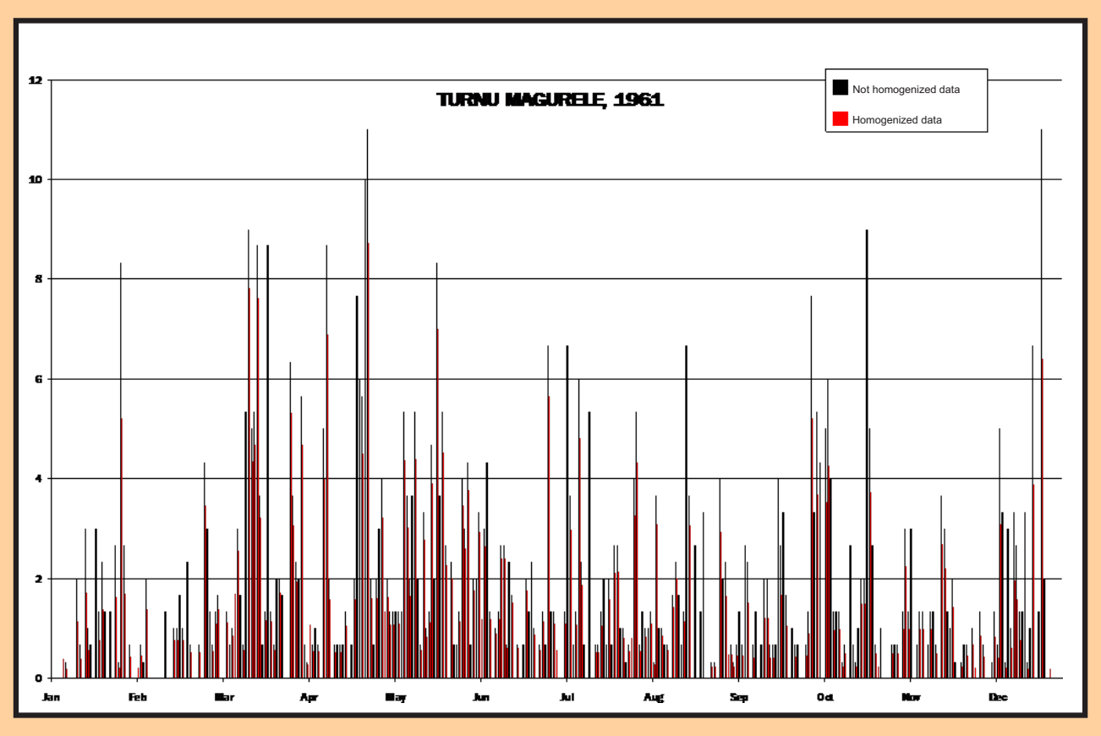
III.3. General metadata

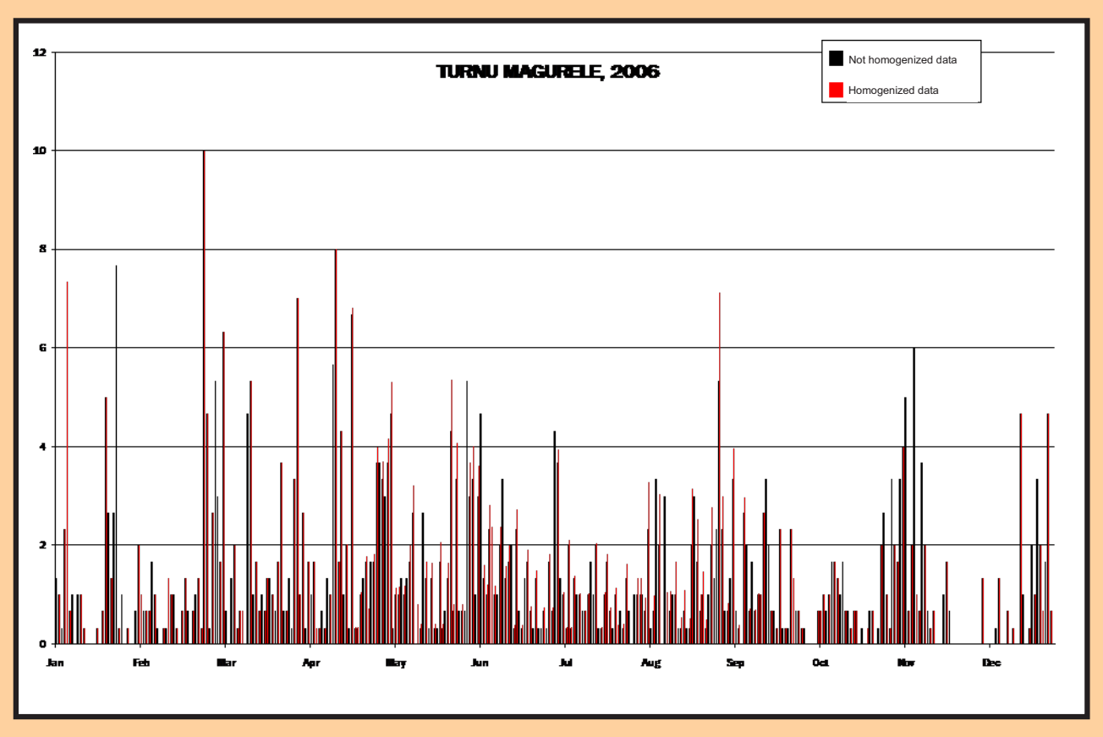
Annales de L'Institut Météorologique de Roumanie (1885 Tome I, 1894 Tome X)

Some information about the meteorological measurements and observation methods used at the end of 19th century and the beginning of the 20-th century are recovered from old documents, scanned and uploaded in the national meteorological metadata-base.

IV. Use of the metadata for quality control of time series

We used MASH v 3.03 in order to accomplish the homogenization of wind daily data.





References

- A. Manea, F. Georgescu, G. Tudorache, (2015), Astronomy and Meteorology - two sciences interconnected in astrophysicist Nicolae Donici's life. Romanian Astronomical Journal, Vol. 25 No.1, ISSN (online) 2285-3758. http://www.astro.ro/~roaj/next.html

- Annales de L'Institut Météorologique de Roumanie pour l'année 1885 – Tome I

- Annales de L'Institut Météorologique de Roumanie pour l'année 1889 – Tome V

- Annales de L'Institut Météorologique de Roumanie pour l'année 1894 – Tome X

-Enrique Aguilar, Inge Auer, Manola Brunet, Thomas C. Peterson and Jon Wieringa „Guide to Metadata and homogeneity" www.bom.gov.au/wmo/climate/ccl/CCl_HM_250603.doc

- Igor Zahumenský Maintenance of Accurate Metadata for all Automatic Weather Station Installations, WMO, CBS/OPAG-IOS/ ET AWS-3/Doc. 3.2(1)

- T. Szentimrey (2008) Development of MASH homogenization procedure for and quality control in climatological databases, Budapest, Hungary, 2006, WCDMPNo. 71: 123-130.

- WMO-No.8 (2010): „Guide to Meteorological Instruments and Methods of Observation"

10th EUMETNET Data Management Workshop - "High quality climate data – the foundation of Climate Services", St. Gallen, Switzerland, 28-30 October 2015

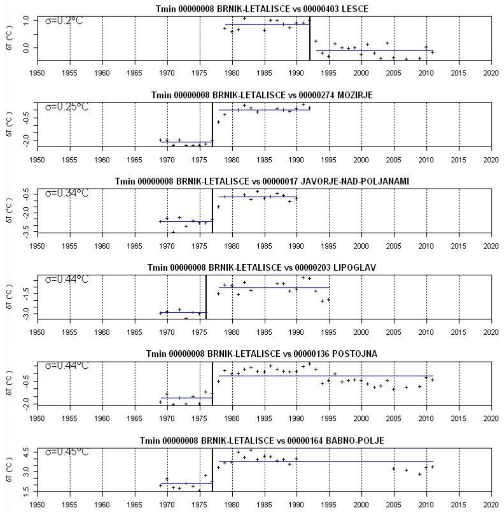
10th EUMETNET Data Management Workshop
St. Gallen, Switzerland, 28–30 October 2015

Metadata in homogenization of Slovenian climatological data

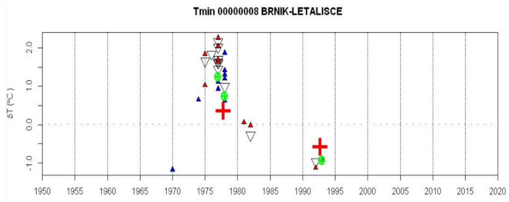
REPUBLIC OF SLOVENIA
MINISTRY OF THE ENVIRONMENT AND SPATIAL PLANNING
SLOVENIAN ENVIRONMENT AGENCY

M. Nadbath, G. Vertačnik E-mail: mateja.nadbath@gov.si

Process of homogenization of Slovenian climatological data was largely supported by metadata.
Breaks in time series of mean air temperature were explained by metadata in 88 %.
We found out that the most common reason of breaks is relocation of meteorological station.
The other are: changes or calibrations of thermometers, new or repainted Stevenson screen and new observer.



An example of R graphics for HOMER's pairwise detection (above) and joint detection (below) for meteorological station Brnik Letališče. Proposed breaks in time series are marked with vertical black lines (above) and red crosses (below).
Explanation for breaks: In year 1978 there was replacement of the meteorological station and in 1993 the automatic station was set up.

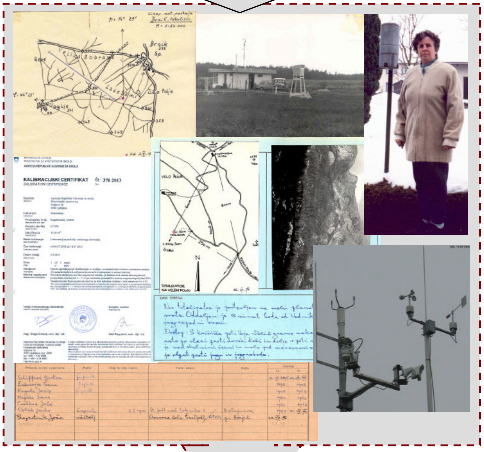


Reasons of breaks in time series of mean air temperature	cases	%
replacement of meteorological station	107	42
new observer	24	10
new / repainted Stevenson screen	33	13
changes / calibration of thermometer	42	17
setting up thermograph / automatic station	16	6
unexplained	30	12

Card-file of meteorological stations in paper form

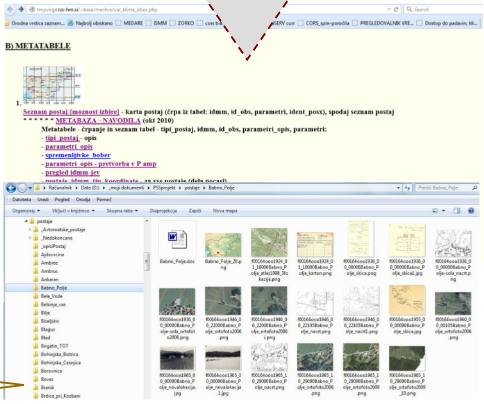
At the beginning of homogenization process the metadata was scattered in different locations and sources; most of it was in paper format only, some was already in digital format.

1. Before digitization the detection of all possible sources of metadata was made.
2. The focus was on meteorological stations used in homogenization.
3. The digitalization of metadata was focused on the period of homogenization process 1961–2011.
4. The data about location or relocation of meteorological station, Stevenson screen, instruments and their calibrations or their changes and observers was collected and digitized.



For 354 Slovenian meteorological stations all known metadata for period 1961–2011 was collected and digitized.
All sketches, photographs and maps of meteorological station were scanned.
All texts : lists of observers, remarks, descriptions, calibration or relocation dates etc. were digitized-typewritten.
The paper archive of metadata was completed with data from digital form and organized in simple digital metabase.

Simple digital metabase of meteorological stations



B) METATABELE

1. Seznam postaj [moznost izbire] - karta postaj (črpa iz tabel: idmm, id_obs, parametri, ident_posx), spodaj seznam postaj
****** METABAZA - NAVODILA (okt 2010)
Metatabele - črpanje in seznam tabel - tipi_postaj, idmm, id_obs, parametri_opis, parametri:
- tipi_postaj - opis
- parametri_opis
- spremenljivke_bober
- parametri_opis - pretvorba v P amp
- pregled idmm-jev
- postaje_idmm_tip_koordinate - za vse postaje (dela pocasi)

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology
MeteoSwiss

A quantitative approach to optimise the quality control system for surface data at MeteoSwiss

V. Knechtl, D. van Geijtenbeek and C. Sigg
MeteoSwiss

Aim

- Improve the performance of the quality control system
    - Reduce the amount of false alarms relative to the amount of corrected values
    - And at the same time, keep the amount of corrected measurement errors at the same level

Introduction: The quality control system

- The analysed quality control system is based on three types of formulas
    - Extrema testing formulas (Fig. 1a)
    - Variability testing formulas (Fig. 1b)
    - Inter-parameter consistency testing formulas (Fig. 1c)

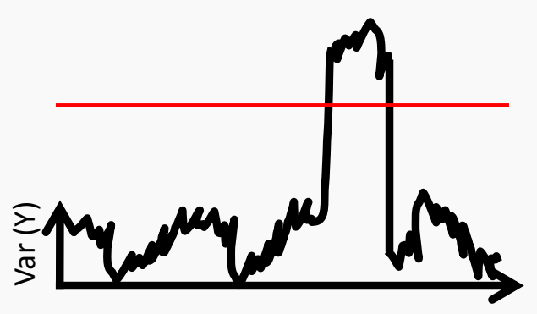


Fig. 1a

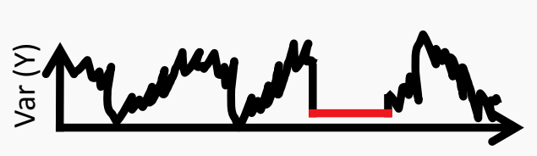


Fig. 1b

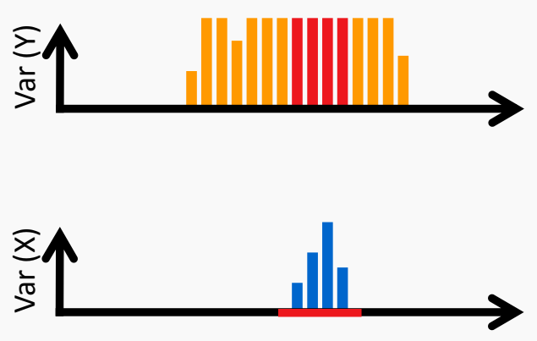


Fig. 1c

- The quality control system at MeteoSwiss consists of around 180 different rules. A rule is composed of a formula and a set of input parameters. For each rule, a station and time dependent set of limits is defined, e.g. $|T_{2m} - T_{2m,\,redundant}| > 5\,°C$.
- The specific limits for each rule depend on the measurement parameter, location and acquisition time of the measurement.
- The quality control system operates for 154 automated stations. 103 out of 154 stations are examined by the experts (Fig. 2).

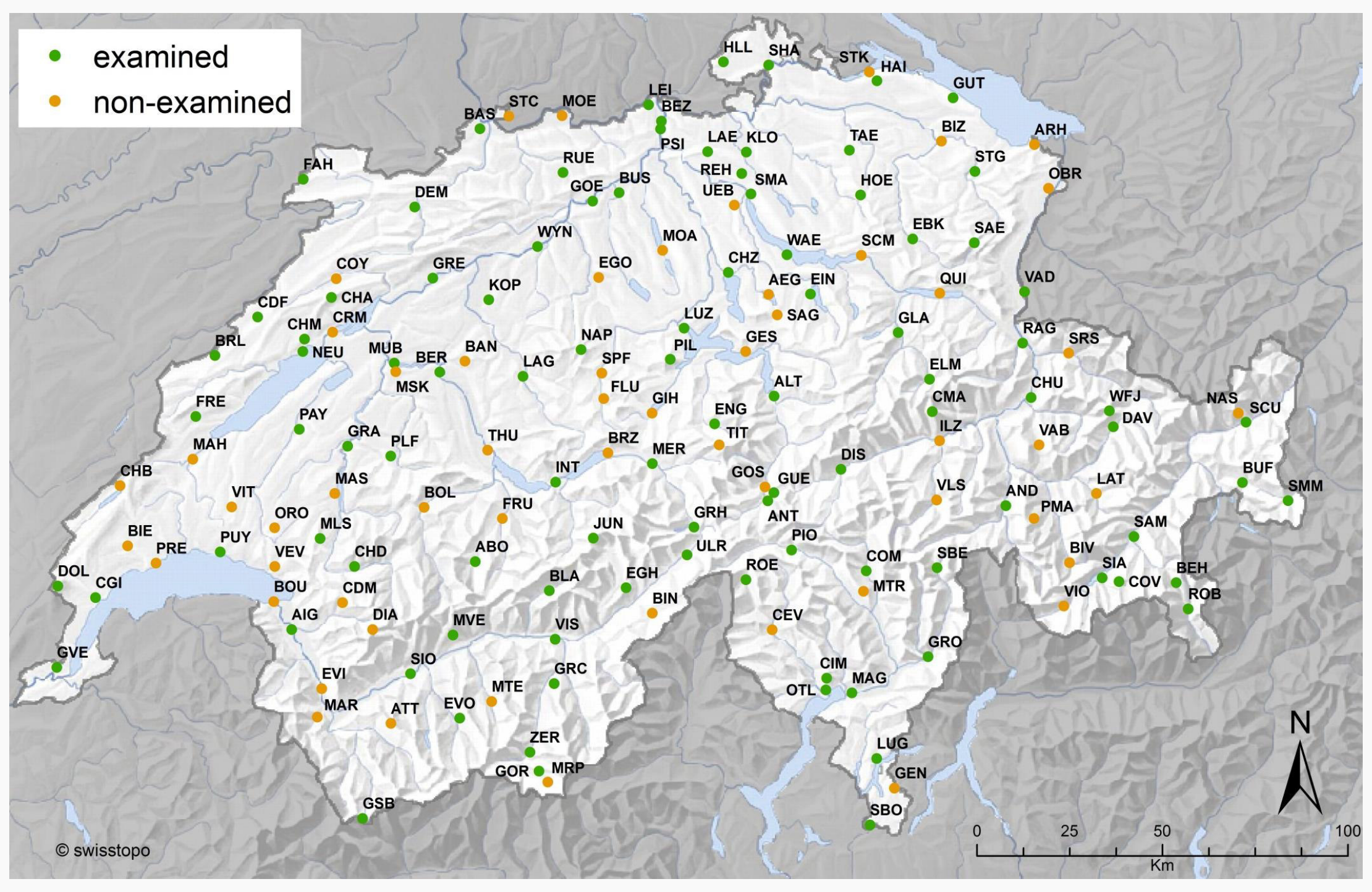


Fig. 2: MeteoSwiss station network (automated stations with 10 min time resolution excl. precipitation only stations).

Method

Definitions

- The rule performance is evaluated with the following measures:

	Value corrected	Value not corrected
Value flagged	True Positive (TP)	False Positive (FP)
Value not flagged	False Negative (FN)	True Negative (TN)

- Corrected: measurement value manually corrected by an expert
- Flagged: measurement value flagged by the quality control system

Approach

- In the first step, the amount of FPs is compared to the amount of TPs for each rule.
- In the second step, rules with a high FP / TP ratio are tested with a range of new limits that reduce the amount of flags compared to the initial limit.
    - A new limit is chosen if the FP / TP ratio can be improved without substantially decreasing the amount of TPs.
    - The rule is removed from the quality control system if the FP / TP ratio cannot be improved.

Data

- The evaluation is done with 18 months of measurement data from 103 stations (see Fig. 2). Furthermore, test output data is used to associate flagged measurements to the 180 rules.

Results

- 31 rules are identified that have a high FP / TP ratio and at the same time, the ratio does not improve with other limits. These rules are removed.
- 6 rules are identified for which the FP / TP ratio can be improved with minimal decrease in TPs.
- These changes result in a decrease of the FP / TP ratio from 6.5 to 4.5 and in a mean decrease of flags per day from 1354 to 949 (Fig. 3).

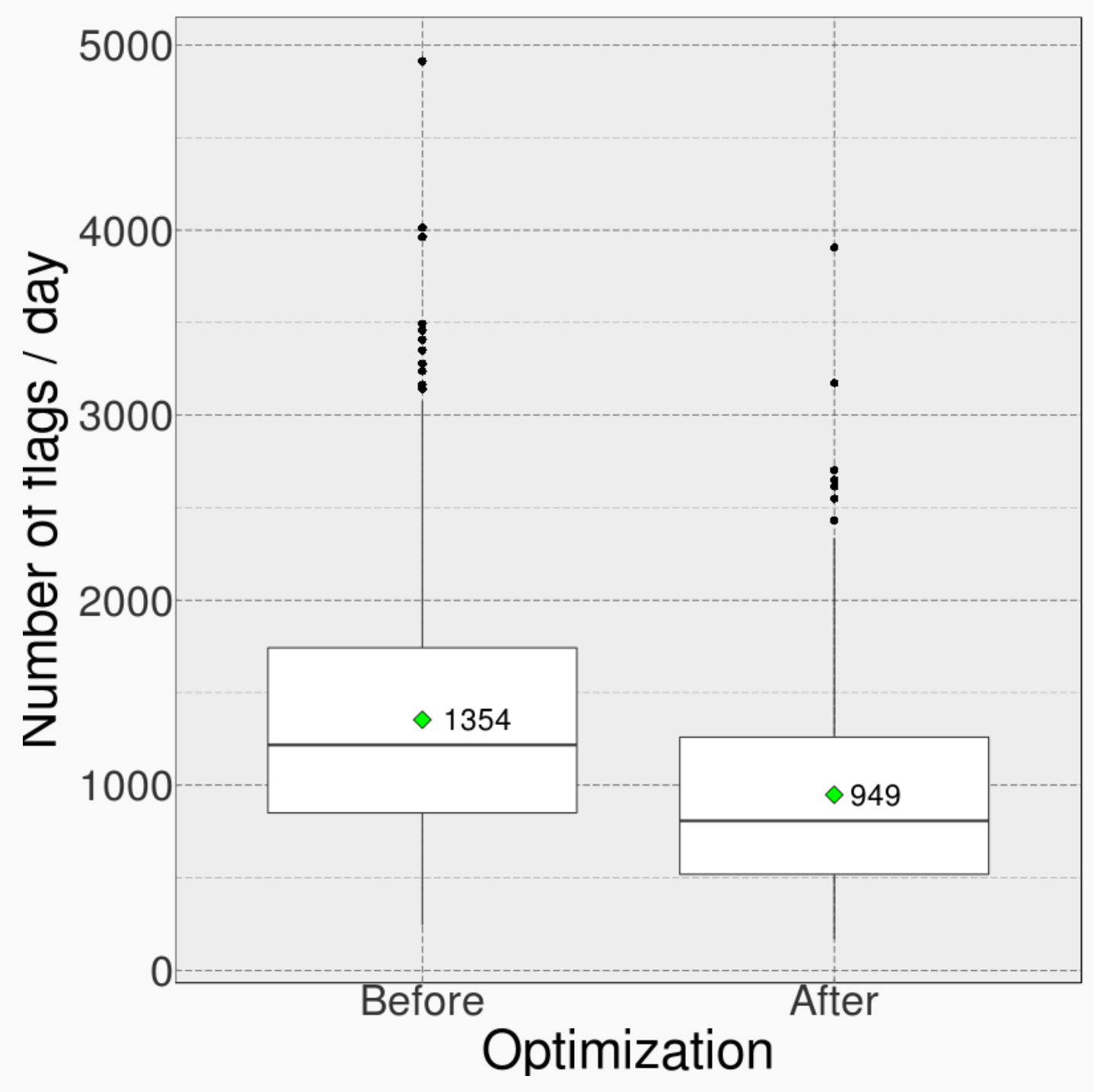


Fig. 3: Boxplots showing the amount of flags per day before and after the changes made on the quality control system. The green dot shows the mean number of flags / day.

Conclusion

- The data quality is improved due to
    - Less flagged data to be inspected. Thus, more time to inspect and correct erroneous measurements which leads to better corrections.
    - Reduced amount of FPs on non-examined data which increases the value of the quality information flags.

Example

- Consistency test between 2m temperature and the redundant 2m temperature:

$$|T_{2m} - T_{2m,\,redundant}| > Limit$$

- Example of a rule violation

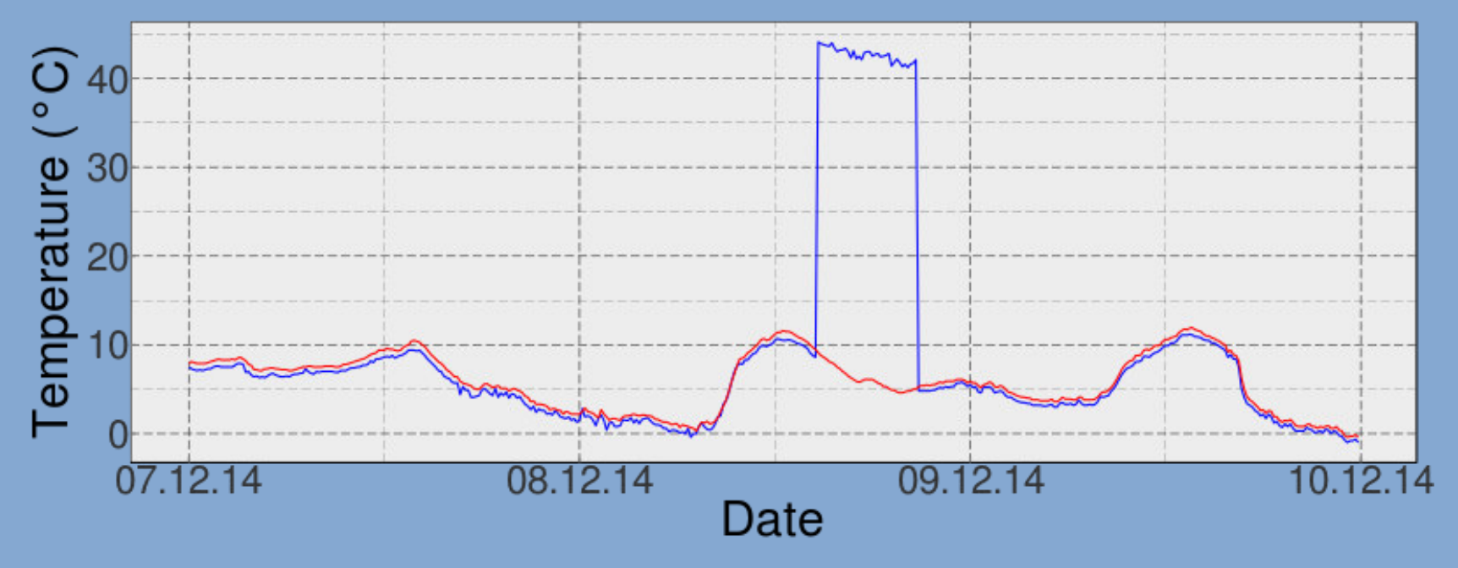


Fig. 4: Main temperature measurement in blue, redundant temperature measurement in red.

- The initial limit is 5 °C. The FP / TP ratio is 2.4. 3269 measurements are flagged within the analysed period. 970 measurements are corrected by the expert, 2299 are FPs (orange dot, Fig. 5).

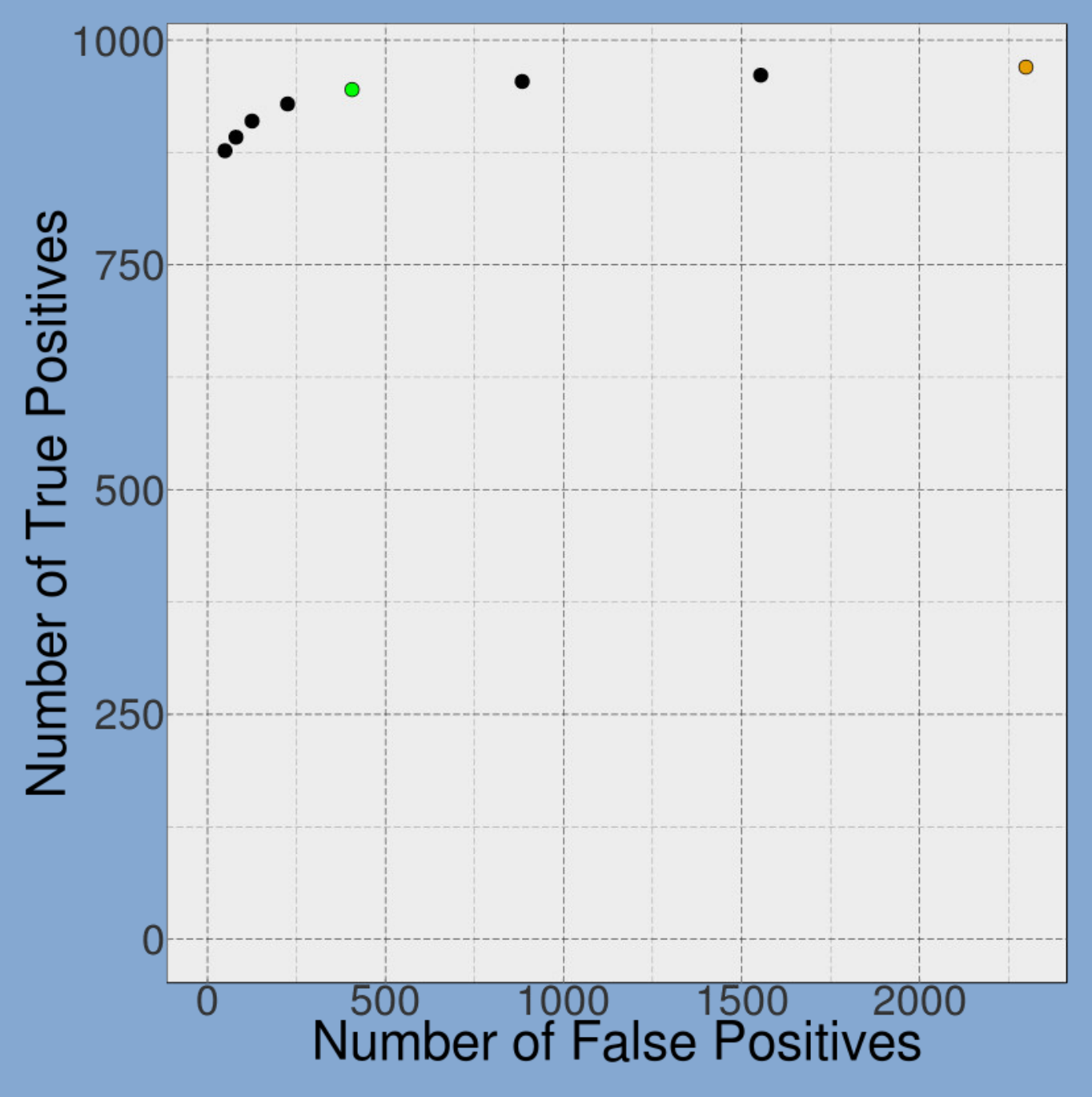


Fig. 5: Number of TPs vs. number of FPs depending on the limit. The orange dot is the initial limit of 5 °C. The green dot is the new limit of 5.9 °C. The limit is increased stepwise by 0.3 °C from 5 °C to 7.1 °C.

- The new limit is 5.9 °C. For this limit, the amount of FPs is reduced from 2299 to 406 without substantially reducing the amount of TPs. The new FP / TP ratio is 0.4.

Contact

Valentin Knechtl: valentin.knechtl@meteoswiss.ch
Deborah van Geijtenbeek: deborah.vangeijtenbeek@meteoswiss.ch
Christian Sigg: christian.sigg@meteoswiss.ch

New quality control interface for rain gauges observations across Belgium

Michel Journée, Charles Delvaux and Cédric Bertrand

Royal Meterorological Institute of Belgium, michel.journee@meteo.be

ABSTRACT

A new interface has been recently developed at the Royal Meteorological Institute of Belgium (RMI) to support the routine daily quality control (QC) of rain gauges observations in Belgium from our centralized database. The QC interface includes various tools allowing the QC staff to highlight suspicious situations that need to be further checked. The QC procedure is thus semi-automated in the sense that the final decision to validate or to correct a value is left to the QC staff.

1. GENERAL SPECIFICATIONS

DATA

- 106 automatic rain gauges, 5-min data (weighing and tipping bucket rain gauges)
- 16 automatic rain detection sensor, 10-min data (precip. duration)
- 220 manual rain gauges with daily data

→ about 33000 data/day to be controlled

→ attribute quality flags to each data (valid, suspicious or erroneous)

→ provide estimation for missing and erroneous data

TOOLS

- centralized interface for all precipitation data
- data visualization on maps
- time series visualization
- comparison against radar estimations
- derivation of a cloud-free index from Meteosat Second Generation (MSG) data
- statistical plausibility tests

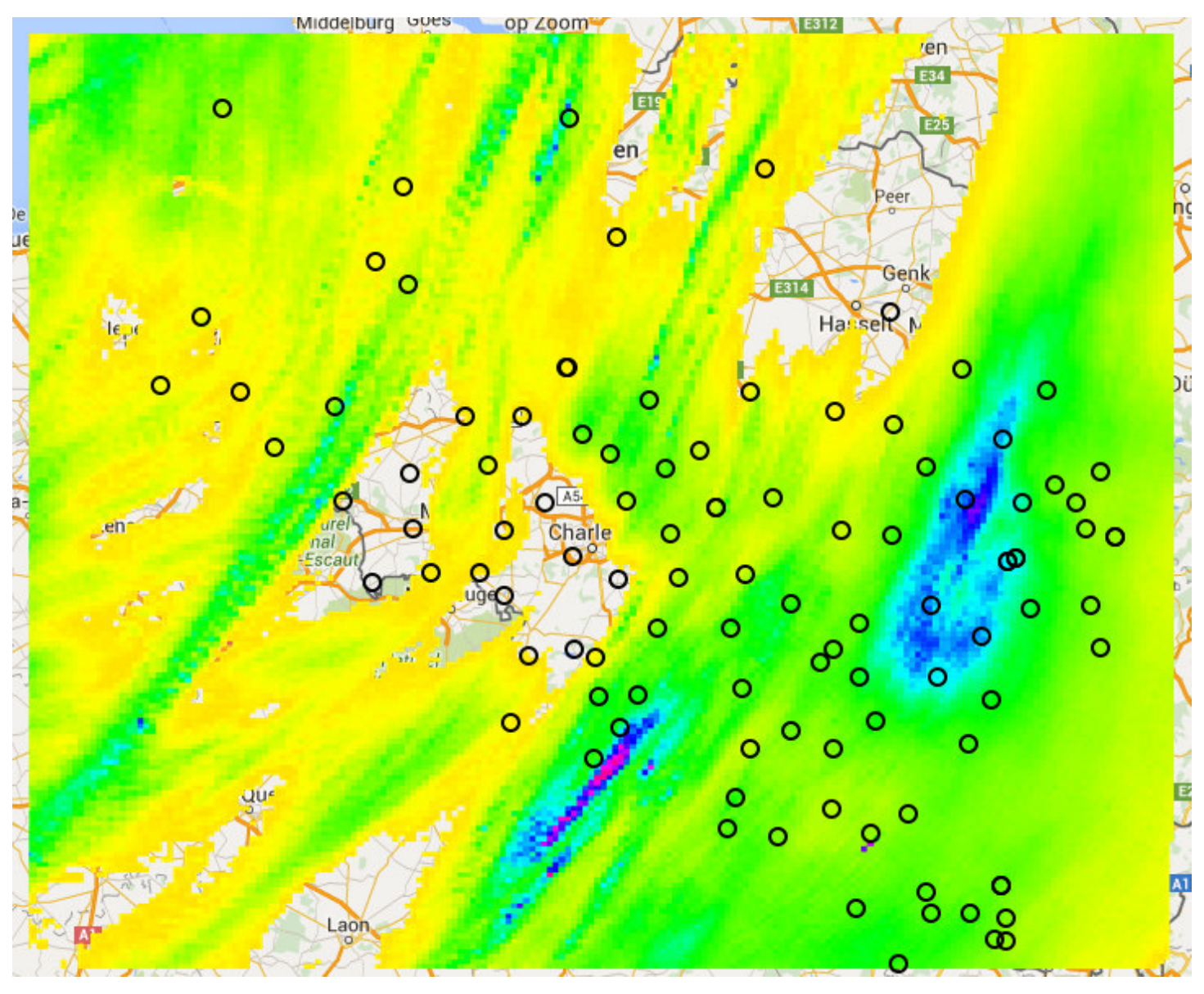
RESSOURCES

- web interface
- Google maps API
- Oracle and SQLite databases
- R+ packages (gstat, rgdal, etc.)

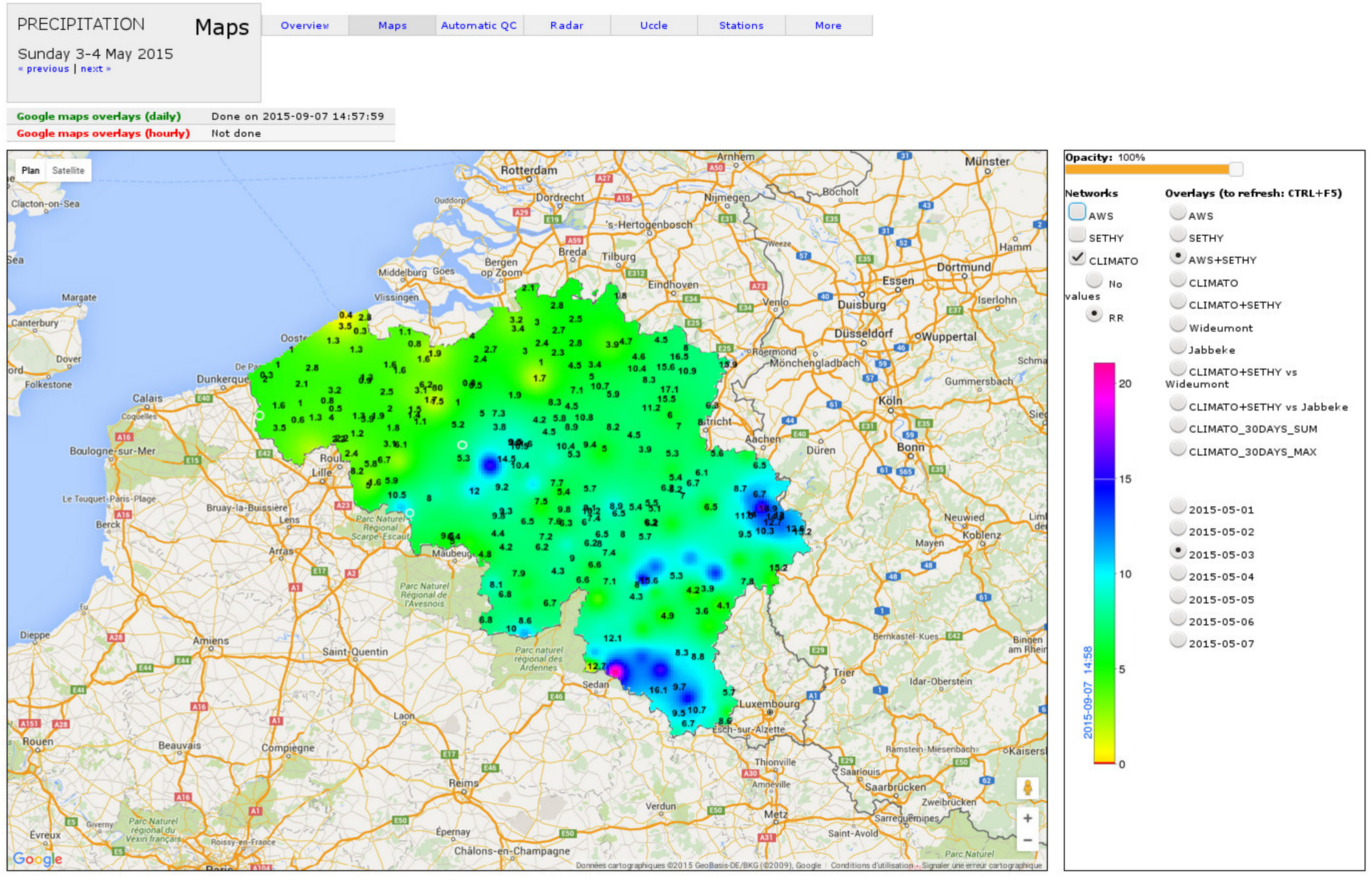
2. SPATIAL ANALYSIS

The QC operator has the choice between various overlays in the Google maps API:

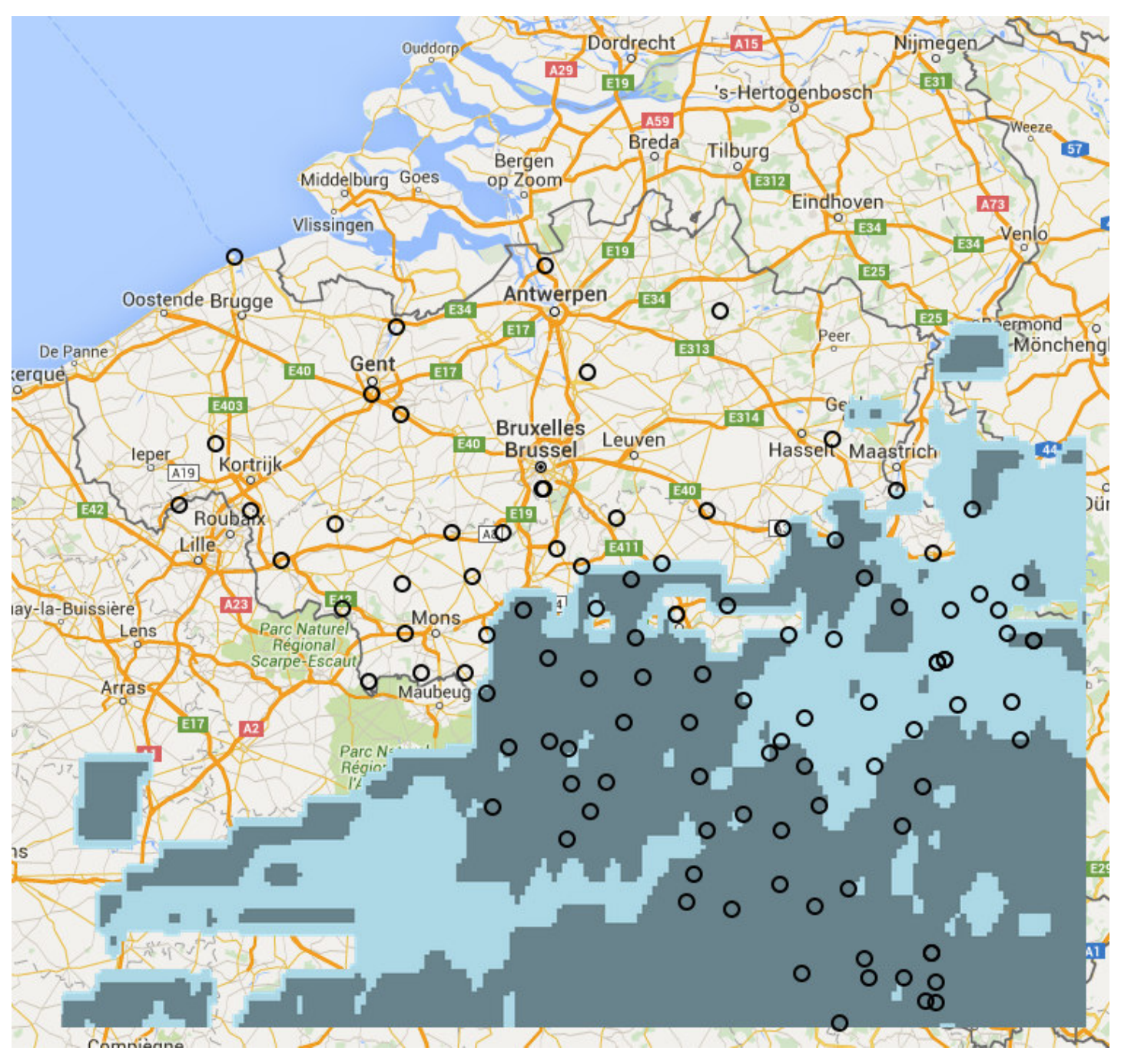
- spatial interpolation of automatic rain gauges (hourly + daily values)
- spatial interpolation of manual rain gauges (daily values)
- spatial interpolation of all rain gauges (daily values)
- meteorological radars estimation (2 radars, hourly + daily values)
- spatial interpolation of the regression residuals between rain gauges values and radar estimates (daily values)
- cloud-free index from Meteosat Second Generation, MSG (hourly values)



Radar estimates



Rain gauges spatial interpolation

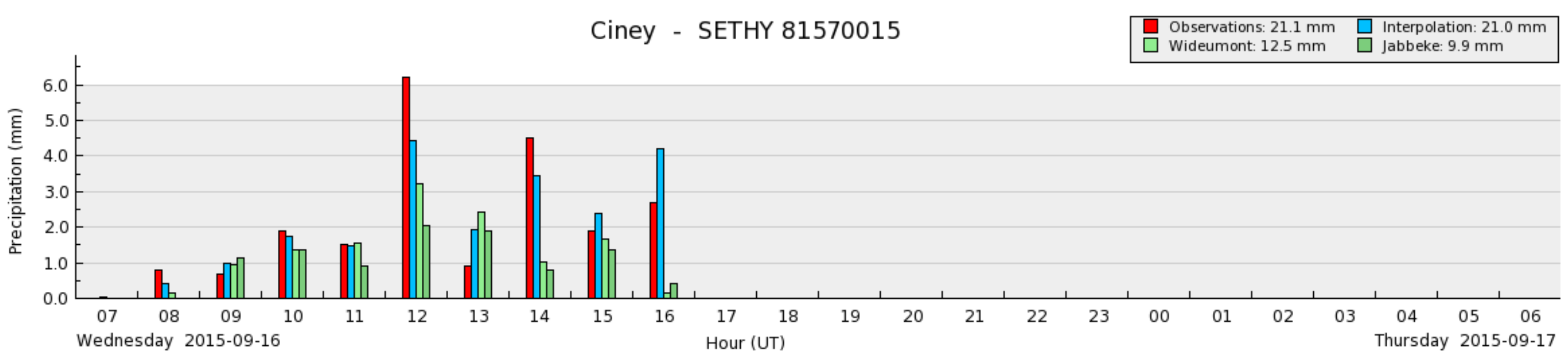


MSG cloud-free index

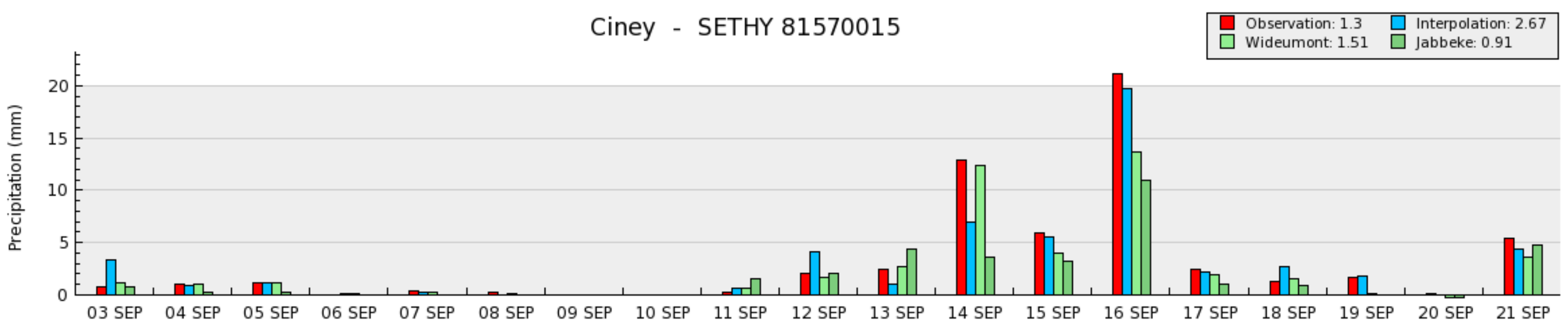
3. TIME SERIES ANALYSIS

At a station, the QC operator can perform various time series comparisons:

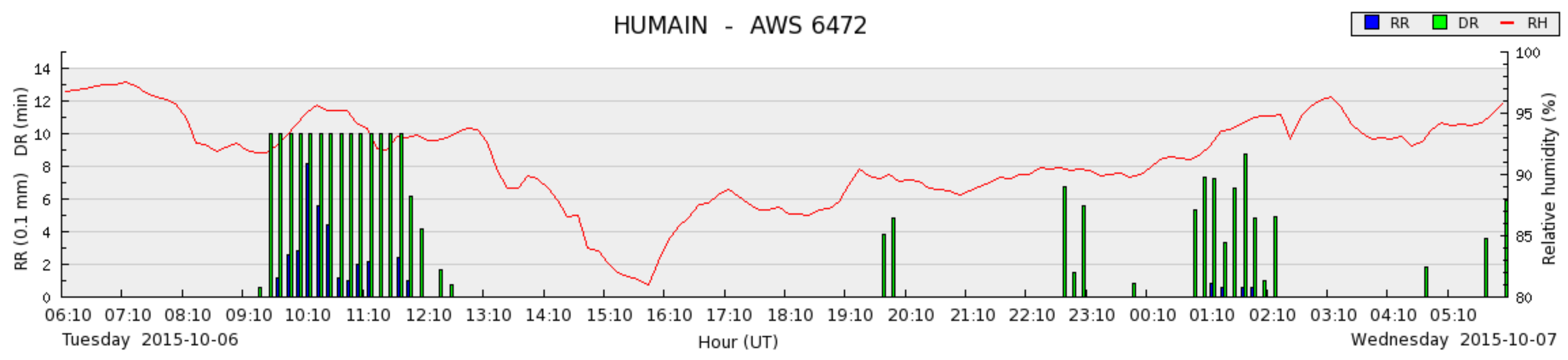
- hourly quantities against radar estimates and closest neighbors interpolation
- daily quantities against radar estimates and closest neighbors interpolation
- 10-min quantities against 10-min durations and 10-min average relative humidity



hourly precipitation quantities vs radar estimates and spatial interpolates



daily precipitation quantities vs radar estimates and spatial interpolates



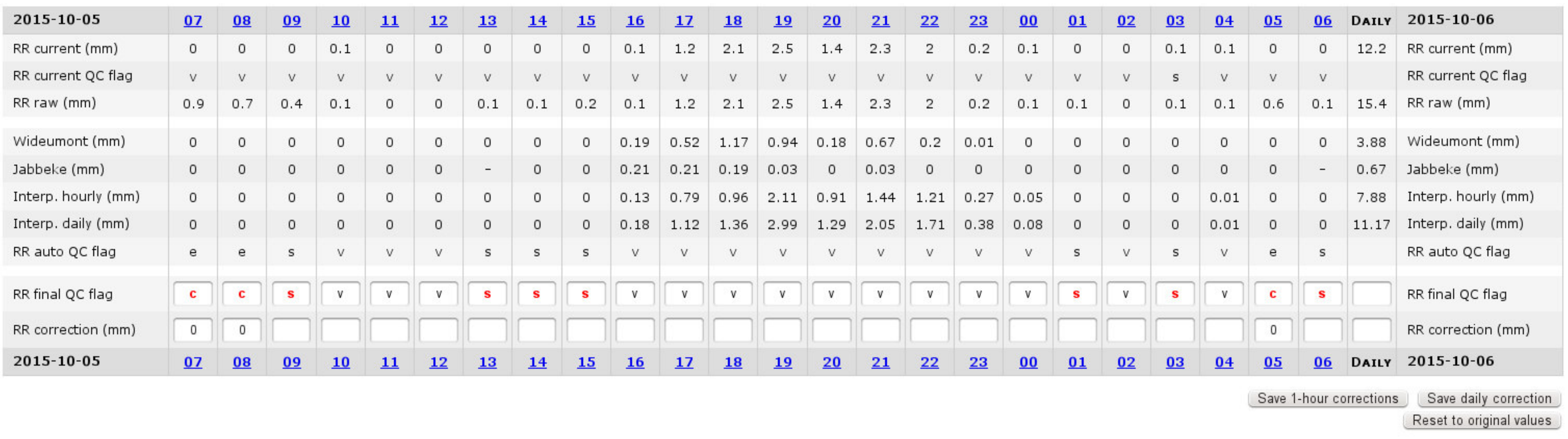
10-min precip. quantities vs precip. duration and relative humidity

4. STATISTICAL QUALITY TESTS

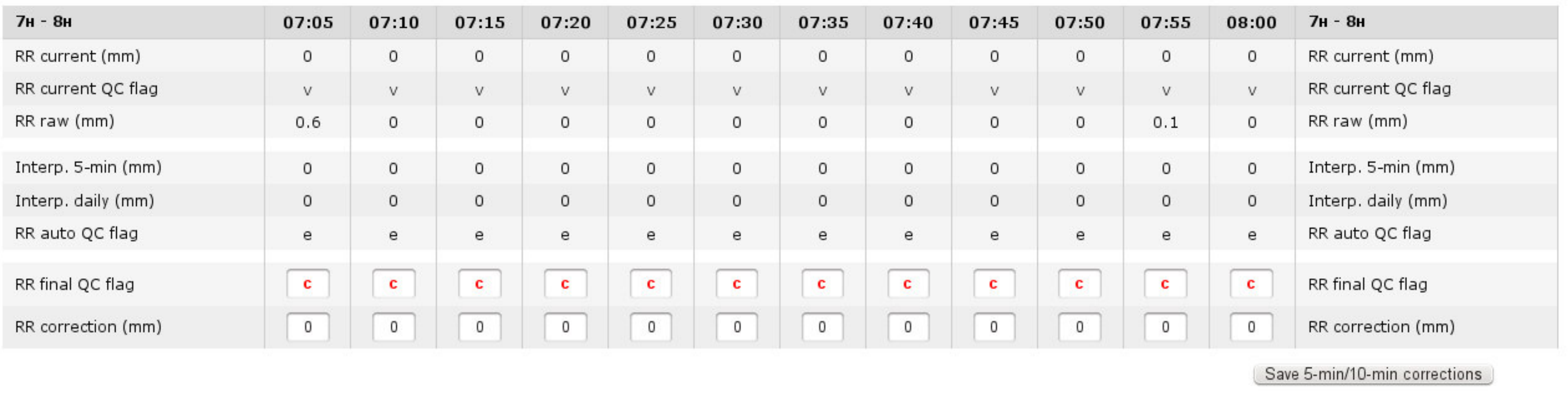
Various types of statistical plausibility tests are performed at all time scales (5min, 1hour and 1day) before analysis by the QC operator:

- Identification of missing data
- physical limits test: non-negative values + upper limit
- spatial consistency test:
  - comparison against a spatial interpolation of neighboring stations' values.
  - special tests for isolated precipitation, isolated dryness and maintenance operations
- comparison against the estimates from the closest meteorological radar
- compatibility tests between quantities and durations.

→ Quality flags (valid, suspicious or erroneous) + spatial interpolation estimations are provided for all data.



2015-10-05	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	DAILY	2015-10-06
RR current (mm)	0	0	0	0.1	0	0	0	0	0	0.1	1.2	2.1	2.5	1.4	2.3	2	0.2	0.1	0	0	0.1	0.1	0	0	12.2	RR current (mm)
RR current QC flag	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	v	s	v	v	v		RR current QC flag
RR raw (mm)	0.9	0.7	0.4	0.1	0	0	0.1	0.1	0.2	0.1	1.2	2.1	2.5	1.4	2.3	2	0.2	0.1	0.1	0	0.1	0.1	0.6	0.1	15.4	RR raw (mm)
Wideumont (mm)	0	0	0	0	0	0	0	0	0	0.19	0.52	1.17	0.94	0.18	0.67	0.2	0.01	0	0	0	0	0	0	0	3.88	Wideumont (mm)
Jabbeke (mm)	0	0	0	0	0	0	-	0	0	0.21	0.21	0.19	0.03	0	0.03	0	0	0	0	0	0	0	0	-	0.67	Jabbeke (mm)
Interp. hourly (mm)	0	0	0	0	0	0	0	0	0	0.13	0.79	0.96	2.11	0.91	1.44	1.21	0.27	0.05	0	0	0	0.01	0	0	7.88	Interp. hourly (mm)
Interp. daily (mm)	0	0	0	0	0	0	0	0	0	0.18	1.12	1.36	2.99	1.29	2.05	1.71	0.38	0.08	0	0	0	0.01	0	0	11.17	Interp. daily (mm)
RR auto QC flag	e	e	s	v	v	v	s	s	s	v	v	v	v	v	v	v	v	v	s	v	s	v	e	s		RR auto QC flag
RR final QC flag	c	c	s	v	v	v	s	s	s	v	v	v	v	v	v	v	v	v	s	v	s	v	c	s		RR final QC flag
RR correction (mm)	0	0																					0			RR correction (mm)
2015-10-05	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	DAILY	2015-10-06



7H - 8H	07:05	07:10	07:15	07:20	07:25	07:30	07:35	07:40	07:45	07:50	07:55	08:00	7H - 8H
RR current (mm)	0	0	0	0	0	0	0	0	0	0	0	0	RR current (mm)
RR current QC flag	v	v	v	v	v	v	v	v	v	v	v	v	RR current QC flag
RR raw (mm)	0.6	0	0	0	0	0	0	0	0	0	0.1	0	RR raw (mm)
Interp. 5-min (mm)	0	0	0	0	0	0	0	0	0	0	0	0	Interp. 5-min (mm)
Interp. daily (mm)	0	0	0	0	0	0	0	0	0	0	0	0	Interp. daily (mm)
RR auto QC flag	e	e	e	e	e	e	e	e	e	e	e	e	RR auto QC flag
RR final QC flag	c	c	c	c	c	c	c	c	c	c	c	c	RR final QC flag
RR correction (mm)	0	0	0	0	0	0	0	0	0	0	0	0	RR correction (mm)

UNIVERSITAT DE BARCELONA

Servei Meteorològic de Catalunya

QUALITY CONTROL OF RAINFALL DATA AT A DAILY SCALE FOR 1855-2014 IN CATALONIA

A. Llabrés-Brustenga*, M. Herrero, M. Perpinyà, A. Rius

Meteorological Service of Catalonia, Berlin, 38-46, 08029 Barcelona
*Correspondence Author: allabres@meteo.cat

Catalonia presents a wide diversity of rainfall patterns:

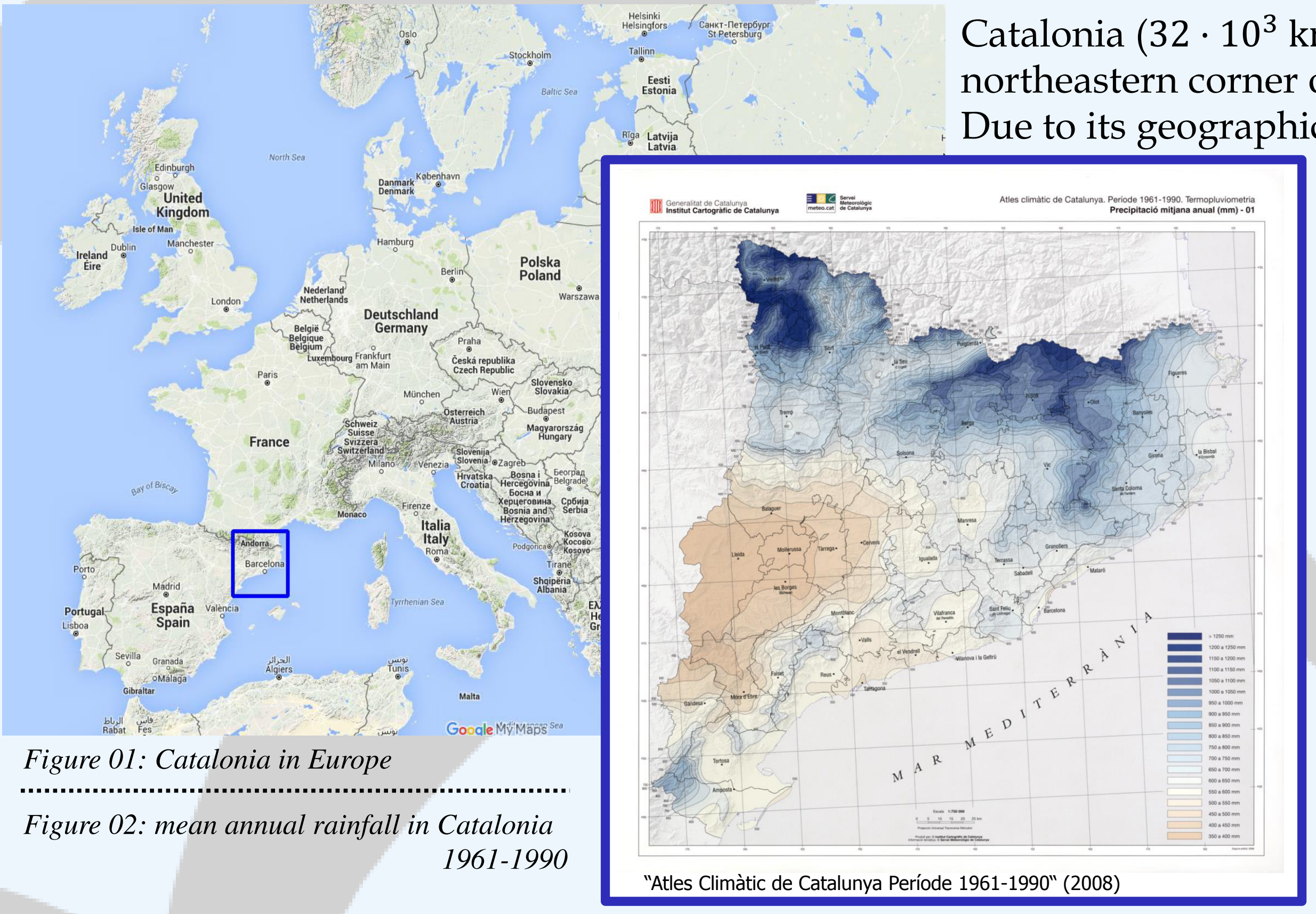


Figure 01: Catalonia in Europe

Figure 02: mean annual rainfall in Catalonia 1961-1990

"Atles Climàtic de Catalunya Període 1961-1990" (2008)

Catalonia (32 · 10^3 km^2) is located in the northeastern corner of the Iberian Peninsula. Due to its geographical position between the Pyrenees and the Mediterranean Sea, its complex orography and its latitudinal situation (which involves influences by both cold polar and warm tropical air masses), the region presents a wide variety of climates. These circumstances add difficulty to the quality control of rainfall data.

The project in its context:

The quality control of daily rainfall data is the first stage of a wider project which is aimed at obtaining maximum expected rainfall and IDF curves at high resolution for Catalonia. This wider study will require a reliable database of daily rainfall observations. Consequently, a strict quality control at a daily scale is essential.

Rainfall data, availability and sources:

The Meteorological Service of Catalonia is working towards collecting and storing all weather data that have been generated in Catalonia during its instrumental history (since the late 18th century). These weather data have been managed by many different organizations, especially during the 20th century. Consequently, the initial data set is a mixture of different measuring methodologies (including manual and automatic measuring methods), quality control procedures and metadata coding. The set of currently available daily rainfall data spreads over the period from 1855 to the present. Measuring has been carried out at 1725 different weather stations. In the present project, the same quality control methodology has been applied to the entire set in order to unify criteria, focusing on the temporal completeness of the series.

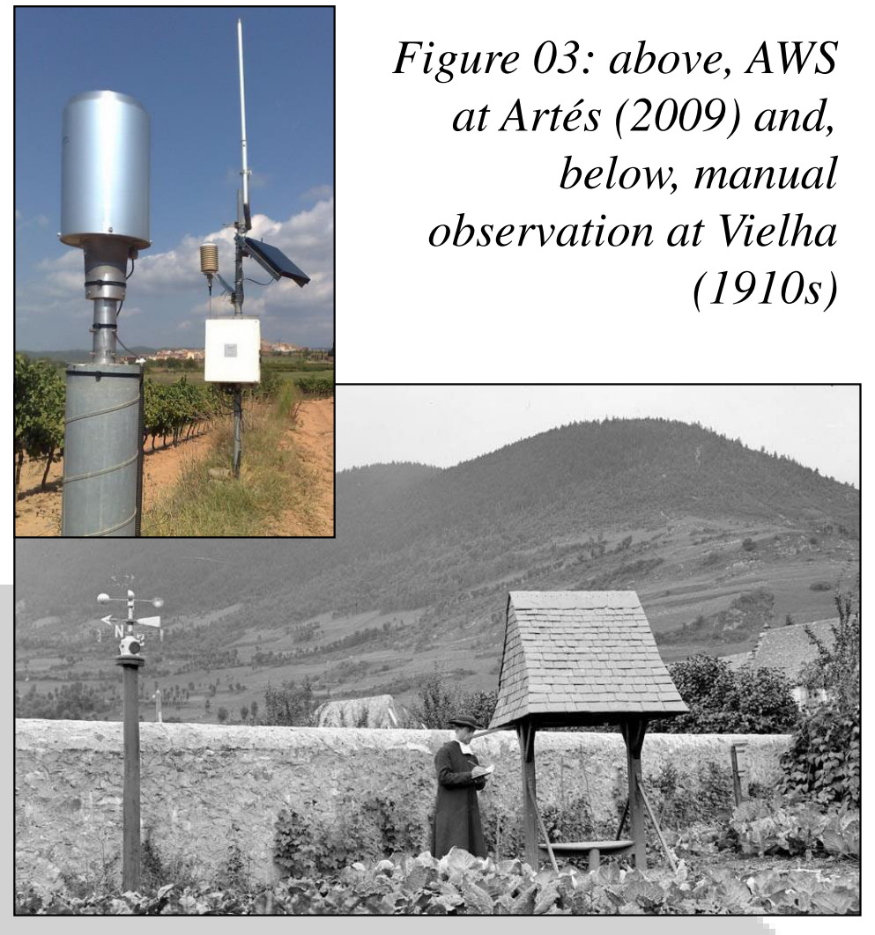
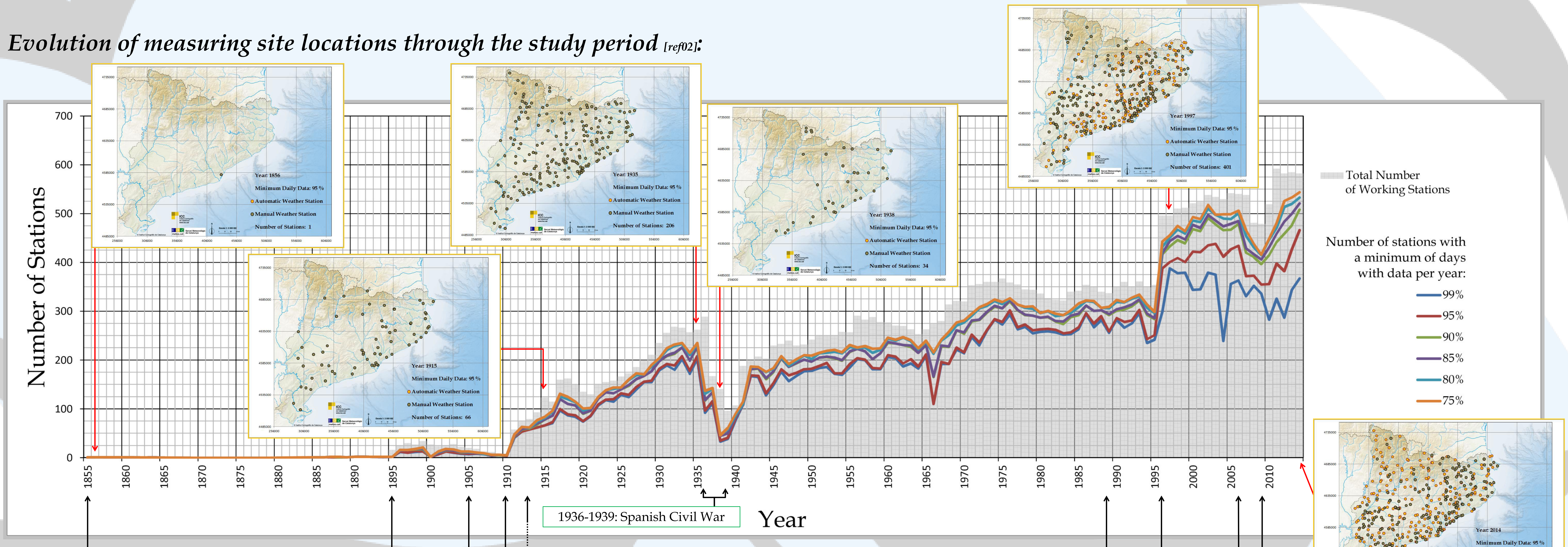


Figure 03: above, AWS at Artés (2009) and, below, manual observation at Vielha (1910s)

Evolution of measuring site locations through the study period [ref02]:



1936-1939: Spanish Civil War

1855: beginning of observations at «Universitat de Barcelona», a measuring site coordinated by «Observatorio de Madrid»

Beginning of the Meteorological Network of Catalonia and Balearic Islands

The Astronomical Society of Barcelona, in collaboration with the Central Meteorological Observatory, works to improve the Catalan Rainfall Network

1989: First Automatic Weather Station (AWS), currently integrated in the network XEMA, managed by the Meteorological Service of Catalonia

Establishment of the Spanish Automatic Hydrologic Data Collection System (SAIH)

1905: beginning of observations at «Observatori de l'Ebre», a series with more than 110 years of daily data and 99% of completeness

1913: beginning of observations at «Observatori Fabra», a series with more than 101 years of daily data and 99% of completeness

Progressively, an important number of SAIH stations gets dismantled

Beginning of the Weather Observers Network (XOM) managed by the Meteorological Service of Catalonia

Basic Control [ref03]:

Negative values:

Few cases were found and adjusted as they proved to be digitalization errors.

Outstanding values:

Values over 300 mm were found and checked; some of them were considered correct, others were clearly errors.

Consecutive equal values:

After checking values (over zero) identically repeated for several days, most of them were found to be correct only at a monthly scale.

Outlier Check:

Identification both with the threshold $Q_3 + 3 \cdot IQR$ and $Q_3 + 5 \cdot IQR$ *[ref05]* (data over 1 mm) per year and also depending on the month of the year.

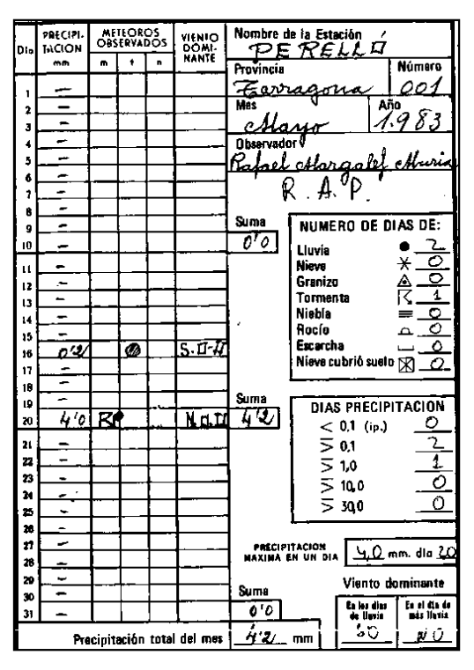
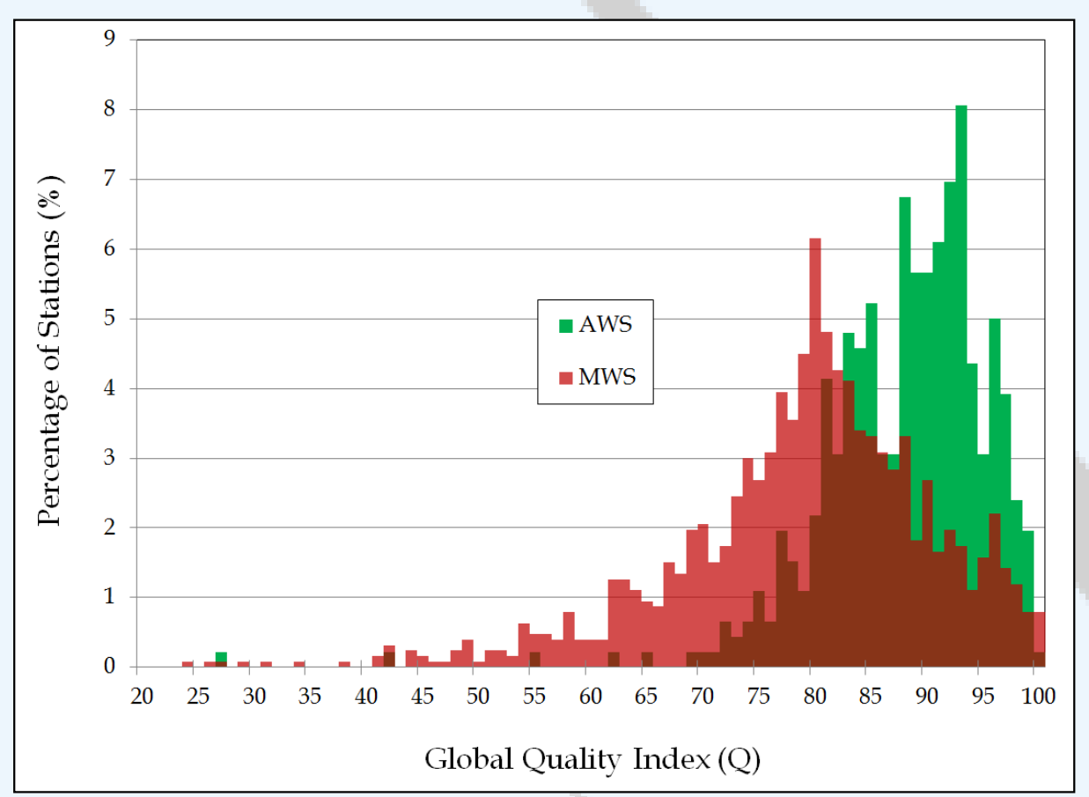


Figure 04:example of original information, handwritten by the observer. It can be used to confirm digitalization errors.

Absolute QC:

Quality Label:

In order to classify each individual series, indices have been developed which account for the main problems usually found in rainfall series. These indices can be averaged into a global index (for the whole series or any period of interest over one year). The proposed indices may weigh differently in the global index according to the known problems of the data set or the purpose of the later study.



The Global Quality Index (Q) takes values between 0 and 100, where 100 indicates the best quality. It can be used to compare the quality of the series in the set. *Figure 05* shows that AWS have a better index more often while Manual stations spread over a wider spectrum as there are much more of these series and they are longer, hence, more problems may appear.

Figure 05: plot of the frequency of appearance (%) of each index in the set of stations (divided into Manual and Automatic Weather Stations) in this case obtained by:

$$Q = \frac{1}{5}\{P_{MinAnnual\%} + P_{\%} + Q_{Zero(month)} + Q_{Gaps(month)} + Q_{Gaps}\}$$

- **Completeness of the series:**

  The completeness of the series is a basic problem *[ref01, 03]* that can be assessed by the percentage of daily data ($P_{\%}$) in the considered period. Regarding the calculation of a global index, the percentage of daily data of the least complete year ($P_{MinAnnual\%}$) has been used.

- **Zero values which are, in fact, lack of observation:**

  a) Systematic untagged accumulations through the week (usually accumulated weekend rainfall, a problem restricted to Manual stations) can be estimated by an analysis of the number of rainfall days depending on the day of the week *[ref05]*.

  b) A high number of months with zero total rainfall while having complete daily data is an indicator of possible missing data tagged as zero values. It may appear both in Manual and Automatic stations even if the source of the problem is different.

$$Q_{Zero(month)} = 100 \cdot \left(1 - \frac{Number\ of\ months\ with\ complete\ zero\ daily\ data}{Total\ number\ of\ months\ with\ complete\ daily\ data}\right)$$

- **Distribution of days with no data:**

  Gaps or periods with no data must be identified. These can then be used to weigh the quality of the series using the following equation:

$$Q_{Gaps} = 100 \cdot \left(1 - \frac{2 \cdot Number\ of\ Gaps + Number\ of\ days\ in\ the\ longest\ Gap}{Total\ number\ of\ days}\right)$$

Q_{Gaps} returns 100 when there are no gaps and gives a value near 50 when one half of the series is missing. For series with the same number of days with data, series with more frequent gaps will return lower values than series with fewer, longer gaps.

In addition, the systematic absence of observation in specific months of the year can be assessed by the Coefficient of Variation of the daily data percentage of the series by month ($Q_{Gaps(month)}$).

Relative QC:

Relative Quality Control of the set presented here is currently being executed. Detection of doubtful daily data is achieved by selecting supporting series *[ref04]* for each candidate series according to their climatic region, spatial distance, altitude difference, correlation of daily rainfall data and absolute quality of the supporting series.

The main goal of the Relative QC is to label data at a daily scale after distinguishing between correct values, erroneous data and days that cannot be classified for lack of appropriate supporting data.

The first phase of the relative controls is at a monthly scale, where clear cases can already be classified. Doubtful data will undergo a stricter, daily-based control which will compare daily rainfall with supporting data on the same day as well as on adjacent days (to account for temporal shifts in the series).

References:

[ref01]: Abaurrea, J.; Asín, J.; Cebrián, A.C.; Centelles, A. (2004). "Metodología para el control de calidad y homogeneidad de una base de datos de precipitación diaria". Publ Spanish Climatol Soc, 2004 431-440
[ref02]: Anduaga Egaña, A. (2012). " Meteorología, ideología y sociedad en la España contemporánea". CSIC. Estudios sobre la ciencia, 61
[ref03]: Einfalt, T.;Gerlach, N.;Podlasly, C.; Demuth, N. (2008). "Rainfall and Climate Data Quality Control". 11th International Conference on Urban Drainage, Edinburgh, Scotland, UK, 2008
[ref04]: Romero, R.; Guijarro, J. A.; Ramis, C.; Alonso, S.; (1998). "A 30-year (1964-1993) daily rainfall data base for the Spanish Mediterranean regions: first exploratory study". International Journal of Climatology,18: 541-560
[ref05]: Viney, N.R.; Bates, B.C. (2004). "It Never Rains on Sunday: The Prevalence and Implications of Untagged Multi-day Rainfall Accumulations in the Australian High Quality Data Set". International Journal of Climatology, 24: 1171-1192
[ref06]: RClimDex Software: http://etccdi.pacificclimate.org/software.shtml

Automatic Range Checks in Quality Control of Observation-network Data:

Towards Operational Implementation

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Curtis R Wood*, Ari Aaltonen, Anna Frey

*curtis.wood@fmi.fi

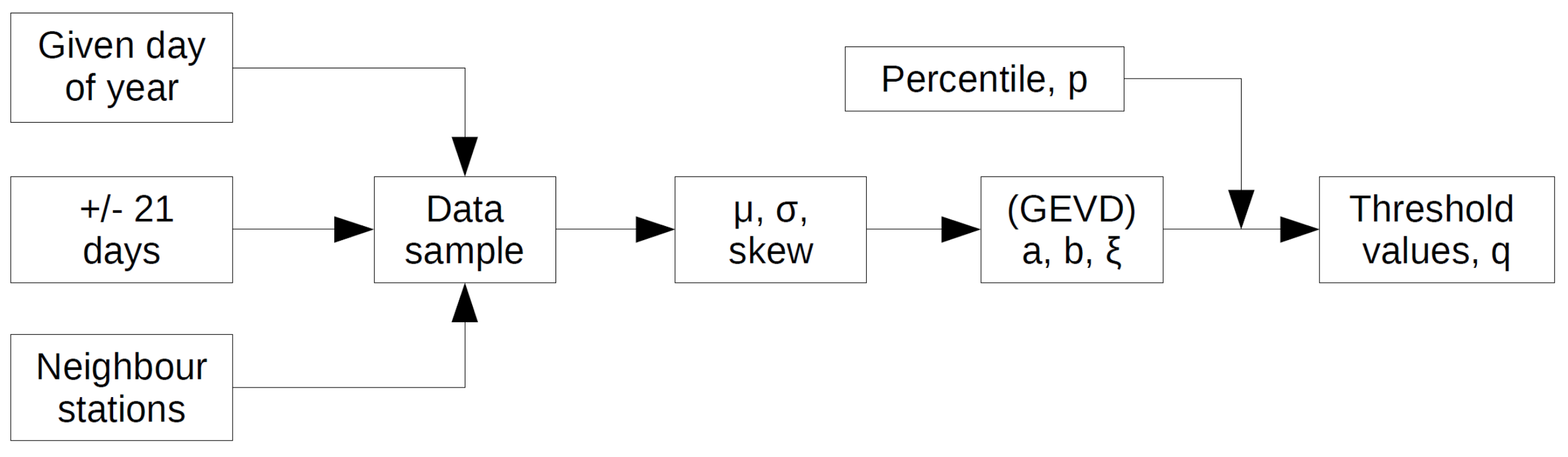
Introduction

Quality-control (QC) systems for data in station networks are needed (Vejen et al. 2002). One part of QC is defining extreme acceptable limit values: **range checks**. Data are flagged that exceed those limit values; flagged data are inspected by a human operator (human quality control: HQC). A former operational method at FMI was that these threshold limits for min and max were manually updated from climate data. For large networks, and in the pursuit of consistency, a reliable and accurate *automated* system is required to reduce HQC and allow limit values to be updated automatically under both climate change and a station's environmental change.

We aim to put this automatic method into operational use.

Methods

This method is very similar to that in Hasu and Aaltonen (2011) where the limit values were defined for all values – e.g. every hour of every day in the year. But for operational use (see later) we use limit values for only daily maximum and minimum. The process is summarized here:



Given day of year

+/- 21 days

Neighbour stations

Data sample

μ, σ, skew

Percentile, p

(GEVD) a, b, ξ

Threshold values, q

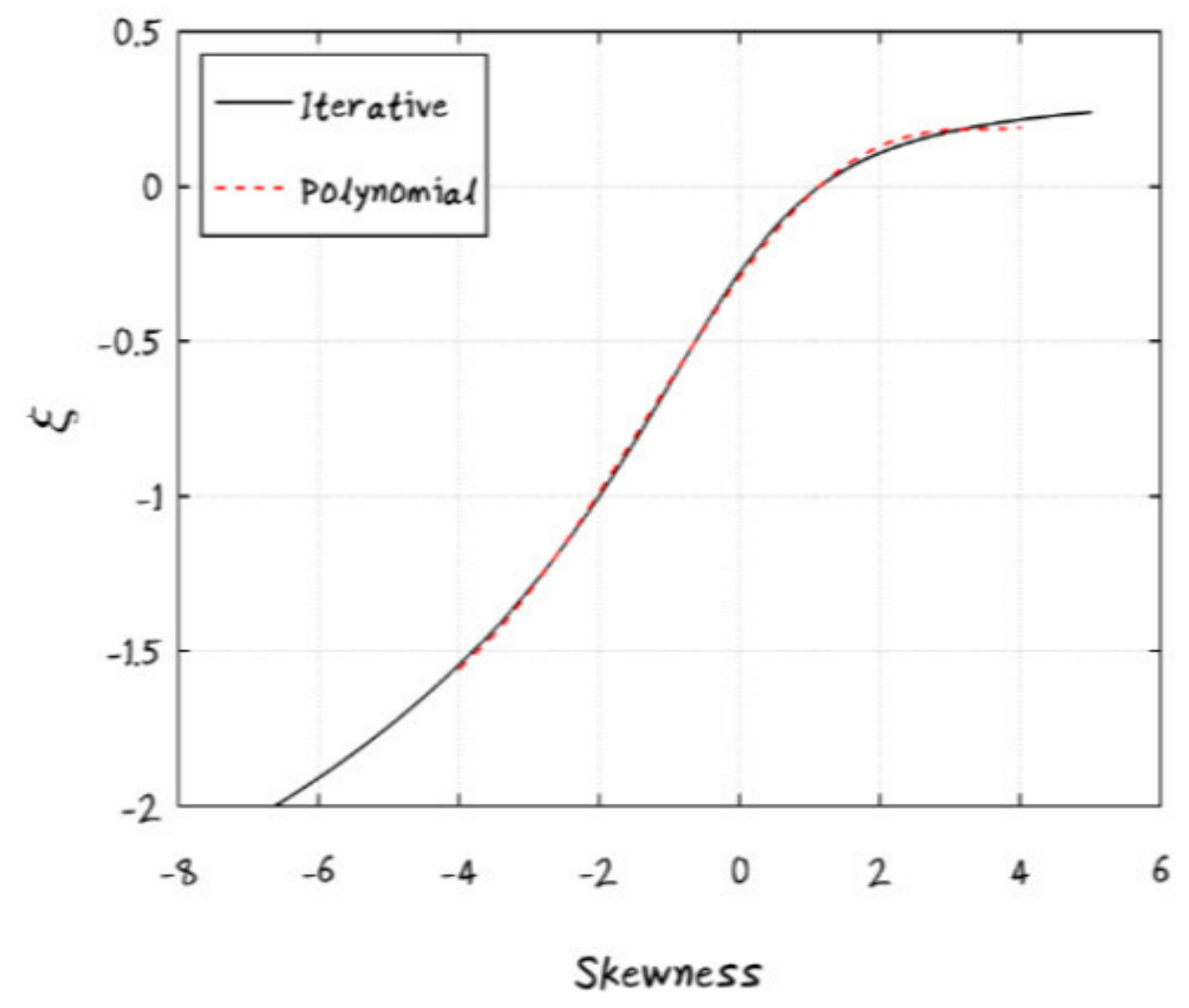
For the day in question – and for all data on this day in this year and previous years – we create two frequency distributions: one of maximum and one of minimum values. We also include ± 21 days, to give us a better sample size; and this can be weighted as Gaussian or evenly. (We could also possibly include neighbour stations to increase sample size.) Anyway, now that we have a sample, we can define statistics such as mean (μ), standard deviation (σ) and skewness (sk = m/σ^3). Using those 1st, 2nd and 3rd statistical moments, we can subsequently fit/define a distribution to this histogram. From the fit/distribution, we can define percentiles *q* (and thence, if required, return periods).

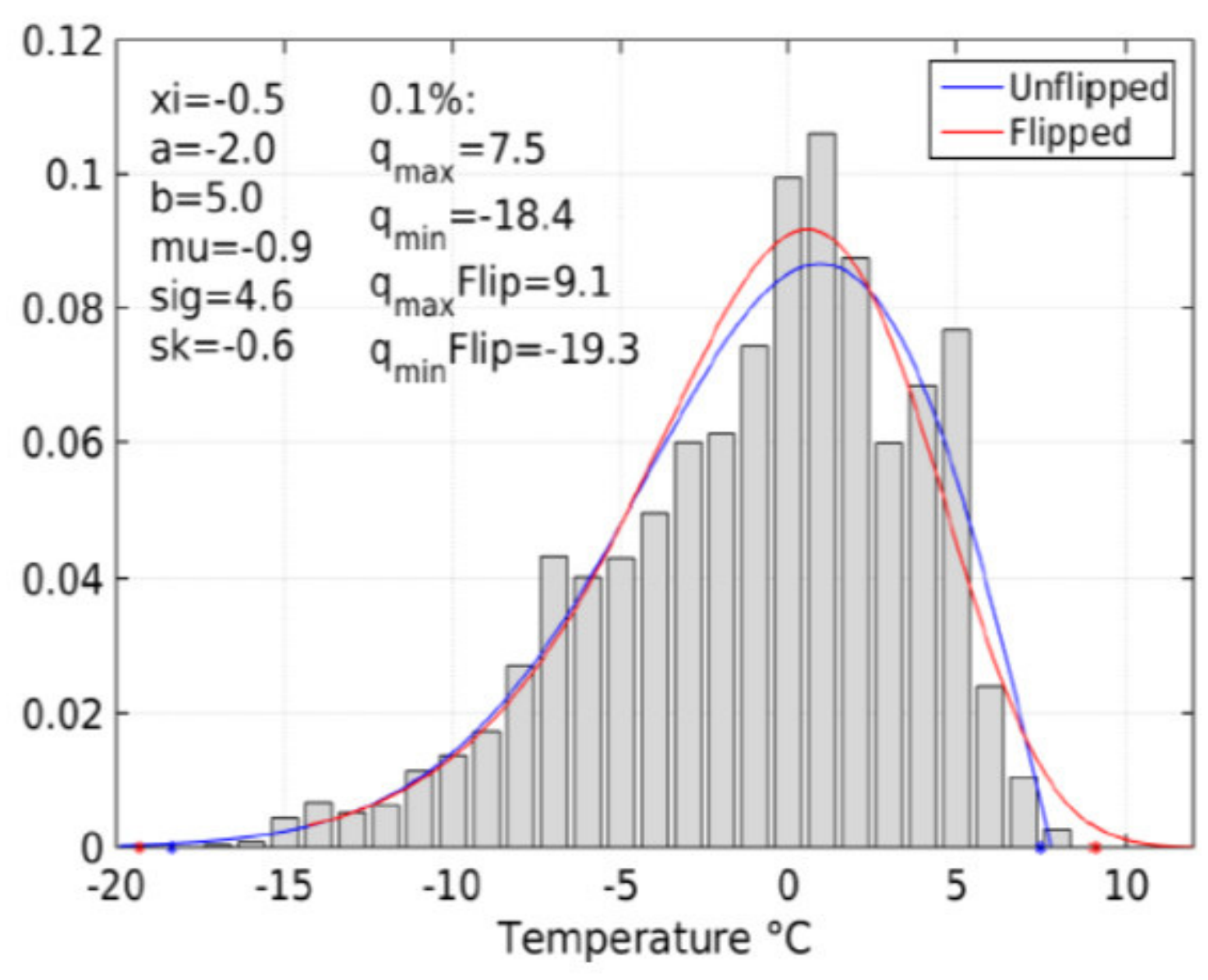
Since we are interested in the tails of the distribution, we use generalized extreme value distribution (GEVD, Coles 2001) which we directly apply to temperature, pressure, wind speed and relative humidity. GEVD comprises types I (Gumbel ξ = 0), II (Fréchet ξ > 0) and III (reverse Weibull ξ < 0). The percentiles for the temperature thresholds are defined as follows (p = 0.001):

$$q_{max} = a + \frac{b}{\xi}\left((-\ln(1-p))^{-\xi} - 1\right) \qquad q_{min} = a + \frac{b}{\xi}\left((-\ln(p))^{-\xi} - 1\right)$$

$$a = \mu + \frac{b}{\xi}(1 - \Gamma(1-\xi)) \qquad b = \sqrt{\frac{\sigma^2\xi^2}{\Gamma(1-2\xi) - \Gamma(1-\xi)^2}}$$

$$\frac{m}{\sigma^3} = sgn(\xi)\frac{\Gamma(1-3\xi) - 3\Gamma(1-\xi)\Gamma(1-2\xi) + 2\Gamma(1-\xi)^3}{(\Gamma(1-2\xi) - \Gamma(1-\xi)^2)^{3/2}}$$



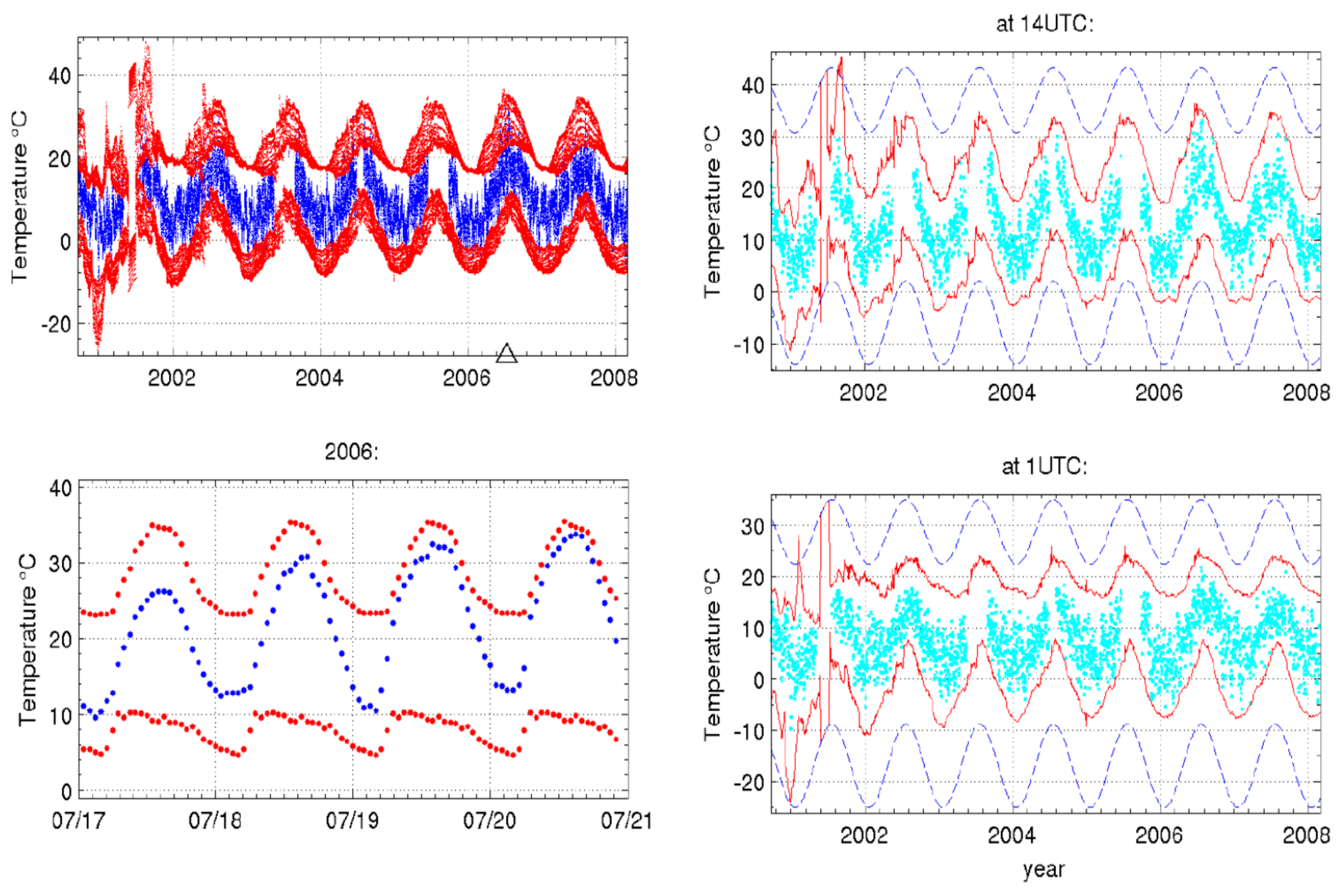


One major caveat becomes apparent from GEVD's being used mostly for the tails on the *skewed* side of the distribution. For our case, we also want information on tail opposite the skewed side. There is no problem for positively skewed distributions (because ξ ≈ 0, hence a continuous GEVD tail on both sides). But for a negatively skewed distribution ξ << 0 (i.e. reverse Weibull), and so the maximum of the fitted distribution is bounded on the upper aide. Thus for those cases we require an unbounded tail (i.e. Fréchet or Gumbell) – an easy way to do this is to mirror/flip the histogram/distribution of the observations so that the sign of skewness changes and the GEVD upper tail is thence unbounded. The example figure (right) is for an arbitrary sample from a station in southern Finland in winter.

Offline testing in MATLAB

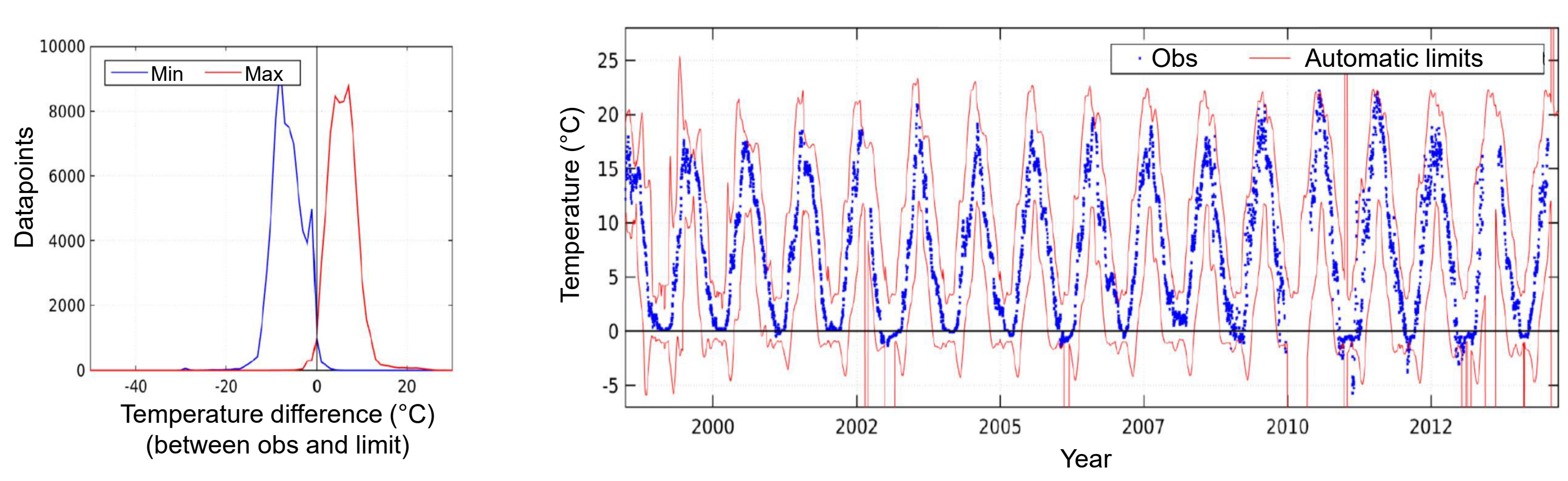
Air temperature at a UK weather station:

(blue dots are observations, red are automated limits)



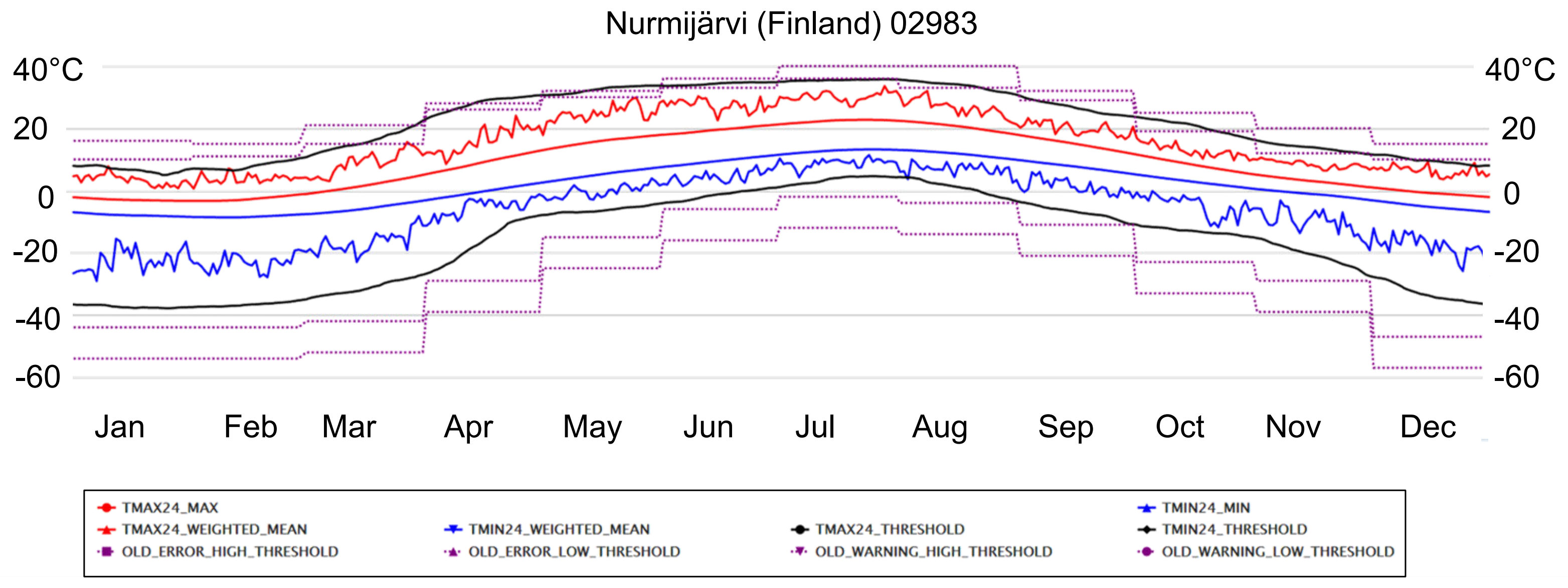
Sea-surface temperature near Helsinki:

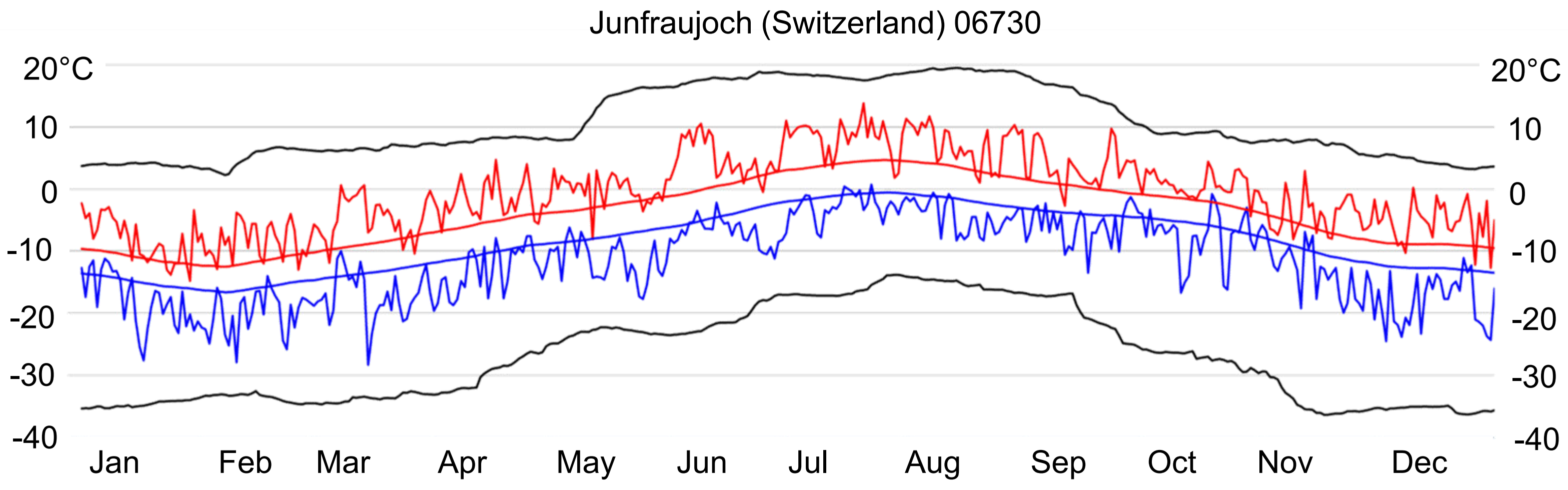
Min limit exceeded in 0.48% of data, and max 0.7% :

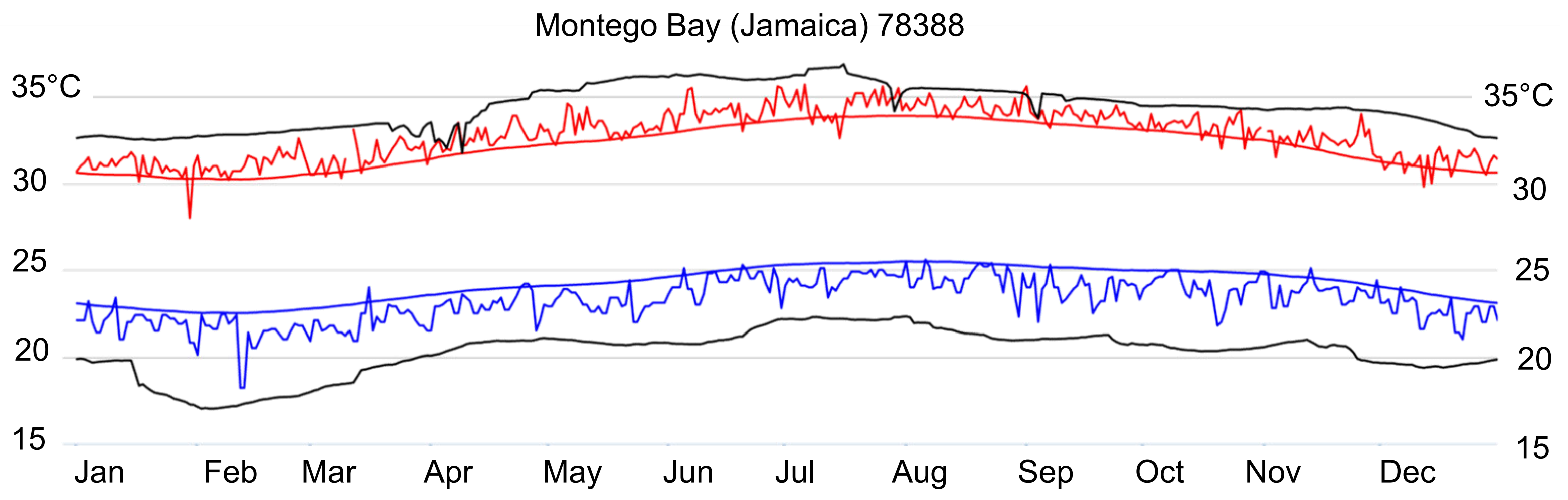


Towards operational use at FMI

These air-temperature examples are directly taken from an FMI intranet site:







Summary

✓Already successfully tested in different countries and for different variables: air temperature, sea-surface temperature, pressure, humidity and wind speed.

✓This will be soon under operational use in FMI.

ACKNOWLEDGEMENTS

Thanks for help from Ismo Karjalainen and Katri Leinonen.

REFERENCES

Coles SG (2001) **An Introduction to Statistical Modeling of Extreme Values.** Springer-Verlag, 209 pages, ISBN 978-1-85233-459-8

Hasu & Aaltonen (2011) **Automatic Minimum and Maximum Alarm Thresholds for Quality Control.** J. Atmos. Oceanic Technol. 28: 74 –84

Vejen F, et al. (2002) **Quality control of meteorological observations—Automatic methods used in the Nordic Countries.** NORDKILM, 109 pages, www.smhi.se/hfa_coord/nordklim/task1/quality_control.pdf

26.10.2015

THE INFLUENCE OF SUBJECTIVE CORRECTIONS ON CLIMATE DATA

Ana Weissenberger, Dubravka Rasol and Helena Lebo Andreis

Meteorological and Hydrological Service, Croatia

MOTIVATION

Until 2008 corrections to the data in electronic form were applied without keeping an untouched electronic copy of the original. Today, it is unknown how much data was corrected in the past, and how those corrections affect the time series used for climatological analyses. Our aim was to find out how much data were corrected, what are the amplitudes of corrections applied and what is the influence of those corrections on the data.

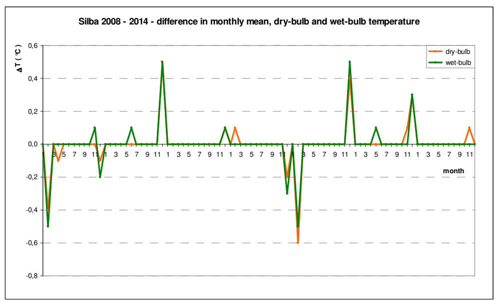
DATA

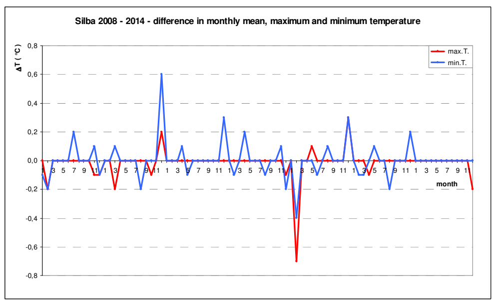
- climatological stations (manual observation)
- 2008 - 2014 dataset
- air temperature (from the three main daily observations), daily maximum and minimum temperature and wet-bulb temperature

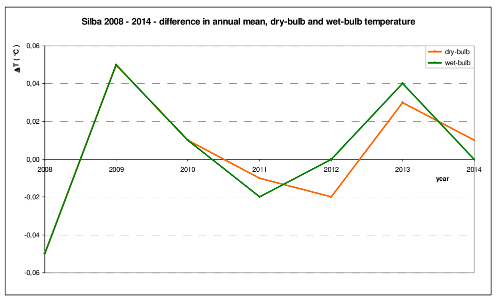
DATASETS WITH RELATIVELY FEW CORRECTIONS

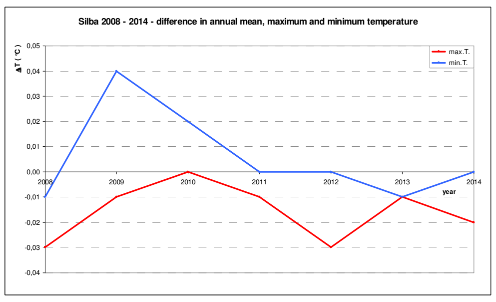
Representative station: Silba

Differences in monthly and annual means calculated from original and corrected data are shown for all four parameters.









Amplitudes of corrections and the percentage of corrected data in the sets are shown.

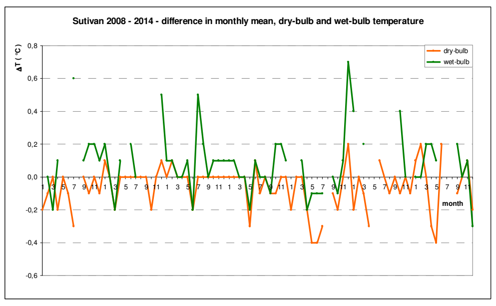
In high quality datasets, corrections of more than 5℃ are rare and those higher than 10℃ were not found.

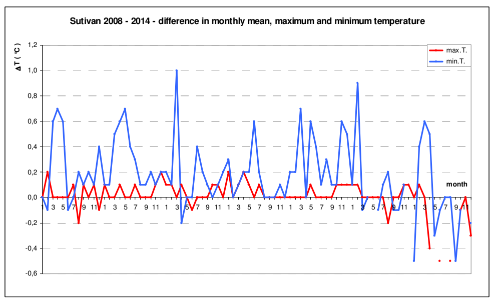
Silba 2008-2014 – total number of data				
	Dry-bulb temp.	Wet-bulb temp.	Max. temp.	Min. temp.
Original	7597	7597	2537	2537
Corrected	7671	7671	2557	2557
Deleted	0	0	0	0
Interpolated	74	74	20	20
Corrections (℃)				
0.1 - 2.0	5	15	7	32
2.1 - 5.0	5	9	11	26
5.1 - 10.0	3	0	1	0
> 10.1	0	0	0	0
Corrections - total	13	24	19	58
Percentage of interpolated and corrected data in the set	1,1 %	1,3 %	1,5 %	3,1 %
Percentage - total	1,5 %			

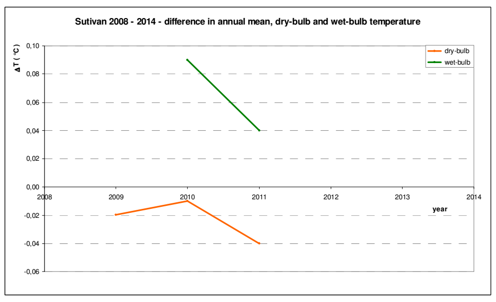
DATASETS WHERE DATA ARE OFTEN CORRECTED OR REMOVED

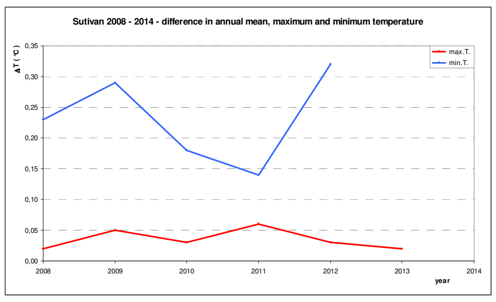
Representative station: Sutivan

Differences in the means for data from this group are not significantly greater, but obviously occur constantly, and breaks are visible where data were deleted.









In this group, considerably larger amount of data was either corrected or deleted.

The highest correction found was 23,0℃.

Running this analysis for all climatological stations in Croatia (109) is planned for future work.

Sutivan 2008-2014 – total number of data				
	Dry-bulb temp.	Wet-bulb temp.	Max. temp.	Min. temp.
Original	7671	7671	2557	2557
Corrected	7485	6357	2465	2465
Deleted	186	1314	92	92
Interpolated	0	0	0	0
Corrections (℃)				
0.1 - 2.0	49	241	73	115
2.1 - 5.0	183	277	38	198
5.1 - 10.0	3	6	4	6
> 10.1	0	0	0	0
Corrections - total	235	524	115	319
Percentage of interpolated and corrected data in the set	3,1 %	8,2 %	4,7 %	12,9 %
Percentage - total	6,4 %			

Removing the urban temperature bias from a village met station record – an empirical approach

Manuel Dienst, Jenny Lindén, Jan Esper – Johannes Gutenberg University Mainz | m.dienst@geo.uni-mainz.de

Introduction

Homogenous long term instrumental temperature data is a key source for studying climate trends[1]. However, many meteorological stations were historically installed in or near urban settlements, thus potentially influenced by an urban warming bias[2,3]. While the urban bias of cities is rather well understood, the effect in small towns and villages is less studied. The objective of this study was to assess and correct for the urban bias in previous station locations in the village of Haparanda, Sweden.

Research site and methods

- Study area: Swedish town Haparanda *(X: 368864, Y: 7305267, UTM 32)* with temperature record starting in 1860
- Analysis of metadata to determine all past locations of meteorological station (see map for station location history)
- Mounting of temperature sensors in all these places to assess different microclimates
- Correction of time series using correction factors derived from differences of each urban location to a sensor outside town at the current AWS

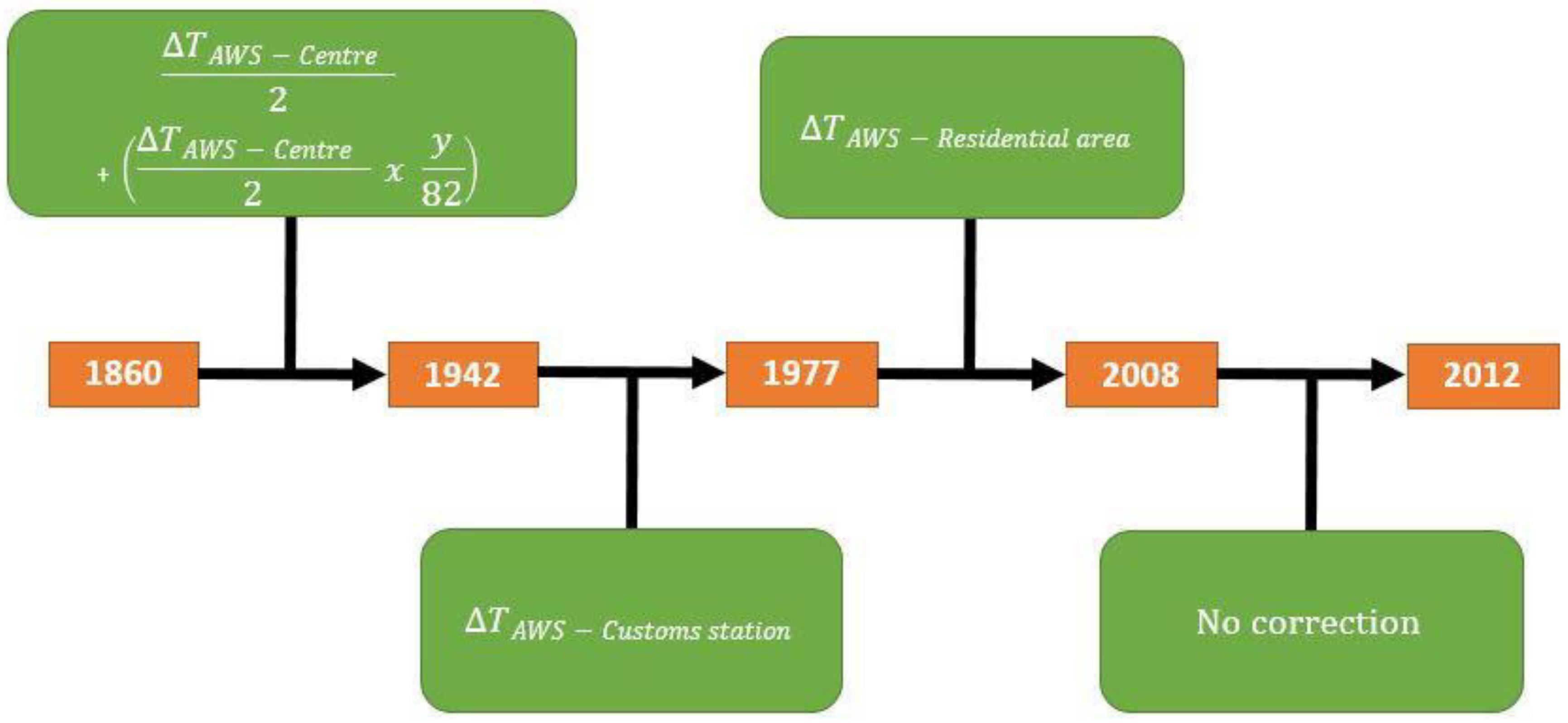


Figure 2: Correction factors used for eliminating the bias for each station relocation. Regarding the first period, half the urban influence was expected from the beginning, while the other half was added gradually over time to account for the growth of the town in the following years.

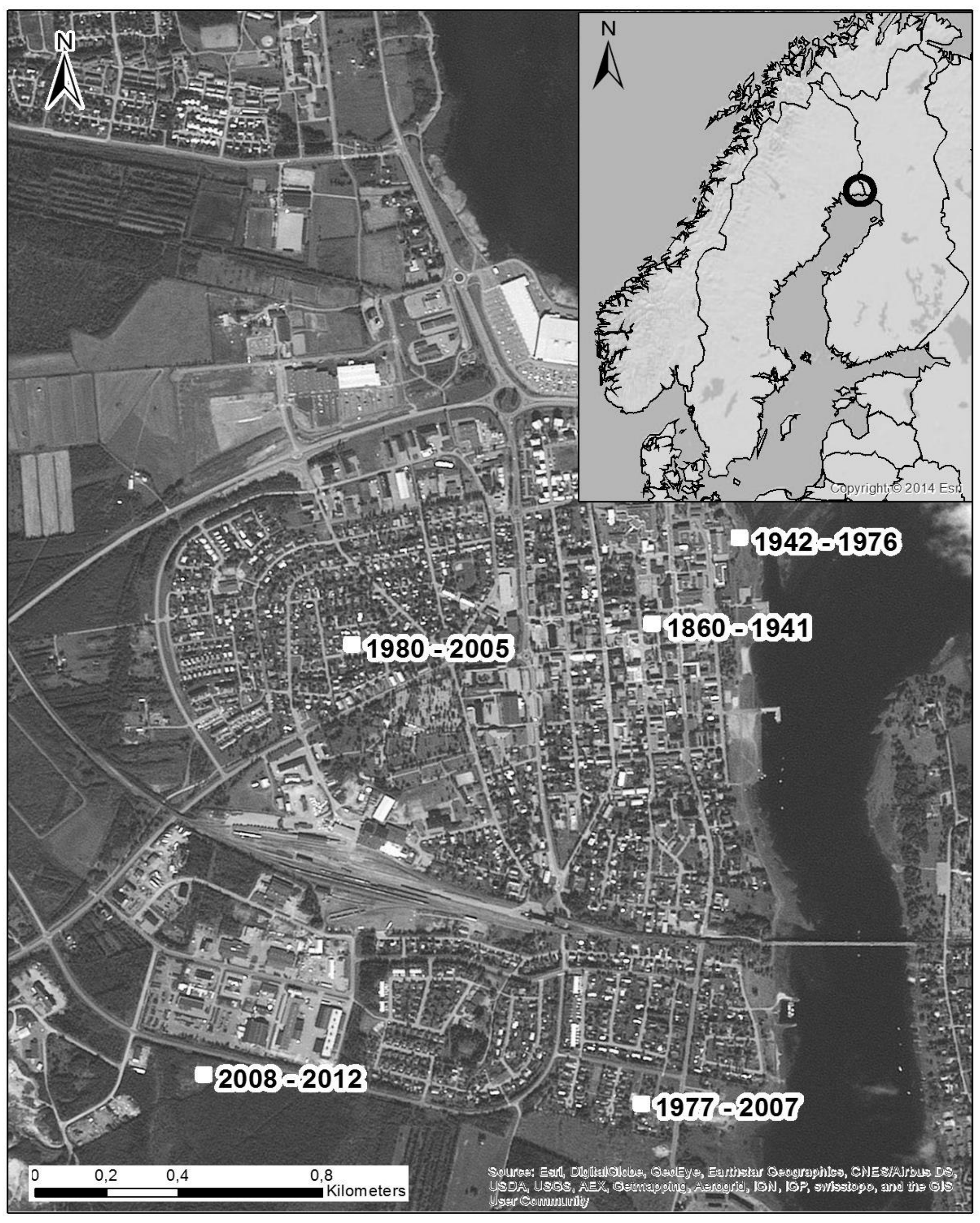


Figure 1: Siting of Haparanda in Sweden (upper right) and satellite picture of the town itself. Shown are all past and the present met station locations. The parallel station (1980 – 2005) is depicted as well.

Results

- Two years of measurements show a warming bias in all previous station locations (largest in the urban center; yearly ∆TM= 0.43 K) compared to the AWS site
- Bias strongest in the first station location, decreasing towards recent times → correction increases temperature trend (figure 3)
- Bias strongest in spring, weakest in winter
- Urban bias strongest at night (Minimum temperatures)

	SPRING	SUMMER	AUTUMN	WINTER	YEAR
	Uncorrected [°C/10y]				
T_{MAX}	0.17	0.04	0.10	0.12	0.11
T_{MIN}	0.26	0.15	0.14	0.14	0.18
T_{MEAN}	0.23	0.11	0.14	0.15	0.16
	Corrected [°C/10y]				
T_{MAX}	0.20	0.05	0.11	0.11	0.12
T_{MIN}	0.31	0.17	0.18	0.16	0.21
T_{MEAN}	0.24	0.12	0.14	0.16	0.18

Figure 4: Temperature trends of corrected and uncorrected time series for various parameters and seasons.

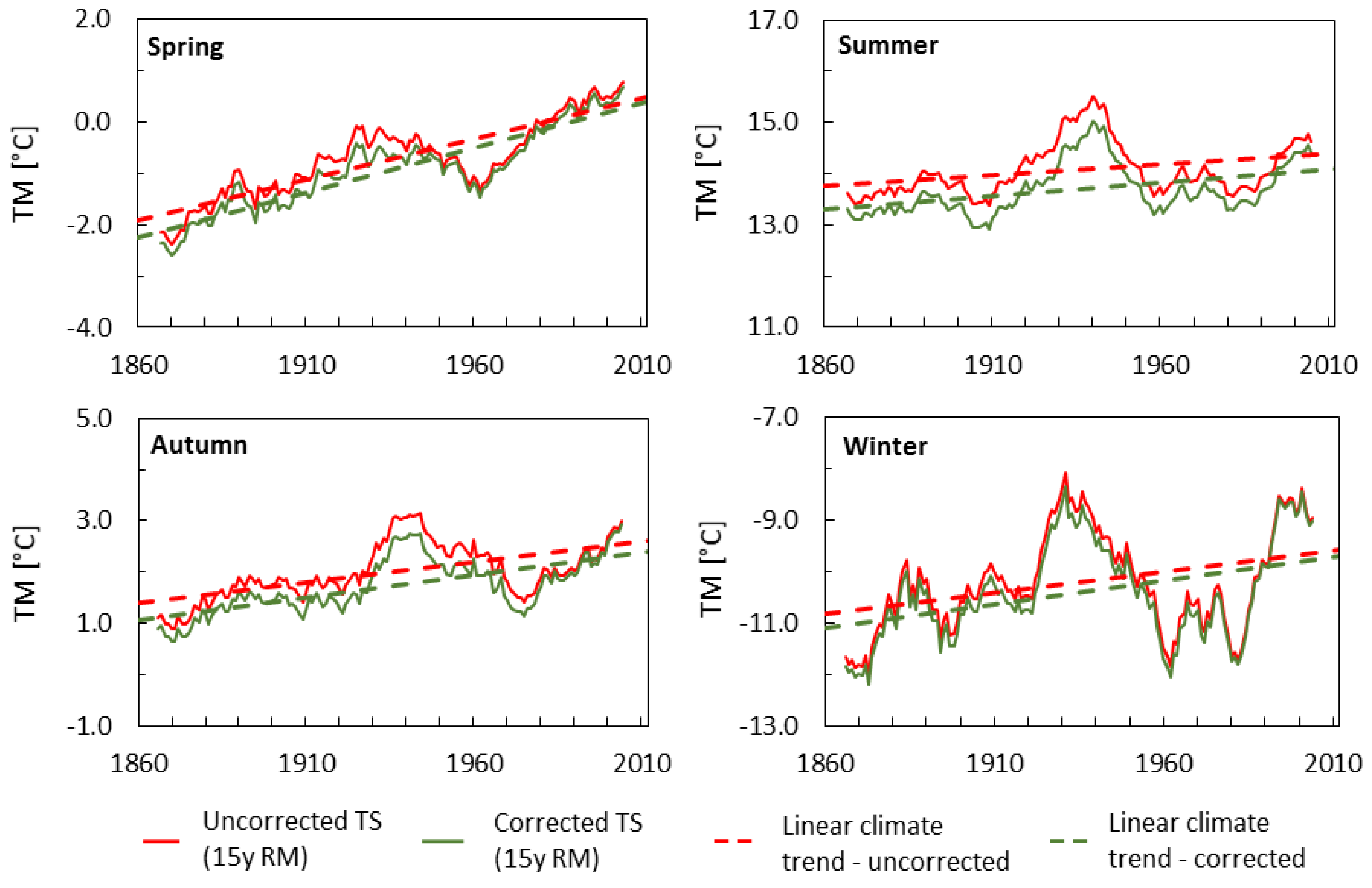


Figure 3: Corrected (green) and uncorrected (red) mean temperature series for each season including the linear trends. The time series was smoothed using a 15 year running mean.

Conclusion

This study clearly indicates that data from the village met station in Haparanda have an urban warming bias in the previous station locations. Applying a simple correction for this bias increased the trend in the long temperature record, especially in spring. As met station relocations from more to less urban surroundings are common, accounting for this urban bias seems to be important for village stations in general as well, and is more likely to result in an increased temperature trend than the opposite.

References:
[1] Auer, I. et al. (2005). "A new instrumental precipitation dataset for the greater alpine region for the period 1800-2002." International Journal of Climatology **25**(2): 139-166
[2] Aguilar, E. et al. (2003). "Guidelines on climate metadata and homogenization." WMO/TD **1186**
[3] Oke, T. R. (2006). "Initial Guidance to obtain representative meteorological observations at urban sites." WMO/TD **1250**: 1-47

Rescue and quality control of European sub-daily data

L. Ashcroft[1], J.R. Coll[1], M. Castellà[1], P. Domonkos[1], E. Aguilar[1], J. Sigró[1] and M. Brunet[1,2]
[1]Centre for Climate Change, University Rovira i Virgili, Av. Remolins 13-15, 43500 Tortosa, Spain
[2]Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK

Questions? Just email lindenclaire.ashcroft@urv.cat or tweet @lindenashcroft.

Introduction

We are rescuing sub-daily European meteorological data to improve regional reanalysis, as part of Uncertainties of Ensembles in Regional Reanalysis (UERRA), an EU/FP7-SPACE-2013 project.

In 18 months we have recovered 5.9 million observations from 99 stations in 13 countries (Figure 1).

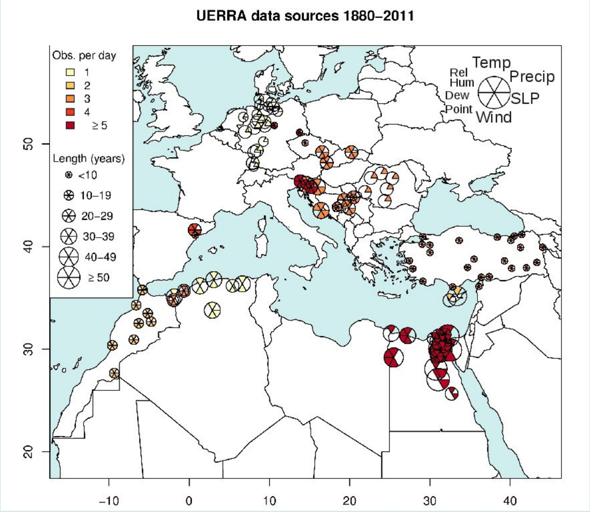
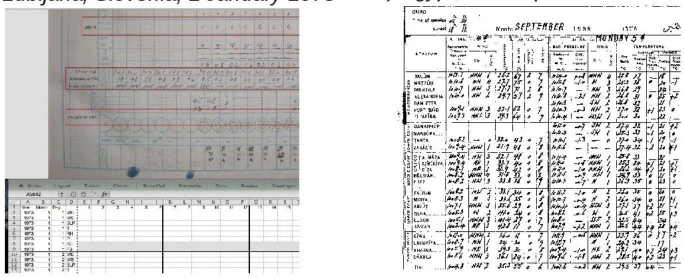


Figure 1. Stations with sub-daily data digitised by UERRA.

What have we learned so far about digitising subdaily data?

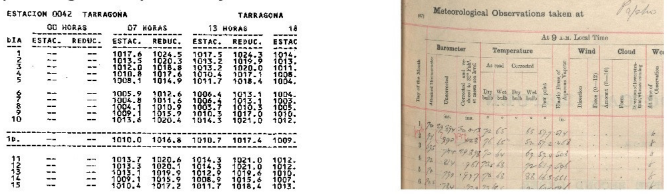
1. All data sources require different preparation: templates are important

- Digitisers need clear instructions about which variables to digitise.
- Spreadsheet templates that match the data source are helpful and reduce errors.
- This is particularly important with sub-daily data, which can be difficult to read (e.g. Figure 2).



a) Lubljana, Slovenia, 1 January 1973

b) Egypt, 4–5 September 1938



c) Tarragona, Spain, July 1977.

d) Paphos, Cyprus, July 1901

Figure 2. Examples of data sources. An example template is shown in 2a).

2. Timely feedback is important to reduce subtle errors and biases

- We conduct systematic visual cross-checking when new sources are digitised.
- Three days of every 4 months for 25% of years are checked, looking for common mistakes.
- Feedback is given to the digitisers to improve future work. Common errors include:
  - > Consistent misreading of confusing values (e.g. a 3 as 8)
  - > Skipped dates, leading to out-of-order data
  - > Incorrect rows or columns digitised (these last two problems are minimised with the use of templates)

These errors are hard to detect using automatic methods.

Table 1: Number of errors found using systematic visual cross-checking.

Number of systematically checked values	Total number of errors found	Gross errors	Subtle and systematic errors	Source errors
20312	418 (2.05%)	8	339	69

3. Automatic quality control is important for gross errors and outliers

- Our programs check for **14 different error types**:

- APRIL 31 — Date order
- Pattern repetition
- T < −91°C, T > 58°C — Record breakers
- Repeated values
- Frequency biases
- Strange distributions
- Climatic outliers
- Jumps and spikes
- $T_{10} > T_{max}$? — Logical failures
- ΔT≈ΔW ≈ RH? — Intervariable comparison
- Calculated vs observed values
- Strange scattering
- Bivariate distribution outliers
- mmHg -> hPa — Unit changes

For example, recovered data from Tarragona, Spain:

- 0700, 1300, 1800 for 1977–1984, 6 variables = 52,020 values.
- Less than 3% of observations were flagged as errors.
- 62% of flagged observations were corrected as typos or removed as clear source errors.

More errors in data digitised without a template.

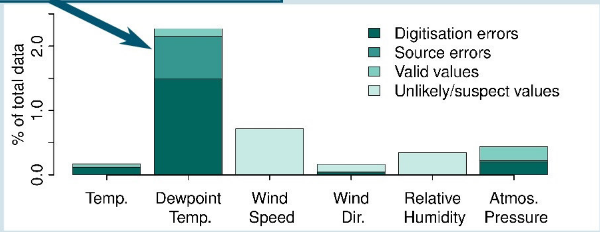


Figure 3. Percentage of total Tarragona subdaily data flagged by the automatic quality control procedure.

Many thanks to the partner National Meteorological Services for providing data for digitisation. We are also indebted to the digitisation team: I. López, D. Azuara, J. Paradinas, A. Domènech, E. Pla, E. Romero, A. Robert, J. Tarragó, C. Bonfill, G. Vandellòs, R. Guerra, A. Àvila, A. Balart, M. Martin, E. Cebrián, P. Sabaté and A. Boqué.

Automatic Range Checks in Quality Control of Observation-network Data:
Towards Operational Implementation

ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Curtis R Wood*, Ari Aaltonen, Anna Frey

*curtis.wood@fmi.fi

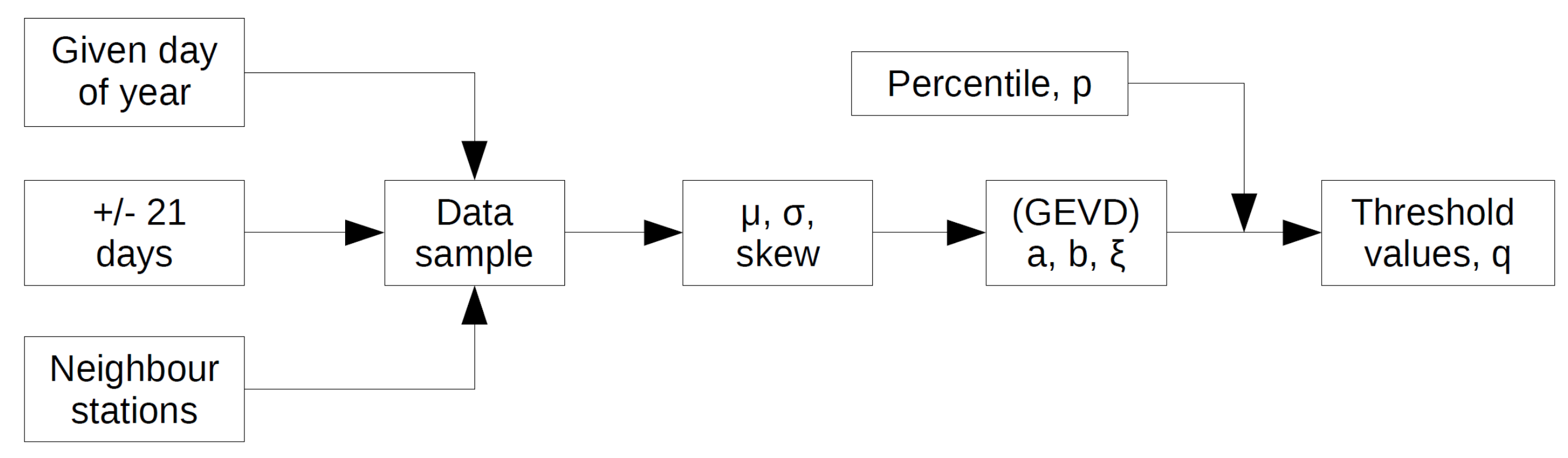
Introduction

Quality-control (QC) systems for data in station networks are needed (Vejen et al. 2002). One part of QC is defining extreme acceptable limit values: **range checks**. Data are flagged that exceed those limit values; flagged data are inspected by a human operator (human quality control: HQC). A former operational method at FMI was that these threshold limits for min and max were manually updated from climate data. For large networks, and in the pursuit of consistency, a reliable and accurate *automated* system is required to reduce HQC and allow limit values to be updated automatically under both climate change and a station's environmental change.

We aim to put this automatic method into operational use.

Methods

This method is very similar to that in Hasu and Aaltonen (2011) where the limit values were defined for all values – e.g. every hour of every day in the year. But for operational use (see later) we use limit values for only daily maximum and minimum. The process is summarized here:



Given day of year
+/- 21 days
Neighbour stations
Data sample
μ, σ, skew
Percentile, p
(GEVD) a, b, ξ
Threshold values, q

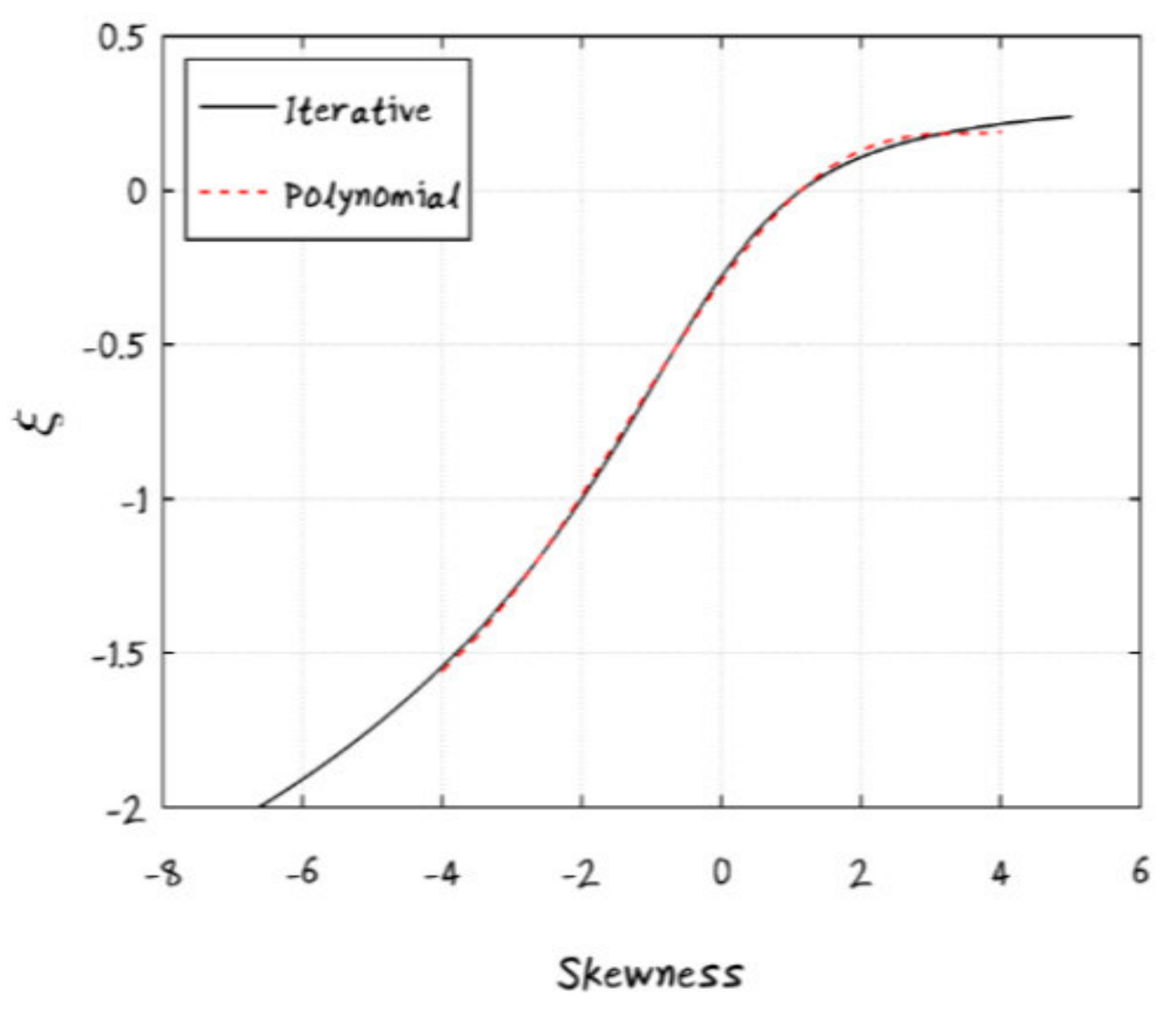
For the day in question – and for all data on this day in this year and previous years – we create two frequency distributions: one of maximum and one of minimum values. We also include ± 21 days, to give us a better sample size; and this can be weighted as Gaussian or evenly. (We could also possibly include neighbour stations to increase sample size.) Anyway, now that we have a sample, we can define statistics such as mean (μ), standard deviation (σ) and skewness ($sk = m/\sigma^3$). Using those 1st, 2nd and 3rd statistical moments, we can subsequently fit/define a distribution to this histogram. From the fit/distribution, we can define percentiles *q* (and thence, if required, return periods).

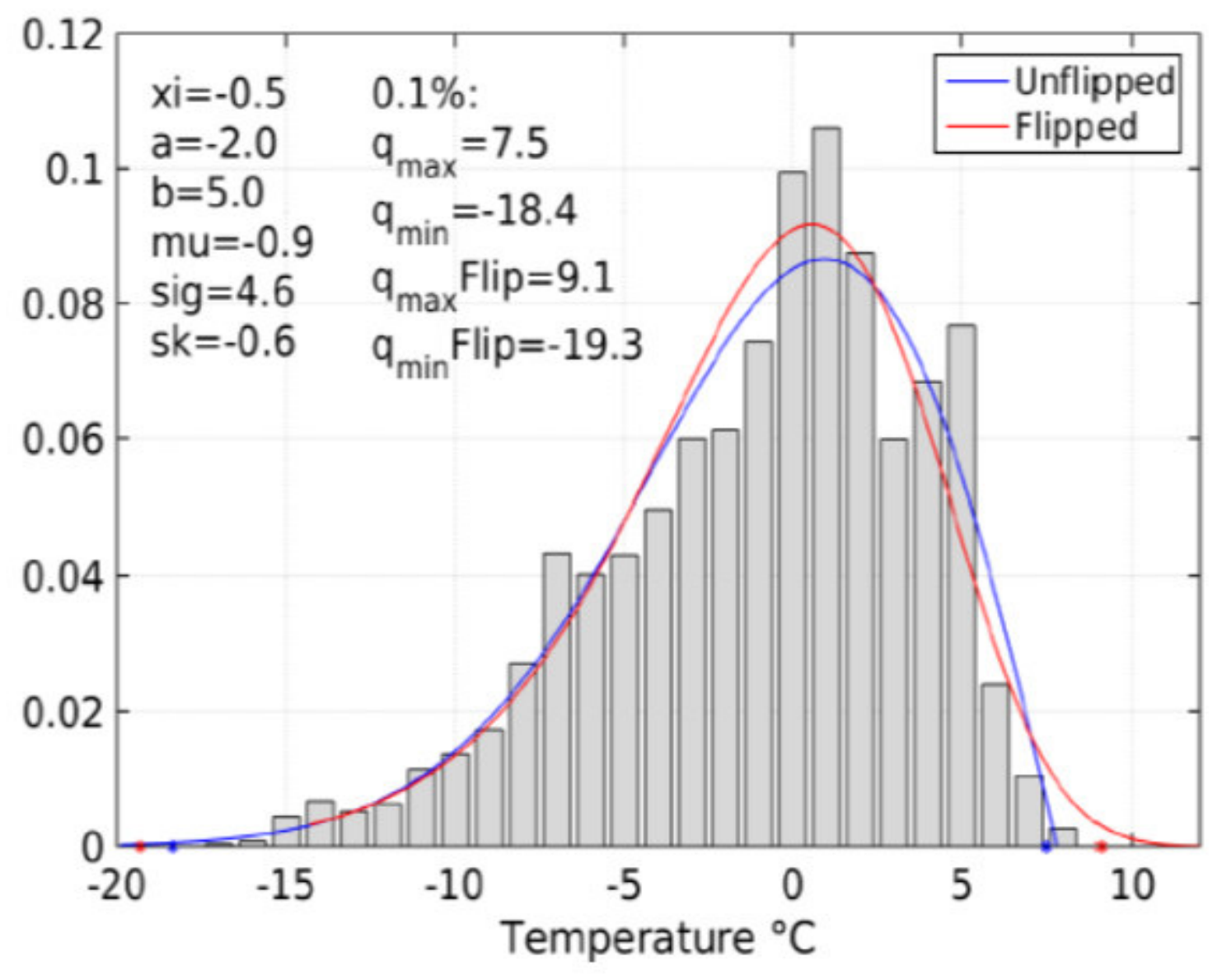
Since we are interested in the tails of the distribution, we use generalized extreme value distribution (GEVD, Coles 2001) which we directly apply to temperature, pressure, wind speed and relative humidity. GEVD comprises types I (Gumbell ξ = 0), II (Fréchet ξ > 0) and III (reverse Weibull ξ < 0). The percentiles for the temperature thresholds are defined as follows (p = 0.001):

$$q_{max} = a + \frac{b}{\xi}\left((-\ln(1-p))^{-\xi} - 1\right) \qquad q_{min} = a + \frac{b}{\xi}\left((-\ln(p))^{-\xi} - 1\right)$$

$$a = \mu + \frac{b}{\xi}(1 - \Gamma(1-\xi)) \qquad b = \sqrt{\frac{\sigma^2\xi^2}{\Gamma(1-2\xi) - \Gamma(1-\xi)^2}}$$

$$\frac{m}{\sigma^3} = sgn(\xi)\frac{\Gamma(1-3\xi) - 3\Gamma(1-\xi)\Gamma(1-2\xi) + 2\Gamma(1-\xi)^3}{(\Gamma(1-2\xi) - \Gamma(1-\xi)^2)^{3/2}}$$



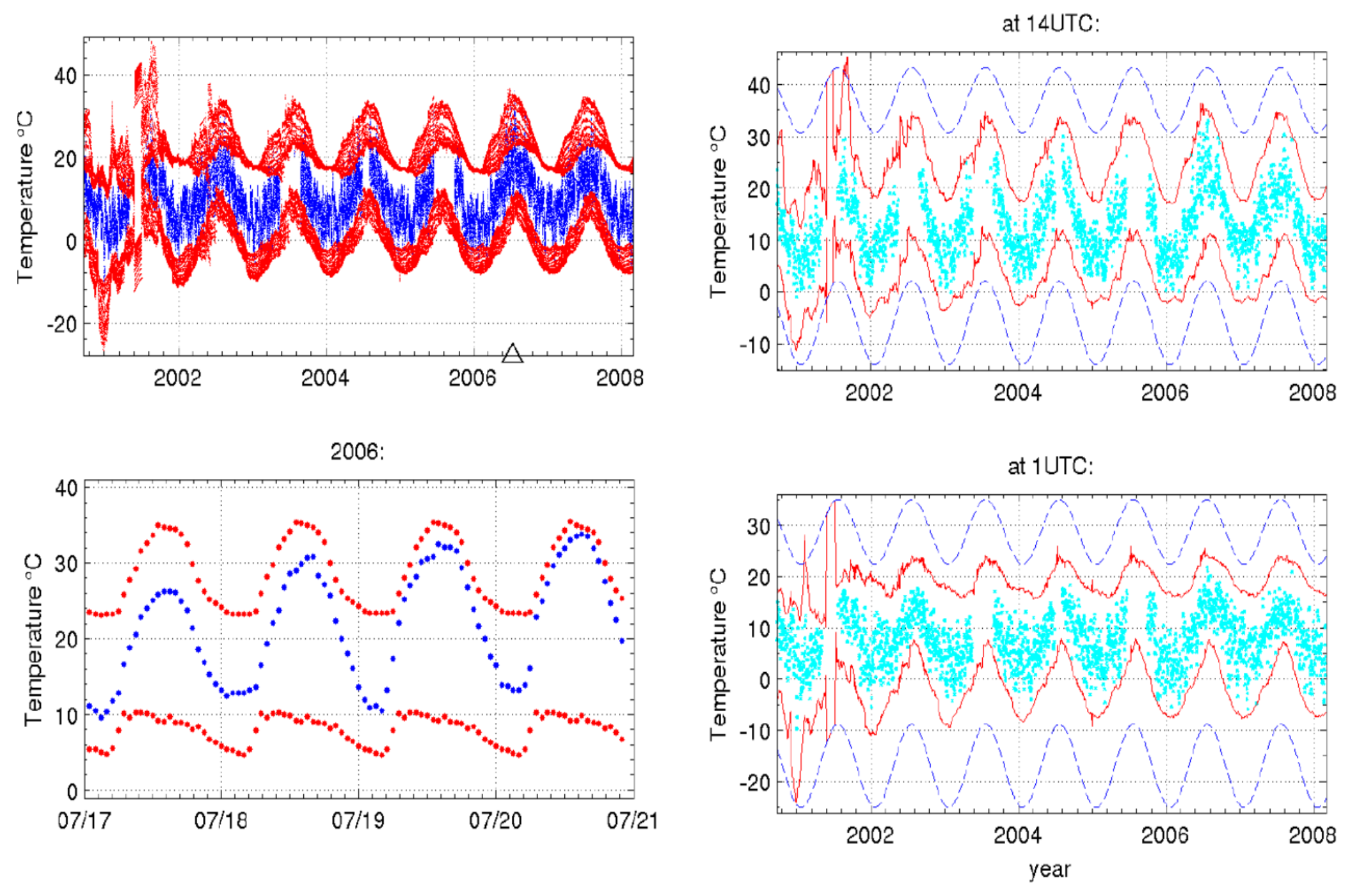


One major caveat becomes apparent from GEVD's being used mostly for the tails on the *skewed* side of the distribution. For our case, we also want information on tail opposite the skewed side. There is no problem for positively skewed distributions (because ξ ≈ 0, hence a continuous GEVD tail on both sides). But for a negatively skewed distribution ξ << 0 (i.e. reverse Weibull), and so the maximum of the fitted distribution is bounded on the upper aide. Thus for those cases we require an unbounded tail (i.e. Fréchet or Gumbell) – an easy way to do this is to mirror/flip the histogram/distribution of the observations so that the sign of skewness changes and the GEVD upper tail is thence unbounded. The example figure (right) is for an arbitrary sample from a station in southern Finland in winter.

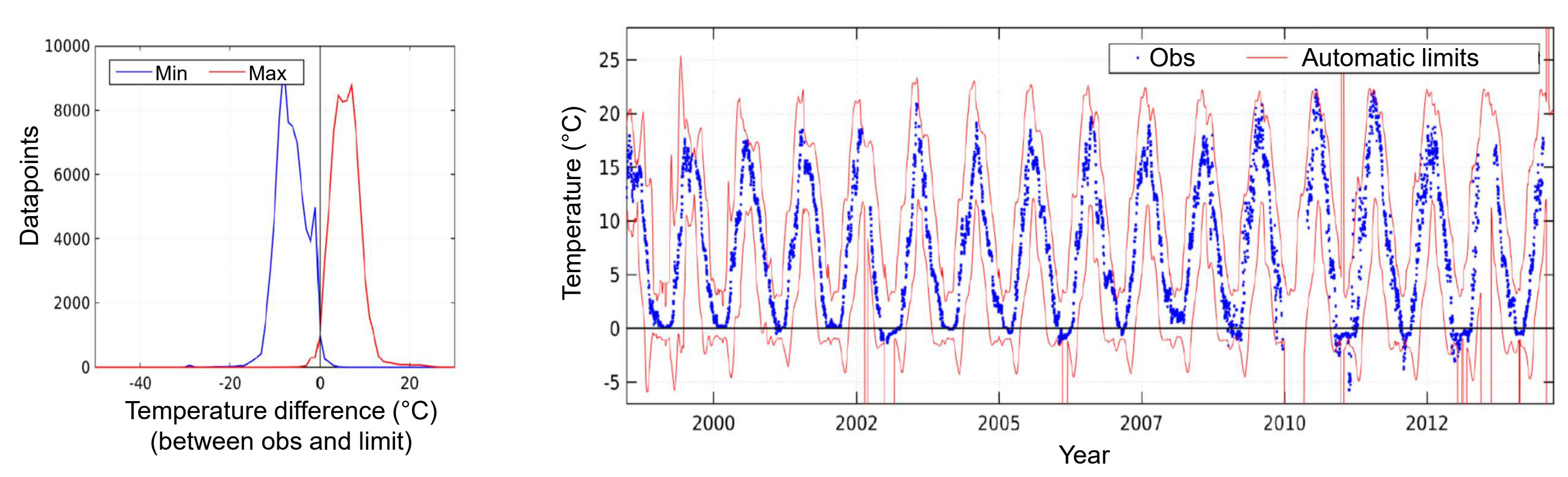
Offline testing in MATLAB

Air temperature at a UK weather station:

(blue dots are observations, red are automated limits)

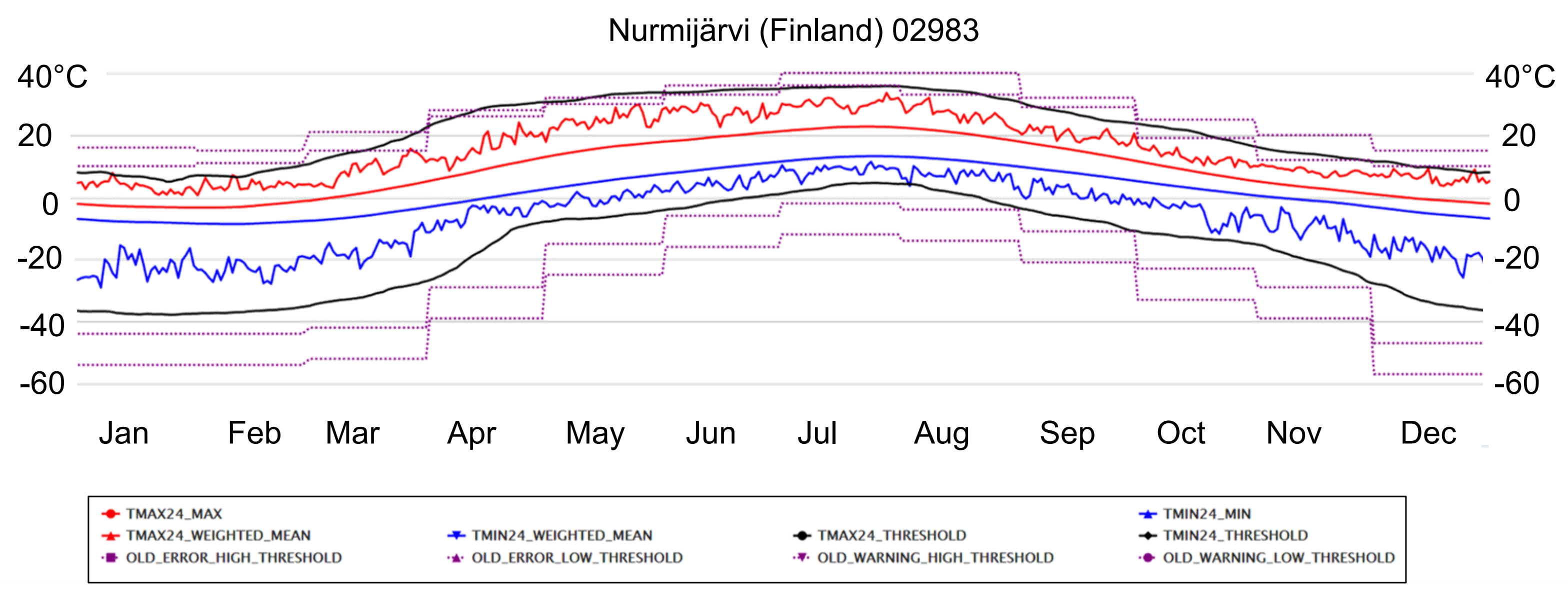


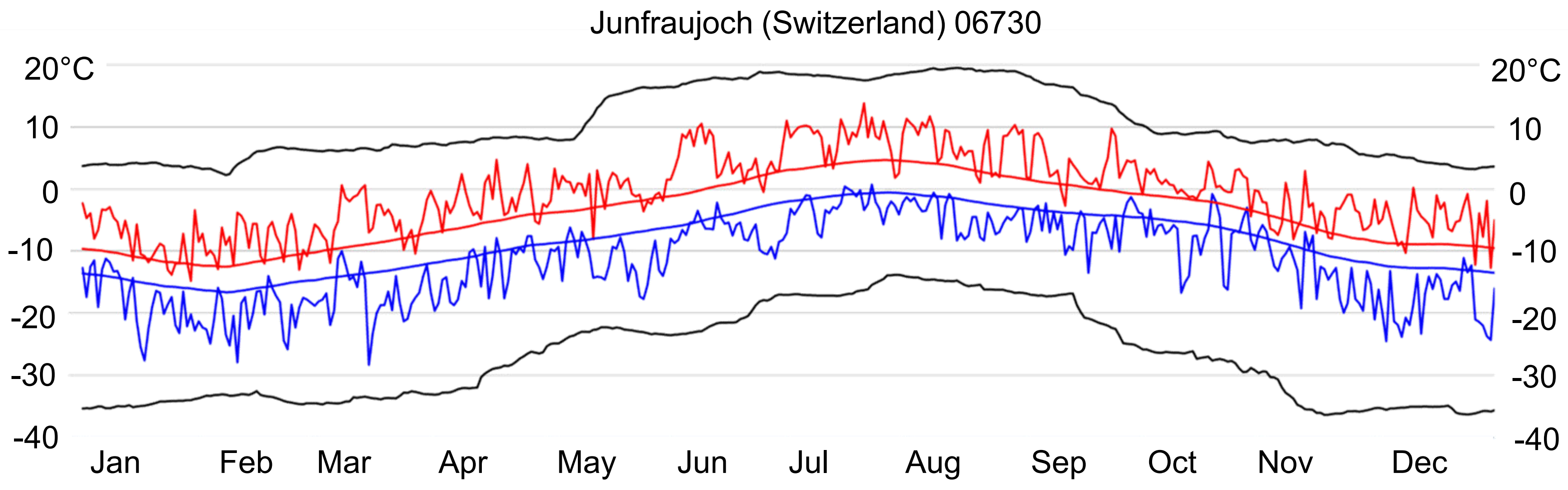
Sea-surface temperature near Helsinki:

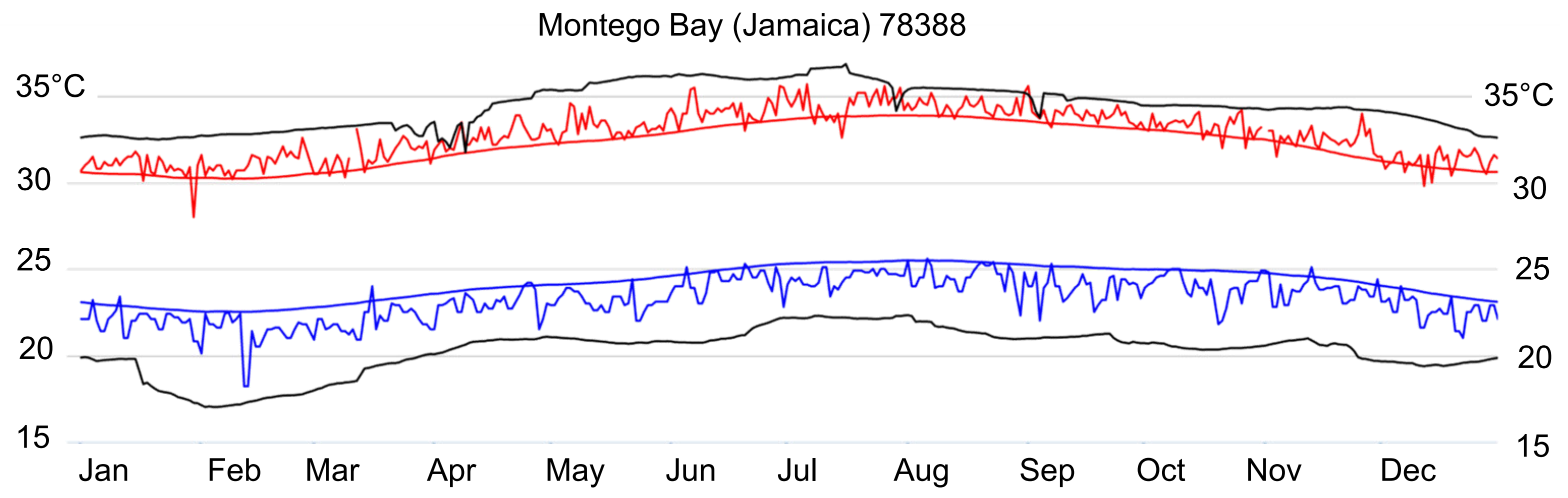


Towards operational use at FMI

These air-temperature examples are directly taken from an FMI intranet site:







Summary

✓Already successfully tested in different countries and for different variables: air temperature, sea-surface temperature, pressure, humidity and wind speed.

✓This will be soon under operational use in FMI.

ACKNOWLEDGEMENTS

Thanks for help from Ismo Karjalainen and Katri Leinonen.

REFERENCES

Coles SG (2001) **An Introduction to Statistical Modeling of Extreme Values.** Springer-Verlag, 209 pages, ISBN 978-1-85233-459-8
Hasu & Aaltonen (2011) **Automatic Minimum and Maximum Alarm Thresholds for Quality Control.** J. Atmos. Oceanic Technol. 28: 74–84
Vejen F, et al. (2002) **Quality control of meteorological observations—Automatic methods used in the Nordic Countries.** NORDKILM, 109 pages, www.smhi.se/hfa_coord/nordklim/task1/quality_control.pdf

26.10.2015

Australian Government

Bureau of Meteorology

A homogenised daily temperature data set for Australia

Blair Trewin and Simon Grainger
Climate Information Services
Australian Bureau of Meteorology
b.trewin@bom.gov.au

1. What observations are available to support long-term temperature analysis in Australia?

The first systematic long-term temperature observations in Australia began in the mid-19th century, although some fragmented short-term observations were made at various locations since shortly after European settlement in 1788.

Until 1901, Australia was governed as six separate British colonies (each with their own government agency responsible for meteorology), and the Bureau of Meteorology was not founded as a federal organisation until 1908. Some of the States (notably Queensland and South Australia) had adopted the Stevenson screen as a standard instrument shelter by the early 1890s, but in New South Wales and Victoria, the Stevenson screen only became standard in 1908. For this reason, and because of a lack of pre-1910 observations of any kind in most of western and central Australia, the ACORN-SAT homogenised temperature dataset starts in 1910.

The Stevenson screen has remained the standard instrument shelter in Australia since 1910. Automatic weather stations have become increasingly widespread since the 1990s and now make up about 80% of the temperature network, but when automatic instruments were introduced, the same screen design was retained.

About 1700 stations have measured temperature in Australia at some time, of which about 750 are currently operating. However, only about 190 have 50 years or more of digitised daily observations.

The total number of stations has not changed greatly since 1910 (although the availability of digitised data has). However, the distribution of stations over the continent has improved since the 1940s, with more stations opening in remote areas of central and northern Australia, and in Tasmania – before 1940, Alice Springs (Figure 1 left) has an effective "footprint" of about 10% of the Australian continent. There are still large data voids in the western interior, with some areas having no data anywhere within a 300-kilometre radius.

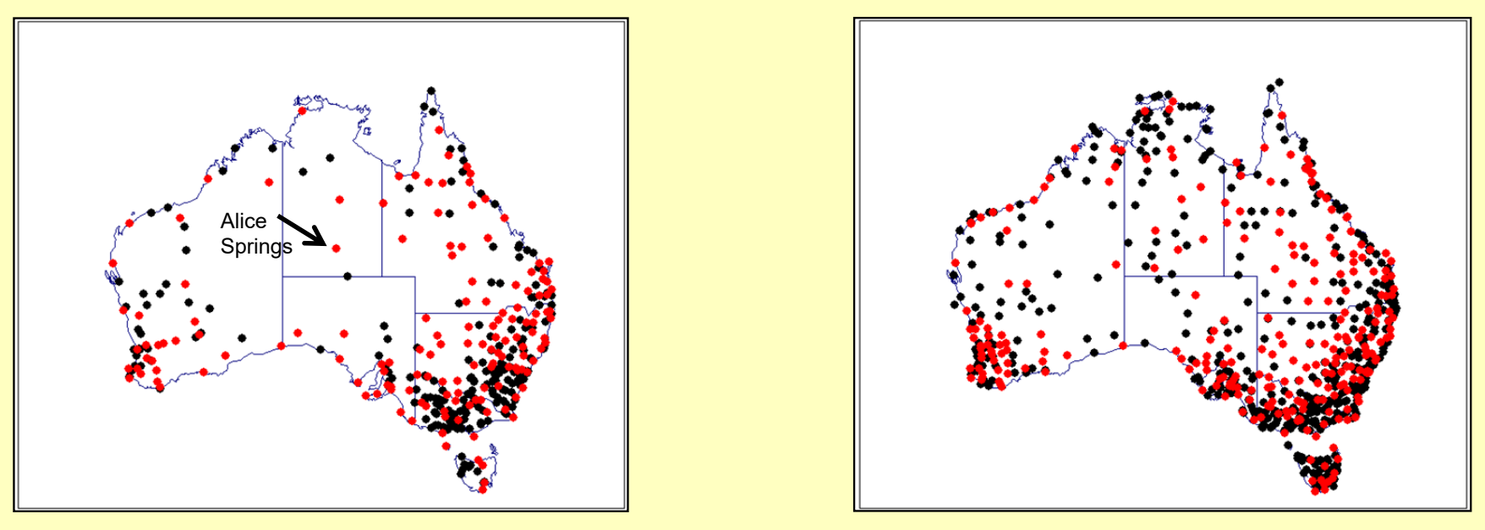


Figure 1. Australian temperature observing network in 1930 (left) and 2010 (right). Sites with 40 or more years of data are shown in red.

2. The ACORN-SAT network and data

112 locations were selected for the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) network (Figure 2). 60 of these 112 locations have digitised daily data extending back to 1910 (in many cases, from a composite of two or more stations in close proximity to each other). A number of key stations in central Australia opened in the 1940s and 1950s. Much pre-1957 Australian daily temperature data remains to be digitised, and about 20 more stations are potential candidates for addition to the data set once their pre-1957 data are available.

Extensive metadata exists in recent years for the ACORN-SAT stations. Metadata becomes sparser as one goes back further in time, especially before 1950, and most pre-1997 metadata exists only on paper documents. As such, although the available metadata is useful, it is not sufficiently complete to be the sole basis of homogeneity assessment.

It has been standard (although not universal) practice since the mid-1990s for parallel observations to be carried out for two years or more when major site moves take place. Before then, it was quite common for major site moves (e.g. town centre to airport) to take place with no parallel observations, and often with no change in station identifier.

The Bureau of Meteorology now has a comprehensive data quality control system, which became fully operational in 2008. Less complex methods were used before then, and much pre-1975 data has only received very limited quality control. An independent quality control process was carried out for the ACORN-SAT data, with the intention of subjecting all historic data to a level of quality control comparable to that used operationally post-2008, to the extent that the availability of reference data (e.g. neighbouring stations) allowed.

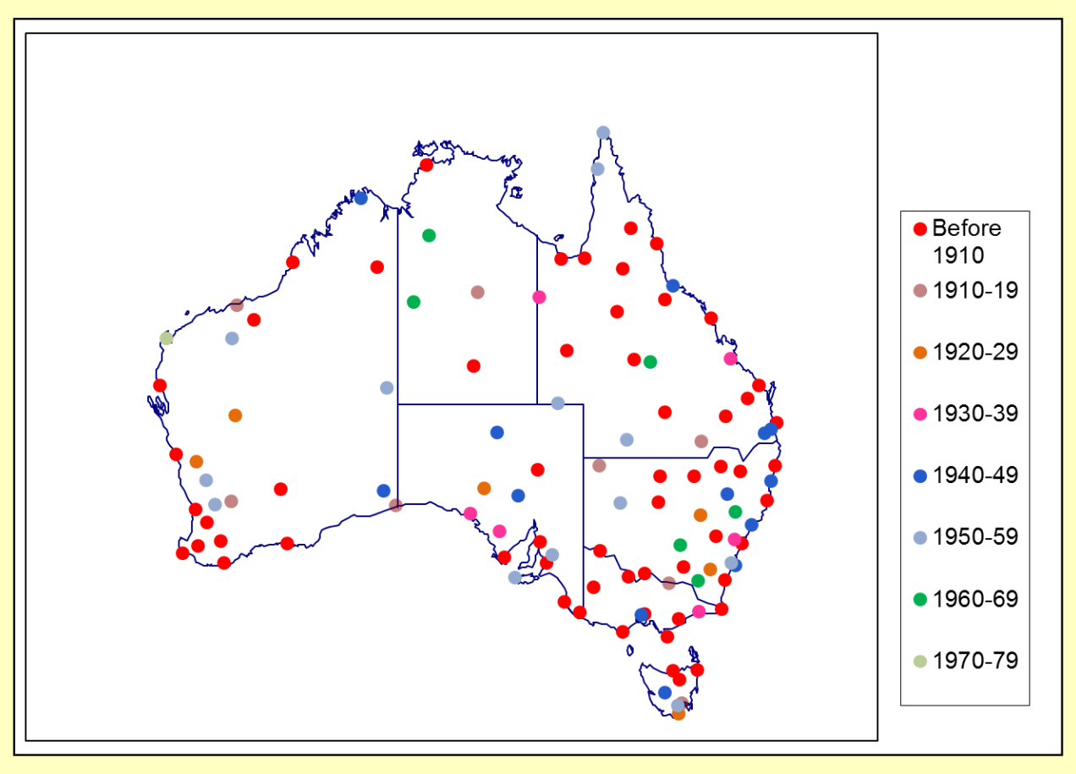


Figure 2. The ACORN-SAT network of 112 stations, with opening dates shown.

3. Why do we need a complex adjustment algorithm?

The impact of an inhomogeneity on temperature observations is not necessarily uniform throughout the year. Inhomogeneities may have a seasonal cycle; for example, if a site moves from a coastal location to one which is more continental, the impact on maximum temperatures is likely to be largest during the summer when sea breezes are a more significant influence.

Figure 3 shows an example of a case where there is no significant inhomogeneity in the annual mean but a substantial one in seasonal values. In this case, the site moved from a location in the centre of a small town with similar local land-surface conditions all year, to cropland about 2 km north of the town. During the summer, there are no crops, the land surface is bare soil and stubble, and maximum temperatures at the new site are hotter; on the other hand, during winter (which is the growing season at this site), the site is surrounded by green crops and is cooler than the town.

Furthermore, some inhomogeneities will have differing impacts under different weather conditions. In particular, the impact of urban development and of local topography on temperature is normally greatest on clear, calm nights, which are generally colder nights (especially in the winter half of the year). This is illustrated by Figure 3, which shows the relationship between winter minimum temperatures at two sites, one near the coast and one 4 km inland, during 7 years of parallel data. On the coldest nights, the inland site is typically 2-3°C colder than the coastal site, but on the warmest nights (which are usually cloudy and/or windy), this difference falls to below 1°C.

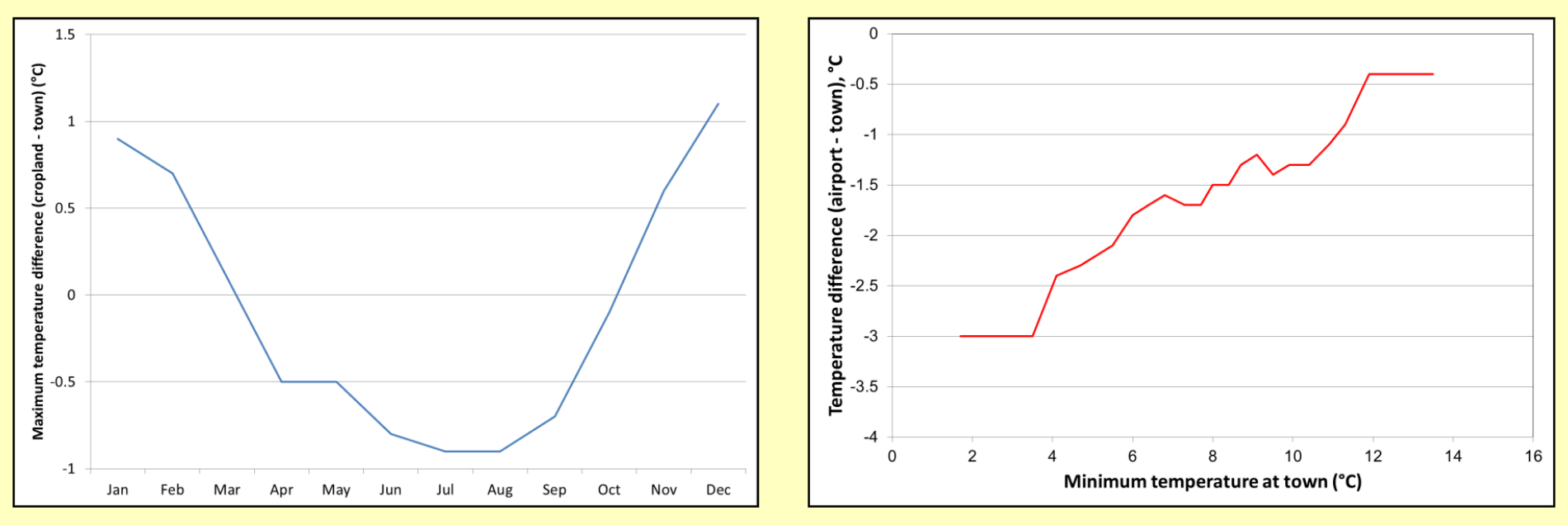


Figure 3. (left) Mean monthly maximum temperature differences during 1999-2000 overlap period between new (cropland) and old (town) sites at Snowtown. (right) Winter minimum temperature difference between airport (inland) and coastal (town) site during 1996-2003 overlap at Port Macquarie.

If seasonal and weather-type impacts on inhomogeneities are not accounted for, an adjustment procedure may produce a series which is homogeneous in annual means, but still inhomogeneous at the seasonal or daily level. Time series which are homogeneous at the daily level are especially critical for the analysis of extremes.

4. Implementation of the percentile-matching algorithm

Initially, inhomogeneities are detected with a combination of metadata, and use of the pairwise homogenisation algorithm of Menne and Williams (2009) (with minor modifications). The algorithm is applied separately to annual and seasonal data, with a potential inhomogeneity being flagged if it appears in either the annual data or at least two seasons (within a two-year window).

Where a potential inhomogeneity is identified, there are two possible methods for adjustment:

(a) Where overlap data exist between the old and new site for at least 1 year.
(b) Where overlap data do not exist or exist for less than one year.

(a) Where overlap data exist, then data from the two stations during the overlap period are used to develop a transfer function between the two sites, separately for each of the 12 calendar months (using data from a 3-month window centred on that month, e.g. May-July for June). The transfer function is defined using the 5th, 10th, …, 95th percentile points, with the temperature difference for values below the 5th percentile assumed to be the same as that for the 5th percentile (and similarly for values above the 95th percentile). Figure 3 (right) shows an example of such a transfer function.

(b) Where no useful overlap data exist, then a two-step process is used, using the best-correlated nearby stations (usually 10) as reference stations. For each reference station, transfer functions are developed (normally using the 5 years before/after the inhomogeneity), first to match the old station to the reference station and then to match the reference station to the new station. This allows the development of a transfer function from the old station to the new station. The final transfer function is the median of each of the individual transfer functions derived from each reference station (Figure 4).

Adjustments are only carried out where the adjustment is at least 0.3°C in the annual mean, 0.3°C in at least two seasons or 0.5°C in one season. An evaluation found that adjustments smaller than this, on average, brought no useful skill over making no adjustment at all. The evaluation also found that the percentile-matching algorithm, when applied for adjustments larger than 0.3°C, provided modest improvements over methods based on monthly means on mean-based indicators such as root-mean-square (RMS) error, and more substantial improvements for indicators based on extremes, especially for low minimum temperatures.

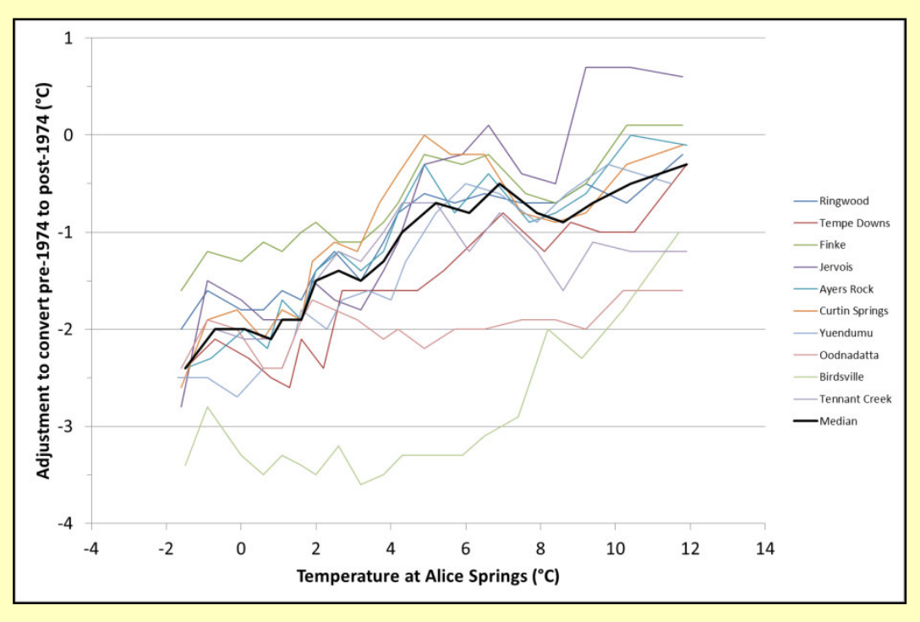


Figure 4. Transfer function for minimum temperature for a 1974 site move at Alice Springs. The final function is in black.

5. What difference does homogenisation make to national and regional time series?

One indicator of the impact that homogenisation has on the temperature data is to compare the homogenised data with unhomogenised data. The data set most readily available for this purposes is the gridded Australian Water Availability Project (AWAP) data set (Jones et al., 2009), which uses all available stations without homogenisation.

At the national level, the ACORN-SAT and AWAP data sets are very similar in the post-1960 period (Figure 5), with the ACORN-SAT data being somewhat cooler in the earlier years of the record, for example by 0.2-0.3°C for both maxima and minima in the 1910-1920 period.

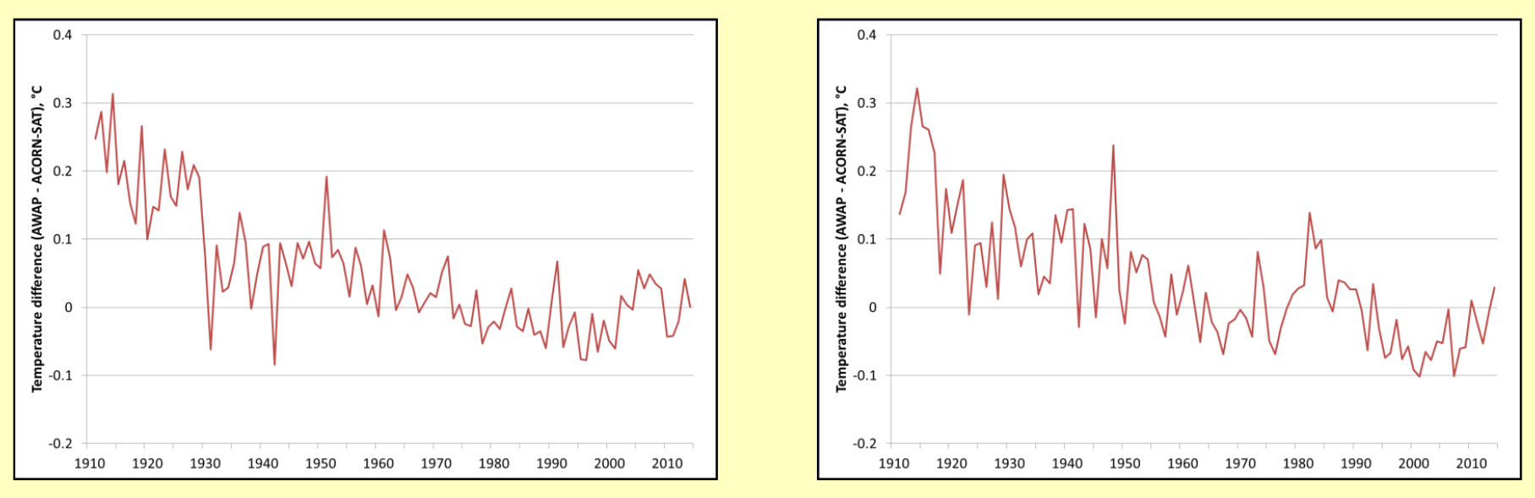


Figure 5. Differences in annual mean Australian temperatures between the AWAP and ACORN-SAT data sets: (left) maximum temperatures, (right) minimum temperatures.

At a regional level, the relationship between homogenised and unhomogenised data can be different from that which exists nationally. Two examples (Figure 6) are in South Australia, where the ACORN-SAT data are warmer in the early years, and in New South Wales, where the unhomogenised minimum temperature data are cooler in recent years. In the South Australian case, major contributors are a 1939 site move in the ACORN-SAT data set from Farina to Marree (which is further north and at a lower elevation, and hence hotter), along with a large inhomogeneity in the 1920s at a non-ACORN-SAT station in a remote area which has a large impact on the AWAP analysis. In the New South Wales case, a number of stations moved from in-town to cooler out-of-town locations at about the same time in the mid-1990s.

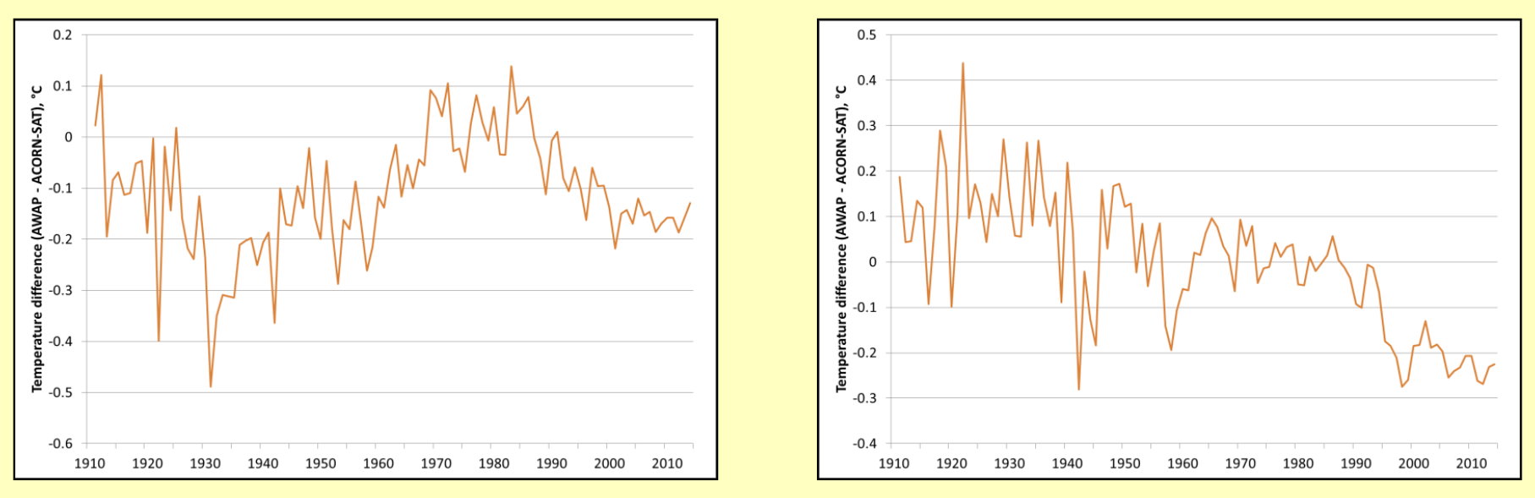
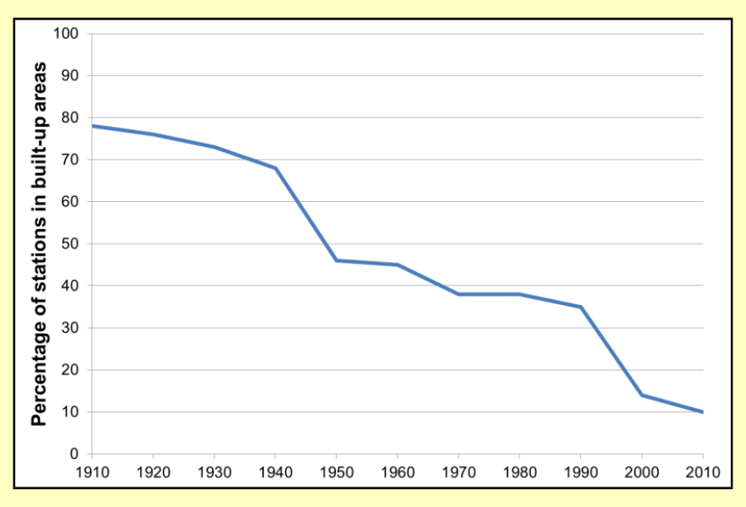


Figure 6. Differences in annual mean temperatures between the AWAP and ACORN-SAT data sets: (left) Maxima, South Australia, (right) Minima, New South Wales.

6. The differing impacts of different inhomogeneity types

Inhomogeneities in climate data sets occur for a wide range of reasons. For long-term temperature data sets in Australia, the most common are site moves, changes in local site condition (such as vegetation changes or building development), changes in observation times, and changes in instruments (although there is no evidence of any significant inhomogeneity arising from the change from manual to automatic instruments).

In the ACORN-SAT dataset, about half the inhomogeneities are supported by metadata, with the remainder being detected by statistical methods alone. Of those inhomogeneities for which a cause can be identified, the mean size of the resultant adjustment is shown in Table 1 below.



Cause	Maximum	Minimum
Merge (with overlap)	−0.08	−0.45
Move (no overlap)	−0.02	−0.13
Screen condition	−0.1	*−0.19*
Site environment	*−0.29*	*+0.32*
Observation time (min only)		+0.16
Others	+0.05	−0.04

(Left) Figure 7. Percentage of all operating ACORN-SAT stations located in urban areas (regardless of size). (Right) Table 1. Mean adjustment (℃) in ACORN-SAT dataset by cause. (Cases where there are fewer than 10 adjustments of this type are shown in italics).

Alternatively, if we separate out those adjustments arising from moves from in-town to out-of-town locations from all other adjustments, the mean adjustment for moves from in-town to out-of-town locations is −0.14°C for maximum temperature, and −0.61°C for minimum temperature. The mean of all other adjustments is near zero for both maximum and minimum temperature (+0.01°C for maxima, +0.04°C for minima).

There is a strong systematic tendency over time for sites to move from in towns to out of towns, which largely accounts for the homogenised minimum temperature data having a stronger warming trend than the unhomogenised data. In 1910, about 80% of the operating ACORN-SAT stations were in towns (some of them very small), whereas in 2015 fewer than 10% are still in towns (Figure 7). Moves out of towns were particularly common in the 1940s, when the demands of military and civil aviation led to many sites being established at airports, and in the 1990s, when the introduction of automatic weather stations was often used as an opportunity to move sites from in-town to out-of-town locations. For maximum temperature, the largest contributor to the difference between homogenised and unhomogenised data has been moves at a small number of key stations with large geographic footprints, most notably a 1932 move at Alice Springs from an enclosed small courtyard to a more open site.

7. How sensitive are adjustments to reference stations or periods?

Some other methods for daily adjustment of temperatures (e.g. Della-Marta and Wanner 2006; Mestre et al. 2011) found that reference stations need to be highly-correlated (0.8 or higher) with the candidate station in order to be useful for adjustment. In contrast, trials using the PM algorithm with reference stations randomly selected from all stations correlated at greater than 0.6, 0.7 or 0.8 found that, whilst the skill of the adjustment decreased with decreasing correlation of the reference stations, a 0.6 minimum correlation threshold still provided useful skill over making no adjustment.

To explore this question further, a series of trials was carried out. In these trials, which were based on a 1995 move from a site in the town (near the coast) to the airport (4 km inland) at Port Macquarie, for each value of N from N=1 to N=10, 50 sets of N reference stations were randomly selected from all those stations correlated with the candidate station at 0.6 or better, and the standard deviation of the annual mean adjustment from the 50 trials was calculated.

It can be seen (Figure 8) that the uncertainty in the annual mean adjustment increases with a decreasing number of reference stations, but that the increase in the uncertainty (relative to N = 10, which is the value normally used in the operational version of ACORN-SAT) is not large until N < 5. The uncertainty in the adjustment where N = 1 is 3 to 4 times larger than that for N = 10, suggesting that using a single reference station is vulnerable to inhomogeneities at that reference station. This also indicates that the use of the median of 10 reference stations in the operational version provides some robustness against inhomogeneities at individual stations.

Figure 8 (right) shows estimated adjustments from individual stations. It is apparent that there is no clear relationship (either in magnitude or spread) between the estimated adjustment size and the correlation with the candidate station. However, there are clearly individual outliers (e.g. stations 2 and 41 for minima, or 20 for maxima), indicative of inhomogeneities at those reference stations.

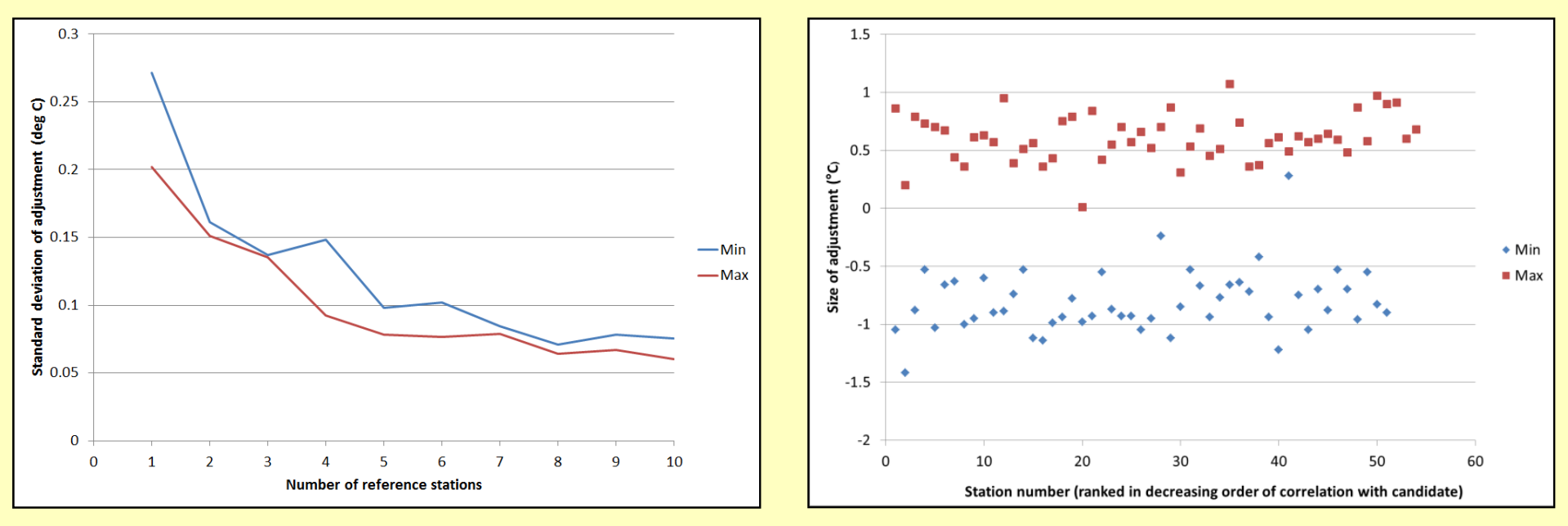


Figure 8. (left) Standard deviation of annual mean adjustment from trials using 50 sets of N reference stations, as a function of N. (right) Estimated adjustments from individual reference stations.

8. Uncertainty in mean Australian temperature

ACORN-SAT time series are being used to investigate the sources of uncertainty in mean Australian temperature. For example, are the differences in annual mean Australian temperature between ACORN-SAT and AWAP (Figure 5) due to systematic differences between the datasets, or can they be explained by variability due to the changing observation network and observational uncertainty?

Following the methodology of Shen et al. (1998), estimates for the uncertainty in area-averaged temperature due to observation error and network distribution can be obtained using ACORN-SAT data. The observation error is obtained through a fitted function of inter-location time series correlations with station separation. The zero-distance fitted correlation is less than 1.0 (Figure 9).

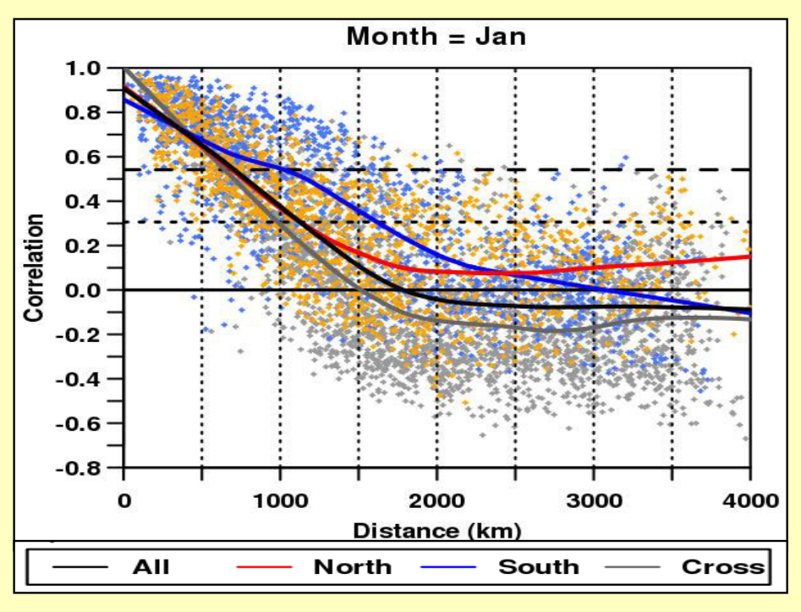


Figure 9. Inter-location correlation as a function of distance for January maximum temperature. The fitted correlation for all stations, and split by region, are shown as solid lines.

The consequence of zero-distance correlation less than 1.0 is that the measured location time series variance is the sum of the "true" variance plus the "observational" variance (Daley 1993; Jones and Trewin 2000). The observation error is therefore the measured variance scaled by the 1.0 minus the zero-distance correlation. This is shown in Figure 10 as an average over the present-day ACORN-SAT network.

The network distribution will be estimated by fitting the spatial variance, estimated using AWAP, to the changing ACORN-SAT network (Shen et al. 1998). This will allow an estimate of the uncertainty in the mean Australian temperature for each year in the record.

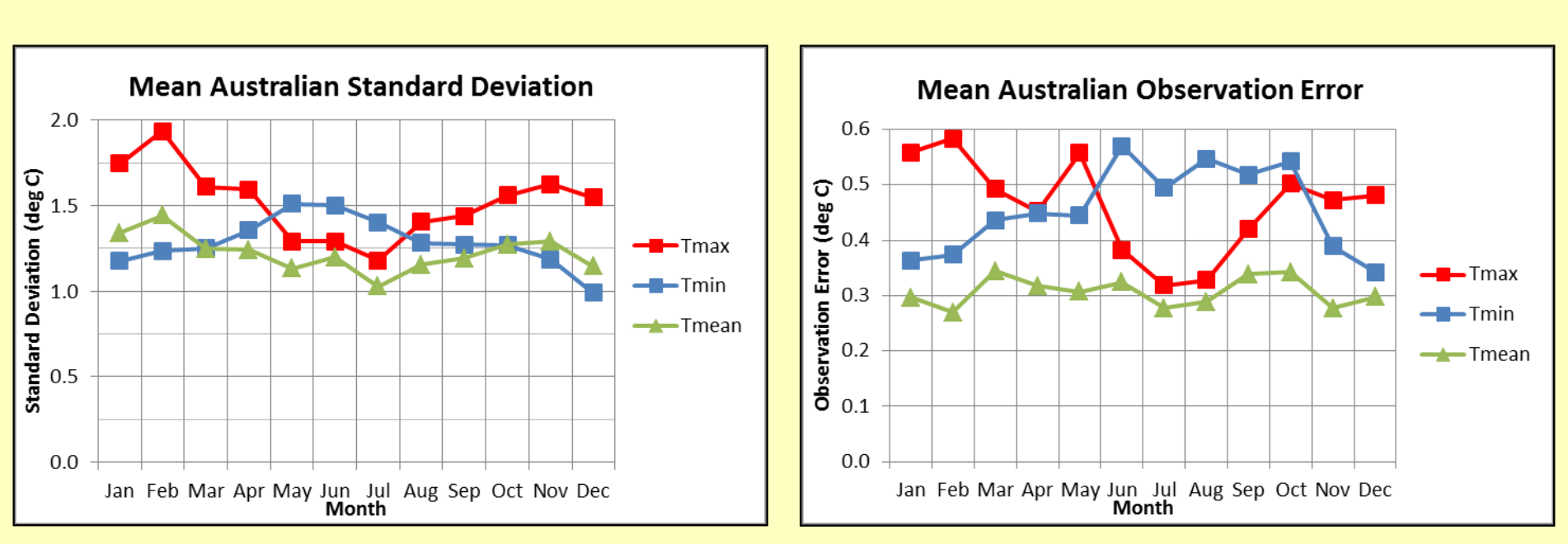


Figure 10. Area-averaged monthly standard deviation of temperature at the ACORN-SAT locations (left panel). Area-averaged monthly observation error estimated from the fitted zero-distance inter-location correlation.

9. Further information and references

The ACORN-SAT data set may be freely downloaded from the Australian Bureau of Meteorology website at http://www.bom.gov.au/climate/change/acorn-sat/. There is also extensive documentation on the methodology (including the two Technical Reports listed below) and the stations used available through that website.

References:

Daley, R., 1993. *Atmospheric data analysis.* Cambridge University Press, 457 pp.

Della-Marta, P.M. and Wanner, H. 2006. A method of homogenizing the extremes and mean of daily temperature measurements. J. Clim., 19, 4179-4197.

Fawcett, R.J.B., Trewin, B.C., Braganza, K., Smalley, R.J., Jovanovic, B. and Jones, D.A.. 2012. On the sensitivity of Australian temperature trends and variability to analysis methods and observation networks. CAWCR Technical Report 50, Centre for Australian Weather and Climate Research, Melbourne.

Jones, D. A., and B. Trewin, 2000. The spatial structure of monthly temperature anomalies over Australia. Aust. Meteor. Mag., 49, 261-276.

Jones, D.A., Wang, W. and Fawcett, R. 2009. High-quality spatial climate data-sets for Australia. Aust. Met. Oceanogr. J., 58, 233-248.

Menne, M.J. and Williams, C.N. 2009. Homogenization of temperature series via pairwise comparisons. J. Clim., 22, 1700-1717.

Mestre O., Gruber C., Prieur C., Caussinus H., and Jourdain S. 2011. SPLIDHOM : a method for homogenization of daily temperature observations. J. Appl. Met. Climatol., 50, 2343-2358.

Shen, S. S., T. M. Smith, C. F. Ropelewski, and R. E. Livezey, 1998. An Optimal regional averaging method with error estimates and a test using tropical Pacific SST data. J. Clim., 11, 2340-2350.

Trewin, B. 2012. Techniques involved in developing the Australian Climate Observations Reference Network – Surface Air Temperature (ACORN-SAT) dataset. CAWCR Technical Report 49, Centre for Australian Weather and Climate Research, Melbourne.

Trewin, B.C. 2013. A daily homogenised temperature data set for Australia. Int. J. Climatol., 33, 1510-1529.

Homogenization of Precipitation Series in Greece

Mamara A.[1,2], Argiriou A.A.[1], Anadranistakis M.[2]

1 Laboratory of Atmospheric Physics, University of Patras, University Campus, GR-265 00 Patras, Greece
2 Hellenic National Meteorological Service, GR-167 77 Hellenicon, Greece

Introduction

Due to the increased interest about climate variability and change, number of methods have been developed in order to deal with the inhomogeneity problems of long term climatic series, since these series reflect the real changes of the climate and provide the most accurate information about the climate evolution. In the early 20th century Conrad (1925) tried to homogenize two precipitation series using the Heidke criterion (Heidke, 1923) and some years later, in the middle of 20th century, (Kohler, 1949) introduced the double mass curve for the break point detection (Venema et al., 2012). Since then many methods have been proposed for the homogenization of climatic series and nowadays around 20 different methods are widely used (Domonkos, 2011).

The main goal of this work is to detect abrupt or more gradual changes in precipitation series from almost the whole Greek meteorological stations network, using three different homogenization methods and to improve precipitation series. It should be pointed out that this is the first attempt to homogenize Greek precipitation series from a large network.

Target Area

Greece occupies the southernmost part of the Balkan Peninsula, jutting out into the Eastern Mediterranean Basin, having the Aegean Sea on the east, the Ionia Sea on the west and the Libyan Sea on the South. Its territory lies approximately between latitudes 34◦and 42◦N and longitudes 19◦ and 30◦E and comprises the mainland (the interior sector of the country), the islands (almost 6000 islands, islets and rocky islets) and the Aegean basin.

Methodology

Three different homogenization methods were applied for the detection and correction of inhomogeneities in precipitation series.

MASH - Multiple Analysis of Series for Homogenization (Szentimrey, 2008) is a relative homogenization method based on multiple comparisons between the climatically similar time series and does not assume homogeneous reference series. The time step of comparisons may be annual, monthly or seasonal and the break point detection is based on hypothesis test for a given significance level, here equal to 1 %.

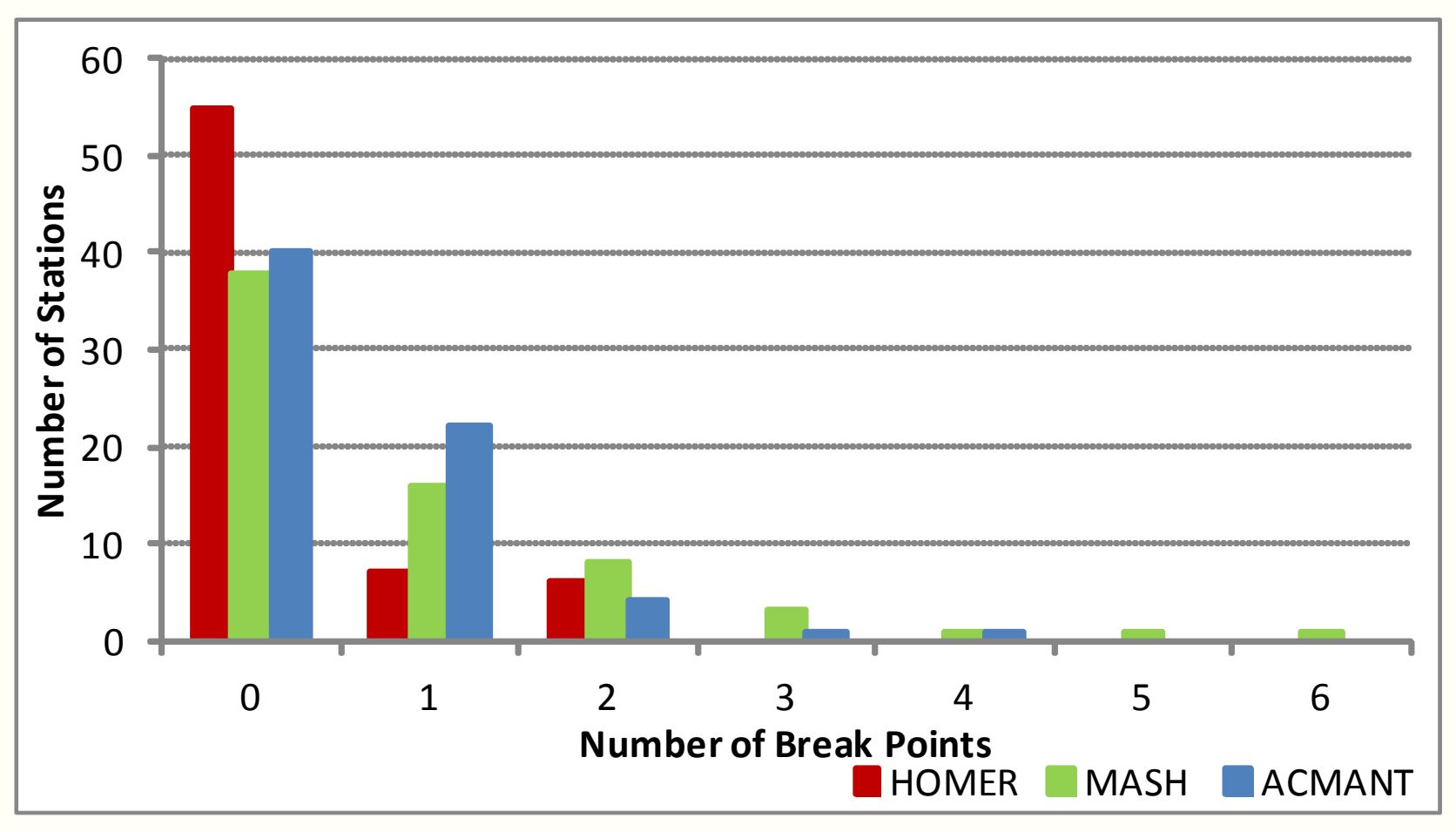
HOMER is a new relative method for homogenizing monthly and annual temperature and precipitation data (Mestre et al., 2013). It was developed in the frame of the European COST Action ES0601 called HOME, devoted to evaluate the performance of homogenization methods used in climatology. It incorporates the best characteristics of some other methods such as PRODIGE (Caussinus and Mestre, 2004), ACMANT (Domonkos, 2011), CLIMATOL (Guijjaro, 2011), that performed good results in benchmark tests (Venema et al., 2012). The detection on break points (for precipitation series) is based on a combination of Dynamic programming and penalized likelihood criteria and on joint segmentation. The correction is based on ANOVA two factors model.

ACMANT is a fully automatic homogenization method developed by (Domonkos, 2011). The most important characteristics are a) harmonization of examinations in different time-scales , b) usage of the optimal segmentation and Caussinus-Lyazrhi criterion in the detection of inhomogeneities, c) usage of ANOVA for the final corrections of inhomogeneities. It uses a pre-homogenization phase where the large errors from the reference composites are filtering.

Results

The examined precipitation series in Greece are fairly homogeneous. Inhomogeneities were found only at 20 % of stations (HOMER), 41 % (ACMANT) and 44 % (MASH).

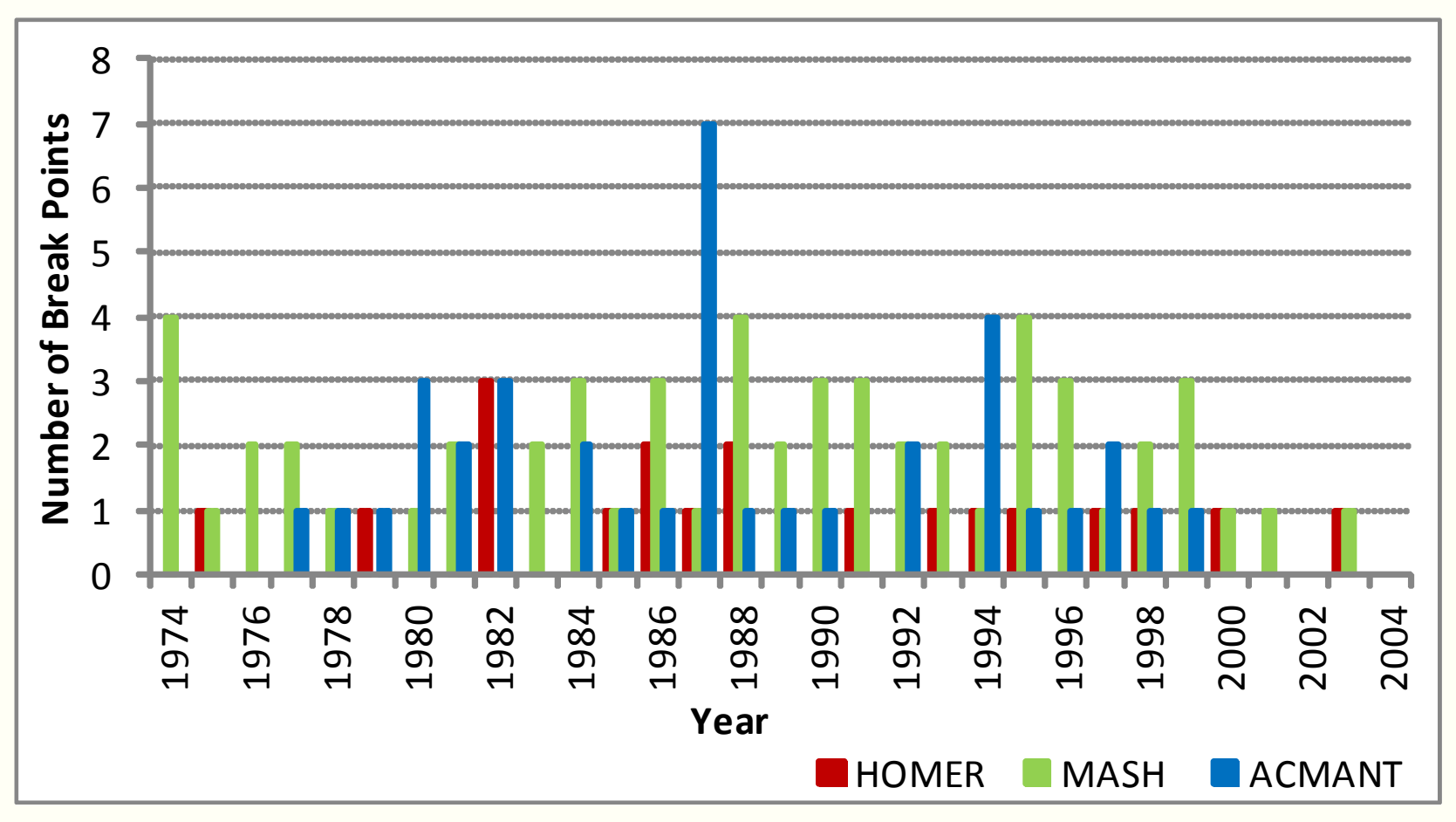
Number of Breaks (outliers are nt included)



Zero break points for 55 stations homogenized with HOMER, 38 with MASH, 40 with ACMANT.

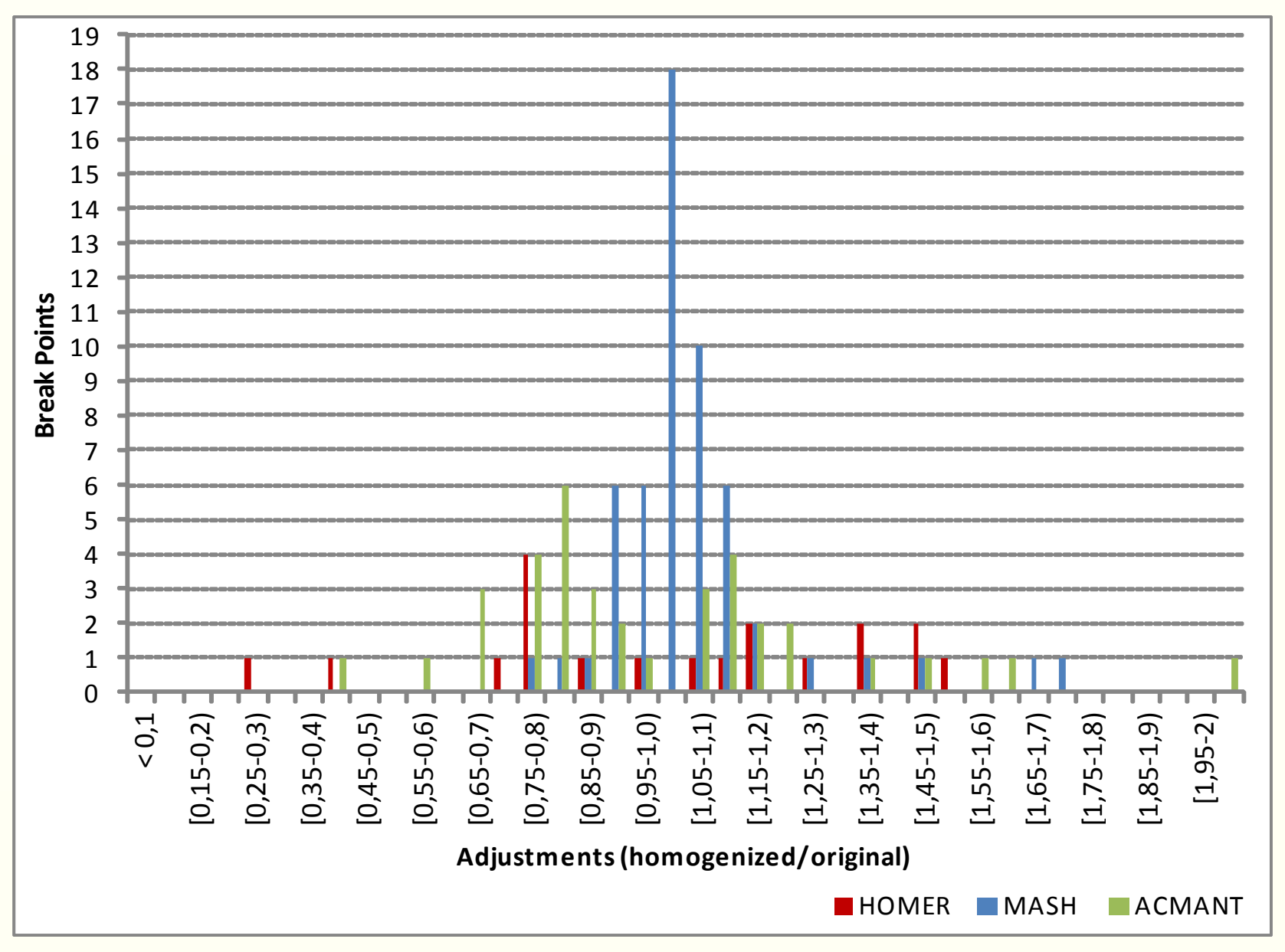
One break for the majority of inhomogeneous series.

Temporal Distribution of Breaks (outliers are not included)



40-50% of breaks have been detected during ‘80s. Due to aviation needs HNMS's weather station network upgraded between middle of ‘70 and early ‘90.

Adjustments on annual time series (homogenized/ original)



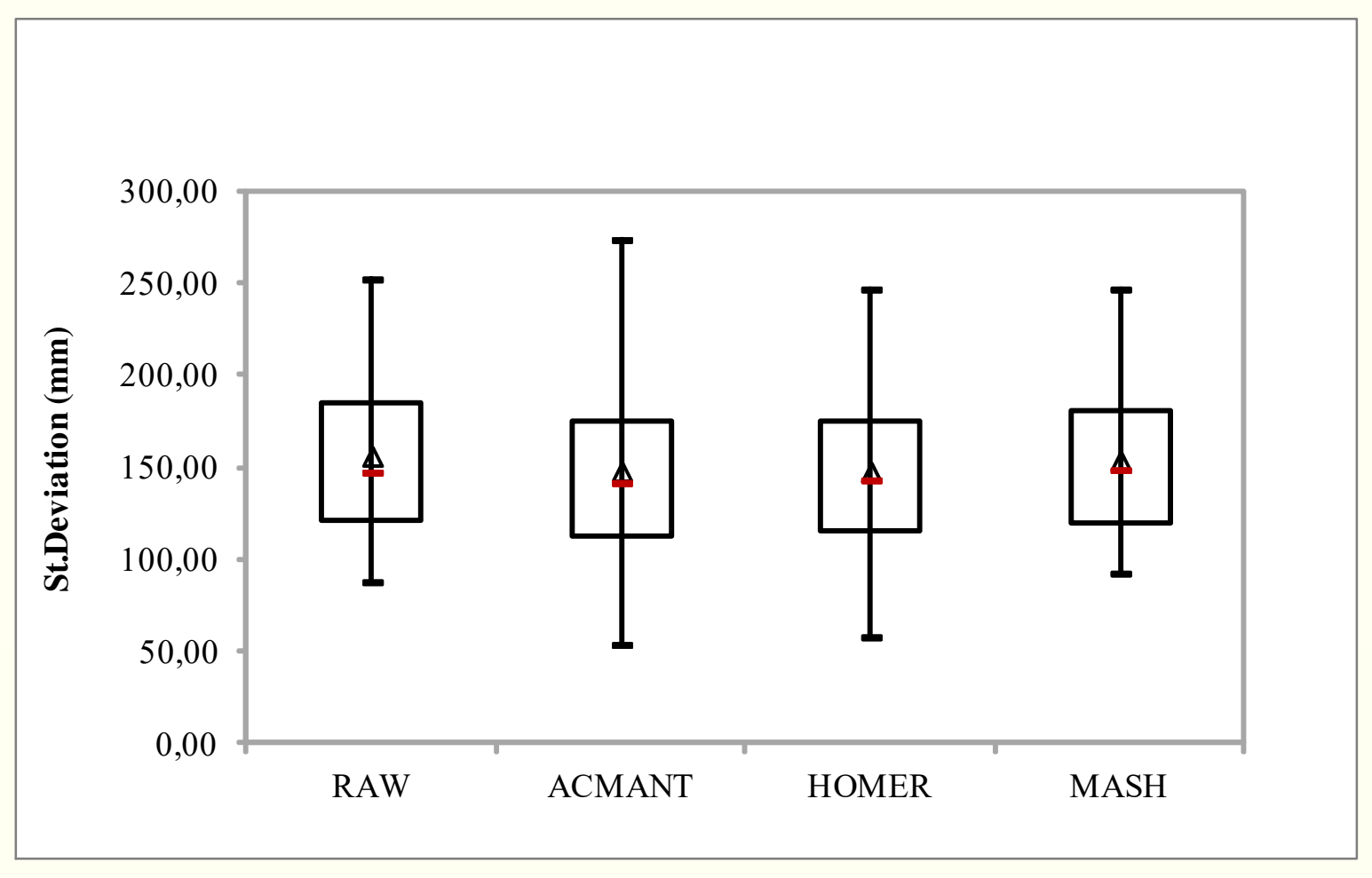
Corrections with HOMER and ACMANT were more or less equally positive and negative

while most corrections with MASH were positive.

Impacts of Homogenization

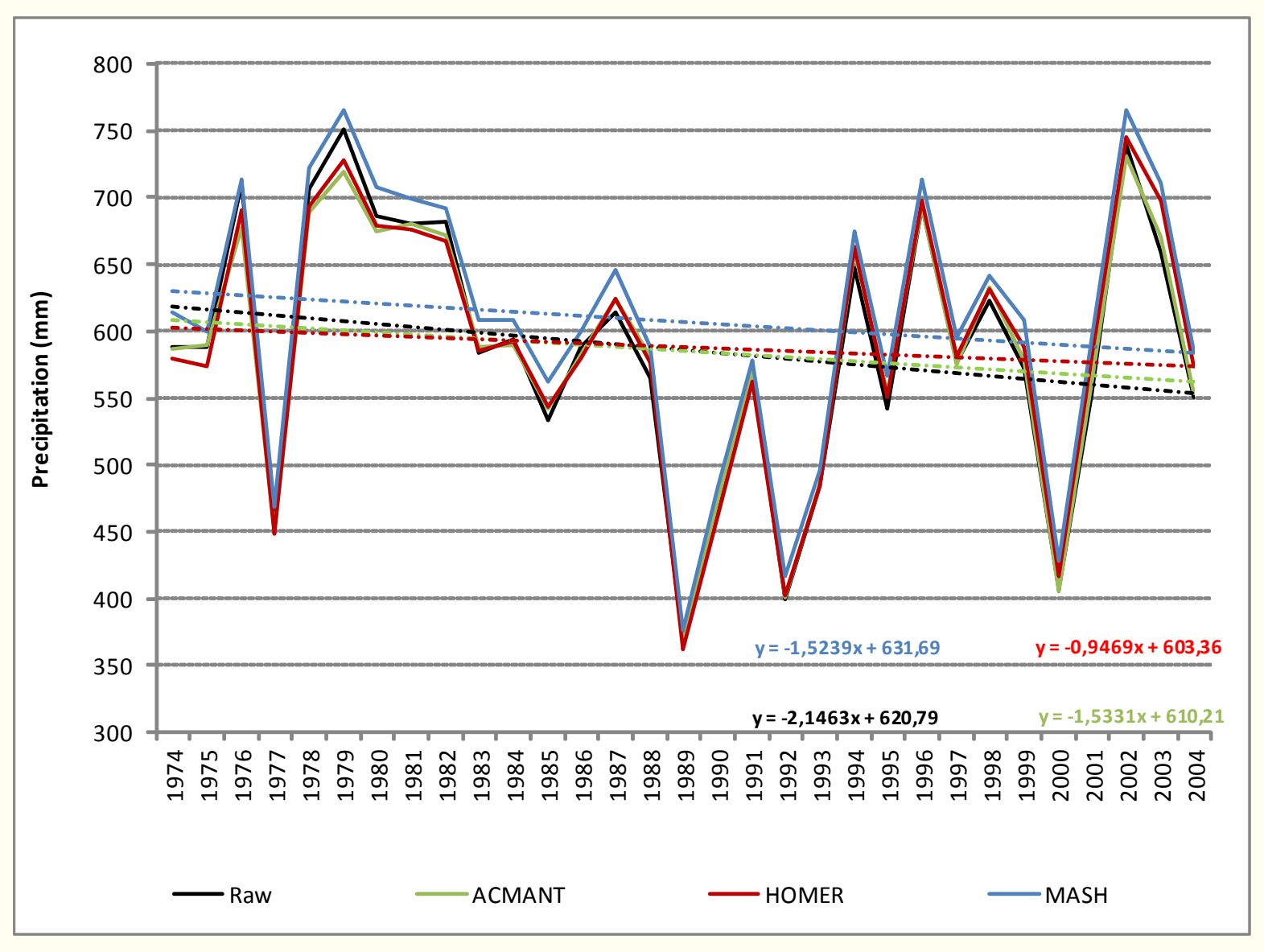
The results were different than those found from the homogenization of temperature series (Mamara et al., 2013;2014)

Boxplot of standard deviation of average annual precipitation series



Absence of substantial deviation between raw and homogenized annual series.

Average annual precipitation series for the whole network for the period 1974-2004



There is no clear improvement of precipitation series after homogenization.

Regional trends per year of annual and semi-annual series

Annual	Raw			ACMANT			HOMER			MASH		
	Test Z	Signific.	Q	Test Z	Signific.	Q	Test Z	Signific.	Q	Test Z	Signific.	Q
Climatic Region 1	-1,80	+	-3,31	-0,95		-1,66	-0,65		-1,02	-1,77	+	-3,01
Climatic Region 2	-0,92		-1,48	-0,14		-0,38	-0,31		-0,35	-0,65		-0,87
Climatic Region 3	-0,71		-3,03	-0,58		-2,30	-0,37		-1,52	-0,58		-2,46
Climatic Region 4	-0,68		-1,63	-0,51		-1,77	-0,51		-1,13	-0,51		-1,35
Climatic Region 5	-0,37		-1,14	-0,07		-0,18	-0,07		-0,52	-0,20		-0,33
October-March	**Raw**			**ACMANT**			**HOMER**			**MASH**		
	Test Z	Signific.	Q	Test Z	Signific.	Q	Test Z	Signific.	Q	Test Z	Signific.	Q
Climatic Region 1	-1,02		-1,82	-0,44		-1,00	-0,17		-0,31	-0,82		-1,68
Climatic Region 2	-0,48		-0,65	0,44		0,51	0,37		0,47	0,14		0,14
Climatic Region 3	-0,85		-2,90	-0,95		-2,81	-0,65		-2,24	-0,88		-2,83
Climatic Region 4	-0,78		-1,17	-0,75		-1,34	-0,61		-1,30	-0,54		-0,86
Climatic Region 5	-0,41		-0,69	-0,03		-0,17	-0,07		-0,66	-0,24		-0,49
April-September	**Raw**			**ACMANT**			**HOMER**			**MASH**		
	Test Z	Signific.	Q	Test Z	Signific.	Q	Test Z	Signific.	Q	Test Z	Signific.	Q
Climatic Region 1	-1,26		-1,32	-0,75		-0,89	-0,65		-0,80	-1,33		-1,36
Climatic Region 2	-0,99		-0,89	-0,82		-0,76	-0,92		-0,59	-0,88		-0,81
Climatic Region 3	-0,37		-0,37	-0,48		-0,42	-0,51		-0,52	-0,61		-0,75
Climatic Region 4	-0,07		-0,03	0,03		0,05	0,07		0,04	0,00		0,03
Climatic Region 5	-0,82		-0,47	-0,58		-0,29	-0,88		-0,51	-0,95		-0,57

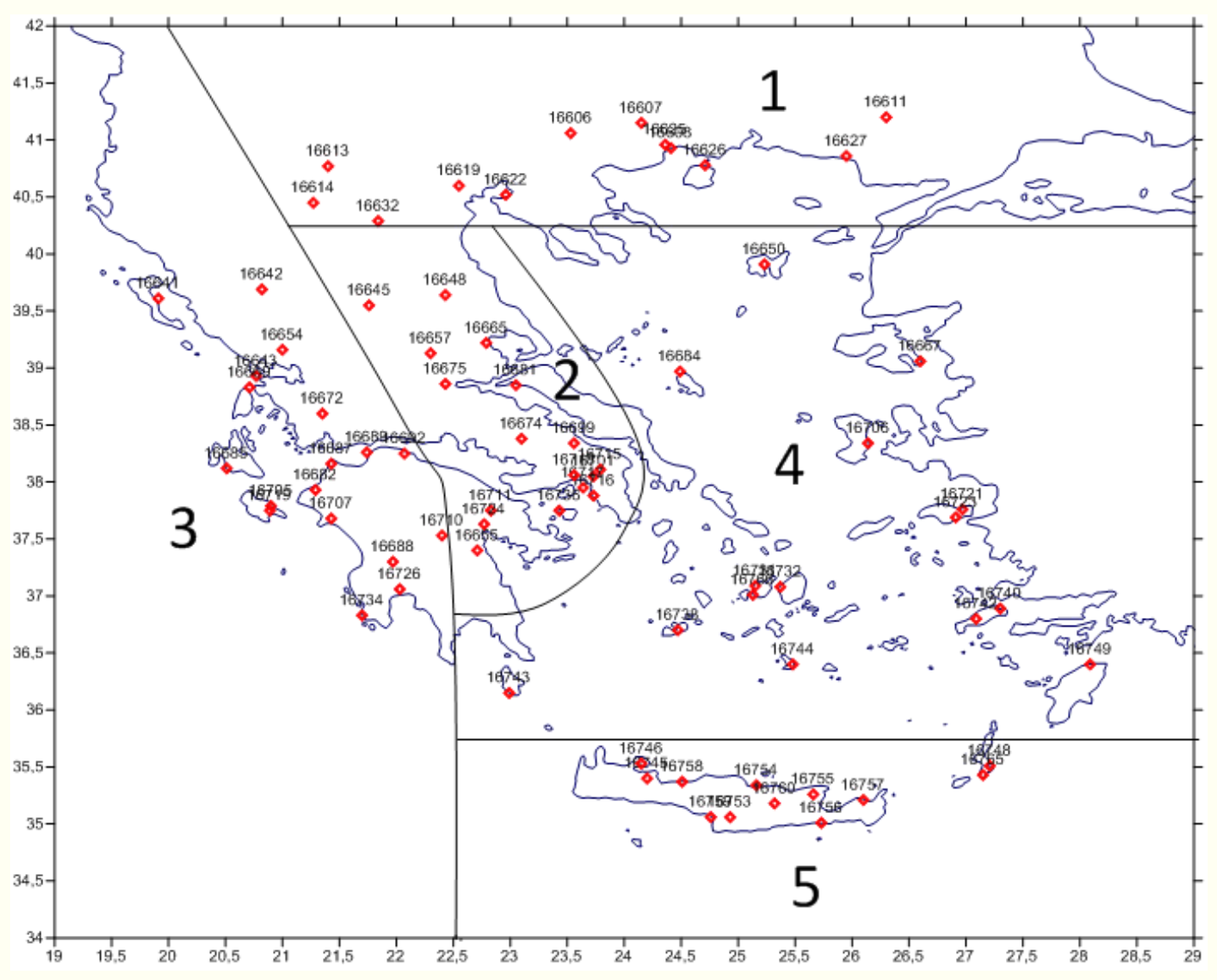
Mann Kendall test was used (Kendall, 1979)

+ stat.significance at 90%

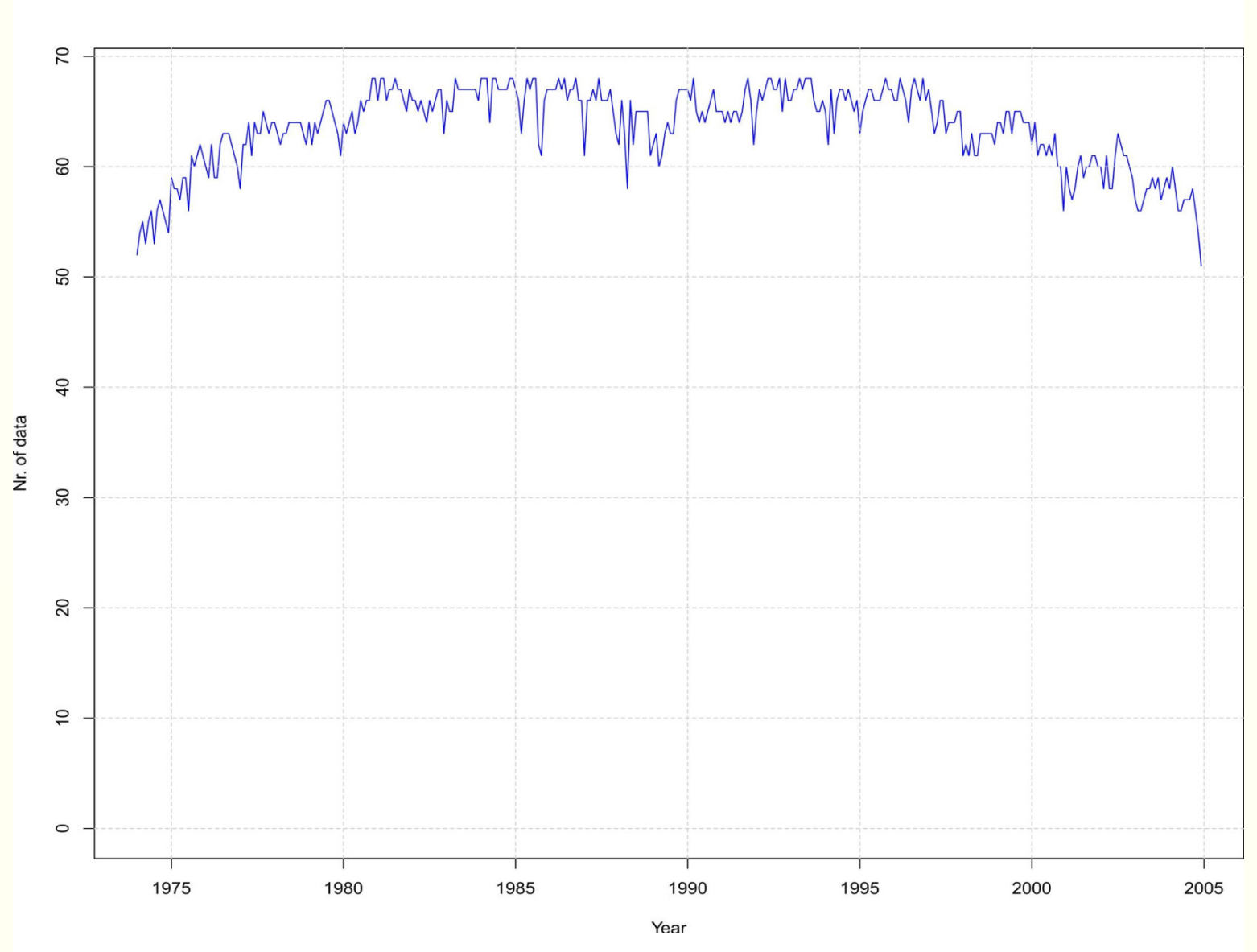
- Negative trends of no significance are prevailing
- After homogenization, trend (October-March) in Region 2 and trend (April-September) in Region 4 are positive.
- In general, homogenization reduced trend size with exception Region 3 (April-September) where trend size increased .

Data

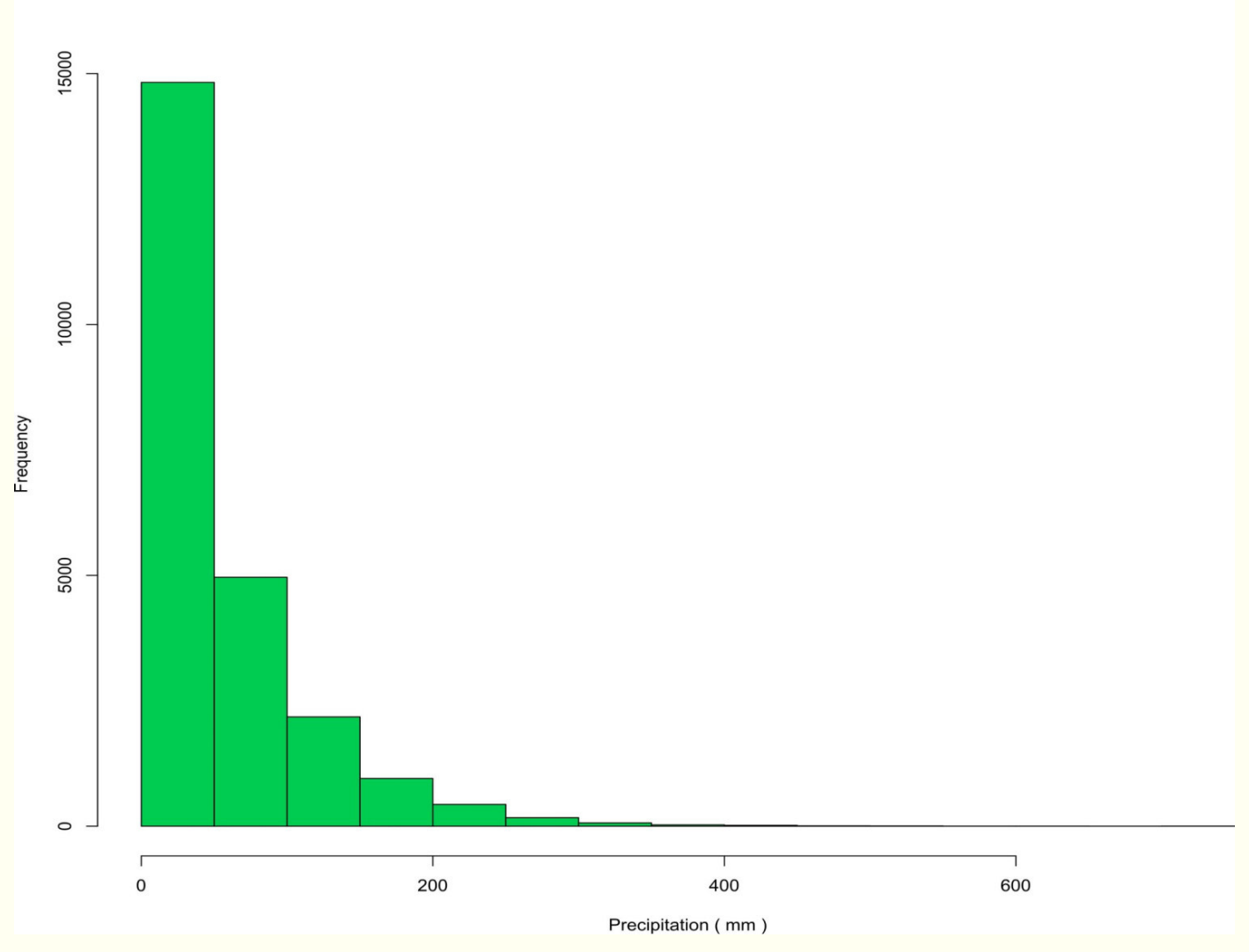
- 74 weather stations belonging at the Hellenic National Meteorological Service HNMS
- 6 series consolidated due to station relocation. Thus 68 precipitation series were examined.
- Examined period 1974-2004.
- 5 climatic zones.



Number of available data series per year

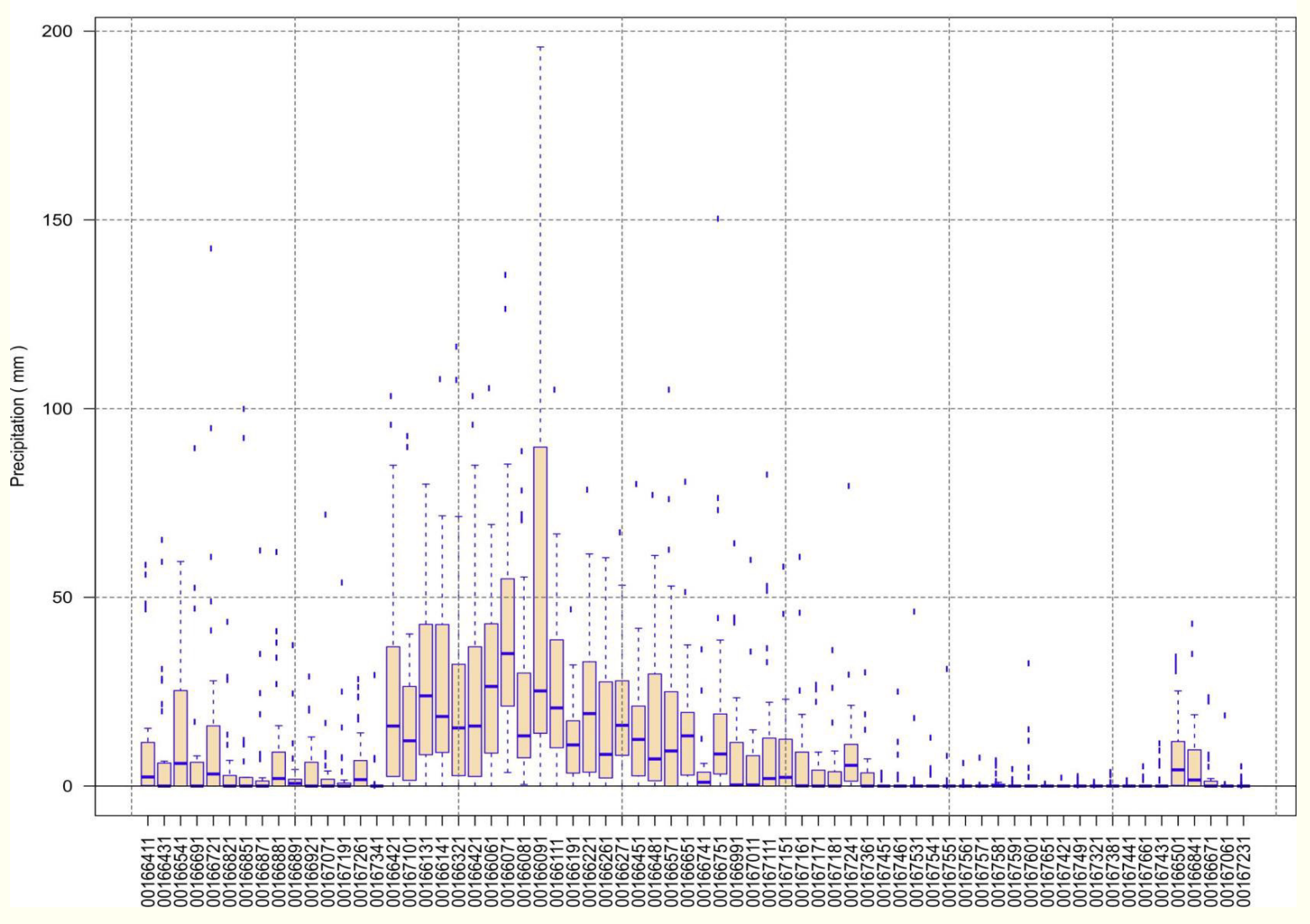


Histogram of frequency of precipitation (mm)



Too many zero values

Box plot of precipitation values per station in July



Most of the median values are 0.0

References

Caussinus H, Mestre O. 2004. Detection and correction of artificial shifts in climate series. Journal of the Royal Statistical Society Series, C - Applied Statistics 53: 405–425. DOI: 10.1111/j.1467-9876.
Conrad V. 1925. Homogenitatsbestimmung meteorologischer Beobachtungsreihen. Meteorologische Zeitschrift, 482–485.
Domonkos P. 2011. Adapted Caussinus-Mestre Algorithm for Networks of Temperature series (ACMANT). International Journal of Geosciences, 2: 293-309, DOI: 10.4236/ijg.2011.23032.
Guijarro JA. 2011. User's guide to climatol. An R contributed package for homogenization of climatological series. Report State Meteorological Agency. Balearic Islands Office Spain. http://www.climatol.eu/climatol-guide.pdf (Accessed 11September2015).
Heidke P. 1923. Quantitative Begriffsbestimmung homogener Temperatur- und Niederschlagsreihen, Meteorologische Zeitschrift, 114–115.
Kendall MG.1976. Rank Correlation Methods. 4th edn. Griffin. London.
Kohler MA. 1949. Double-mass analysis for testing the consistency of records and for making adjustments. Bull. Amer. Meteorol. Soc. 30: 188–189.
Mamara A, Argiriou AA, Anadranistakis M. 2013. Homogenization of mean monthly temperature time series of Greece. International Journal of Climatology, 33: 2649–2666. DOI: 10.1002/joc.3614.
Mamara A, Argiriou AA, Anadranistakis M. 2014. Detection and correction of inhomogeneities in Greek climate temperature series. International Journal of Climatology, 34: 3024–3043.
Mestre O, Domonkos P, Picard F, Auer I, Robin S, Lebarbier E, Bohm R, Aguilar E, Guijarro J, Vertachnik G, Klancar M, Dubuisson B, Stepanek P. 2013. HOMER: a homogenization software – methods and applications. Idojaras. Quarterly Journal of the Hungarian Meteorological Service, 117, 1: 47-67.
Szentimrey T. 2008. The manual of multiple analysis of series for homogenization (MASH). Hungarian Meteorological Service, Budapest, Hungary.
Venema et al. 2012. Benchmarking monthly homogenization algorithms for monthly data. Climate of the Past, 8: 89-115.

Contact

Mamara Anna email: annamamara@yahoo.gr, phone: (+30) 2109699171
Argiriou Athanassios email:athanarg@upatras.gr, phone: (+30) 2610996078
Anadranistaki Manolis email: anad@hnms.gr, phone: (+30) 2109699050

Climatological correction of Swedish daily precipitation measurements

Master Thesis

Developement and evaluation of an Automatic Precipitation Correction Algorithm intended for implementation on daily observations from the Swedish Meteorological and Hydrological Institute's (SMHI) network of precipitation gauges.

Figure Left: SMHI Gauge with Nipher wind shield. Photo SMHI
Right: Geonor Gauge with Alter wind shield. Photo SMHI

Rain Gauges

Manual Gauge
- SMHI gauge with Nipher windshield
- Long timeseries: 1960-2015
- 12h/24h observations

Automatic Gauge
- Geonor gauge with Alter windshield
- 1995-2015
- 1h observations

Problems to be addressed

- Determining wind-loss coefficients for the two gauge types.

  Annual total precipitation for the seperate sets of gauges differ. Possible reasons: Automatic gauges are often located at more exposed sites; Different windshields generate different aerodynamic force fields around the gauge rim.
- Determining Dynamic Correction Model (DCM) variables.

Required Variables	SMHI	Geonor
At-rim wind speed (m/s)	not available	not available
Precipitation Intensity (mm/h)	not available	available
Precipitation type (liquid/solid)	not available	not available
Temperature	available	available

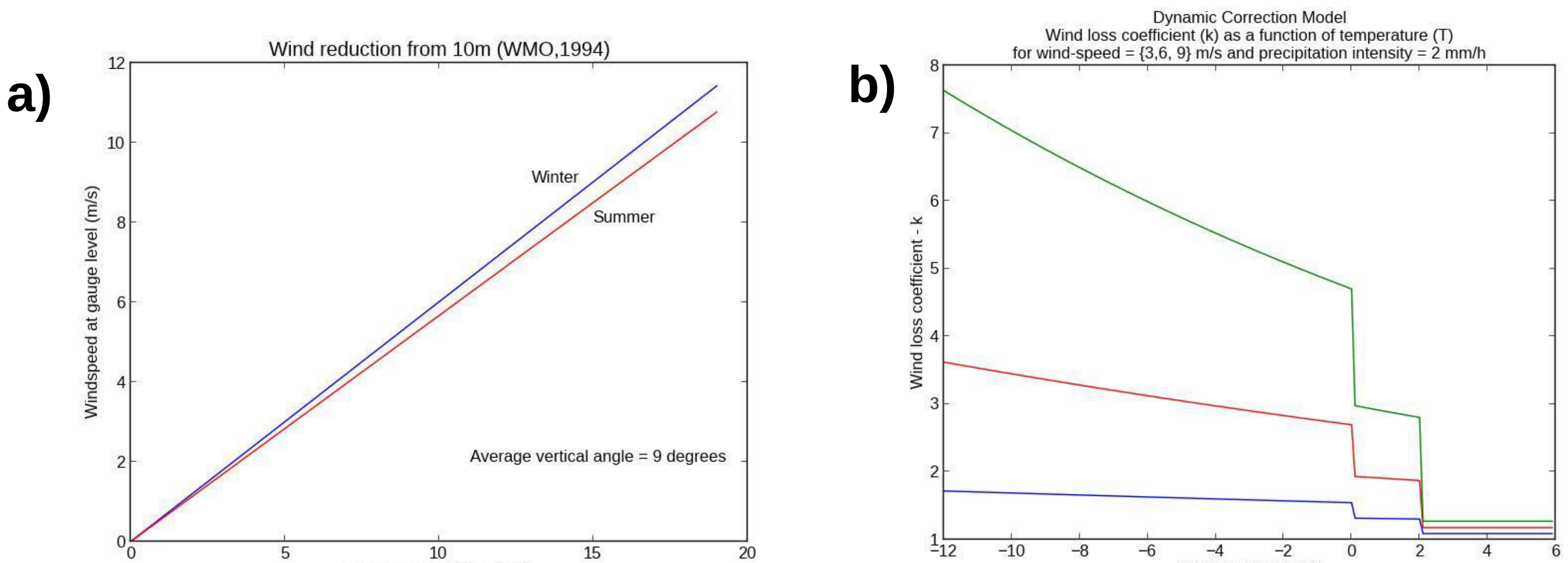


Figure Graphs showing the response of the model when calculating **a)** At-rim wind speed from wind speed at 10m and **b)** aerodynamic correction factor from temperature.

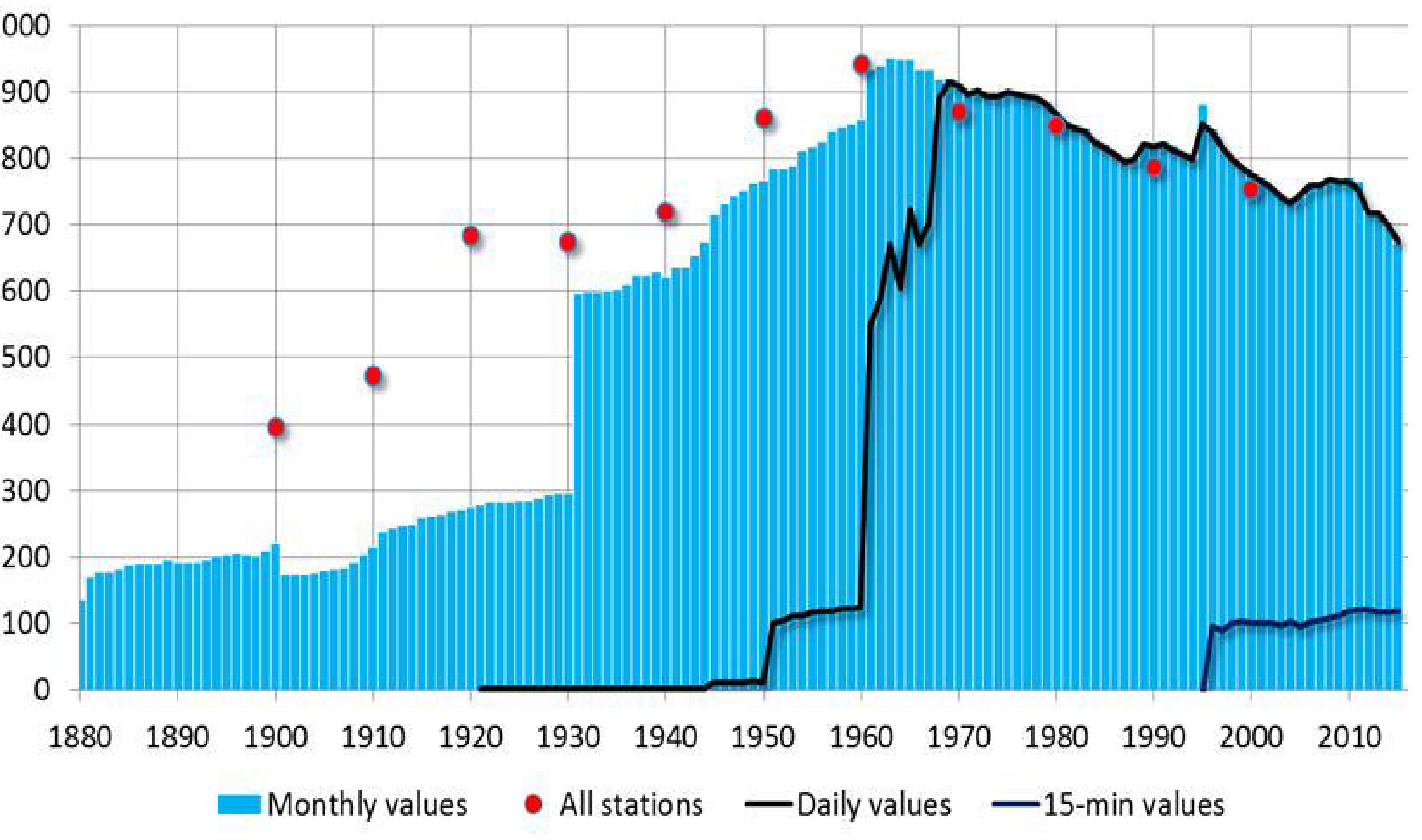


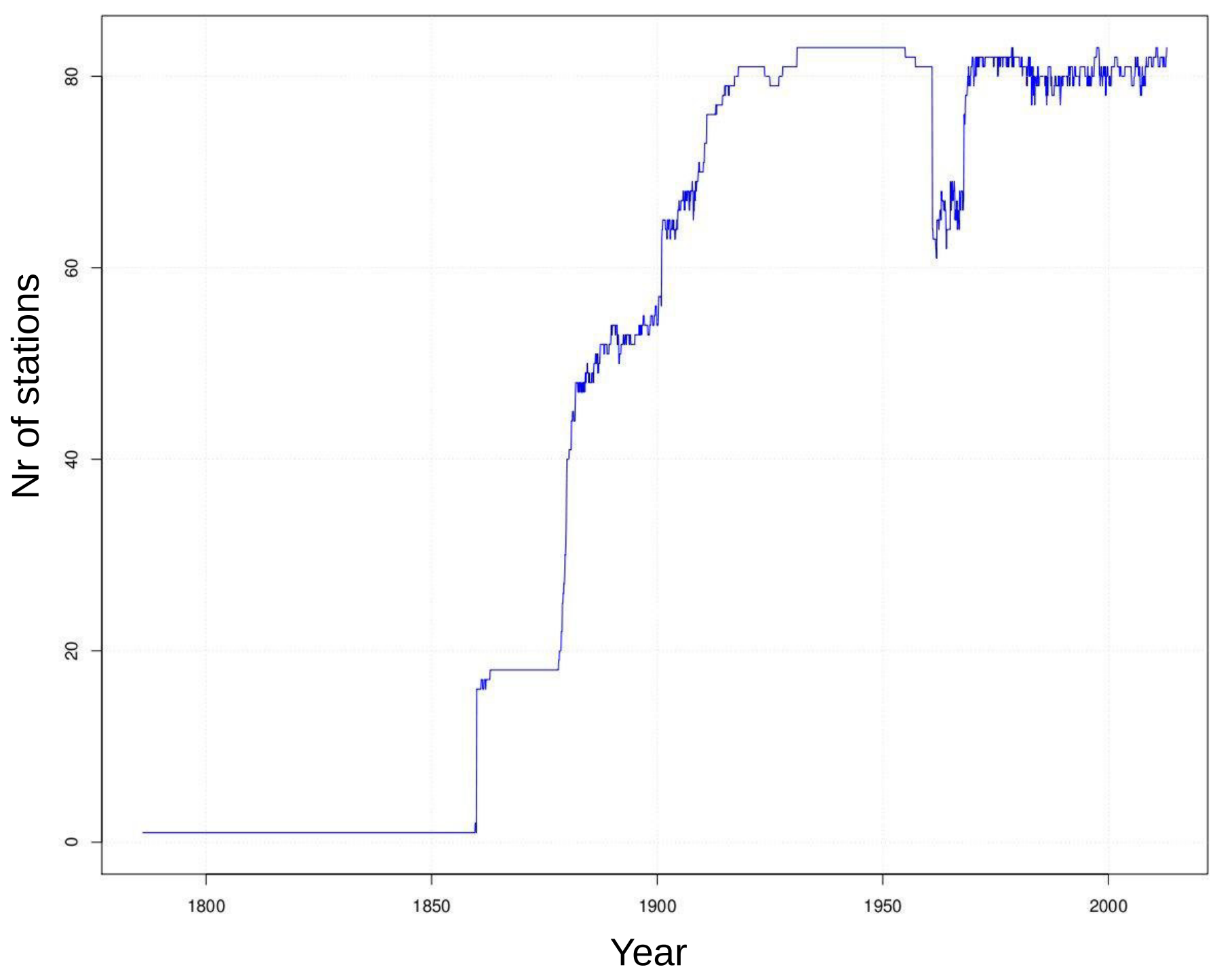
Figure Data availability. Information showing the number of SMHI stations providing monthly, daily and 15-min values of precipitation each year for the period 1880-2010. (Lennart Wern)

Pawel Harney
Stockholms University
Master's Programme in Hydrology, Hydrogeology and Water Resources
pawel.harney@gmail.com

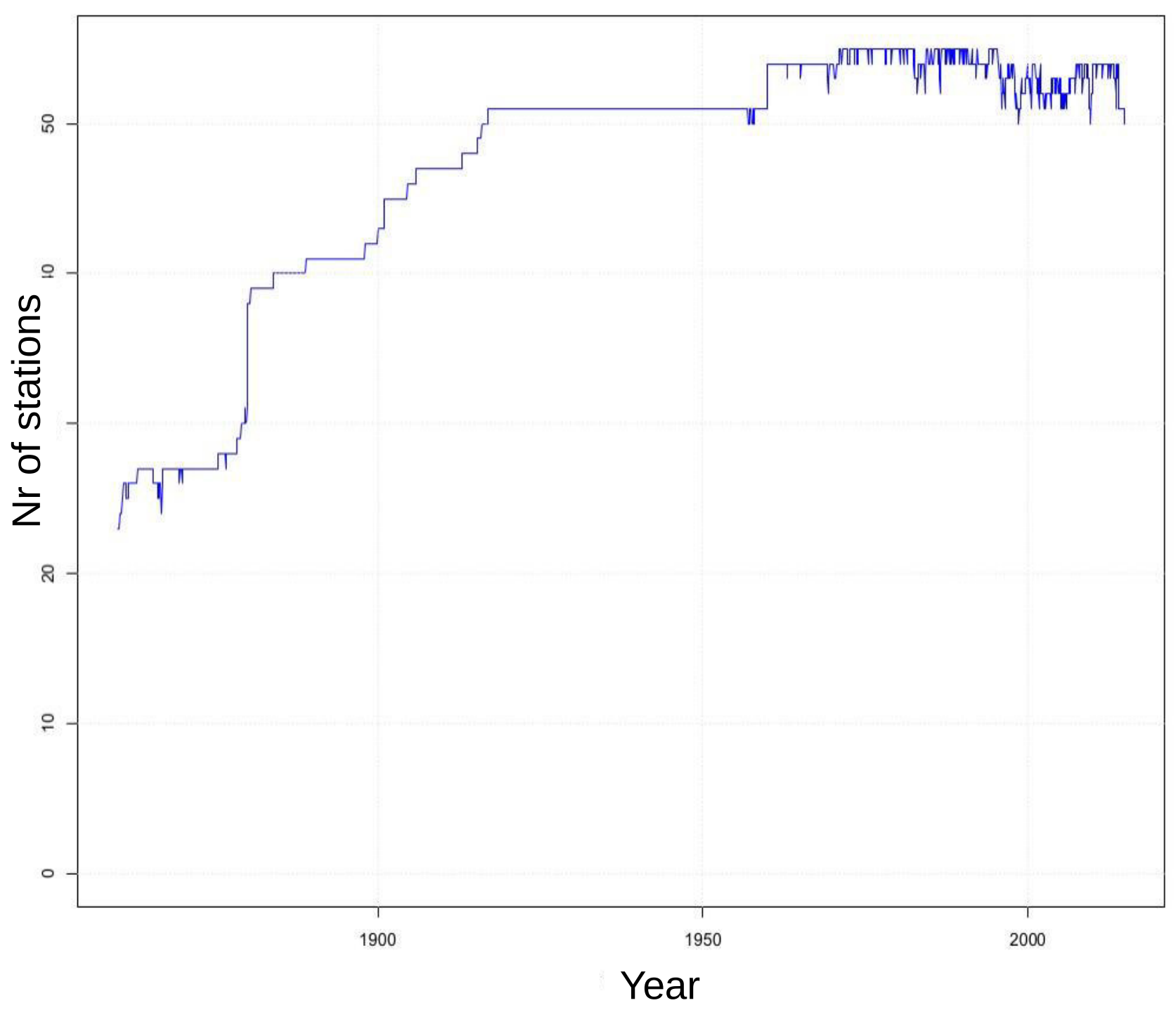
Homogenization of Swedish monthly temperature and precipitation

Swedish monthly temperature and precipitation data have been homogenized using the Homer software. Around 50 temperature and 90 precipitation stations have been used with data from 1860 to 2014.

Both the automatic and interactive mode of the Homer software have been tested. The results are now under evaluation.

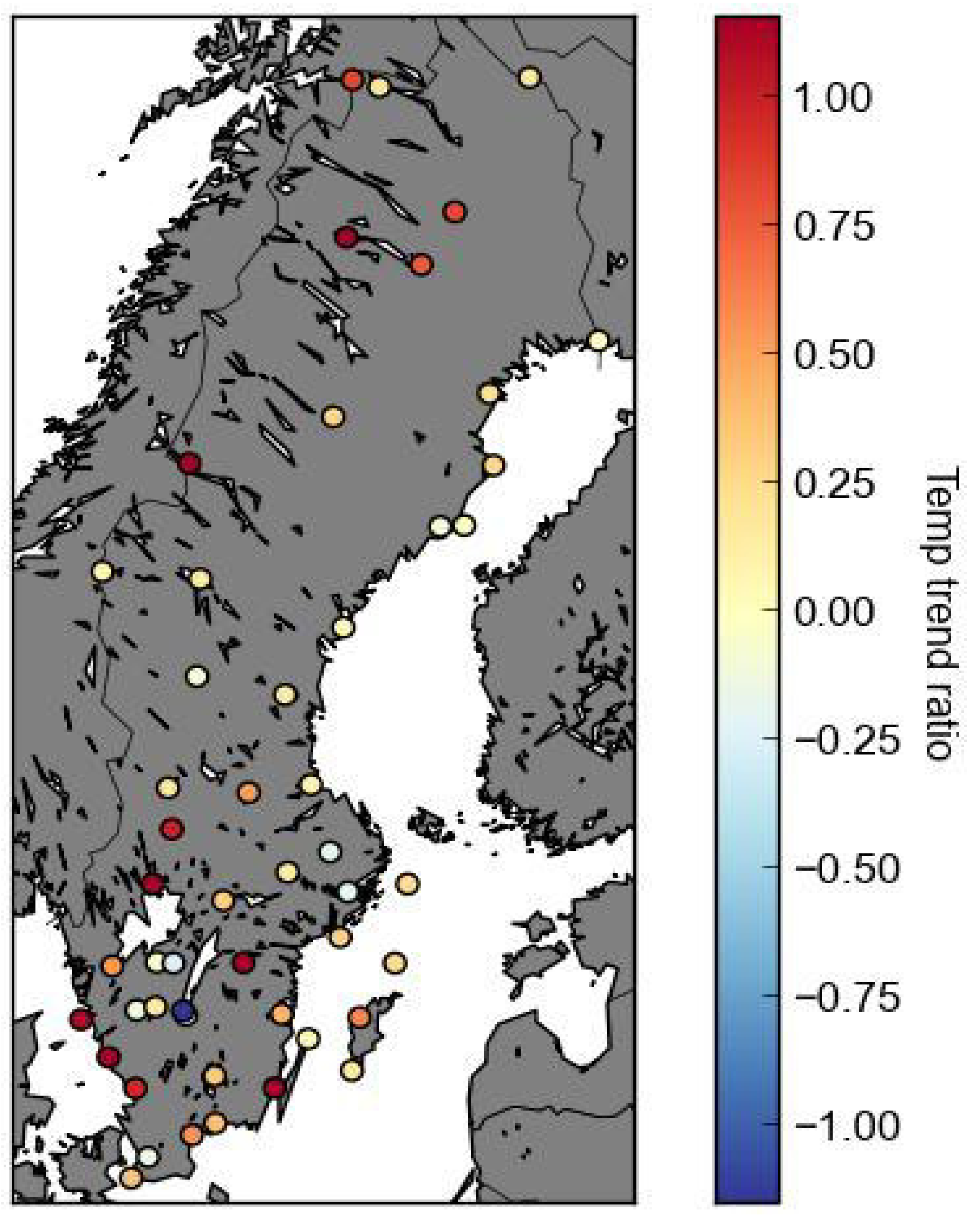


Precipitation station network

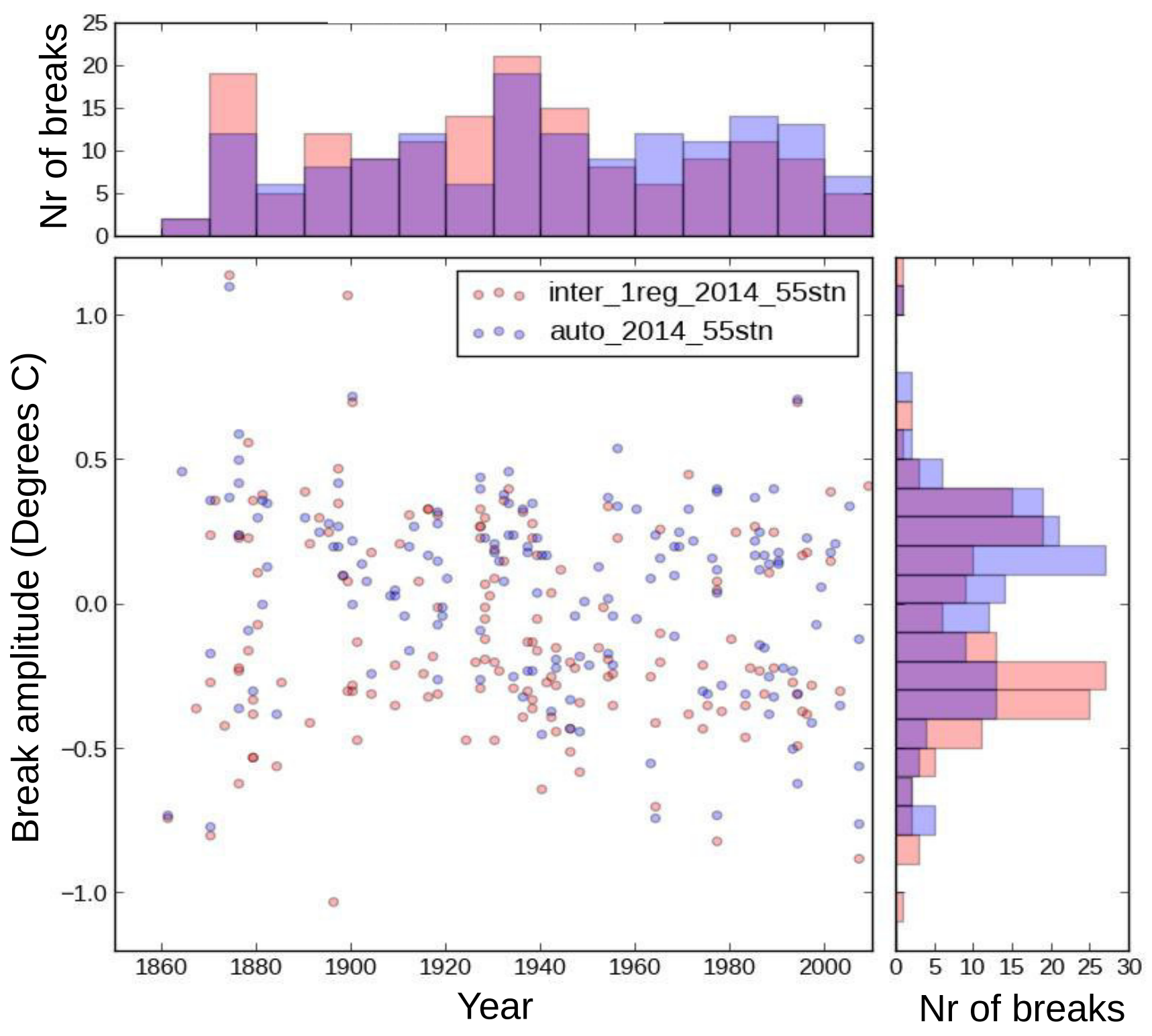


Temperature station network

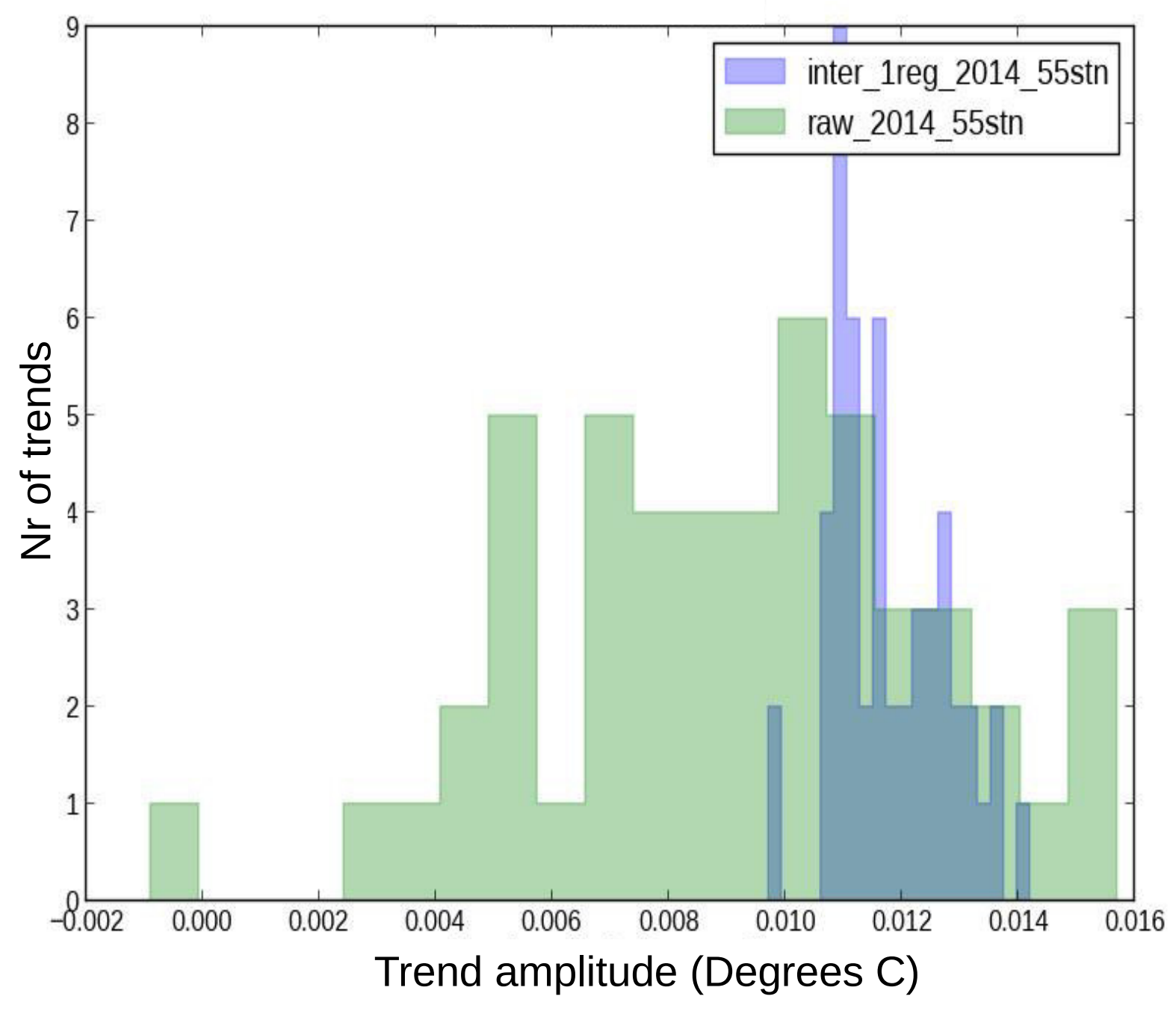
Precipitation station network



Homogenization effect on temperature trend per station ((Ho-Ra)/Ra)



Histogram of 55 temperature stations with time (year) and amplitude (°C) for detected breaks. Red color for interactive mode and blue color for automatic mode in Homer.



Histogram of 55 temperature trends (°C / year) for each station timeseries. Purple color for interactive homogenization with Homer. Green color for temperature trend in raw data.

Outlook

- Use more meta data
- Homogenize daily temperature data

Ph. Dr. Erik Engström
Swedish Meteorological and Hydrological Institute
erik.engstrom@smhi.se

Homogenization and analysis of an expanded long-term monthly rainfall network for the Island of Ireland (1850-2010)

S. Noone[1], C. Murphy[1], J.Coll[1], T.Matthews[2], D.Mullan[3], R.L.Wilby[4], S. Walsh[5]

[1]Irish Climate Analysis and Research UnitS (ICARUS), Department of Geography, National University of Ireland Maynooth, Maynooth, Ireland.
[2]School of Natural Sciences and Psychology, Liverpool John Moores University, Liverpool L3 3AF, England, UK.
[3]School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, Belfast BT7 1NN, Northern Ireland, UK.
[4]Department of Geography, Loughborough University, Loughborough LE11 3TU, England, UK
[5]Climatology and Observations, Met Éireann, Glasnevin Hill, Glasnevin, Dublin 9, Ireland.

1. Introduction

Long precipitation series help contextualise recent climate variability, identify emerging trends, ground-truth climate model projections and understand impacts on sectors such as agriculture, water resources and flood management (e.g. Jones *et al.,* 2006; Wilby and Quinn, 2013).

High-quality observations prior to 1900 are relatively rare. In Ireland most precipitation analyses have been restricted to relatively short records (post 1940s) (eg. Kiely, 1999; Sheridan, 2001; Wang *et al.,* 2006). Few studies have used long records for Ireland findings show increases in winter precipitation and decreases in summer over the period (e.g. McElwain and Sweeney, 2007; Jones and Conway, 1997).

2. Aims and Objectives

- To construct a temporally homogenized, long-term Island of Ireland Precipitation (IIP)
- Expand the existing catalogue of long-term monthly rainfall stations available in Ireland by recovering data for an additional eight stations.

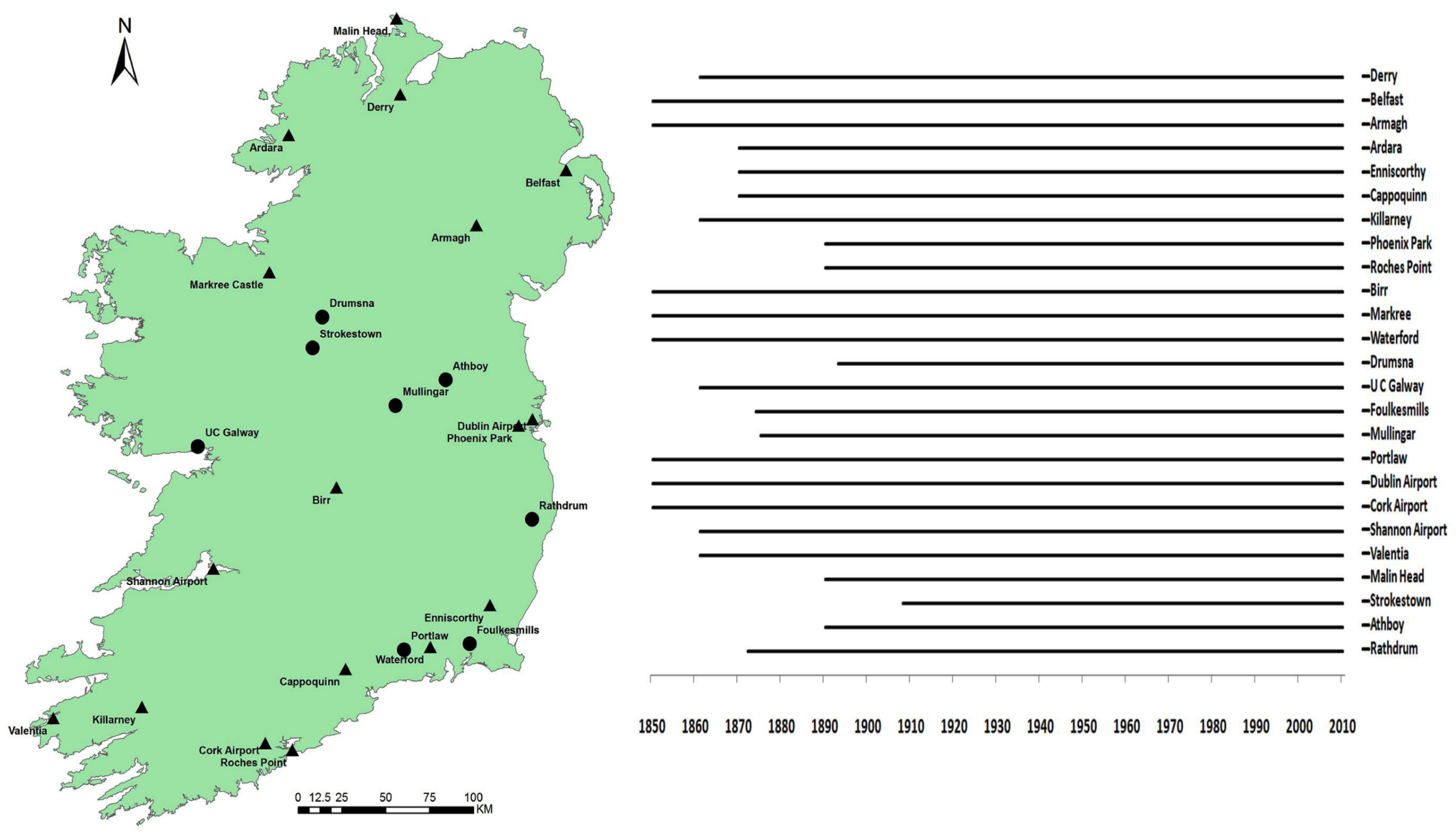
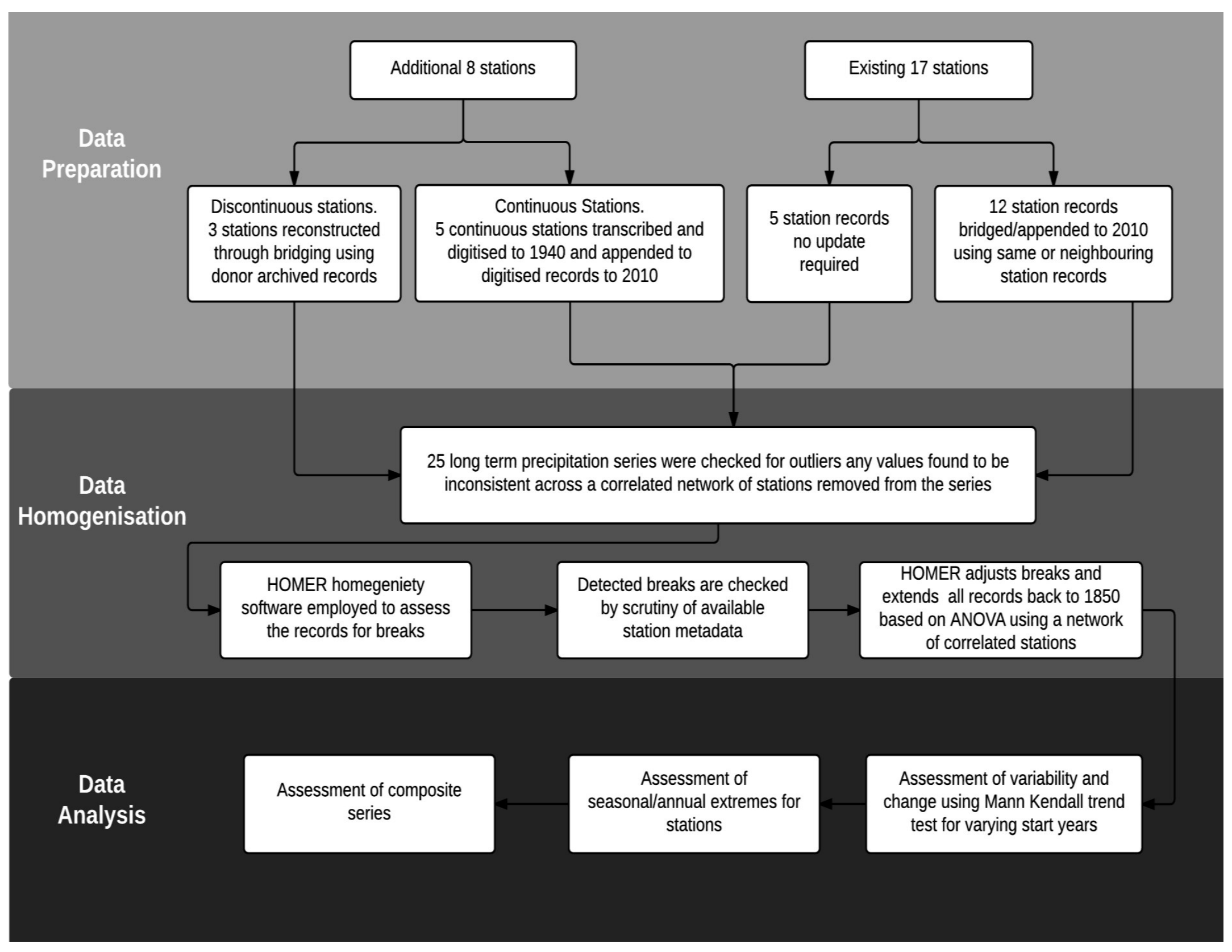


Figure 1 Location and updated record length for 25 stations used in the analysis. Triangles represent existing 17 long term stations; circles represent additional 8 stations

- Homogenize the expanded catalogue of 25 stations using the software and approach of the HOME COST action.
- Use the expanded network to extend and update all stations to a common period of 1850-2010.
- Assess variability and change within this expanded, extended and quality assured network.

3. Methods


Data Preparation

Additional 8 stations

Existing 17 stations

Discontinuous stations. 3 stations reconstructed through bridging using donor archived records

Continuous Stations. 5 continuous stations transcribed and digitised to 1940 and appended to digitised records to 2010

5 station records no update required

12 station records bridged/appended to 2010 using same or neighbouring station records

Data Homogenisation

25 long term precipitation series were checked for outliers any values found to be inconsistent across a correlated network of stations removed from the series

HOMER homegeniety software employed to assess the records for breaks

Detected breaks are checked by scrutiny of available station metadata

HOMER adjusts breaks and extends all records back to 1850 based on ANOVA using a network of correlated stations

Data Analysis

Assessment of variability and change using Mann Kendall trend test for varying start years

Assessment of seasonal/annual extremes for stations

Assessment of composite series

Figure 2 Workflow stages and key steps

- Existing 17 long-term stations were updated to 2010.
- Additional 8 long-term stations were transcribed/digitized from Met Éireann hardcopy archives, 3 discontinuous stations needed bridging/infilling from donor stations.
- Compiled metadata, station notes and methodology
- Outliers were identified and likely cases removed.
- The HOMER (HOMogenization softwarE in R) package interactive semi-automatic method was applied to the data.
- Detected breaks assessed with scrutiny of available metadata.
- Sensitivity of break detection to network density examined over 1941-2010 network of 211 stations.
- Pettitt 's statistic was used to examine HOMER homogenised annual IIP station series for any remaining change points.
- Variability/change was assessed for all 25 homogenized IIP stations.
- Driest/wettest years identified and Non-parametric Mann-Kendall (MK) test was used to detect monotonic trends in seasonal / annual mean series.
- See Figure 2 for workflow and key steps in research.

4. Homogenization Results

- 53 inconsistent monthly values across 16 stations identified and removed.
- 25 breaks detected by HOMER across 14 stations.
- Metadata scrutiny revealed 20 of the breaks are coincident with issues such as changes in gauge size and position, stations closures and moves, previous bridging/infilling.

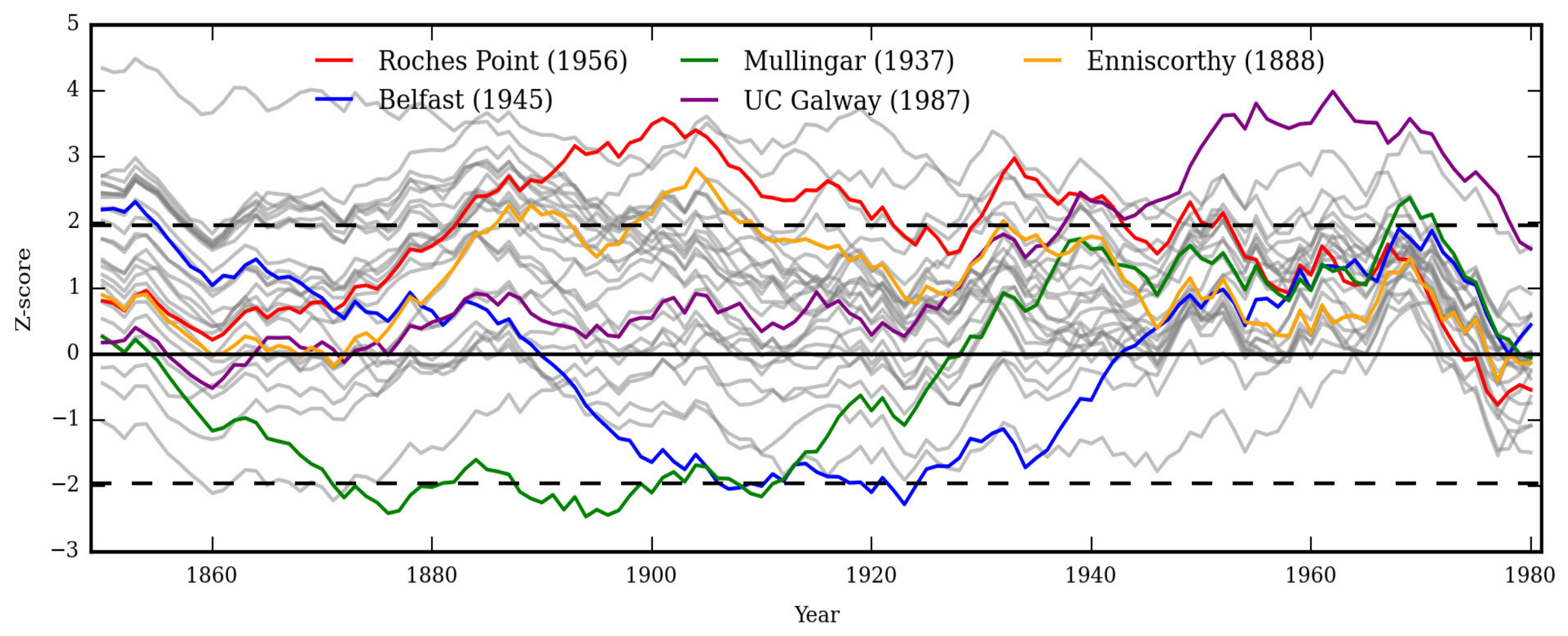


Figure 3 Persistence of trends for stations with breaks unconfirmed by metadata relative to the persistence of trend for fully homogenised records for all 25 stations (grey lines and including 11 stations with no breaks detected). Dotted horizontal lines represent critical values beyond which trends are significant at 0.05 level.

- Five unexplained breaks were further investigated using MK test to examine the persistence of trends for uncorrected series relative to fully homogenised network.
- Large deviations in trend persistence around identified break points, therefore all breaks (confirmed and unconfirmed by metadata) were subject to adjustment (Figure 3).
- Sensitivity of break detection to network density results show good consistency of break detection frequency and timing re-running HOMER for candidate stations using a finer density network for the period 1941-2010.
- The Pettitt test revealed simultaneous break point at two stations (Ardara and Drumsna) in 1933 within same reference network, caution is flagged at these series.
- Little evidence from metadata that breaks were missed by HOMER.

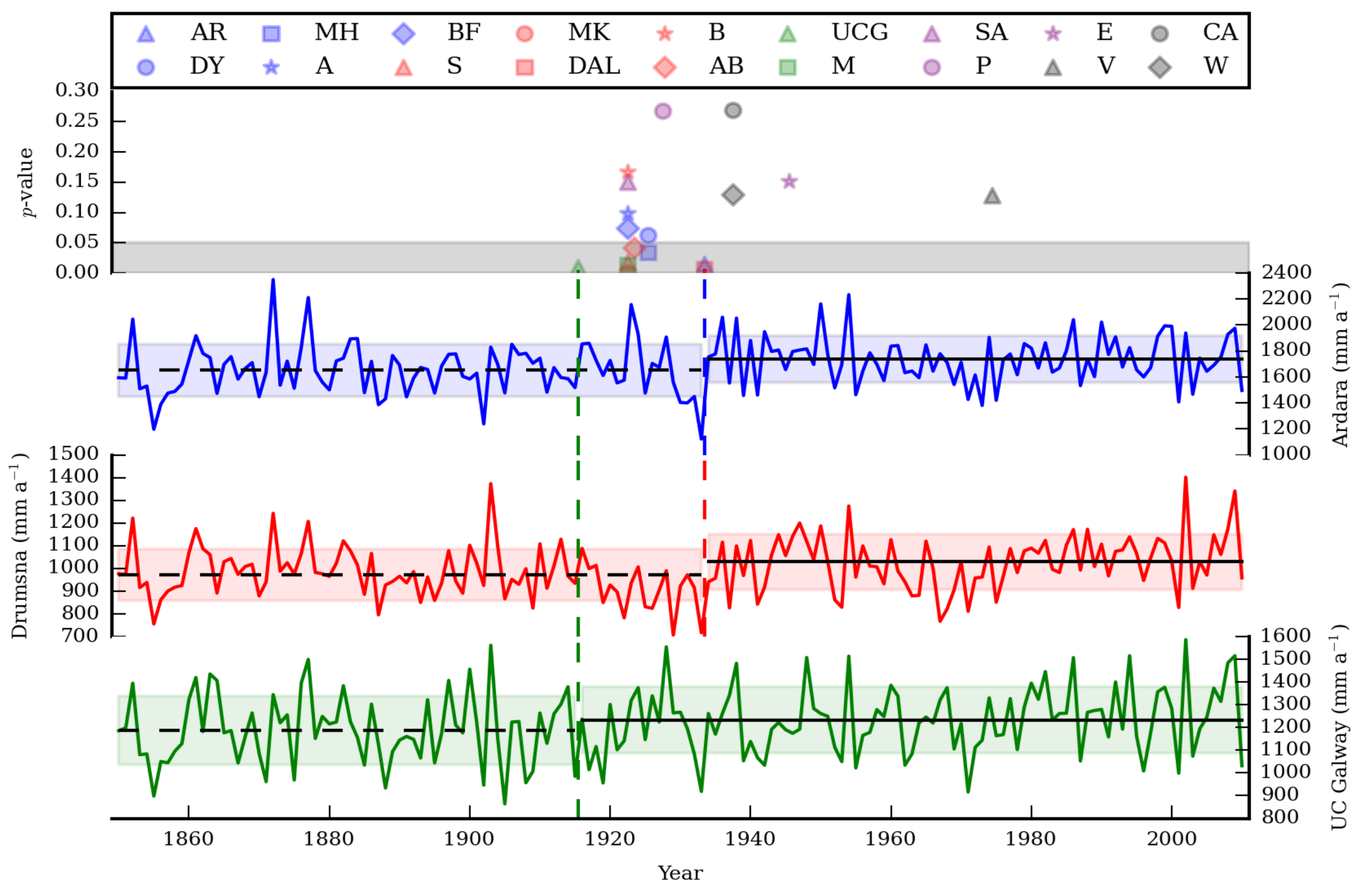


Figure 4 Top: Pettitt test for change points in mean annual precipitation. The grey shaded region indicates the region in which *p*-values are interpreted as significant (0.05 level). Time series plots (lower 3 axes) illustrate the mean annual precipitation at the stations with significant change points outside the 1920s. These change points are highlighted with dotted vertical lines; the colours of which correspond to the respective time series. Note that Ardara and Drumsna both have change points in 1933, hence the blue and red dotted lines are over-plotted. The dotted (solid) black horizontal lines illustrates mean values before (after) the identified change point; the shaded region spans ± 1 standard deviation.

5. Key Findings

- Most stations show non-significant annual positive trends.
- Most IIP stations show positive winter trends and negative summer trends significant (0.05 level)
- North -West stations have strongest positive winter trends significant (0.05 level) for tests commencing before 1890.
- South and east stations have strongest negative summer trends for tests commencing before 1880.
- Shorter records are not representative of longer-term trends.

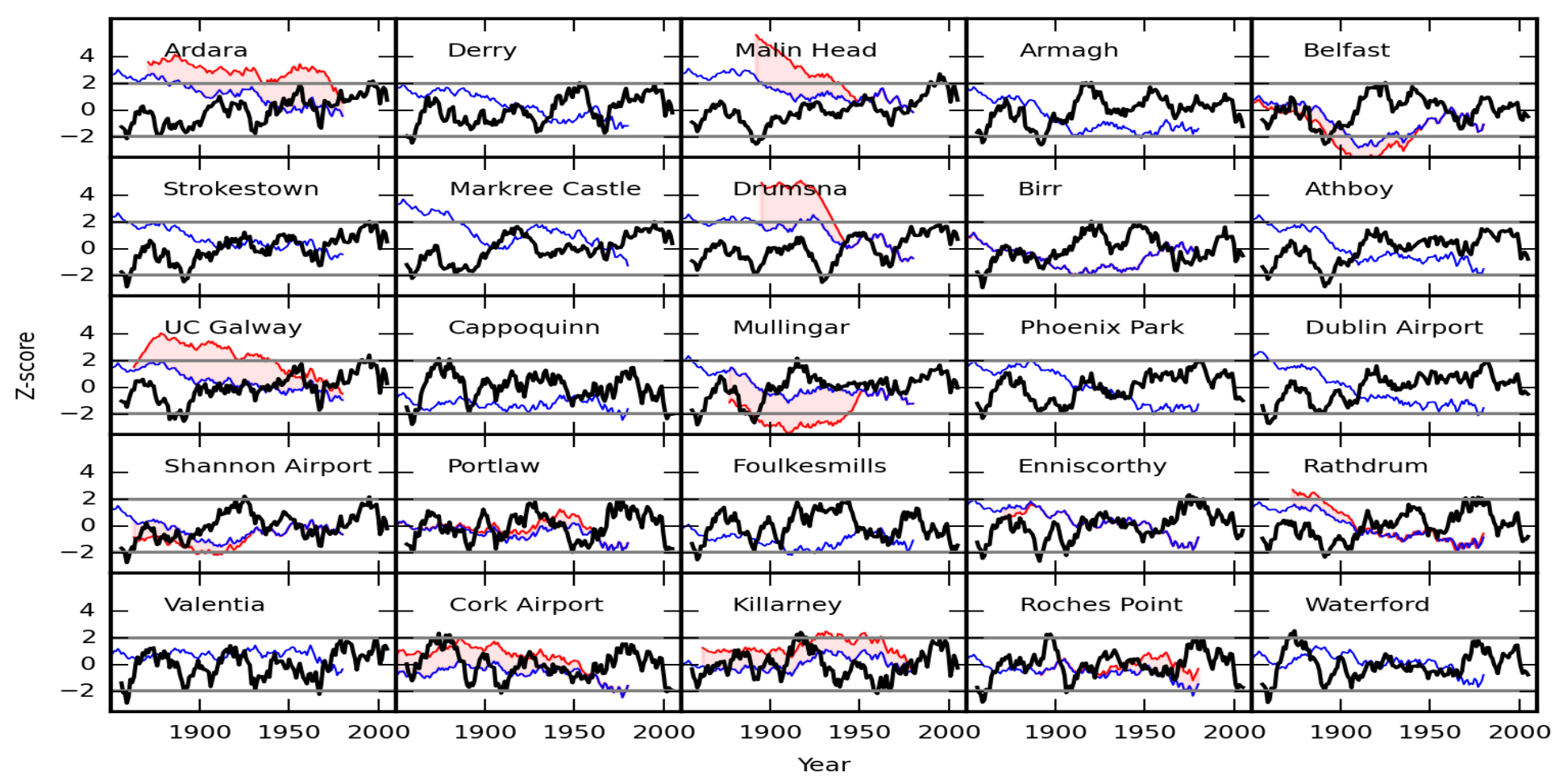


Figure 5 Homogenized winter time series for all stations smoothed with an 11-year moving average (black line). MK Z scores are shown before and after homogenization where applicable (red: unhomogenized; blue: homogenized/no breaks detected) calculated for varying start years (which are given by the x-coordinate). The grey lines indicate ±1.96; absolute values exceeding these bounds are interpreted as significant at the 0.05 level.

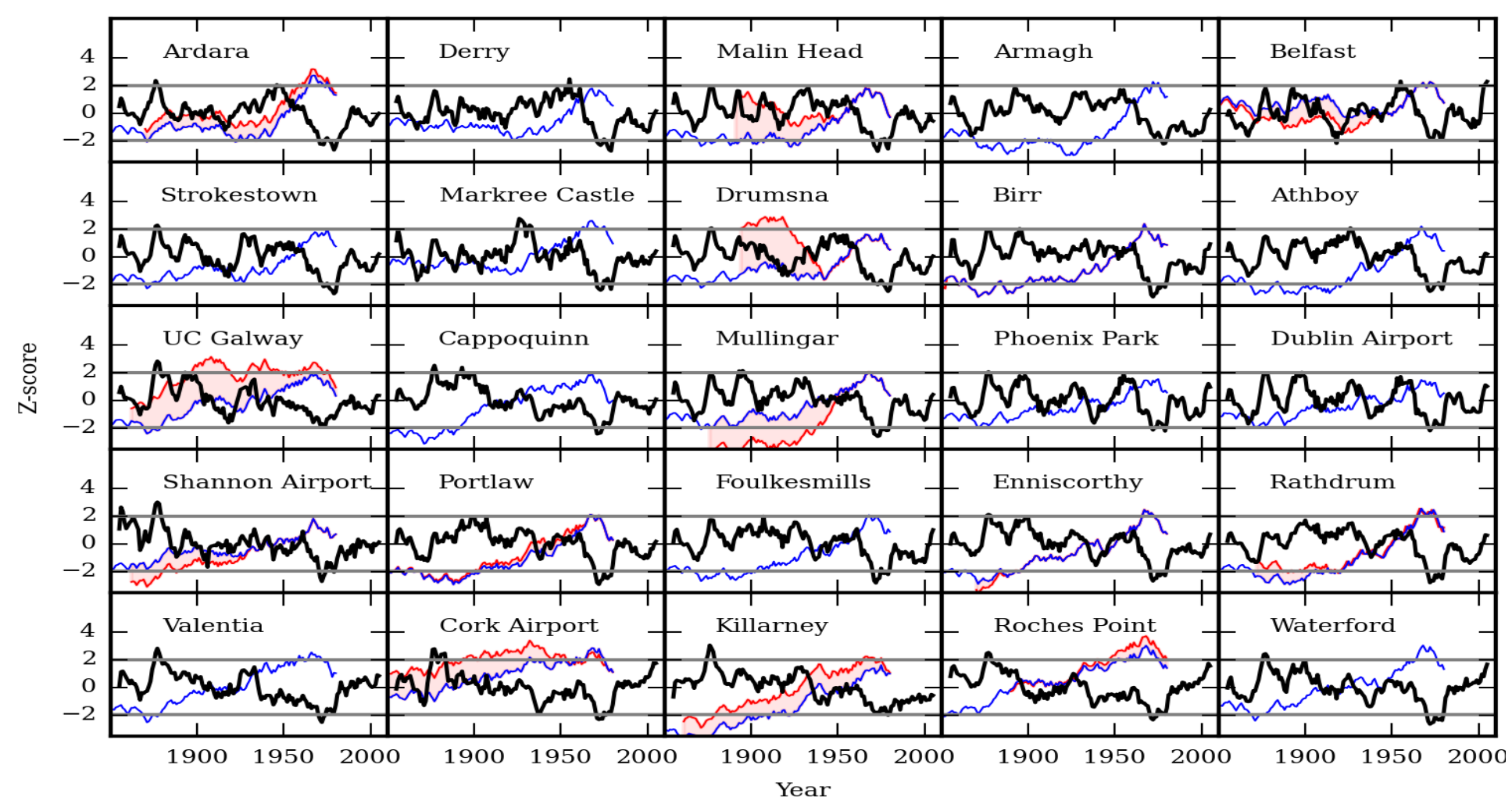


Figure 6 Same As Figure 5 but for summer

6. Acknowlegdements

We thank Met Éireann for access to archived records and provision of digitised data. Met Éireann librarian Mairéad Traynor provided help and advice with archive records. Thanks are also extended to Mary Curly for advice and support in applying HOMER and to Prof. Phil Jones for advice and insight on early CRU dataset work. Mark Bailey and John Butler at Armagh Observatory gave data and advice on Armagh records. Shaun Harrigan offered valuable feedback on previous drafts. SN was funded by the Irish Research Council. CM/TM and JC/SW acknowledge funding provided by the Irish Environmental Protection Agency under projects 2014-CCRP-MS.16 and 2012-CCRP-FS.11 respectively

7. Cited Work

Jones P, Conway D. 1997. Precipitation in the British Isles: an analysis of area-average data updated to 1995. Int. J. Climatol. 17: 427-438

Jones PD, Lister DH, Wilby RL, Kostopoulou E. 2006. Extended riverflow reconstructions for England and Wales, 1865-2002. Int. J. Climatol. 26: 219-231

Kiely G. 1999. Climate Change in Ireland from precipitation and streamflow observations. Adv. Water. Resour. 23: 141-151.

McElwain L, Sweeney J. 2007. Key Meteorological Indicators of Climate Change in Ireland, Environmental Protection Agency, Johnstown Castle Estate, Wexford, Ireland. 31 pp.

Sheridan T. 2001. Analysis of trends at some Irish rainfall stations. Met Eireann Technical Report, Dublin, Ireland.

Wilby RL and Quinn NW. 2013. Reconstructing multi-decadal variations in fluvial flood risk 660 using atmospheric circulation patterns. J. Hydrol. 487: 109-121.

Wang S, MC Grath R, Semmler T, Sweeney C. 2006. Validation of simulated precipitation patterns over Ireland for the period 1961-2000. Int. J. Climatol. 26: 251-26

Presented at the 10th EUMETNET Data Management Workshop in St. Gallen, Switzerland. 28th -30th October 2015

Contact :Simon.J.Noone@mumail.ie

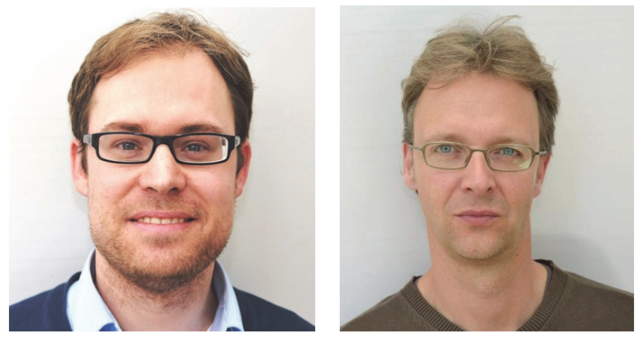
Long series of Swiss seasonal precipitation: Homogenisation, regionalisation and trends

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology
MeteoSwiss

S. C. Scherrer, M. Begert, M. Croci-Maspoli & C. Appenzeller
Federal Office of Meteorology and Climatology (MeteoSwiss), Zürich (Switzerland)



Contact
michael.begert@meteoswiss.ch

Introduction

The knowledge of precipitation trends and variability is vital for many aspects of life and socio-economic sectors. However, confidence in precipitation trends is still limited and merits regular reassessment. Here, **seasonal and annual homogenised precipitation series** in **Switzerland** are investigated for the period 1901-2013 in terms of **trends, interannual variability and the influence of large-scale European flow patterns**.

Data homogenisation

Homogenisation methodology

- Homogenised using the MeteoSwiss homogenisation software THOMAS
- Combination of methods using reference series (with large effort put into selection of homogeneous stations) and meta data analysis
- Correction done on monthly basis as several inhomogeneities (e.g. site relocations) show correction factors with strong seasonal dependencies

Causes for inhomogeneities

cause	frequency
site relocation	41%
unknown	17%
automation	12%
change of instrumentation	12%
inspection/calibration/maintenance	10%
observer change	6%
change of environment	2%

Tab. 1: Causes for inhomogeneities in the 32 precipitation series used for regional trend analysis

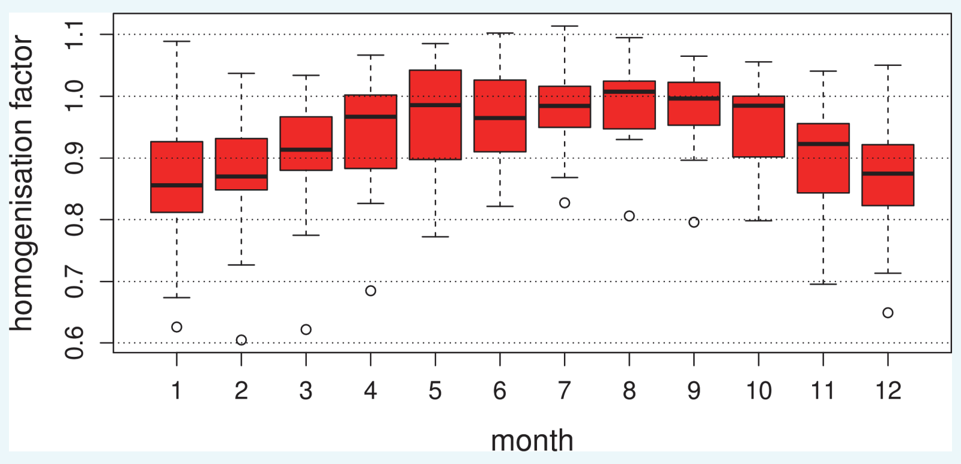


Fig. 1: Monthly homogenisation factors used to correct for automation (transition from Hellmann to tipping bucket)

In total, 87 breaks (e.g. on average 2.7 breaks per series) have been corrected. 90% of the homogenisation factors applied to monthly precipitation sums were between 0.8 and 1.2, in a few cases (especially for mountain stations), factors between 0.6 and 1.6 had to be applied.

Regionalisation

Clustering methodology

- agglomerative hierarchical clustering with complete linkage (DeGaetano 2001)
- dissimilarity measure: Spearman rank correlation
- applied to monthly precipitation anomalies (305 stations, 1961-2006)

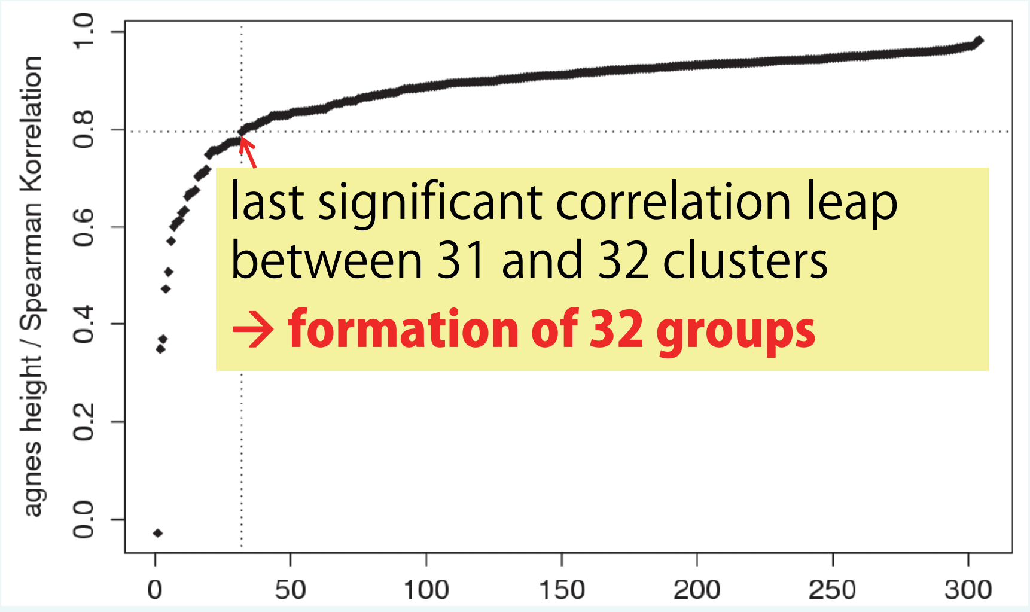


Fig. 2: Dendrogram heights and lowest correlation within the groups, respectively, against the number of clusters in the hierarchical clustering of temperature series. Dashed lines indicate the „optimal" number of clusters and the according correlation.

Map of 32 precipitation clusters

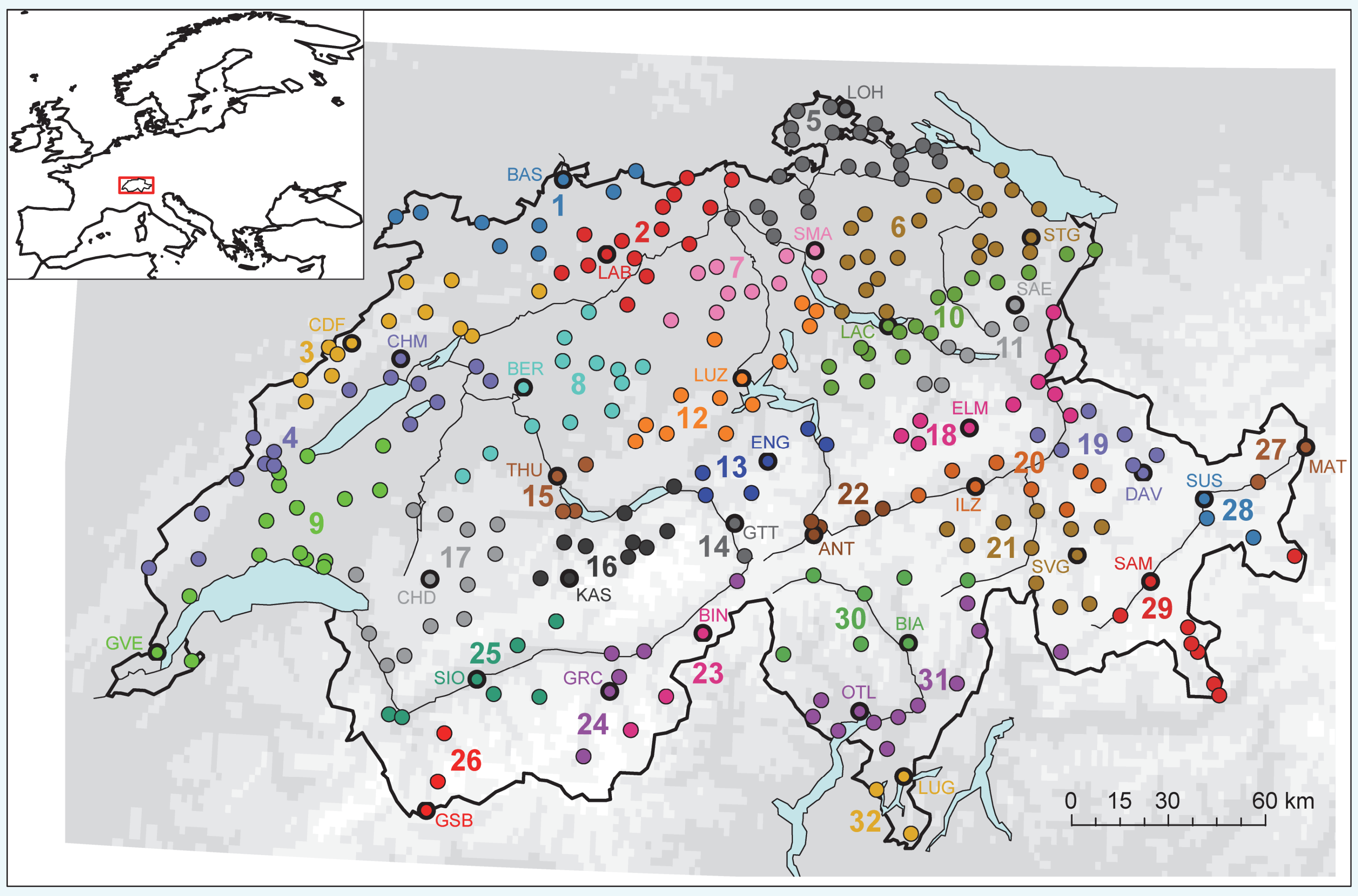


Fig. 3: Map of with location of 305 MeteoSwiss precipitation stations in the colour of the 32 numbered regions as determined with an agglomerative hierarchical clustering algorithm. The 32 stations analysed (one for each region) are marked with bold circles and labelled with a three character code. The grey shading shows basic features of topography. The inset on the top left corner shows the location of Switzerland in Europe.

Trend analysis

Annual trends

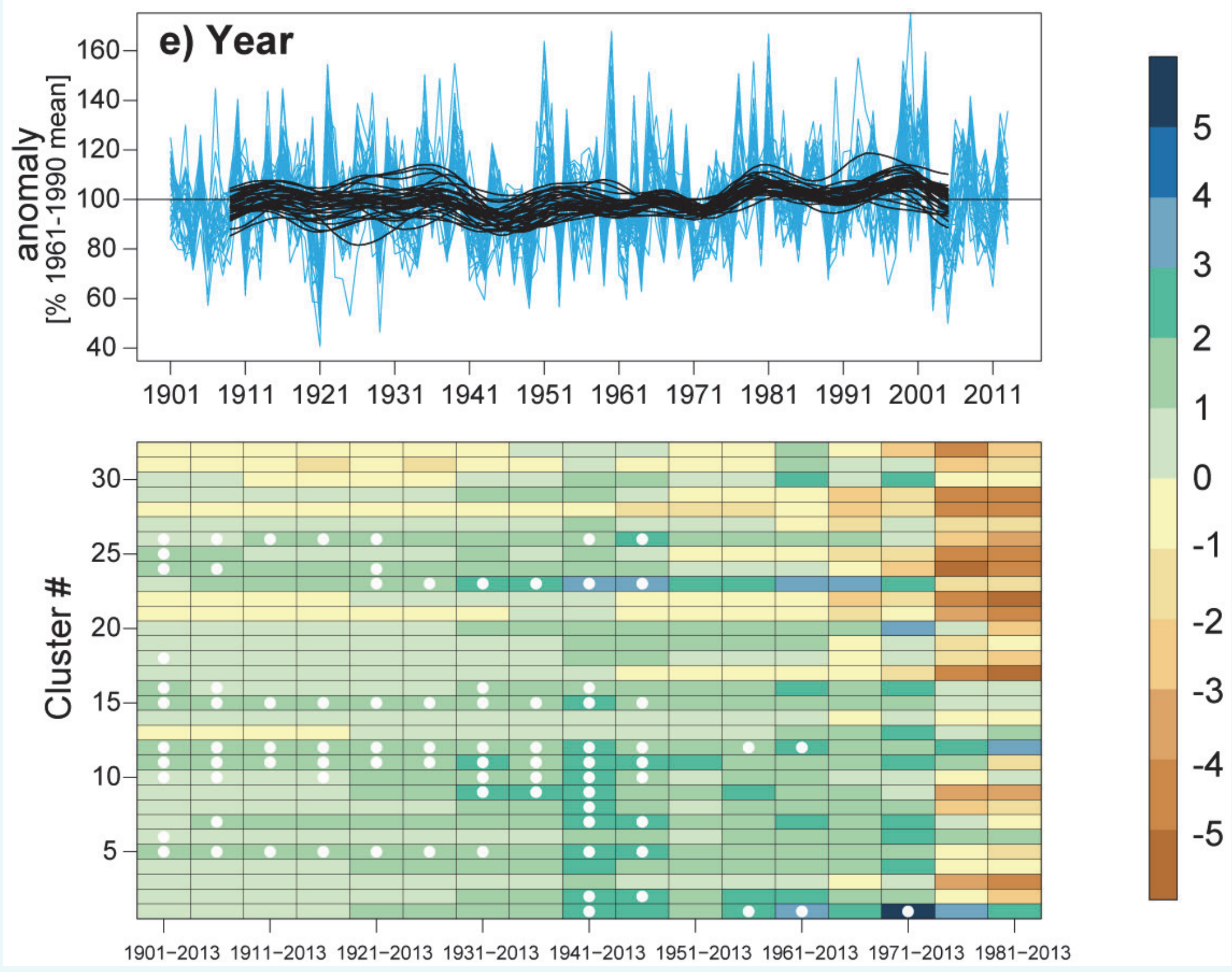


Fig. 4: Time series and time-moving trends for all 32 regions and annual precipitation. Upper panel: precipitation anomalies in % from the 1961-1990 mean (blue lines) and the low pass smoothed anomalies (black lines). Lower panel: linear trends for different time periods. Negative trends: brownish colours, positive trends: green/blue. Significant trends (p<0.05) marked with a white dot.

Winter- and summer trends

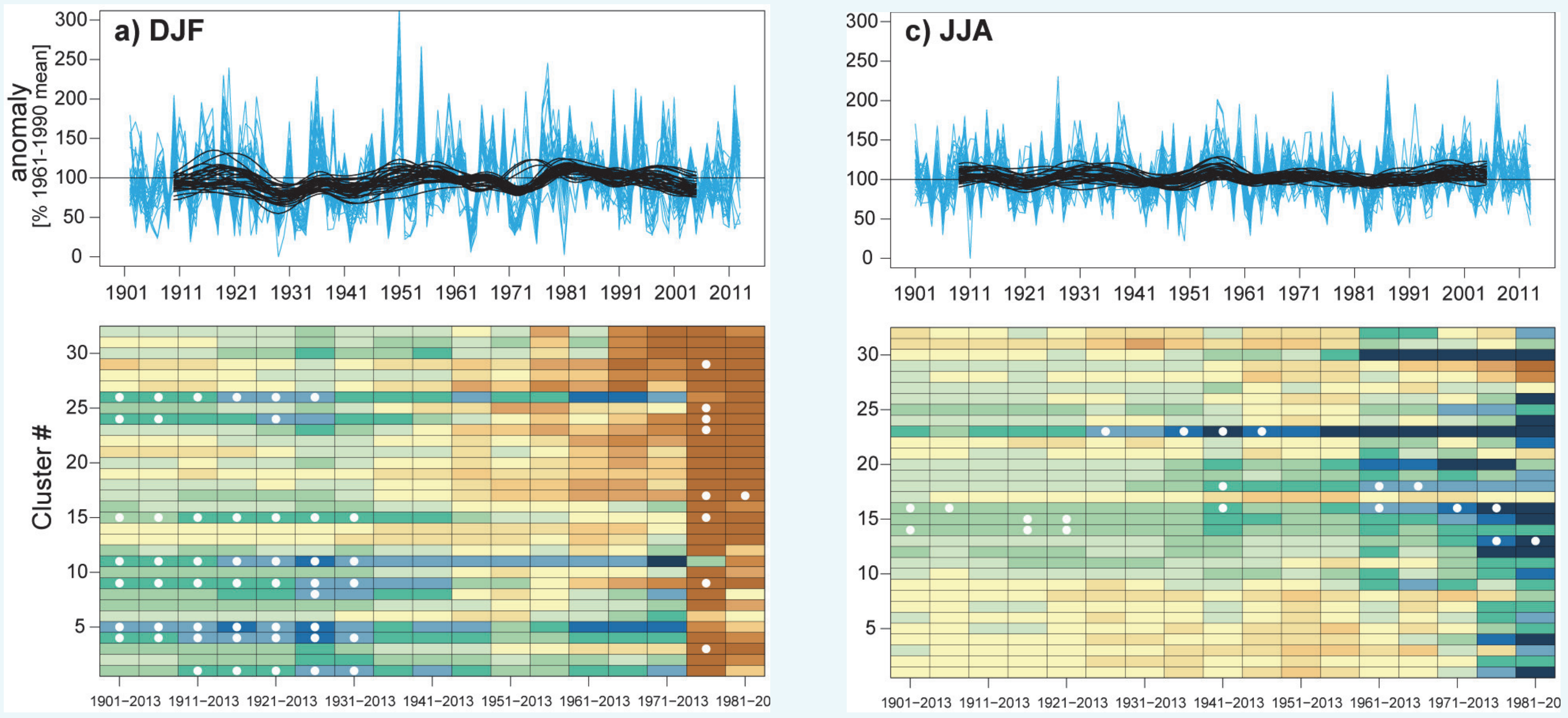


Fig. 5: As Fig. 3 but for winter (DJF, left panel) and summer (JJA, right panel).

Interaction with large scale flow

Climate patterns and Swiss precipitation variability

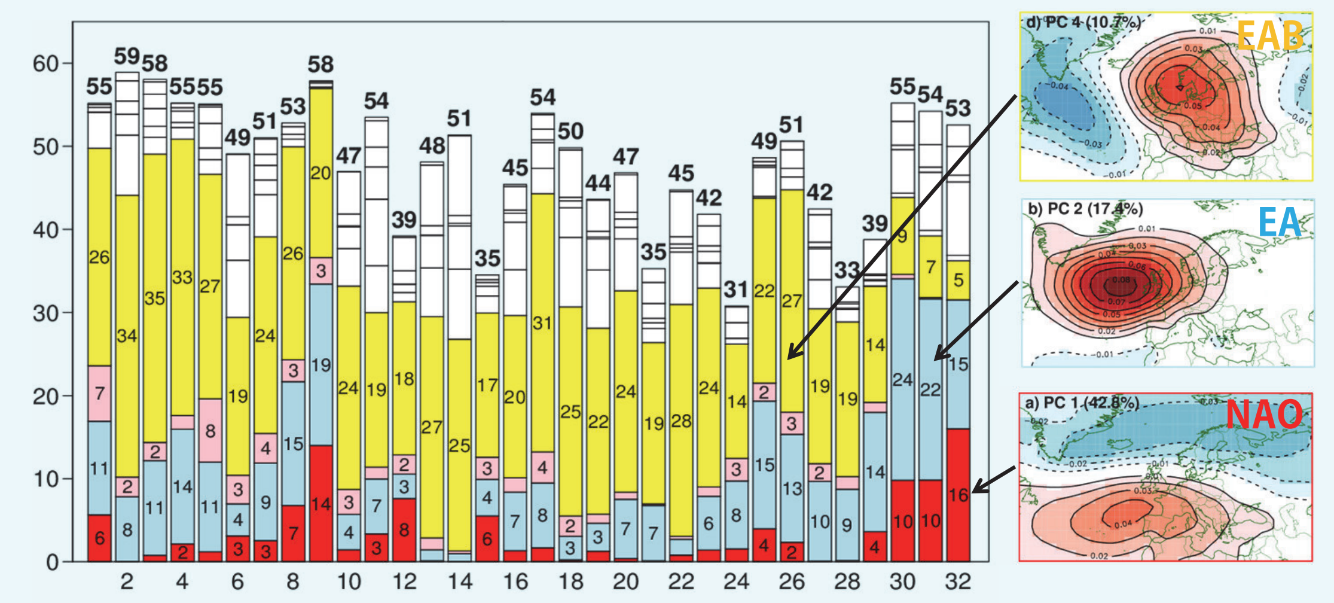


Fig. 6: The variance of the 32 precipitation regions explained by the first 10 sea level pressure patterns for DJF in the 1901-2012 period. Contributions from PC1 (NAO) are shown in red, PC2 (EA) in blue and PC4 (EAB) in yellow. The SLP patterns are shown on the right.

Dry and wet season composites of sea level pressure / wind

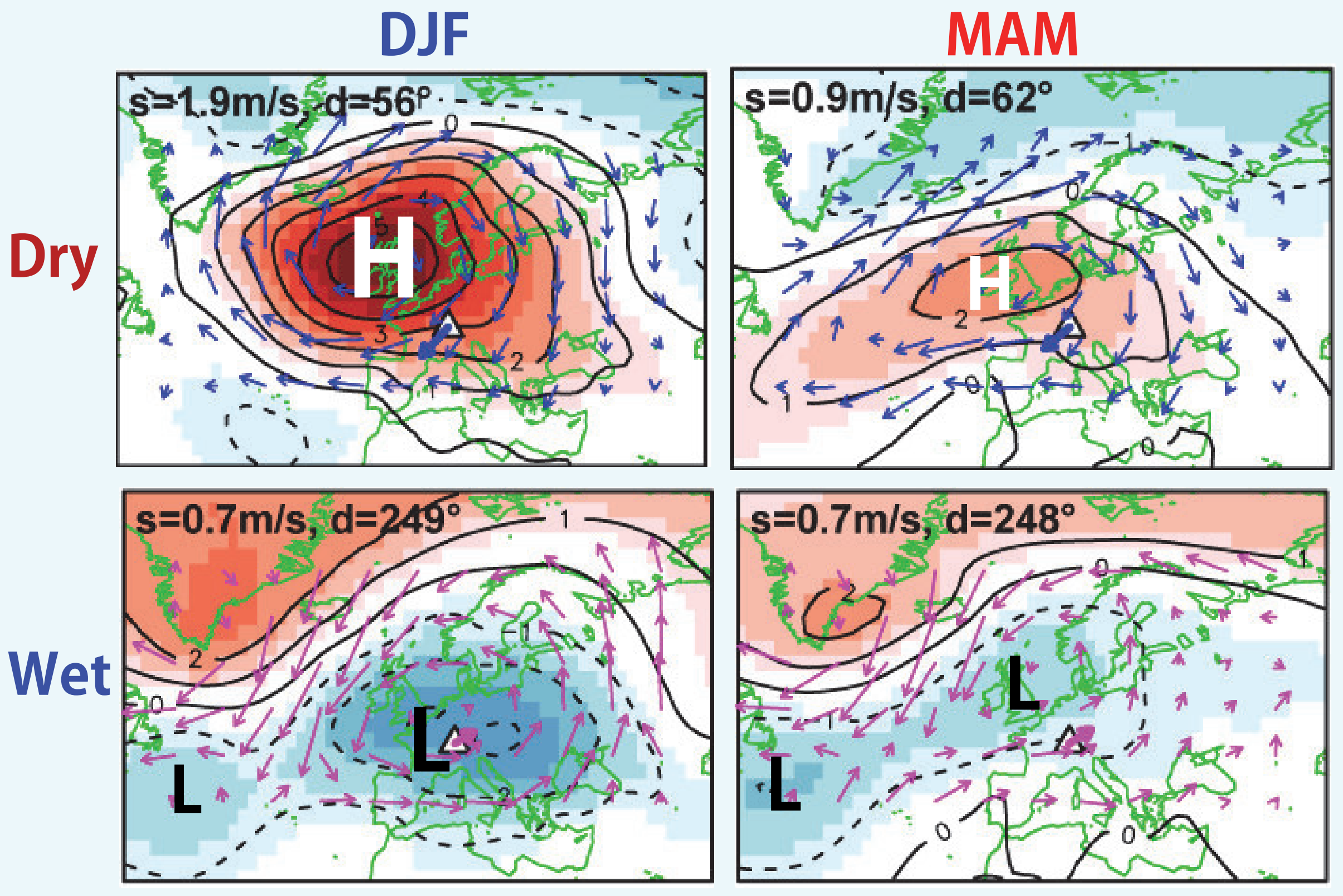


Fig. 7: Composites of sea level pressure and 850 hPa wind anomalies for seasonal "dry" ("wet") composites, (top (bottom) row) defined as the 18 driest (wettest) seasons in the 1901-2012 period for DJF (left) and MAM (right). The arrows show the direction and speed of the anomalous 850 hPa wind. Region considered is the eastern Jura.

Conclusions

- **91% of the series showed inhomogeneities (on average 2.7 per series).**
- **32 distinct Swiss precip regions were defined by objective clustering.**
- **Pos. precip trends for 50+ year series in winter, autumn & year. Pos. trends in 81 (72%) of annual (winter) series in 1901-2013.**
- **Euro-Atlantic blocking (Eastern Atlantic pattern) most important patterns in explaining northern (southern) Swiss precipitation variability in winter. Some different patterns for other seasons.**

Literature

→ DeGaetano AT, 2001. Int. J. Climatol. 21: 791-807. doi:10.1002/joc.645.
→ Begert M, 2008. Arbeitsberichte der MeteoSchweiz, 217 (in German, short version in English available).
→ Begert M, Schlegel T, Kirchhofer, W. 2005. Int J Climatol, 25, 65–80. doi:10.1002/joc.1118

The Parallel Observations Science Team (POST) compiling a global database with parallel measurements to study non-climatic changes

Victor Venema, Renate Auchmann, Enric Aguilar, Petr Stepanek, Ingeborg Auer, Cesar Azorin-Molina, Theo Brandsma,
Michele Brunetti, Manuel Dienst, Peter Domonkos, Alba Gilabert, Jay Lawrimore, Jenny Lindén, Ewa Milewska,
Øyvind Nordli, Marc Prohom, Jared Rennie, Peter Thorne, Blair Trewin, Lucie Vincent, Kate Willett, Mareile Wolff

1

Introduction

The Parallel Observations Science Team (POST) is compiling a comprehensive database with parallel measurements, which is important for a better understanding of non-climatic changes (inhomogeneities) affecting the evaluation of long term changes in daily climate data.

POST is part of the International Surface Temperature Initiative

2

Scientific background

Long instrumental climate records are usually affected by inhomogeneities (non-climatic changes) due to, for example, relocations and changes in instrumentation, instrument height or data collection and manipulation procedures.

These inhomogeneities distort the climate signal and can hamper the assessment of trends and variability. Thus to study climatic changes we need to accurately distinguish non-climatic and climatic signals.

3

Changes in extremes

> "[Inhomogeneous data] affects, in particular, the understanding of extremes, because changes in extremes are often more sensitive to inhomogeneous climate monitoring practices than changes in the mean."

Trenberth, K.E., et al., 2007: Observations: Surface and Atmospheric Climate Change. In: Climate Change 2007: The Physical Science Basis. Cambridge University Press, Cambridge,

4

A global parallel climate dataset

Current parallel data studies are limited to local and regional case studies.

The effect of specific transitions depends on the local climate and the most interesting questions are about the systematic large-scale biases produced by transitions that occurred in many regions.

Thus a large global parallel dataset is highly desirable as it allows for the study of systematic biases in the global record.

5

Historical transitions

We will work on studying the historical transitions that took place in many countries and have the potential to bias the climate record.

These have a different influence depending on the local climate, thus we need data from all climates world-wide.

People contributing parallel datasets are invited to join these studies. The data will be published together with the articles.

If original data cannot be published we will publish indices (extremes, variability).

6

Envisioned projects

We are working on the below transitions.

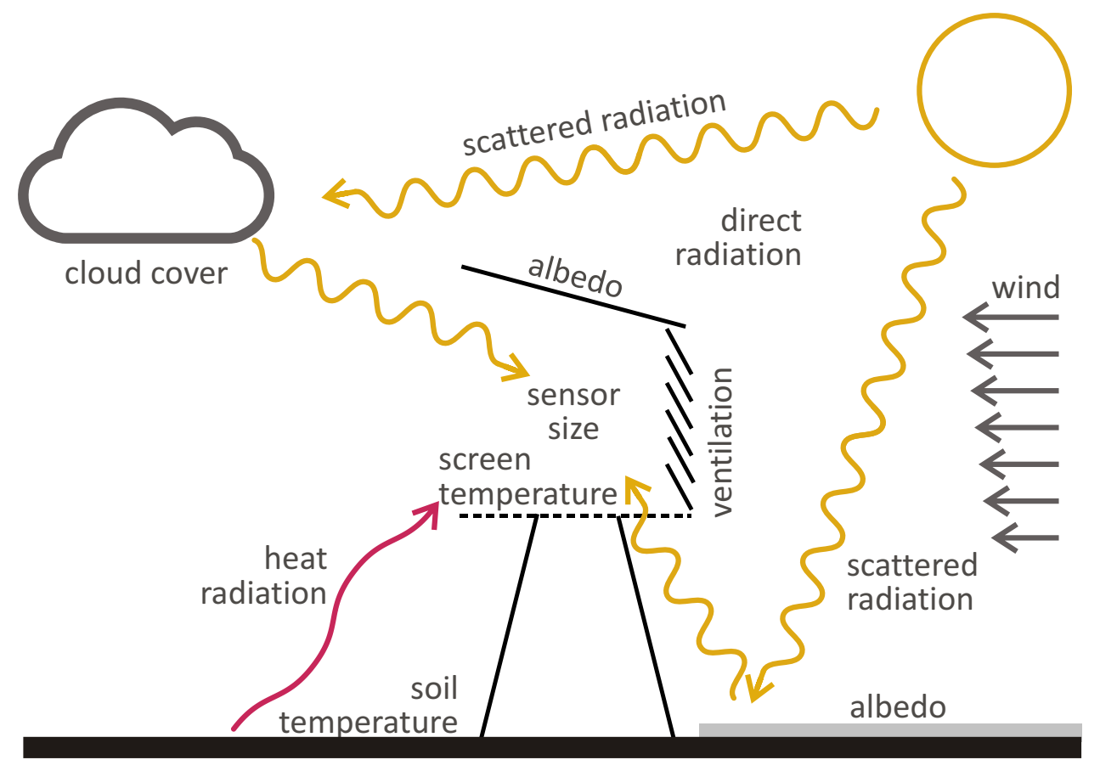
- From conventional observations to AWS
    - Temperature (Enric)
    - Precipitation (Petr)
- To Stevenson screens (Theo)
- Relocations: village, urban stations, airports (Alba)

Ideas (volunteers) for further studies are welcome. For example, on humidity or wind.

7

Exposure improvements

Especially before the Stevenson screen, temperature observations were affected by radiation errors, which depend on insolation, wind, vegetation, soil moisture, etc.



8

Exposure - Wild screen

Wild and Stevenson screen in Basel, Switzerland. Wild screen is affected by scattered solar radiation, warming from ground and infra-red cooling. Analysed in Auchmann and Broennimann (2012).

9

Parallel data & predictors

High-resolution data is important for understanding the physical causes for the differences. For the same reason we are also interested in other measured variables at the same station.

For example, in case of parallel temperature measurements, the influencing factors are expected to be insolation, wind and clouds cover.

In case of parallel precipitation measurements, wind and temperature are potentially important.

10

More information

New members very much welcome!

Please write: Victor Venema,
Victor.Venema@uni-bonn.de
Or Renate Auchmann
renate.auchmann@giub.unibe.ch
To join and for more information
http://tinyurl.com/ISTI-Parallel

11

What data do we need?

POST will gather parallel data in their native format (to avoid undetectable conversion errors, we will convert native data to a standard format ourselves).

We are interested in data from all climate variables at all time scales; from annual to sub-daily.

Metadata that describe the parallel measurements are as important as the data themselves and will be collected as well. For example, the types of the instruments, their siting, height, maintenance, etc.

12

ISTI-POST Members

POST members are actively building and coding the parallel database and involved in the studies. The POST members are the authors of this poster. If you would like to be informed about POST and have the possibility to give feedback, you can become an associate member.

Associate members: Andreas Becker, Stefan Brönnimann, Manola Brunet, Sorin Cheval, Aryan van Engelen, Constantinos Kolokythas, Frank Kaspar, Albert Klein Tank, Franz Gunther Kuglitsch, Monika Lakatos, Anna Mamara, Hermann Mächel, Colin Morice, José Guijarro, Clara Oria, David Parker, Mário Gonzalez Pereira, Sarah Perkins, Michael de Podesta, Elke Rustemeier, Javier Sigro, Tamas Szentimrey, Gregor Vertacnik, Xiaolan L. Wang, Rachel Warren, Steven Worley, Markus Ziese.

13

Support for POST

The WMO has recently called on all members to assist in gathering parallel datasets for an international dataset.

The database is endorsed by the Task Team on homogenization (TT-HOM) of the Commission for Climatology (CCI).

The International Surface Temperature Initiative (ISTI) will host a copy of the parallel dataset, as well as the European Climate Assessment & Dataset project (ECA&D). This will ensure professional and permanent archiving and thus the long-term use of these important datasets.

Meteorologisches Institut
Universität Bonn

Photo:SCREEN project, URV, Tarragona, Spain

IENet: A homogenised precipitation network for Ireland – preliminary analysis

John Coll[1], Peter Domonkos[2] , Mary Curley[3], Enric Aguilar[2], Séamus Walsh[3], John Sweeney[1]

[1] Irish Climate Analysis and Research Units, Department of Geography, Maynooth University, Maynooth, Ireland. [2] Centre for Climate Change, C3, University Rovira i Virgili, Spain. [3] Met Éireann, Glasnevin Hill, Dublin 9, Ireland.

Introduction

Motivation: Climate change studies based only on raw long-term data are potentially flawed due to the many breaks introduced from non-climatic sources, consequently quality controlled and homogenised climate data is desirable for basing climate related decision making on. This reflects a growing demand for climate information or climate services more generally for use across a range of decision-making environments. Seasonal cycles of precipitation in Ireland are projected to become more marked as the climate changes, and regional extremes in summer dry spells and winter precipitation have been recorded in recent years. Therefore to analyse and monitor the evolution of precipitation patterns across Ireland, quality controlled and homogenous climate series are needed.

Aims and objectives: To compare the results of two modern relative homogenisation methods (HOMER and ACMANT) for a medium sized network of 333 series (IENet) from the Met Éireann monthly precipitation series network for the 1941 – 2010 period (Figure 1).

Study area: The study area is the whole island of Ireland, that covers ~84 421 km^2 on the Atlantic margin of northwest Europe, between ~51° and 56° N. Elevations reach up to 1038 m above sea level (a.s.l.) (Corrán Tuathail, Co. Kerry) although much of the island is lowland, partly surrounded by mountains, with a characteristic temperate oceanic climate. On average, annual precipitation ranges from 750 to 1000 mm in the drier eastern half of the country and >3000 mm yr−1 in parts of the western mountains (Rohan 1986).

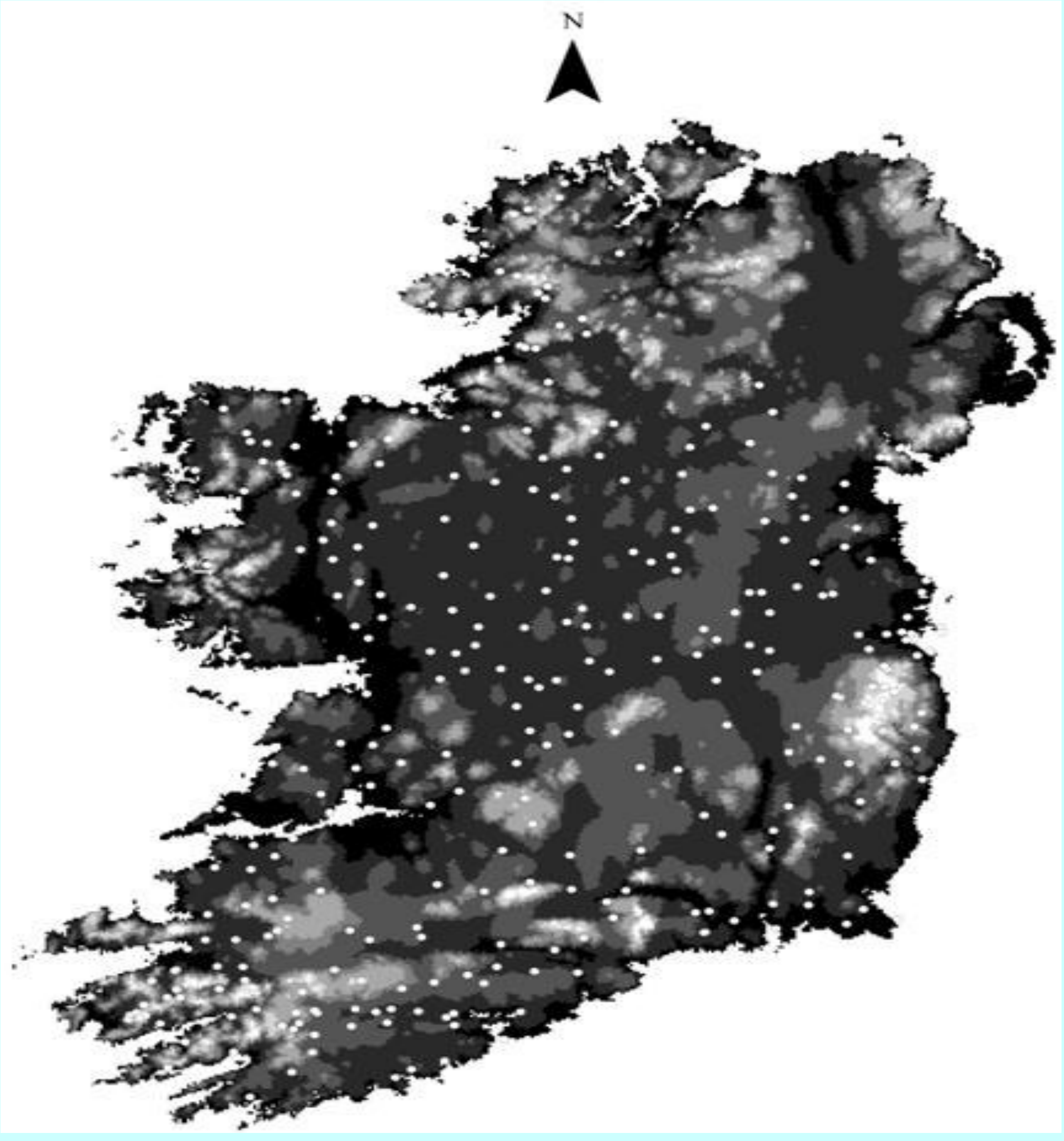


Figure 1. Annotated map of the island of Ireland showing the selected Met Éireann monthly station locations for the current network of 333 precipitation stations (IENet). Precipitation station series which have been homogenised by both the HOMER and ACMANT programmes are denoted by white circles. Topography is represented by contrast shading.

Data

Monthly series: Rainfall has been measured in Ireland since the early nineteenth century with a peak of over 800 rainfall stations in the late 1950s, and currently rainfall is recorded at synoptic and climatalogical weather stations; in addition, there is a wide network of voluntary rainfall observers (Walsh, 2012). For the country-wide IENet series used here, station elevations ranged from 5 – 701 m above sea level with a mean elevation of ~78 m, and intact contiguous records ranged from ~31 to ~70 years for the 1941-2010 period.

Table 1: Similarities between HOMER and ACMANT

Time Resolution	Monthly
Reference series selection	Correlation-based
Detection method	Step function fitting
Detection statistic	Penalised likelihood
Correction method	ANOVA

Table 2: Differences between HOMER and ACMANT

Implementation	HOMER	ACMANT
Time series comparison	pairwise	weighted reference
Harmonisation of detection	network wide	in individual series
Working mode	interactive	automatic

Data homogenisation: HOMER (Mestre et al. 2013) and ACMANT (Domonkos, 2011, 2014) are two modern, multiple break homogenisation methods developed to detect and correct multiple change points using reference series, and were developed during the Action COST-ES0601 (HOME, Venema et al. 2012). HOMER includes the best segments and best features of some other state-of-the-art methods: PRODIGE (Caussinus and Mestre, 2004), ACMANT and Joint Detection (Picard et al. 2011). An outline of common and different properties of the methods is provided (Table 1 and Table 2).

Results

HOMER break detections: Two hundred and thirty five stations were found to be homogenous, but 120 breaks were detected by HOMER across the other 98 stations and multiple breaks were found in 22 records. To date 89 breaks have metadata support and metadata is awaited for some further station records. HOMER correction amplitudes for inhomogeneities in the series ranged from -0.61 to +0.77 (Figure 2).

ACMANT break detections: ACMANT consistently detected more breaks than HOMER in all decades, with the biggest disparities in break detection rates in the 1950s and 1960s (Figure 2). One hundred and seven stations were found to be homogenous, but 224 breaks were detected by ACMANT across the other 226 stations and multiple breaks were found in 132 records.

The spatial locations of multiple break detections for the IENet series were different between HOMER (Figure 3a) and ACMANT (Figure 3b). In addition, there are regional differences, HOMER e.g. detected more multiple breaks associated with stations in the mountains of the SW. Whereas ACMANT detected more multiple breaks in stations across the country compared to HOMER.

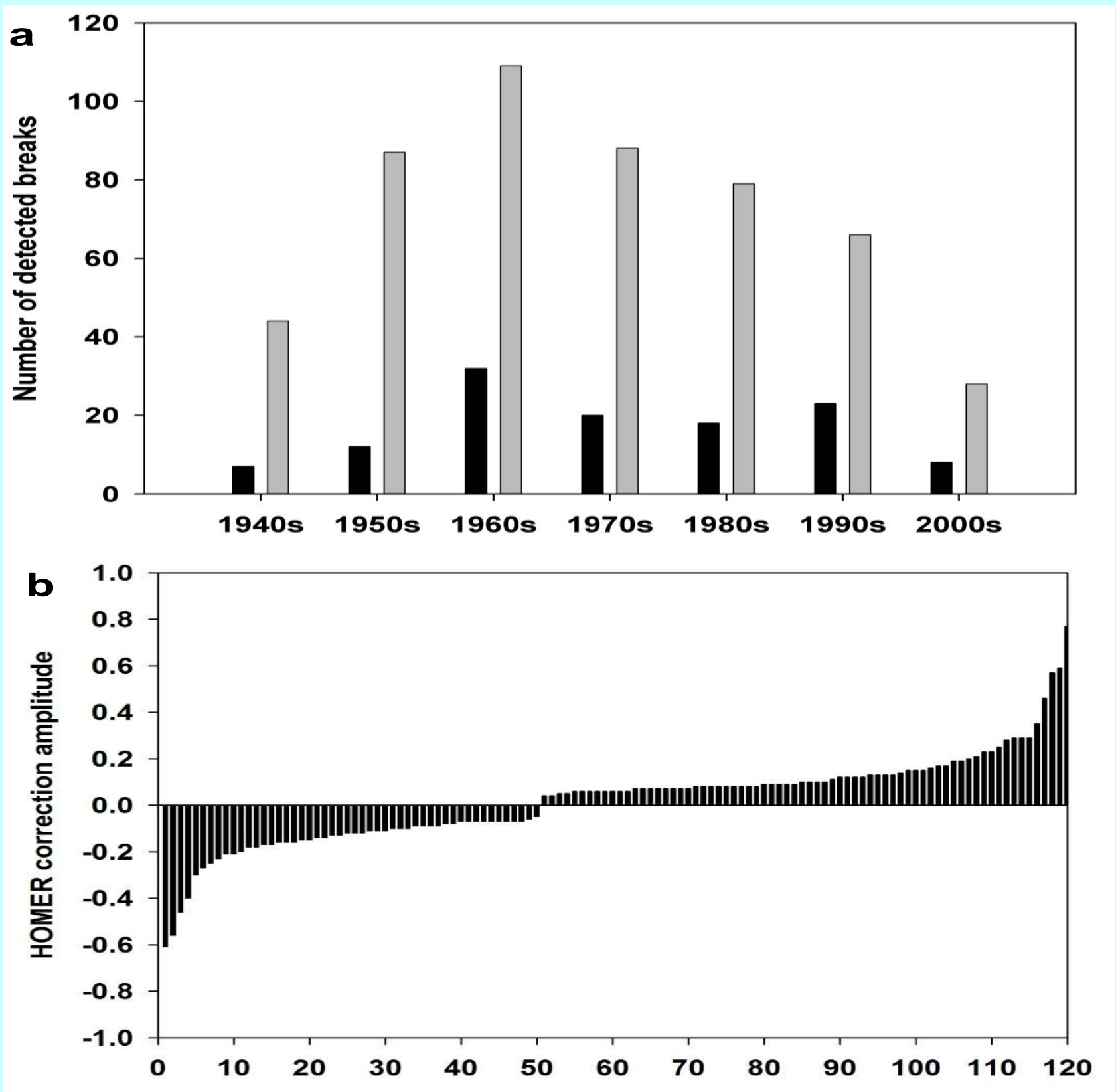


Figure 2. a) Histogram break detection count summary by HOMER and ACMANT for the IENet series by decade. Black bars denote HOMER detection counts, grey bars denote ACMANT detection counts. b) HOMER correction amplitude range for detected breaks.

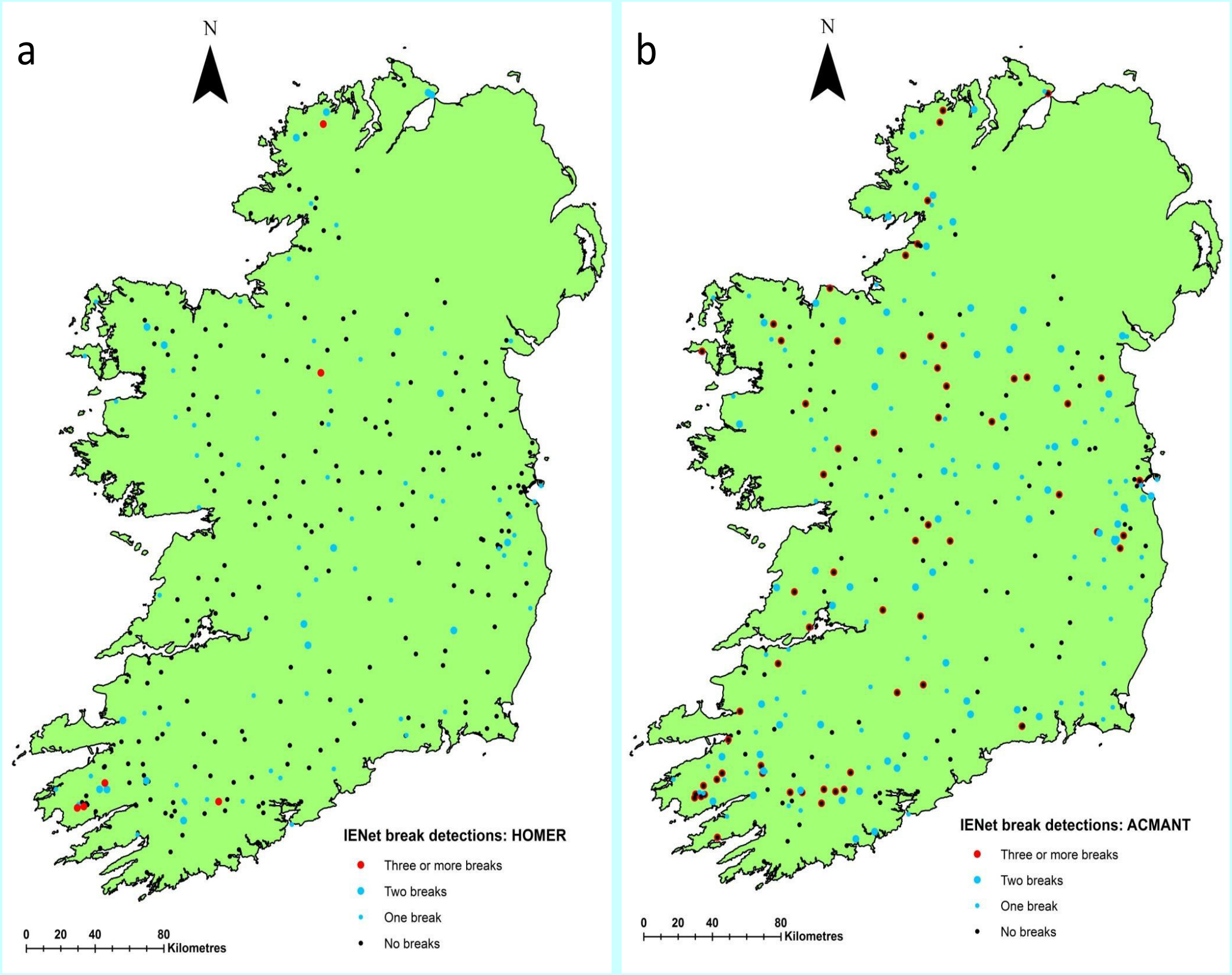


Figure 3. Number of inhomogeneities detected in the IENet precipitation series by a) HOMER and b) ACMANT. Series with no breaks (black circles) and one or more breaks (variable size blue and red circles) are shown throughout.

Discussion

The mean r-value for candidate and reference series for IENet across the networks defined by HOMER was 0.90 and the overall mean network range of r was 0.85 - 0.95. HOMER was first run on all series with known outliers included, and the results scrutinised; series were then re-processed in HOMER following the removal of outliers. However, the distribution of years with breaks detected by HOMER remained the same, indicating that for IENet break detection by the programme is not sensitive to outliers.

Amplitude provides an indication of the magnitude of breaks detected as well as the amount of adjustment needed to correct the inhomogeneity. For the results with HOMER, across all stations and detected breaks the mean correction amplitude was 0.02. A combination of the density of the network and high correlation coefficients between the station series (reflecting a maritime climate) are allowing both HOMER and ACMANT to detect relatively small breaks.

Conclusions and ongoing work

- HOMER consistently detects less breaks than ACMANT for the current IENet precipitation series. ACMANT detects more overall and more multiple breaks than HOMER, and with a different spatial pattern of detections.
- For the 120 breaks identified by HOMER, 89 (~74%) were confirmed by the metadata. However, metadata were not available for all the station series in the current network.
- The spatial characteristics of the IENet precipitation records (a dense network) allied to the climatic characteristics of a maritime region (relatively low amplitudes of variation) result in highly correlated series. These properties of the data and the network are useful for the test application of relative homogenisation methods to observed series.
- The analysis using ACMANT has been extended to a wider network of 700+ series and the results are being processed. However, it is unlikely that full metadata support will be available to check all the ACMANT detections for this larger network. Despite considerable effort it has not been possible to apply HOMER to this larger network due to complications with missing data in many of the series.
- We consider that by using Ireland as a case study, the prospects for evaluating variations in network density on the break detection frequency of methods such as HOMER and ACMANT for real world precipitation time series are excellent.

References

Caussinus, H. and Mestre, O. (2004). Detection and correction of artificial shifts in climate series. *J. Roy. Stat. Soc. Series* C53: 405–425.

Domonkos, P. (2011). Adapted Caussinus-Mestre Algorithm for Networks of Temperature Series (ACMANT). *Int. J. Geos.* 2: 293-309.

Domonkos, P. (2014). Homogenisation of precipitation time series with ACMANT. *Theoretical and Applied Climatology*. 118:1-2. DOI 10.1007/s00704-014-1298-5.

Mestre, O., Domonkos, P., Picard, F, Auer, I., Robin, S. et al (2013). HOMER: a homogenization software – methods and applications. *Idojaras*. 117 (1): 47-67.

Picard, F., Lebarbier, E., Hoebeke, M., Rigaill, G., Thiam, B. and Robin, S. (2011). Joint segmentation, calling and normalization of multiple CGH profiles. *Biostatistics*. 12: 413–428.

Rohan, P.K. (1986). The climate of Ireland. The Stationery Office, Dublin

Venema V, Mestre O, Aguilar E, Auer I and others (2012). Benchmarking Homogenization Algorithms for Monthly Data. *Clim. Past.* ***8: 89-115.***

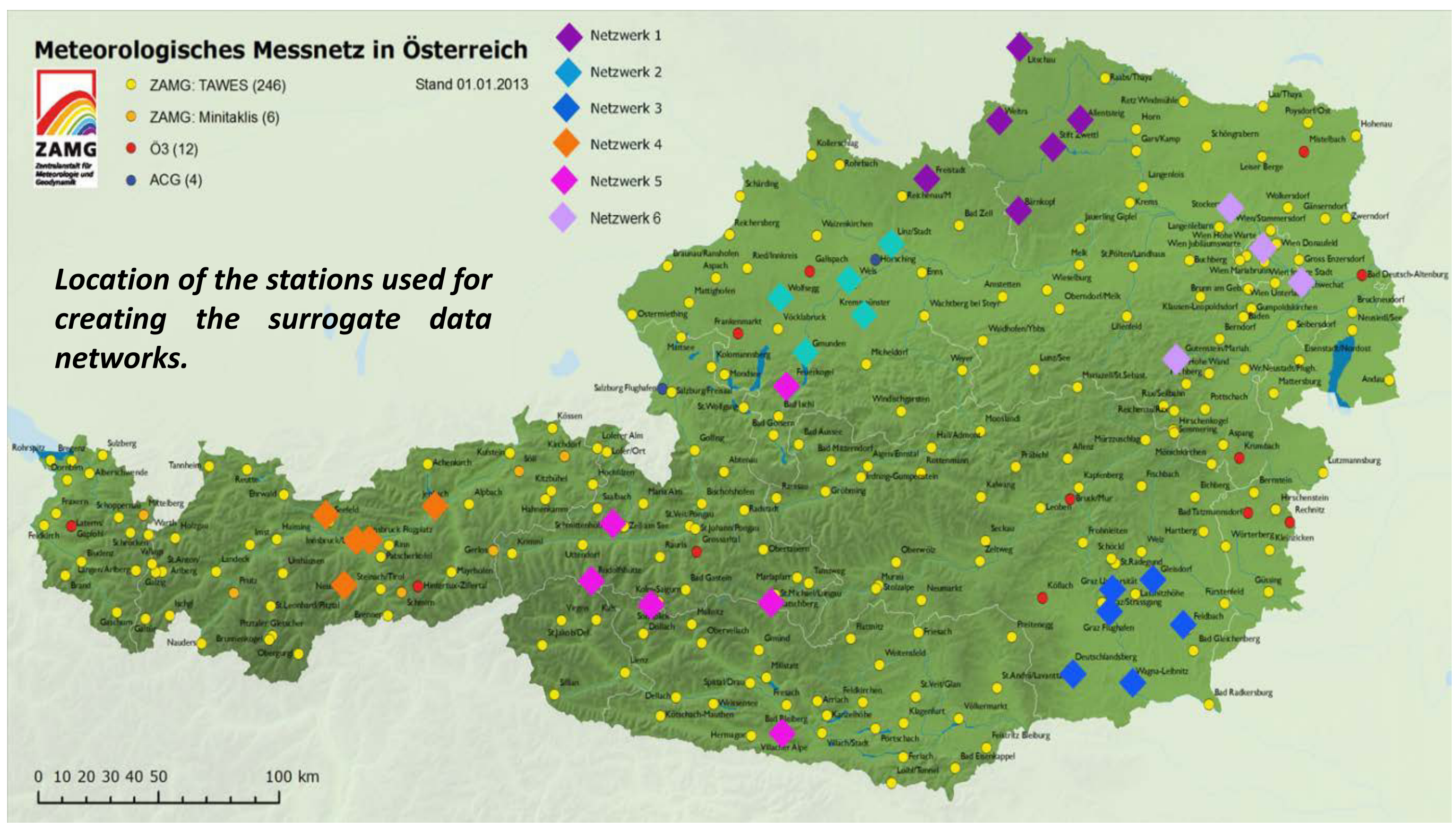
Walsh, S (2012). A summary of climate averages for Ireland 1981-2010. Met Éireann Climatological Note 14. Dublin. 16pp.

Acknowledgements: We thank the Irish Environmental Protection Agency (EPA) for the project funding and support. This research was supported under grant 2012-CCRP-FS.11.

New developments in homogenisation efforts for Austria

Barbara Chimani[(1)], Ingeborg Auer[(1)], Annemarie Lexer[(1)], Manfred Ganekind[(1)], Konrad Andre[(1)], Victor Venema[(2)]
[(1)] Zentralanstalt für Meteorologie und Geodynamik (ZAMG), Austria
[(2)] University Bonn, Germany

After homogenising daily temperature and precipitation data for a set of about 70 Austrian stations from 1948 onwards, other parameters seemed necessary to be homogenised in order to provide a set of meteorological, homogenised data for further research.
The parameter chosen was the relative humidity, as this parameter is of interest not only in meteorological questions but has influence on materials and plants as well. Nevertheless the parameter has some difficulties not that evident in temperature or precipitation: only a small range of values occurs in nature and the instruments are changed quite regularly, leaving only short periods without interruption.
Four methods have been compared in this study:
ACMANT, MASH, HOMOP, PROCLIM.



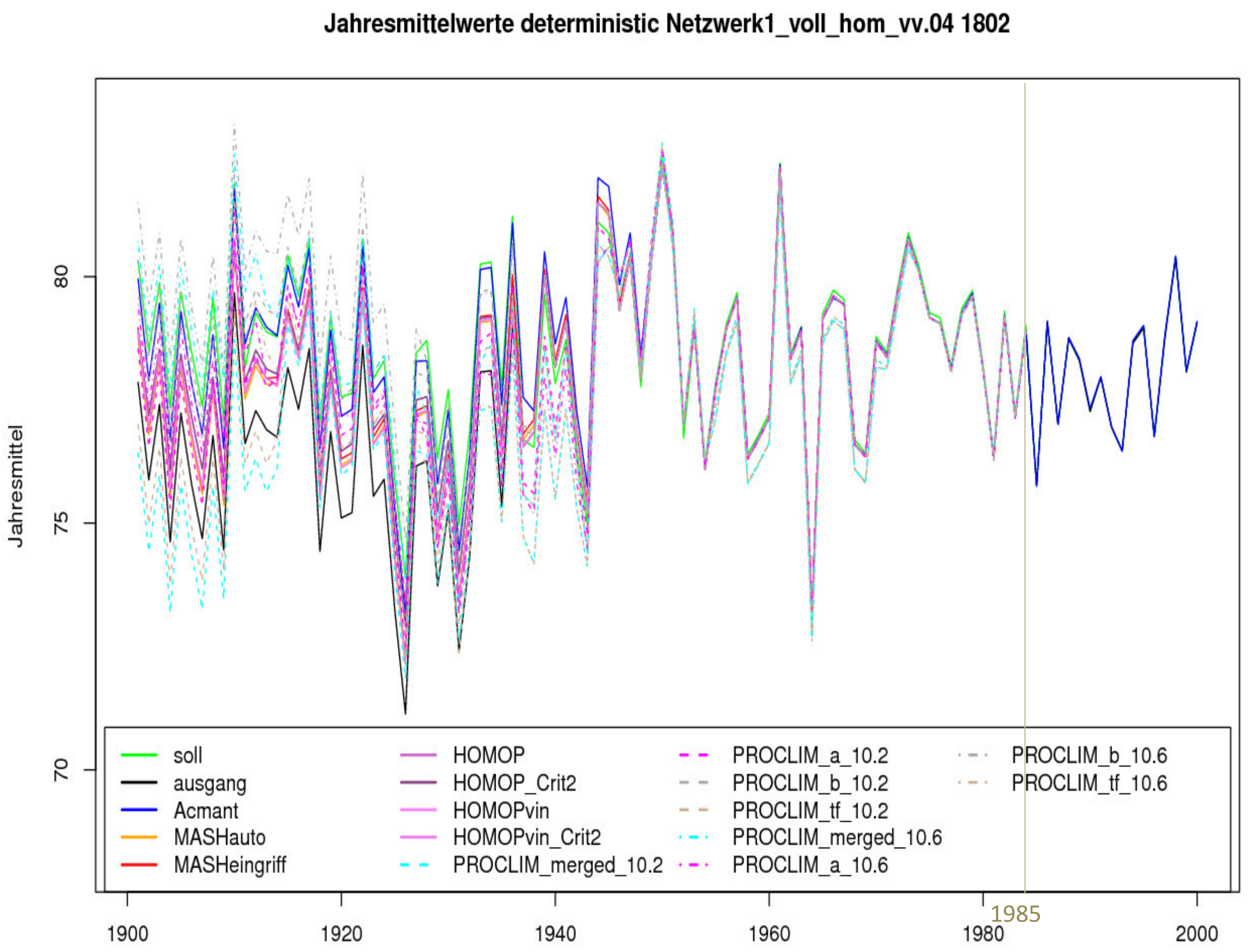
Location of the stations used for creating the surrogate data networks.

1) Creation of a surrogate dataset

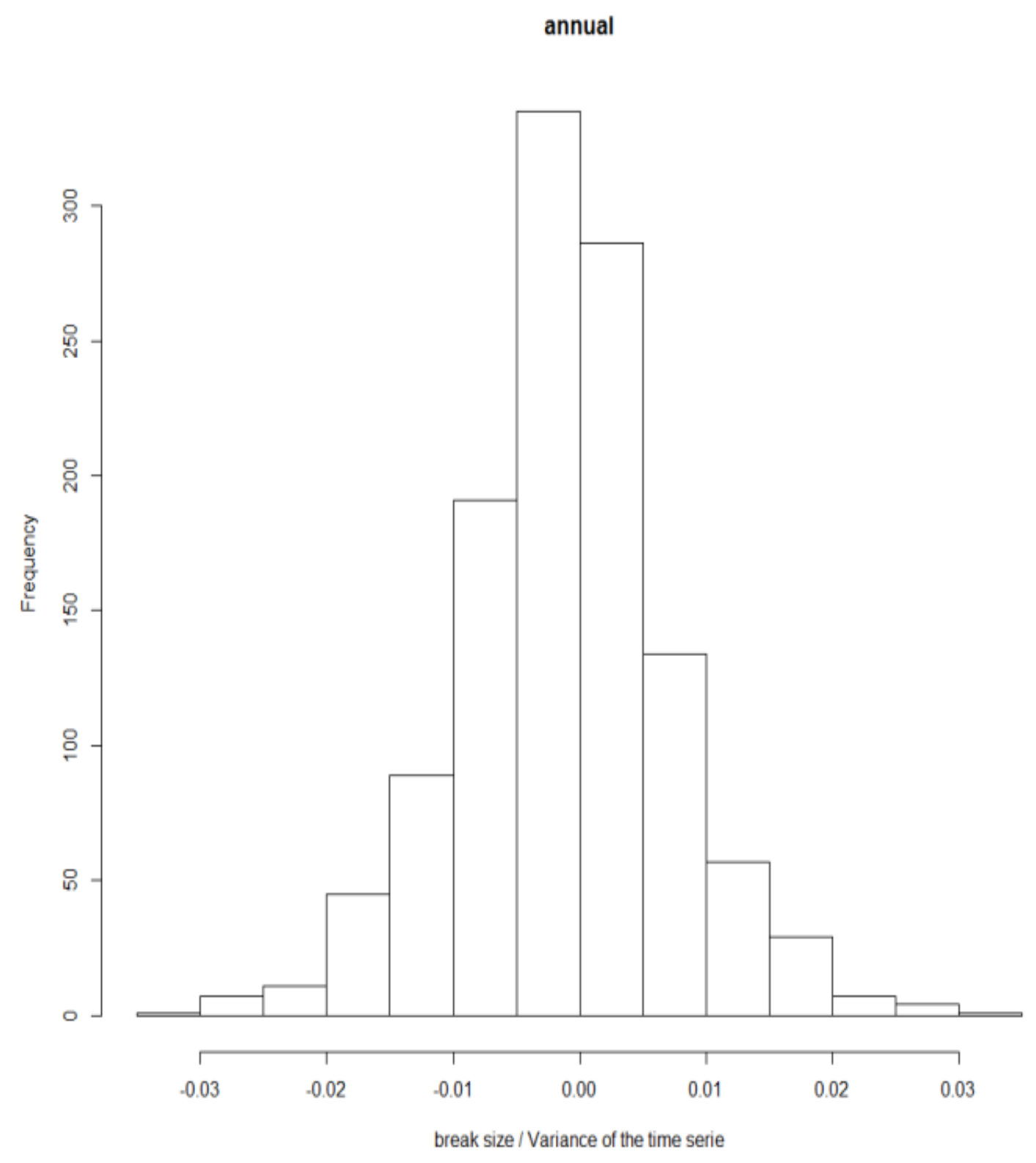
Information on frequency of stations relocation and instrumental changes were collected using the station meta data archive of ZAMG.

Typical characteristics of breaks due to relocations were analysed by the use of parallel measurements for different stations in Austria.

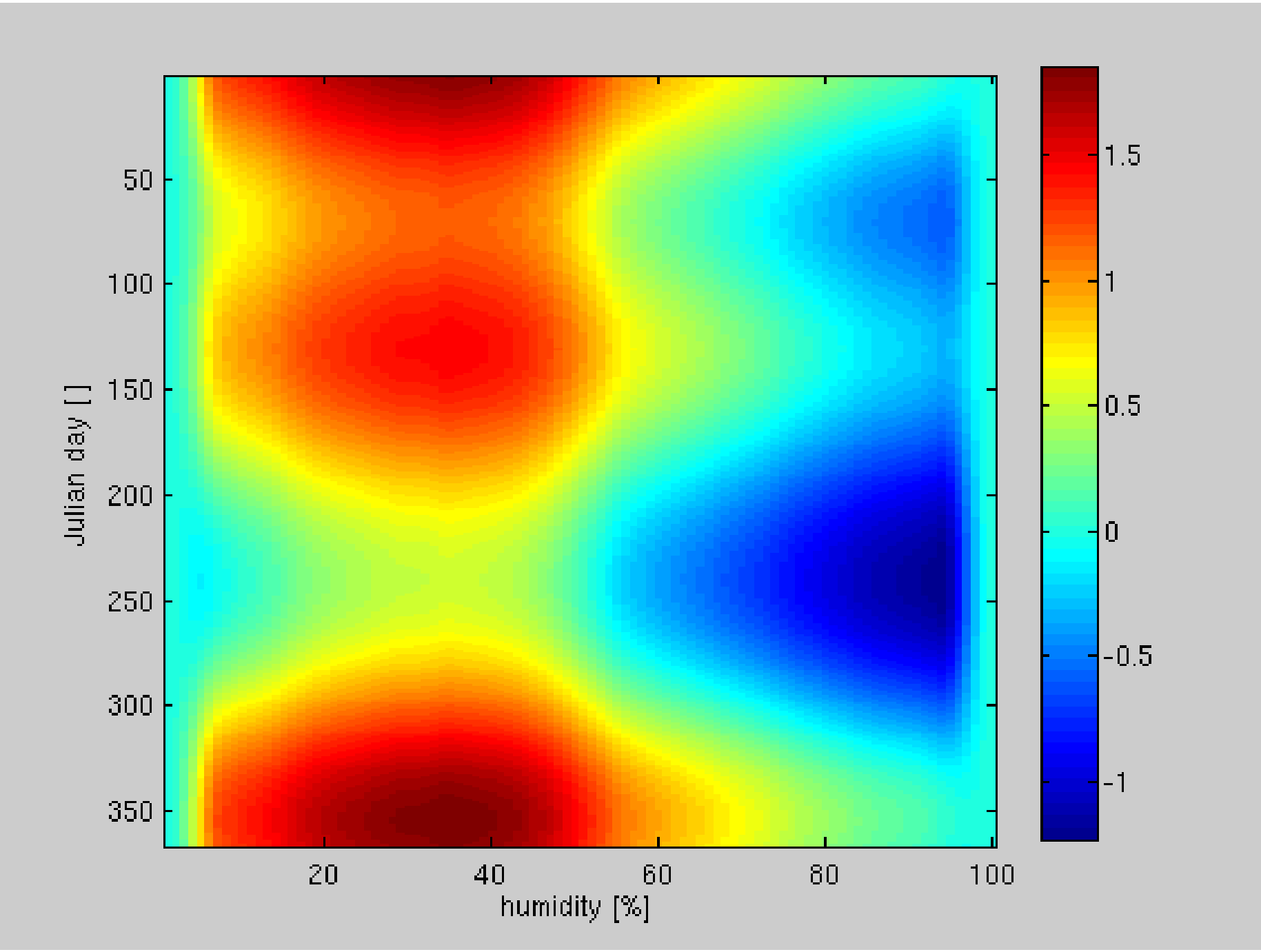
Networks of stations that seemed to be quite homogeneous according to the stations meta information, was checked on homogeneity in monthly resolution with HOMER. This station data was used to create surrogate datasets of 100years of length.



Example of homogenisation results in a deterministic case
Green: homogeneous searched time series
Black: inhomogeneous time series
All the other lines represent the result of the homogenisation methods



Histogram of the quotient of break size to variability of the time series



Example of a break size (colours) in dependence of the true value (x-axis) and the time of the year (y-axis)

2) Break detection

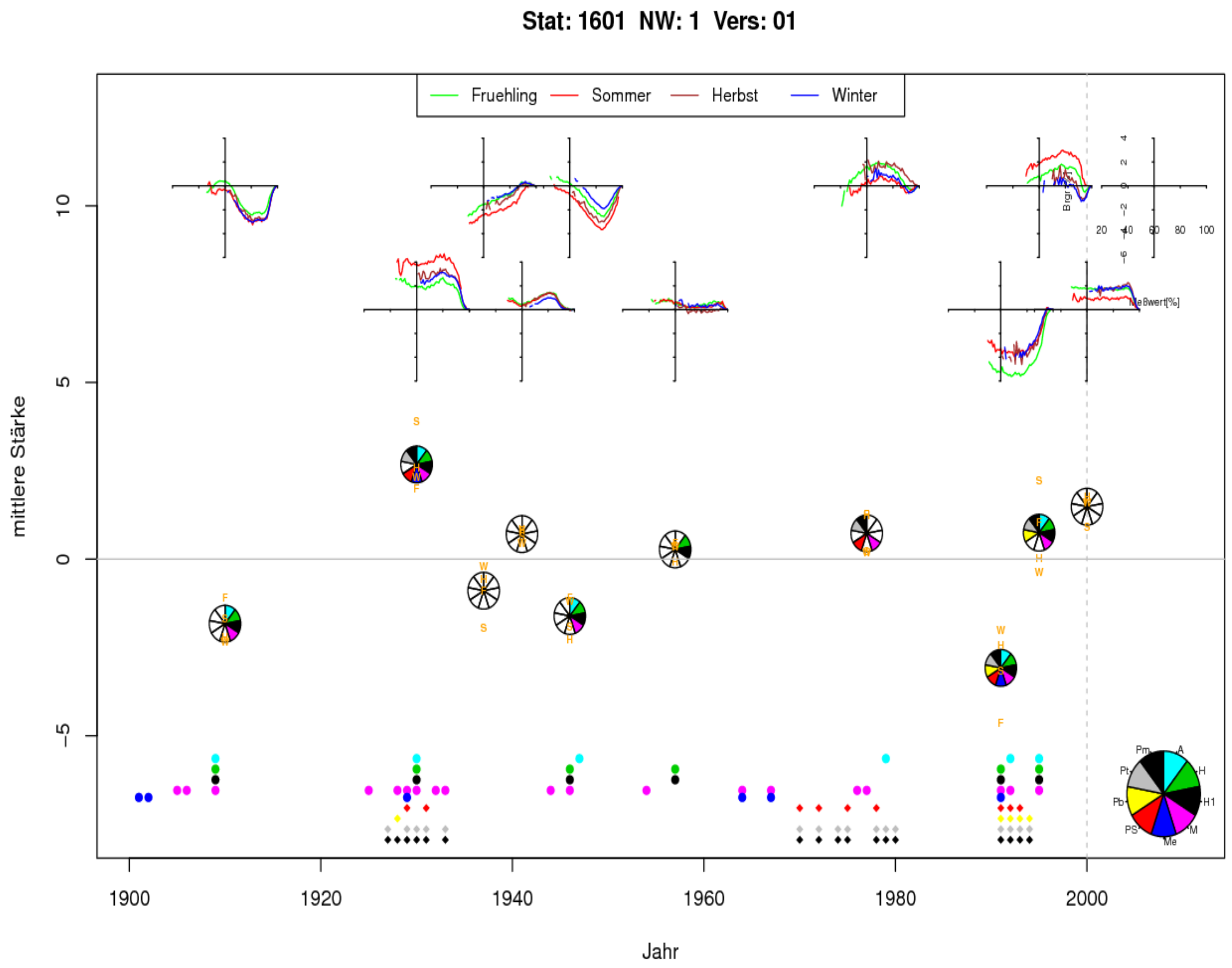
The ability of four different methods, tested during the COST-Action ES0601, was studied using a deterministic (without missing data and white noise) and a realistic (missing data, different start of time series and white noise) dataset. For MASH, HOMOP and PROCLIM different settings were tested.

A break was detected correctly if the year was within the range of ±1 year of the actual break year. All detected breaks within one year counted as one break.

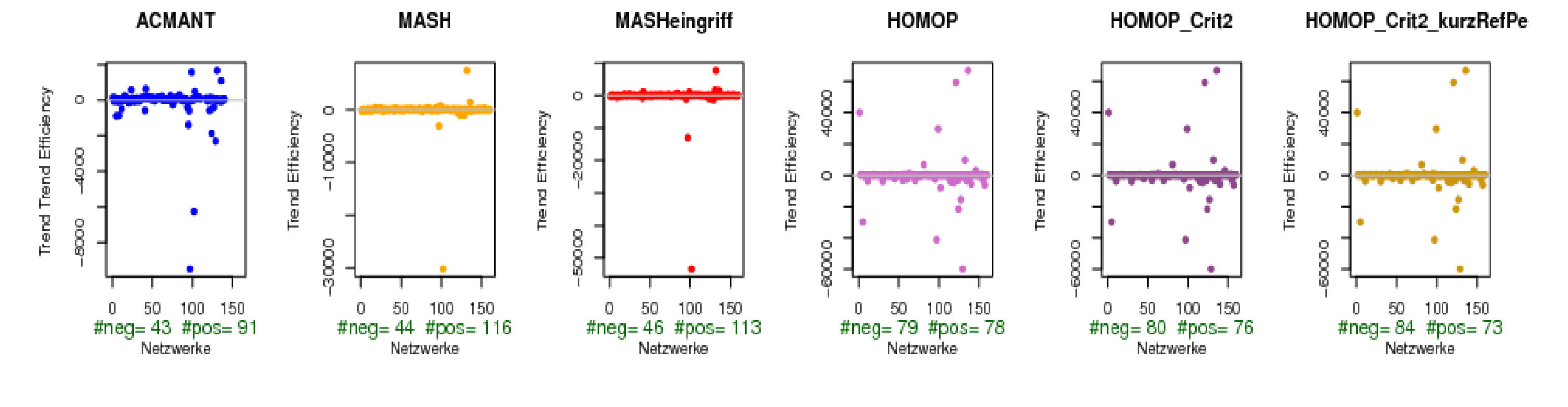
Results

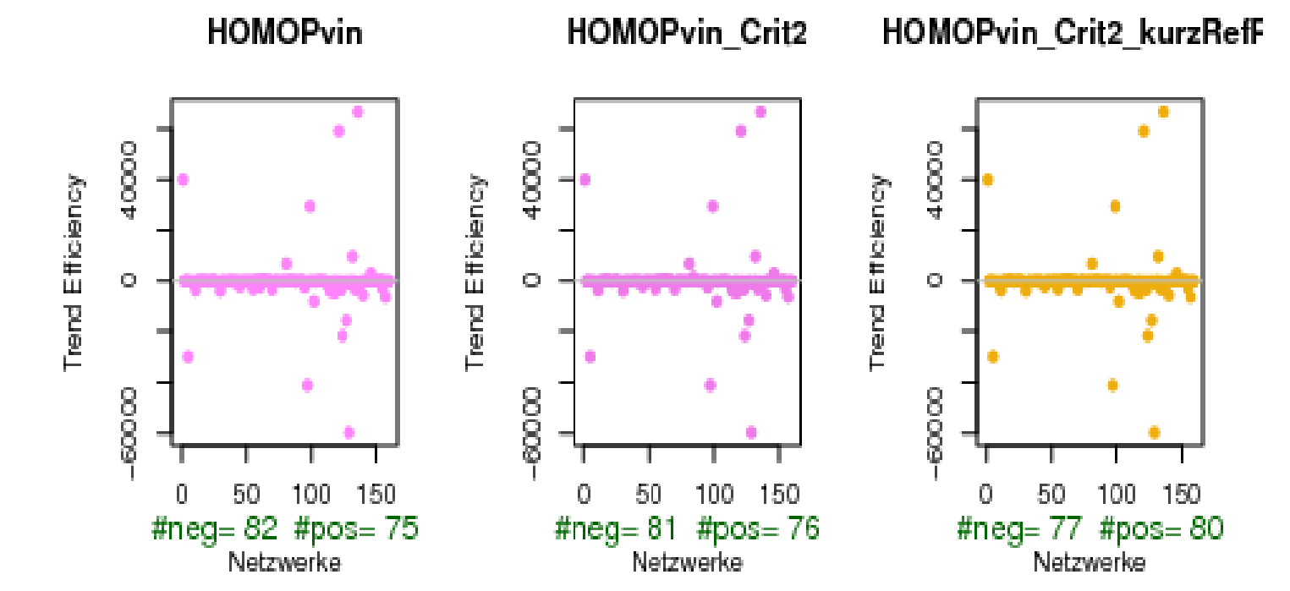
All the breaks of the surrogate dataset are clustered with breaks of the reference time series or are located near the end of the time series. Nearly 3/4 of the breaks are additionally clustered with breaks in the time series itself. More then 1/4 of the breaks was not detected by any of the 4 methods.

The ability to detect the correct breaks is highest for ACMANT and HOMOP. In the case of HOMOP more time series are seen as homogeneous than in ACMANT. Trying to detect breaks in the homogeneous surrogate time series, ACMANT detects more breaks than HOMOP.
For all four methods the ability is lower in the realistic case than in the deterministic.



Result of a break detection for one station in Network1.
Upper row: break signal as a function of measurement for different seasons (Fruehling=spring; Sommer=summer; Herbst=autumn; Winter=winter)
Mean row: Letters mean break signal for 4 seasons and the year. Colours giving the methods with detected this break
Lower row: all detected breaks by the different methods(colours)
X-axis: year of the breaks, Y-axis= mean break signal in mean row.





Trend-Efficiency for the different methods in the realistic case: positive Efficiency= improvement, negative Efficiency= worsening
Each point represents one time series.
In green the number of worsened and improved time series is given.

3) Correction

The ability for break correction in the realistic case was tested for ACMANT, MASH and HOMOP using the according break information. For all but ACMANT different settings have been used.

Results

Only MASH found solutions for each time series, while ACMANT had the most troubles. The networks for which single time series couldnot be solved varied between the methods.

The improvements in RMSE, Variability and Trends were analysed.
While improvements of the RMSE are possible for a part of the time series for all the methods, only ACMANT improves the RMSE in the realistic dataset in about half of the time series.
For the trend all of the methods can improve the trend for about half of the time series. MASH is leading this analyses with about 2/3 of the time series being improved.
The variance is not strongly influenced for ACMANT and MASH in the realistic case.

Interfering with MASH without the knowledge of any metadata, but preferring not to use breaks that extinguish each other within 1 year leads to less homogenisation skill.

4) Change in observing routine

With the start of 1971 the evening observation hour changed from 21°° to 19°°. At the same time a new calculation method for averaging the daily data was implemented. To adjust this break a seasonal multivariate regression using the measurements of this day as well as of the next day and the time of sunrise and sundown can be used. Due to strong local influences only a small part of the stations has enough information before and after the change to calculate the adjustments.

Contact:
Barbara.chimani@zamg.ac.at

Acknowledgement:
This research is funded by the Austrian Climate Research Program (ACRP).

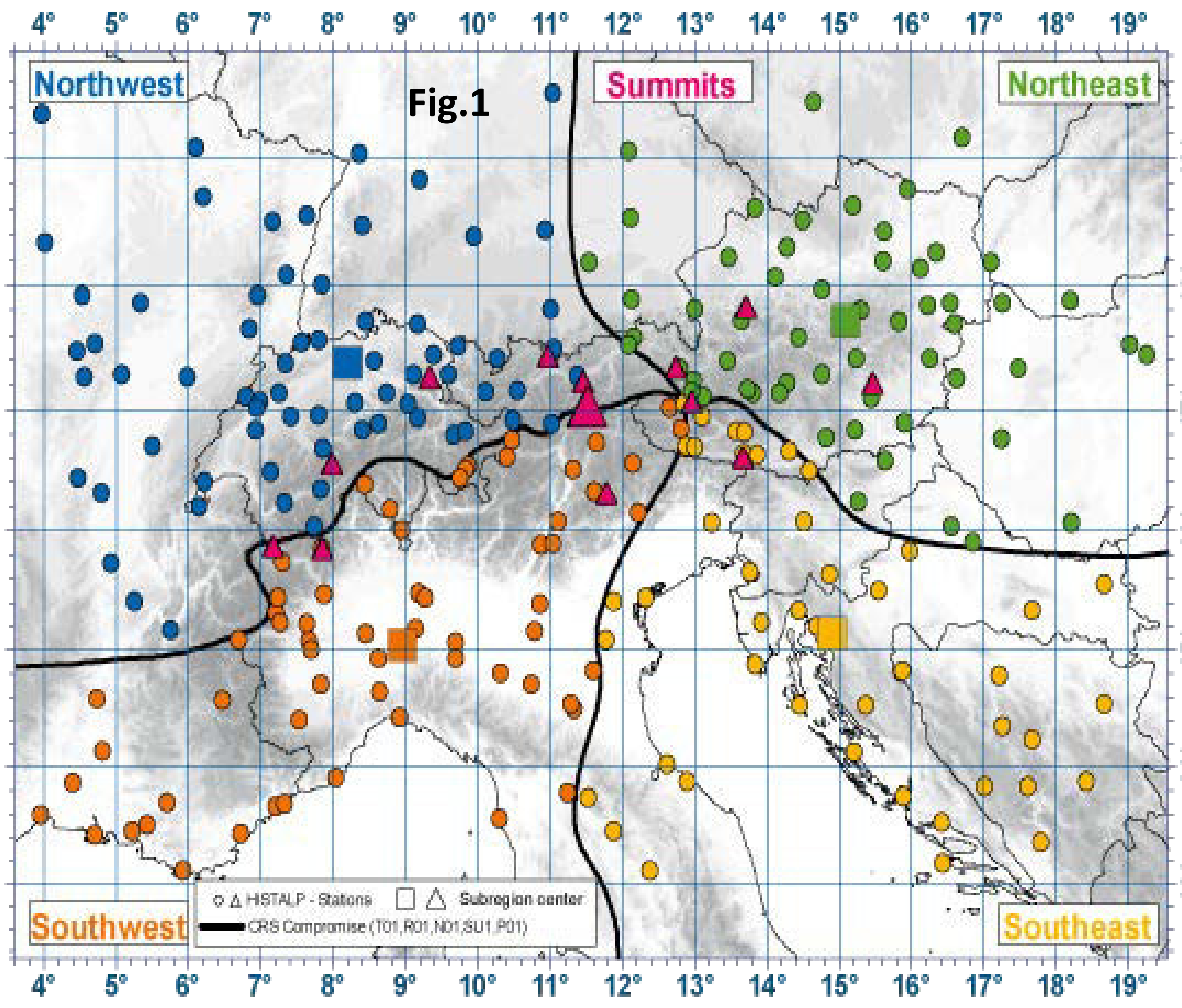
Climate data supporting climate services
Okt 2015

New developments in HISTALP

Barbara Chimani, Manfred Ganekind, Ingeborg Auer, Angelika Höfler
Zentralanstalt für Meteorologie und Geodynamik (ZAMG), Austria

The first steps towards HISTALP were taken in 1997 and in 2002 the name HISTALP was established. A lot of time was invested in stations selection, data quality control and homogenisation. This lead to a widely recognised instrumental dataset used in a wide variety of climate studies.
Since the creation of the HISTALP-dataset more than 10 years have passed. The data was regularly updated, thanks to all the data providers. But in the mean time the homogenisation routines improved, leading to more objective results and due to the time span additional breaks might have occurred. Therefore a new homogenisation of the dataset was decided in 2013. Temperature and precipitation have been homogenised and are under quality control at the moment.

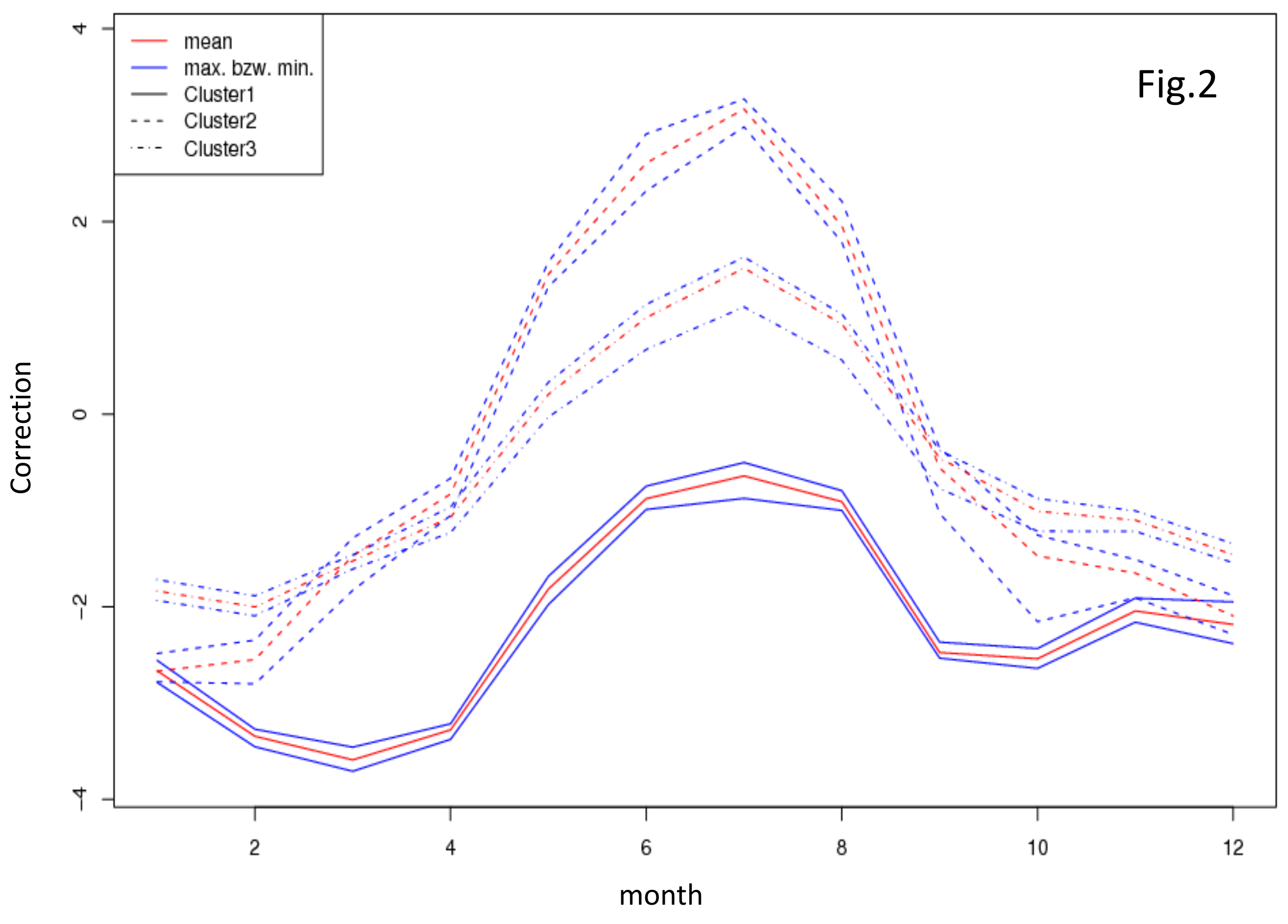
Scientific Developments



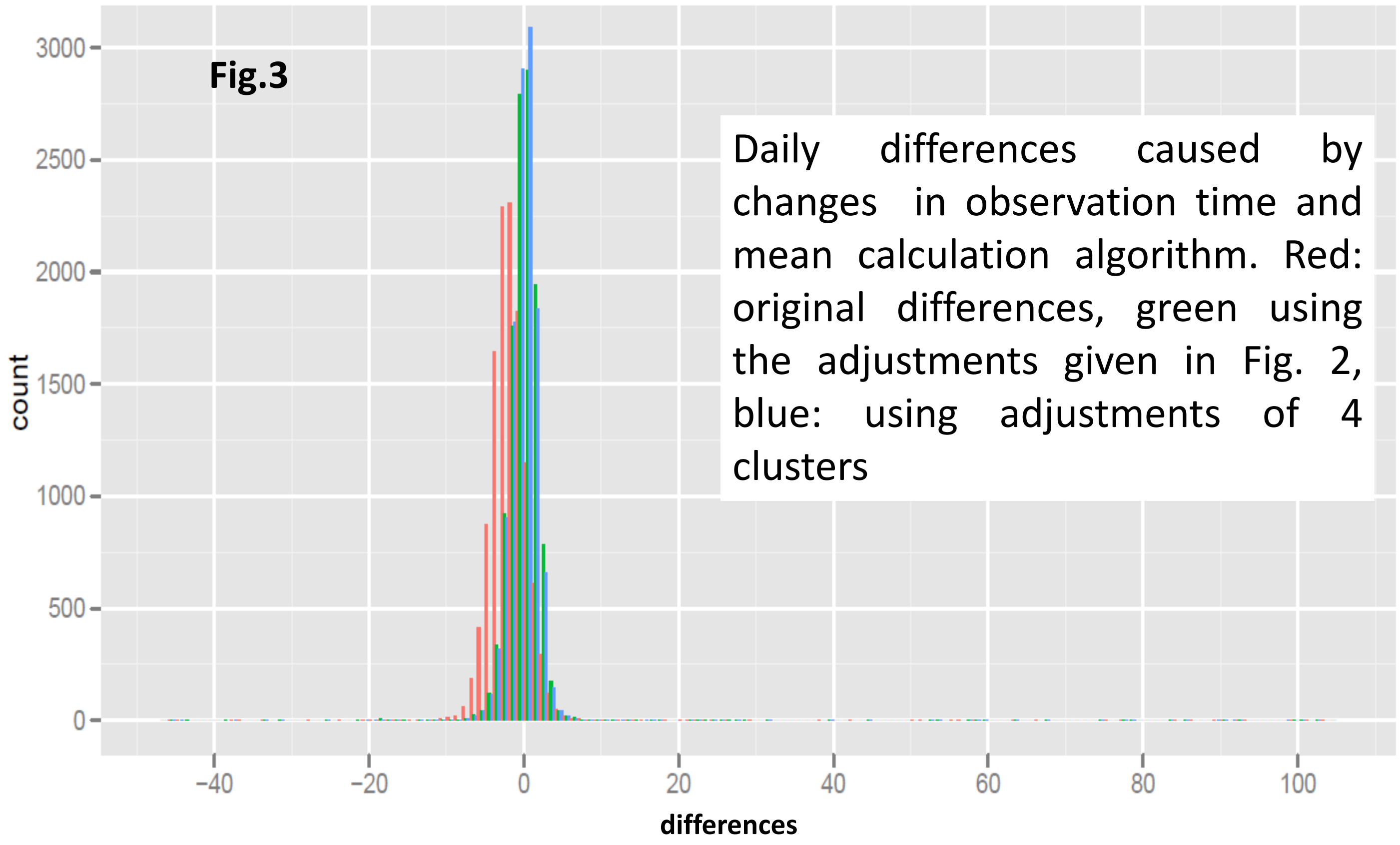
For homogenisation of the monthly temperature and precipitation data HOMER was used. High correlated reference stations of the HISTALP-Dataset were used. While small networks have been used for homogenisation of temperature data, 4 big networks have been applied for precipitation. Each of these networks covered one of the low level HISTALP-climate zones (Fig. 1) including some nearby stations of the neighbouring climate zone, for spatial consistency.

For temperature a correction factor for bias caused by a change of observation time will be applied for Austrian stations.
Instead of using the geographical situation of the stations to define the correction factor, a cluster method was applied. 3 Clusters were chosen and uncertainties were calculated by leaving out each station of the cluster once. (Fig.2)

monthly correction of observation time bias and uncertainties

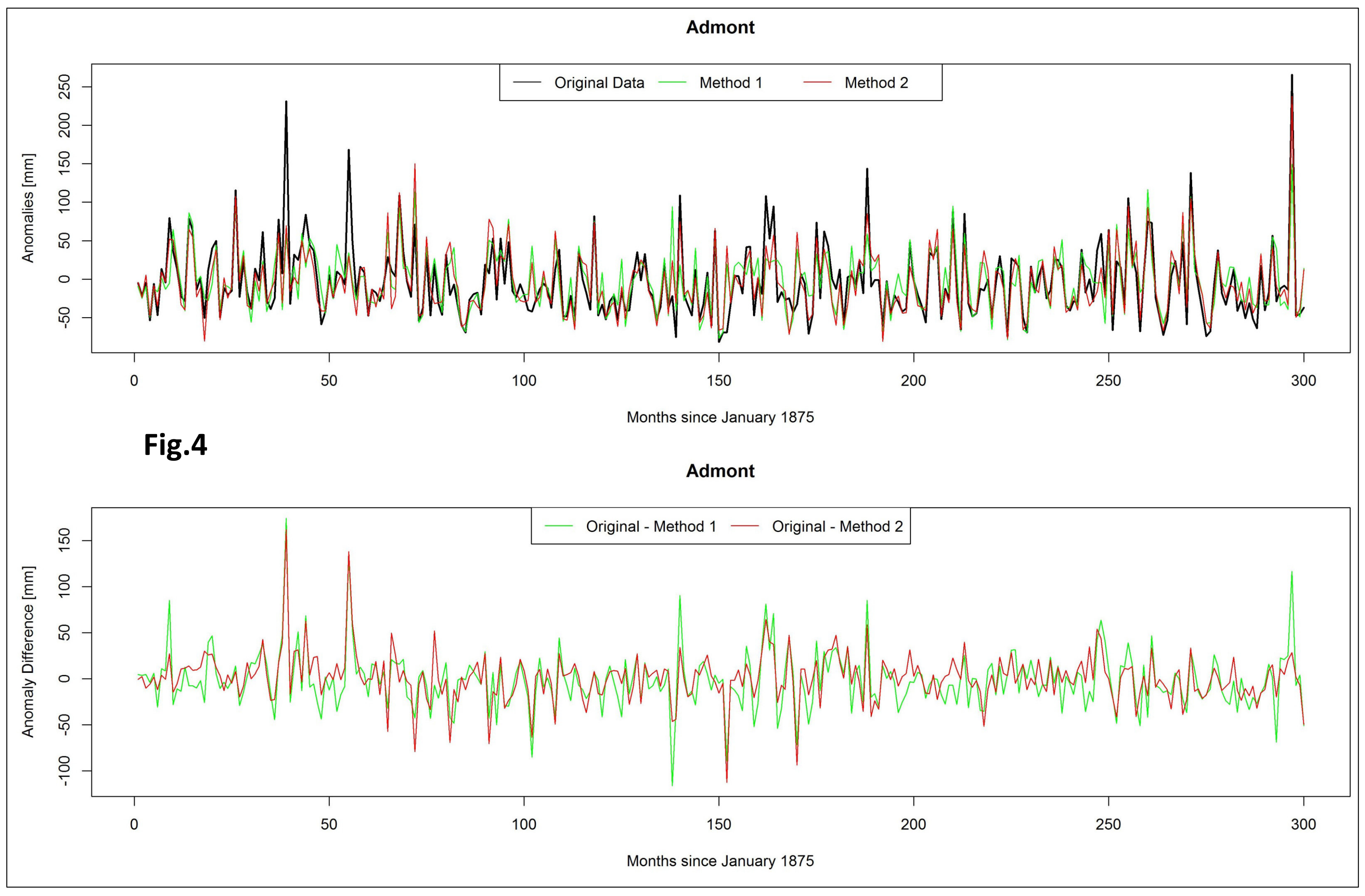


Differences between current means, original mean using 9pm measurement and corrected 9pm means



Daily differences caused by changes in observation time and mean calculation algorithm. Red: original differences, green using the adjustments given in Fig. 2, blue: using adjustments of 4 clusters

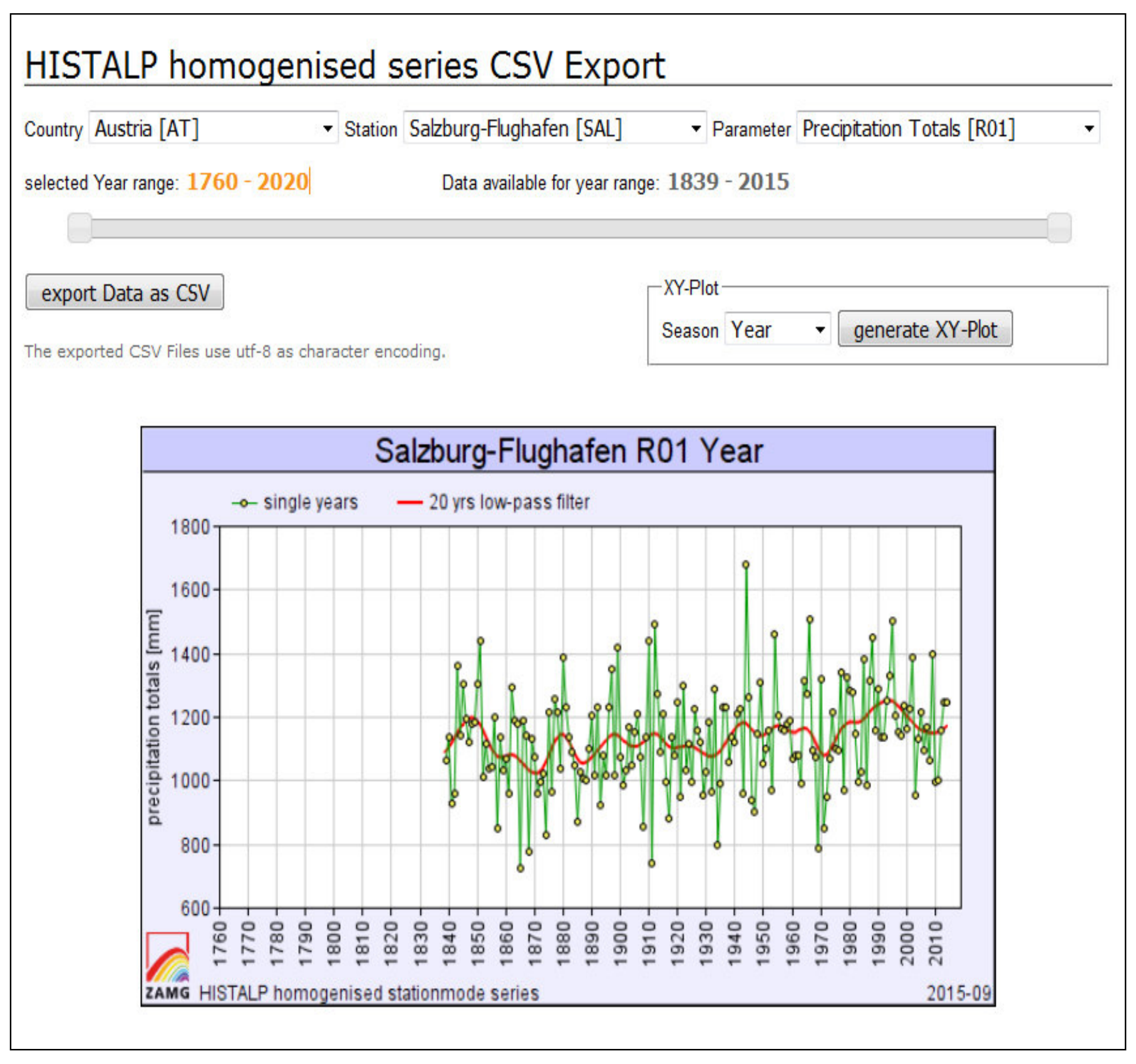
Different EOF-methods are tested for reconstruction of missing measurements at the beginning of the precipitation time series to create a temporal consistent dataset of monthly precipitation analyses for the whole GAR in order to create a set of monthly spatial analyses with a consistent stations dataset.



Literature:
A more complete list of citations can be found at : www.zamg.ac.at/histalp

Auer I, Böhm R, Jurkovic A, Lipa W, Orlik A, Potzmann R, Schöner W, Ungersböck M, Matulla C, Briffa K, Jones PD, Efthymiadis D, Brunetti M, Nanni T, Maugeri M, Mercalli L, Mestre O, Moisselin J-M, Begert M, Müller-Westermeier G, Kveton V, Bochnicek O, Stastny P, Lapin M, Szalai S, Szentimrey T, Cegnar T, Dolinar M, Gajic-Capka M, Zaninovic K, Majstorovic Z, Nieplova E, 2007. HISTALP – Historical instrumental climatological surface time series of the greater Alpine region 1760-2003. International Journal of Climatology 27: 17-46
Mestre O, Domonkos P, Picard F, Auer I, Robin S, Lebarbier E, Böhm R, Aguilar E, Guijarro J, Vertachnik G, Klancar M, Dubuisson B, Stepanek P, 2013: HOMER: a homogenization software – methods and applications, IDÖJÁRÁS- Quarterly Journal of the Hungarian Meteorological Service, 117, 47-67

Developments of Webpage

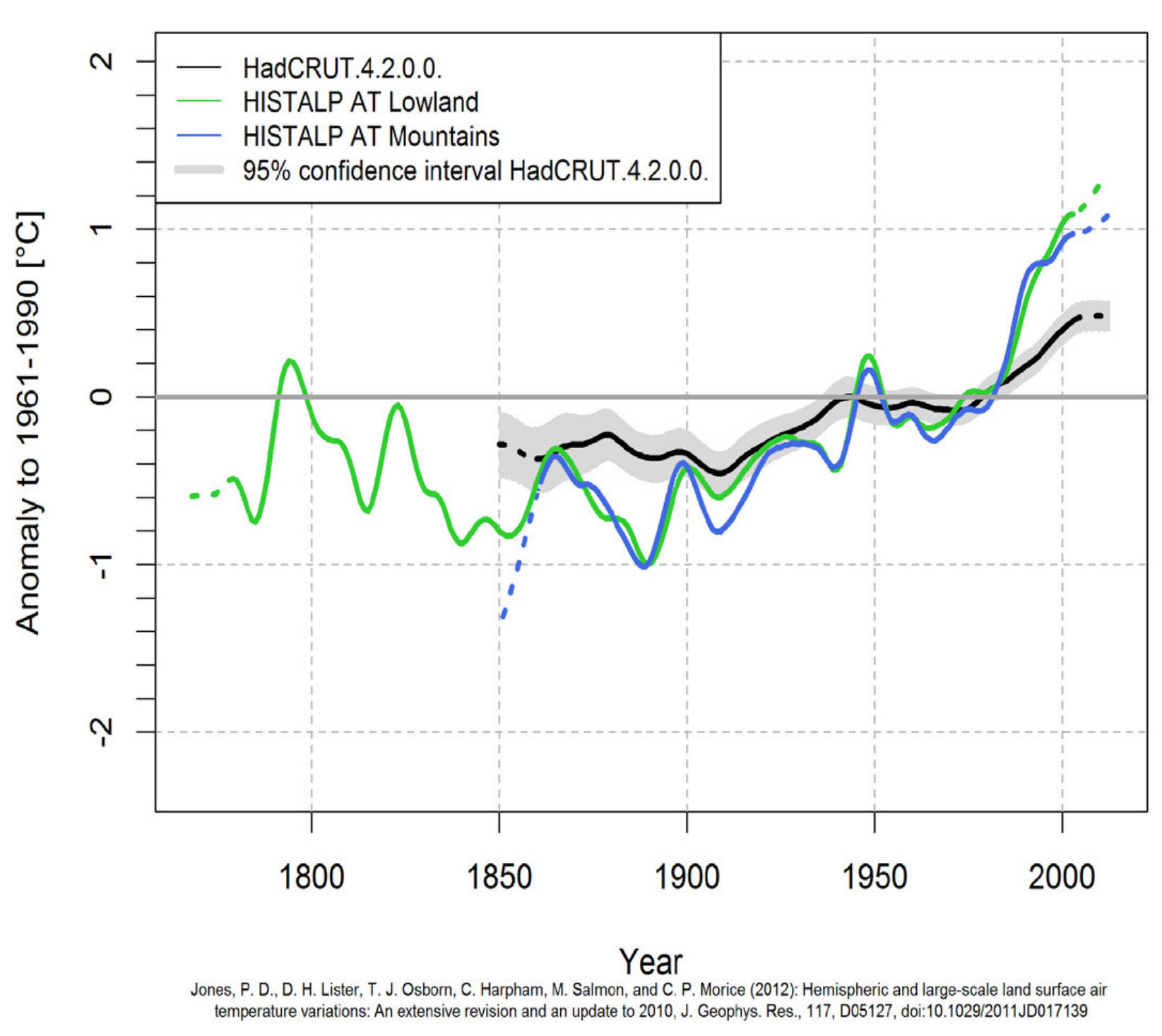


HISTALP homogenised series CSV Export

Country: Austria [AT]
Station: Salzburg-Flughafen [SAL]
Parameter: Precipitation Totals [R01]
selected Year range: 1760 - 2020
Data available for year range: 1839 - 2015
The exported CSV Files use utf-8 as character encoding.
XY-Plot
Season: Year

- Comparison between global climate evolution and Austrian temperature series were included.
- Improved data download area
- Enhanced user friendliness by FAQ-Area
- Reference period for anomalies in newsletter changed to 1961-1990

Annual temperature anomalies (20 year Gaussian filter)



Jones, P. D., D. H. Lister, T. J. Osborn, C. Harpham, M. Salmon, and C. P. Morice (2012): Hemispheric and large-scale land surface air temperature variations: An extensive revision and an update to 2010, J. Geophys. Res., 117, D05127, doi:10.1029/2011JD017139

Future Developments

- Homogenisation of sunshine duration and air pressure
- Recalculation of spatial analyses
- Including of daily data (Austria), but additional homogenised daily data from other HISTALP-areas highly welcome
- Including relative humidity and wind information

Development of long-term daily climate series from stations records across Belgium

O. Bleeckx, C. Delvaux, M. Journée and C. Bertrand

Royal Meteorological Institute of Belgium, Brussels, Belgium – Cedric.Bertrand@meteo.be

Objective

Take advantage of the digitization project "DIGITALISATION Programme-Phase 1" (financed by the Belgian Science Policy Office) ended in 2012 which has made available daily temperature and precipitation records since 1880 for a number of climatological stations in Belgium - records from 1950 and beyond being already archived on digital format at RMI -) to develop additional reference long-term daily climate series. Only one long-term temperature time series (i.e. Uccle) and a few more precipitation series (see Figs. 2 & 3) are currently available for the whole Belgian territory.

Station Locations and Period of Record

623 series of precipitation and 239 series of temperature have been digitized for the period 1880-1949 (which roughly represents five millions of keyed daily data) completing the 1950-2015 records (646 series of precipitation (RR) and 320 series of temperature (TT)). However, most stations have too little data, and some have excessive missing data, poor site characteristics or observations, or are otherwise unsuitable. Fig. 1 displays the evolution of the number of available daily observation for the period 1880-2015. Locations of RR (TT) stations which were at least in operation during 1 day over the time period 1880-2015 are given in Fig. 2 (3).

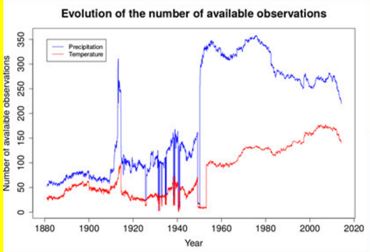


Fig.1: Evolution of the number of available observations per day for the period 1880-2015 in the RMI DB (RR in blue and TT in red)

Creation of long series

Data from about 1060 stations have been found over the time period 1880-2015 and placed in the RMI central database. During metadata processing, a first series of stations were selected for their completeness, length of record, and spatial distribution. Stations with more than 360 months (30 years) of data were taken into account and all individual stations that are located in close proximity to these selected stations (i.e., within a 8-km maximum radius) were considered as meteorologically similar to the selected one. At the end of this process, a total of **21 RR** and **14 TT** series were identified as potential historical long-time series.

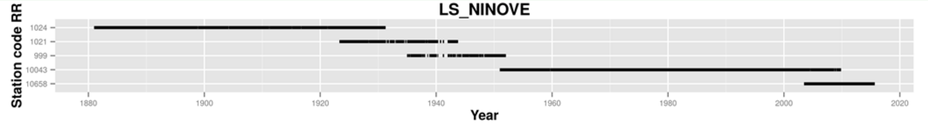


Fig. 4 : Example of long series construction (Ninove station, RR)

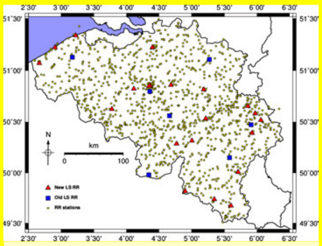


Fig. 2: Locations of RR stations active at least one day during the 1880 – 2015 period (existing long-series in blue and potential new long series in red).

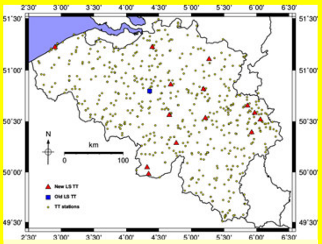


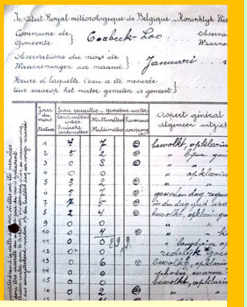
Fig. 3: Locations of TT stations active at least one day during the 1880 – 2015 period (existing long-series in blue and potential new long series in red).

Quality Control Procedures (I)

→**QC of recovered and digitized data:**

A first series of quality control tests and procedures were applied to the digitized data (i.e. 1880-1949) to ensure that the observations recorded on the original documents are accurately represented in the RMI DB. A two steps QC approach was implemented:

1. **Visual QC**: cross-comparison between the data source and the keyed data to verify the fidelity of digitization (e.g., typo errors – forgetting a coma, doubling a number, adding or forgetting a number, omission of the negative sign …-, keying in one element as another element, keying one hour as another hour, keying one hour as another hour, keying the date of the form, and shifting the day of the data up or down to an incorrect day, non-data keyed as a zero value, attributing the form to another station and/or another month, …)



Institut Royal météorologique de Belgique – Koninklijk Mee

Commune de
Gemeente } Corbeek-Loo

Observations du mois de
Waarnemingen der maand } Januari

Heure à laquelle l'eau a été mesurée
Uur waarop het water gemeten is geweest }

Fig.5: Example of form leading to keying error (Korbeek-Lo station)

→ More than 145000 values (16 % of the retained data) were rekeyed !

2. **Automatic QC**: known systematic errors (e.g., unit conversion , one day shift in precipitation amount, …) are flagged and corrected when possible

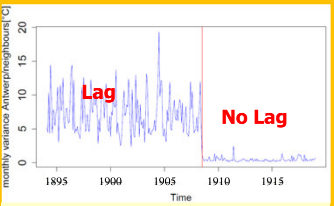


Fig. 6 : Example of one-day time lag correction in the RR series (Antwerp stations)

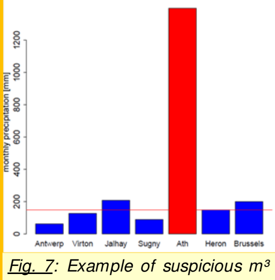


Fig. 7: Example of suspicious m^3 to mm conversion (Ath station, August 1938)

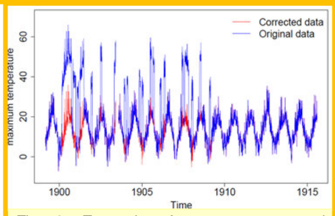


Fig. 8: Example of temperature unit (i.e. 30 °C offset) conversion error (Oostende – Abbe Dewitte-- station)

Quality Control Procedures (II)

Without outliers being property treated, homogenization and analysis may render misleading results. Automated quality assessment procedures have been developed and are applied to all stations records (i.e. from 1880 to 2015) to isolate and flag potentially errant values as well as for ensuring internal consistency and temporal and spatial coherence of the data

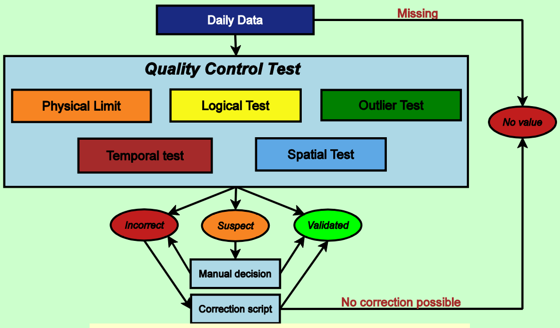


Fig. 9: Sketch of the automated QA procedures

- ***Physical Limit*** : -40 °C < Temperature < 50 °C
  0 mm ≤ Precipitation < 250 mm
- ***Logical Test*** : max temperature ≥ min temperature
- ***Temporal Test*** : analyze the rate of daily temperature change in order to detect anomalies
- ***Spatial Test*** : comparison using neighboring values. Threshold depends on the network density
- ***Outlier Test*** : Temperature must be included within a lower and upper bounds, using an expectancy envelope which assume that the annual temperature variations follow a sinusoidal wave

Conclusions and perspective

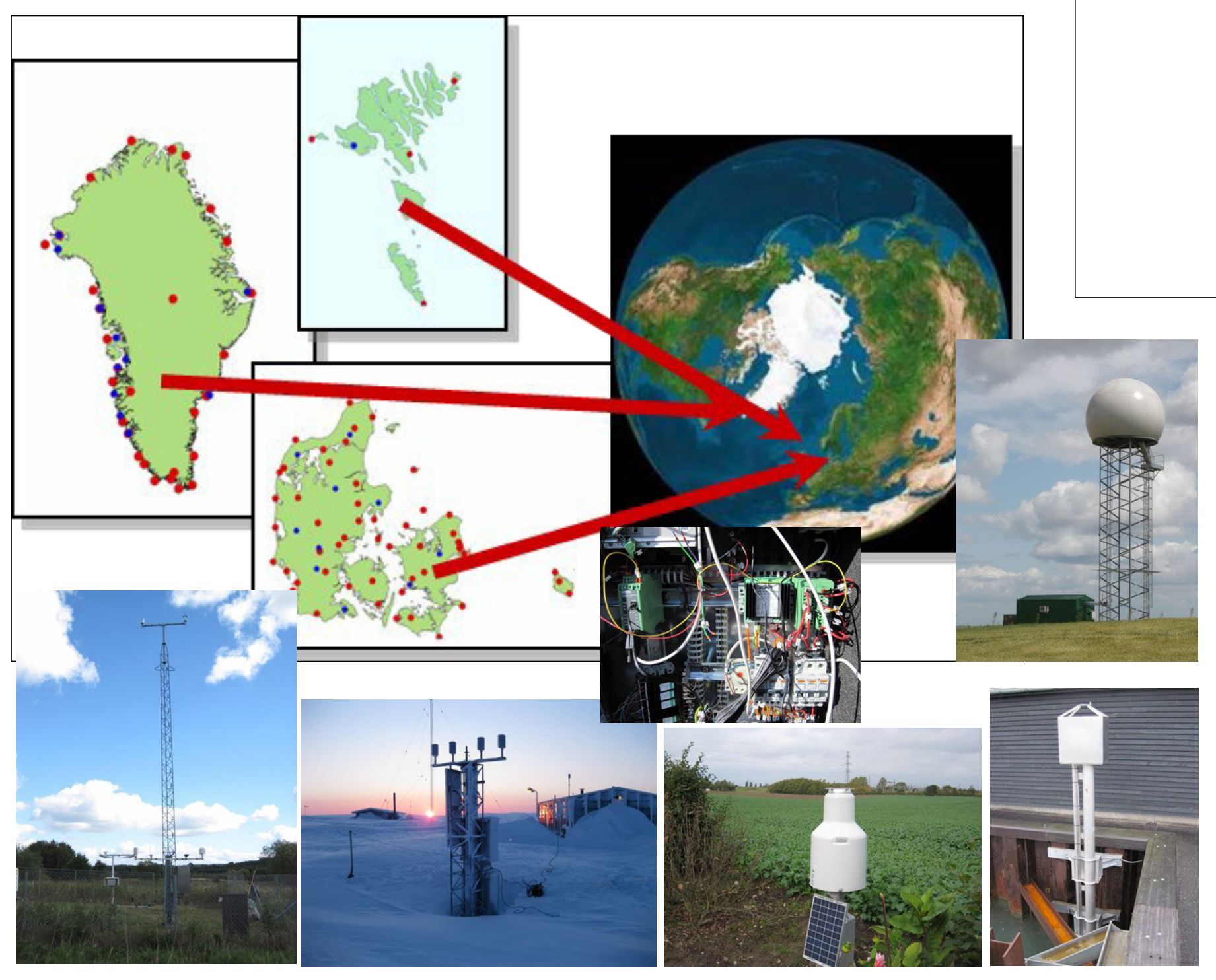
Thanks to a recent digitization project and fruitful excavations in the RMI's archive, it was possible to identify 21 RR and 14 TT potentially new long-term daily time series TT in Belgium. When the QC of the new historical time series will be fully completed, all the series (i.e. "old" and "new" - 28 RR and 15 TT -) will be homogenized first on a monthly basis using the HOMER software and in a second time on a daily basis if realistic.

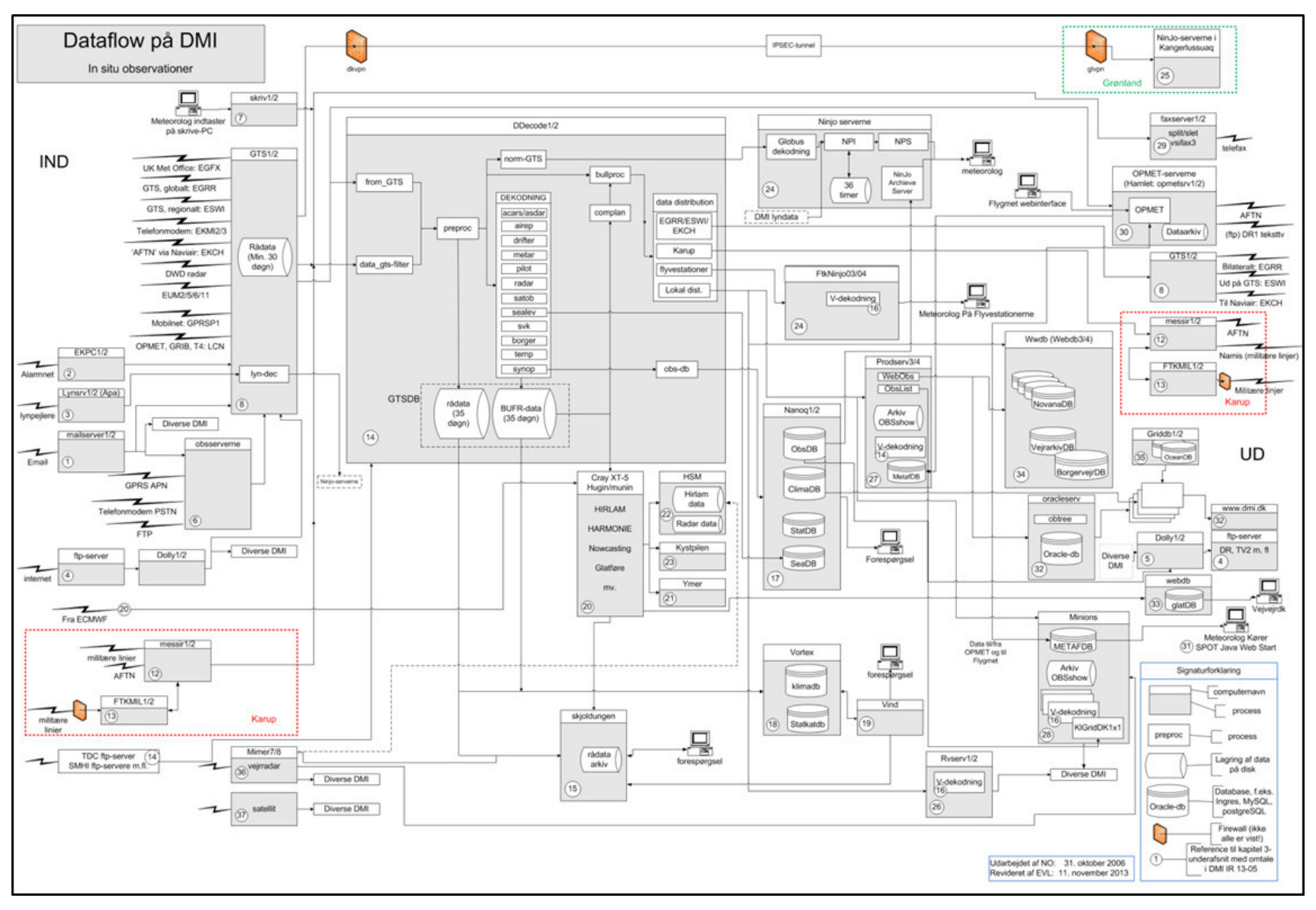
PAST CLIMATE MONITORING AT THE DANISH METEOROLOGICAL INSTITUTE (DMI)

SENIOR CLIMATOLOGIST JOHN CAPPELEN, WEB: research.dmi.dk/staff/all-staff/jc/ EMAIL: jc@dmi.dk

10th EUMENET DATA MANAGEMENT WORKSHOP, ST. GALLEN, SWITZERLAND 28-30 OCTOBER 2015

Observation of weather/ sea and data flow





At the very start...

In 1751 meteorological observations in the Round Tower in Copenhagen were initiated.

From 1751 to 1766 the thermometer was situated inside a room in the little observatory on the top of the tower (!) In 1767 the thermometer was situated outside the observatory facing north.

DMI are in possession of the meteorological observations from the Round Tower from the period 1767 – 1819. From 1818 meteorological observations from the old Botanical Garden in Copenhagen also are available.

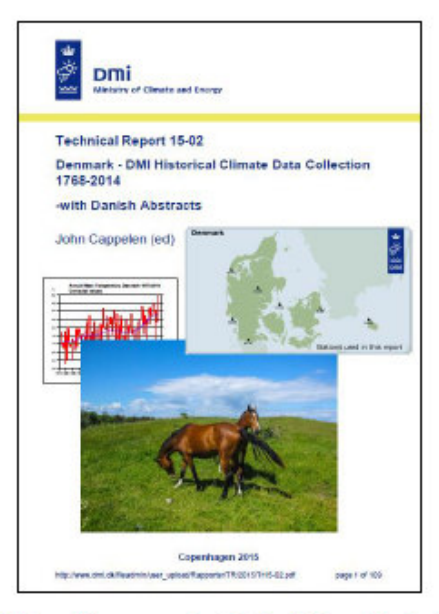
Source: Ane Marie Closter, Rikke Margrethe Closter, John Cappelen, Jens Hesselbjerg Christensen, Kirsten Christoffersen (FBL) and Claus Kern-Hansen (2006):
Temperature measurements in Copenhagen from 1767 to 1860. DMI Technical Report No. 06-13.
Link:
http://www.dmi.dk/dmi/tr06-13.pdf
http://www.dmi.dk/dmi/rundetaarn_data.zip

1760's marked the beginning of the very first systematic observations in Greenland….

From 1784 a composite SW Greenland temperature series is processed….not least because of H.H. Lamb.

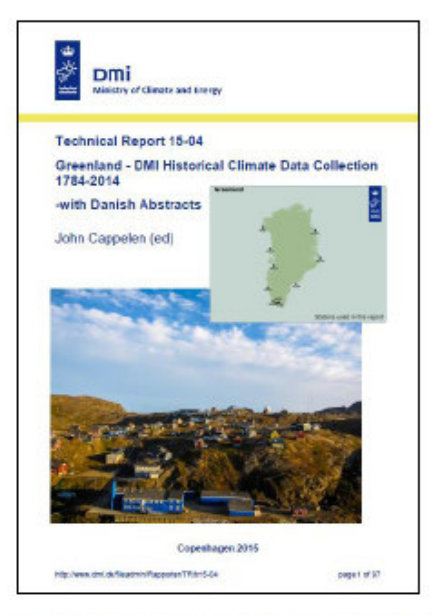
Historical Climate data collections – long term series

Denmark, the Faroe Islands and Greenland. Selected stations/parameters, Observations, Daily, Monthly, Yearly, Country-wise (DK), Storms (DK), Descriptions of the Climate in text



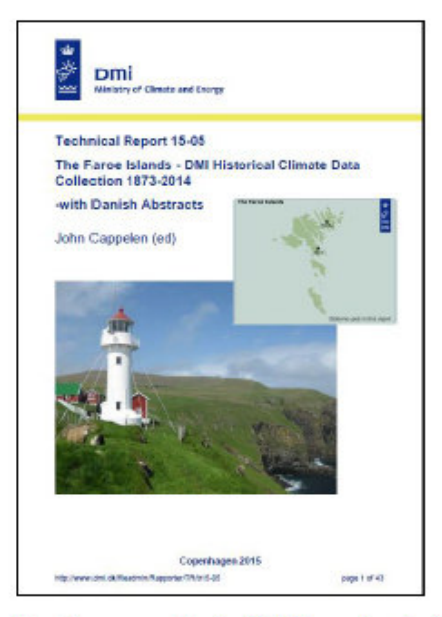
Technical Report 15-02
Denmark - DMI Historical Climate Data Collection 1768-2014
-with Danish Abstracts
John Cappelen (ed)

http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-02.pdf (report)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15_02.zip (data)



Technical Report 15-04
Greenland - DMI Historical Climate Data Collection 1784-2014
-with Danish Abstracts
John Cappelen (ed)

http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-04.pdf (report)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15_04.zip (data)



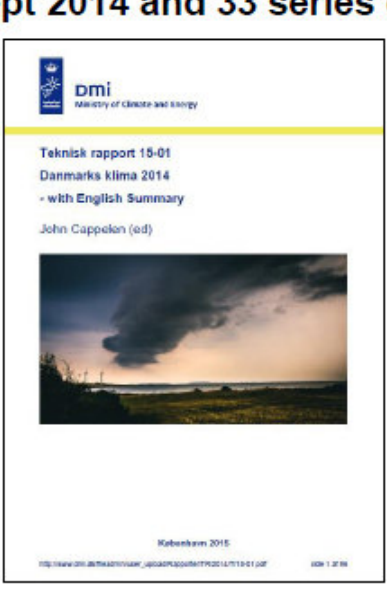
Technical Report 15-05
The Faroe Islands - DMI Historical Climate Data Collection 1873-2014
-with Danish Abstracts
John Cappelen (ed)

http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-05.pdf (report)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-05.zip (data)

Other important annual reports

The annual climate of Denmark.

Extreme value analysis of daily and five days accumulated precipitation for 38 selected series from 1872. Five series contains data within the whole period 1872-sept 2014 and 33 series contains data within the period 1961-sept 2014.

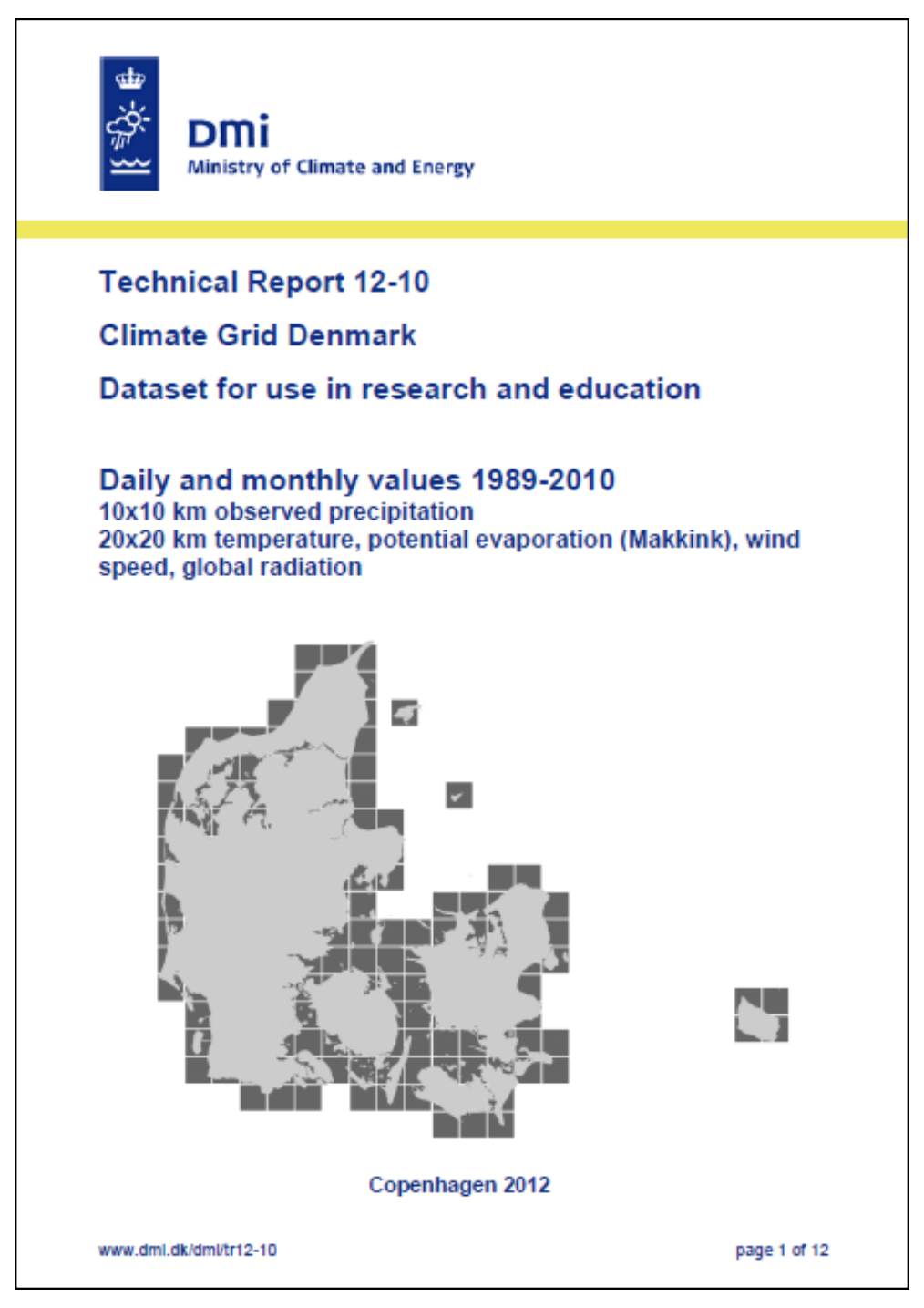


Teknisk rapport 15-01
Danmarks klima 2014
- with English Summary
John Cappelen (ed)

http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/Tr15-01.pdf (report)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-01.zip (data)

Teknisk rapport 14-15
Ekstremnedbør i Danmark 1872-september 2014 og 1961-september 2014
- Ekstremværdianalyse af døgnnedbør og femdøgnsum i Danmark
1872-september 2014 (5 stationer)
1961-september 2014 (33 stationer)
John Cappelen og Peter Grunnet Wang

http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2014/TR14-15.pdf (report)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2014/tr14-15.zip (data)



DMI
Ministry of Climate and Energy

Technical Report 12-10
Climate Grid Denmark
Dataset for use in research and education

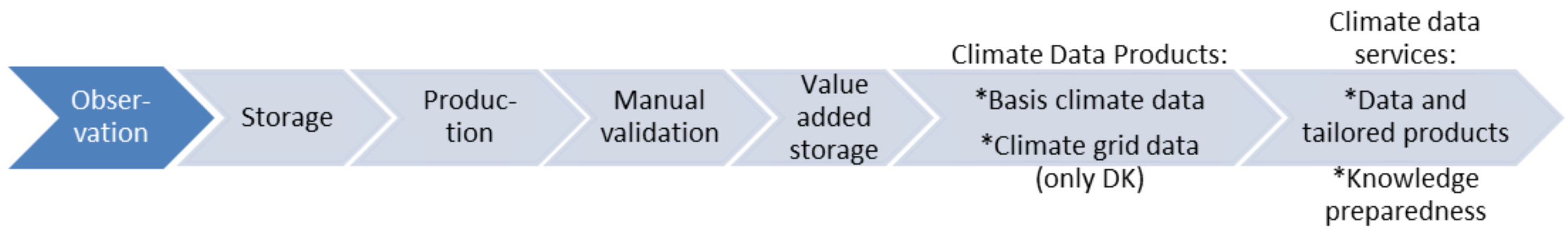
Daily and monthly values 1989-2010
10x10 km observed precipitation
20x20 km temperature, potential evaporation (Makkink), wind speed, global radiation

Copenhagen 2012

www.dmi.dk/dmi/tr12-10 page 1 of 12

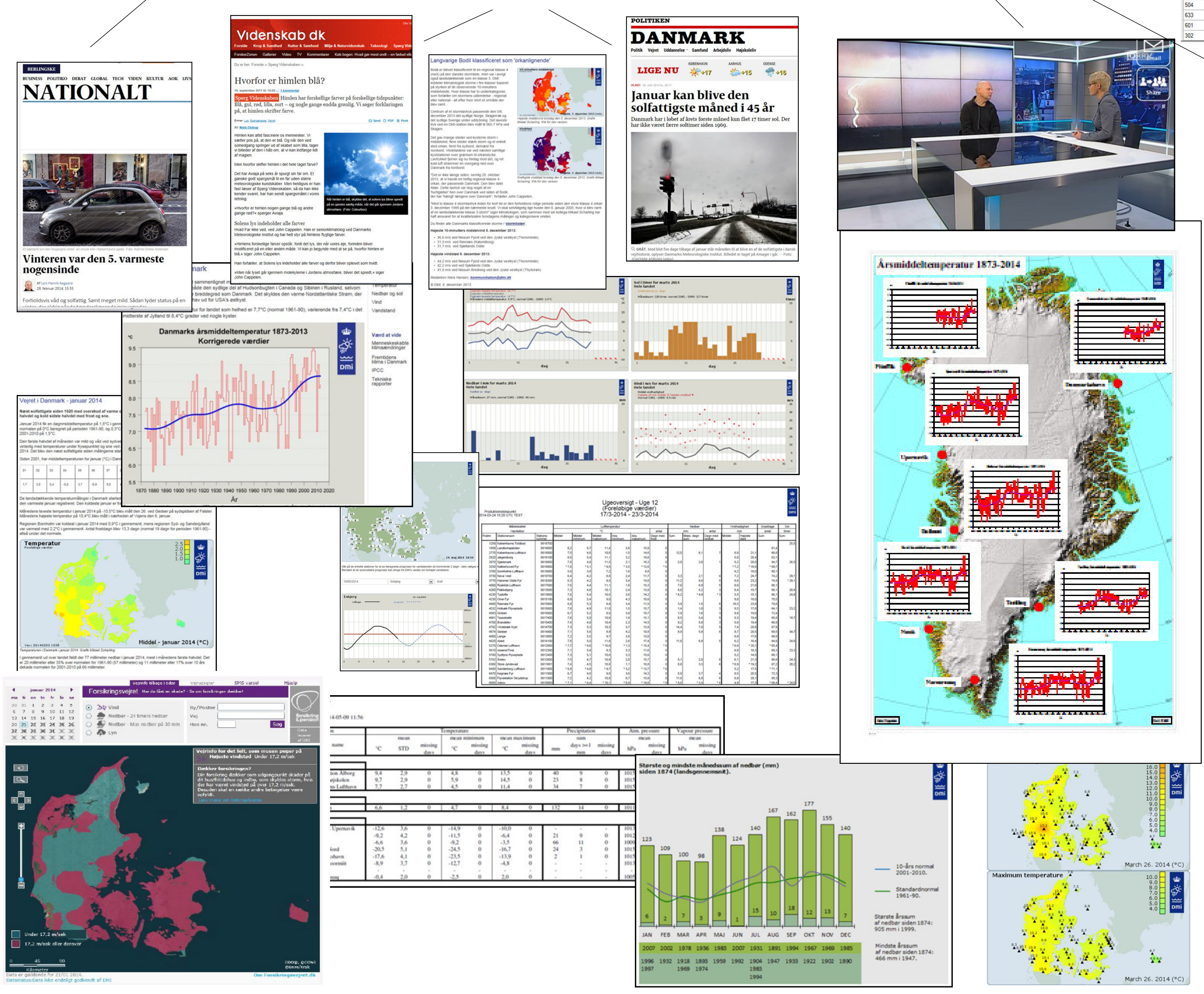
Monitoring the climate; Denmark, Greenland and the Faroe islands

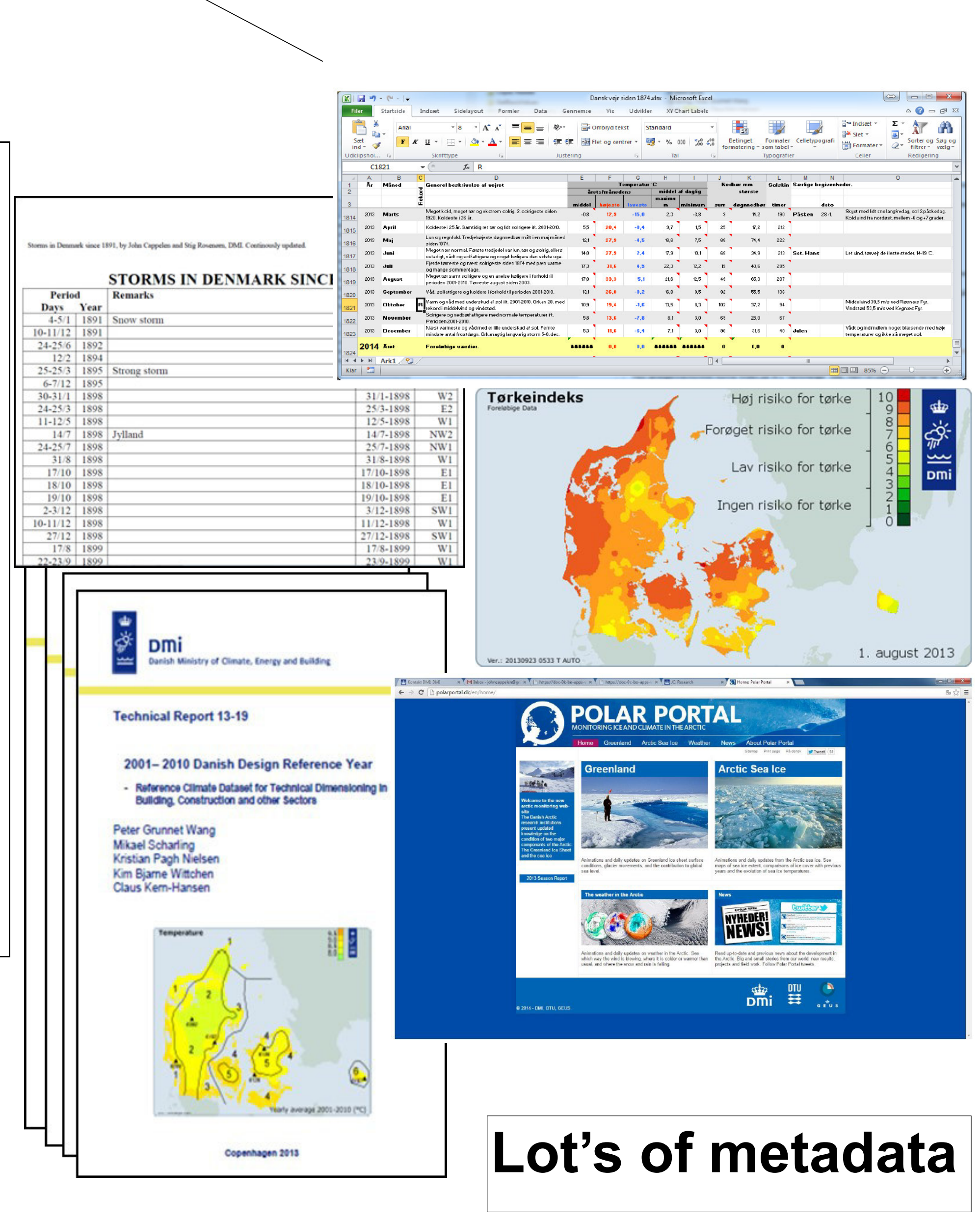
- Climate data production
- Climate data services



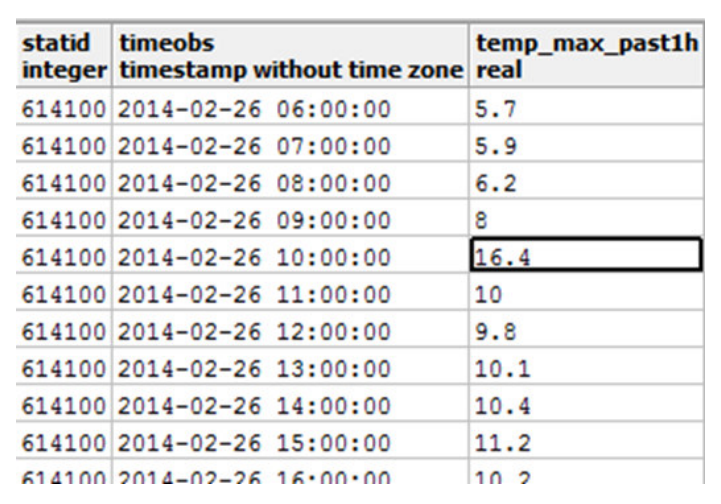
Obser-vation → Storage → Produc-tion → Manual validation → Value added storage → Climate Data Products: *Basis climate data *Climate grid data (only DK) → Climate data services: *Data and tailored products *Knowledge preparedness

Climate monitoring and services

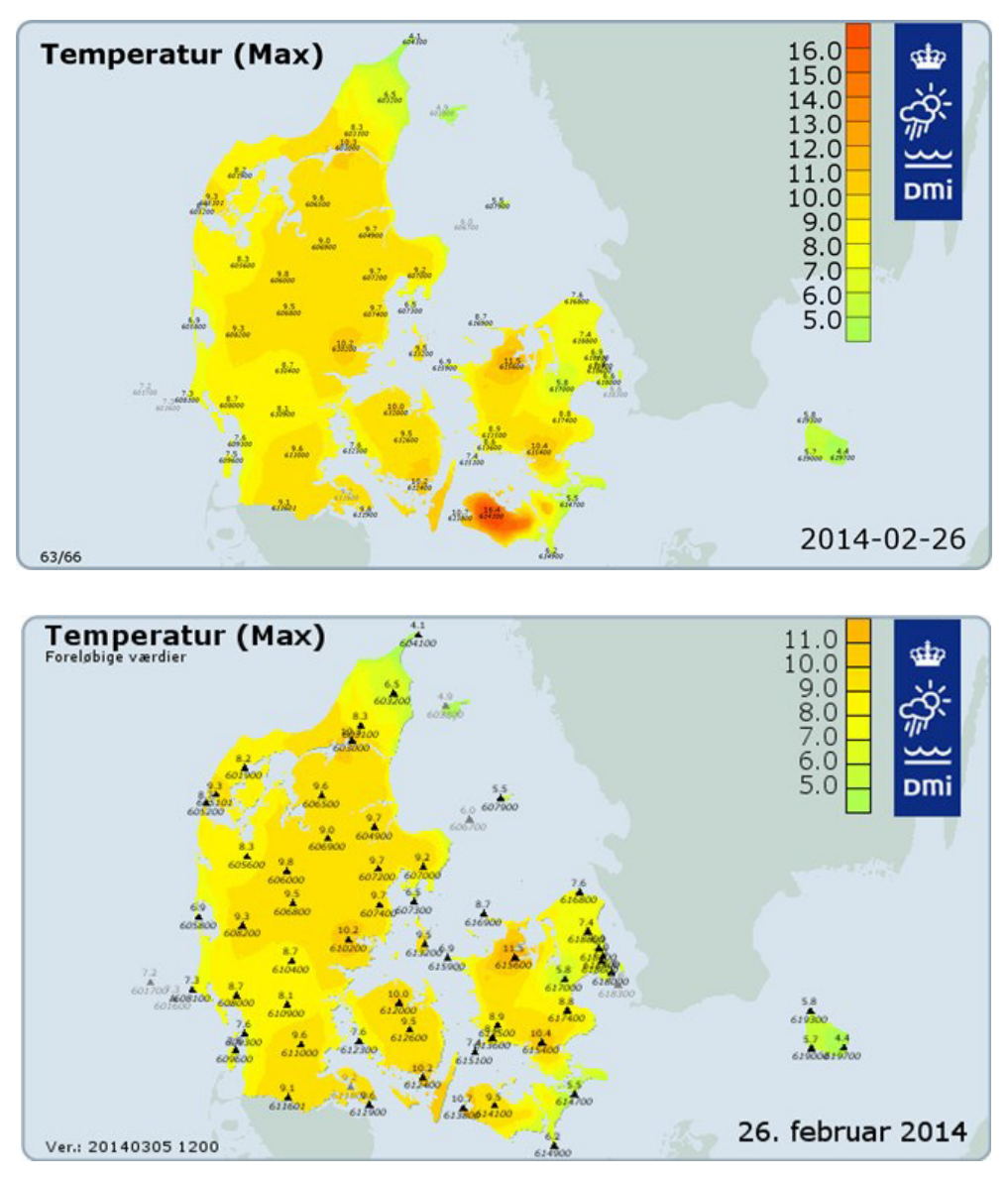


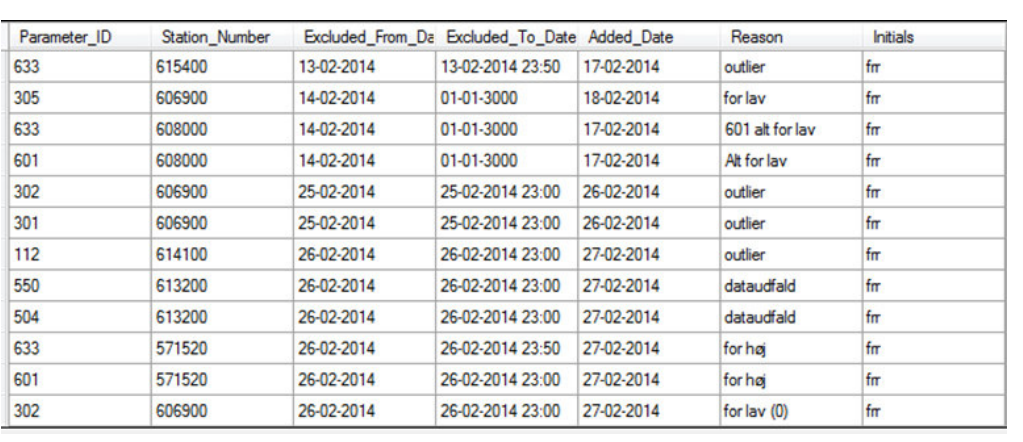


Manual validation of climate data



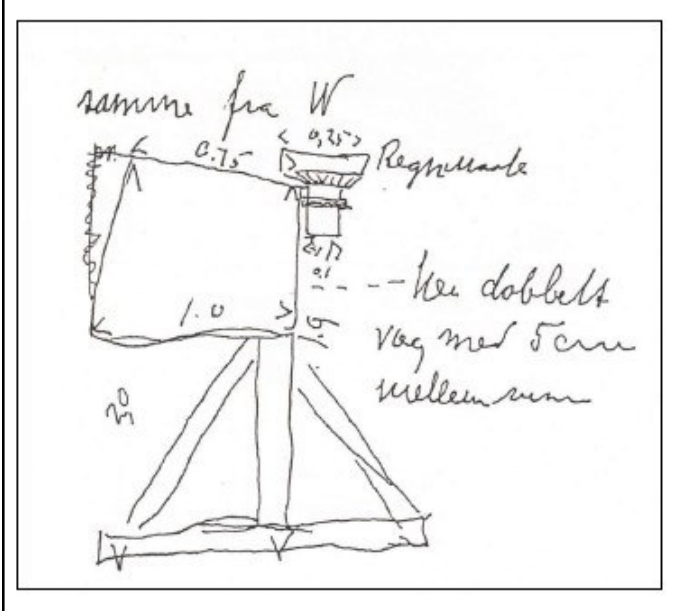
statid integer	timeobs timestamp without time zone	temp_max_past1h real
614100	2014-02-26 06:00:00	5.7
614100	2014-02-26 07:00:00	5.9
614100	2014-02-26 08:00:00	6.2
614100	2014-02-26 09:00:00	8
614100	2014-02-26 10:00:00	16.4
614100	2014-02-26 11:00:00	10
614100	2014-02-26 12:00:00	9.8
614100	2014-02-26 13:00:00	10.1
614100	2014-02-26 14:00:00	10.4
614100	2014-02-26 15:00:00	11.2
614100	2014-02-26 16:00:00	10.2





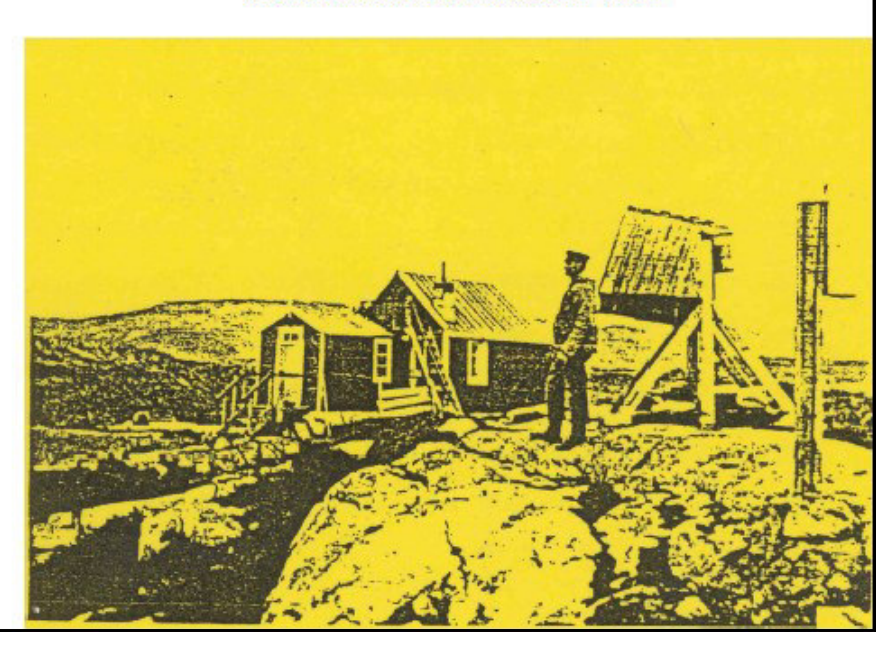
Parameter_ID	Station_Number	Excluded_From_Date	Excluded_To_Date	Added_Date	Reason	Initials
633	615400	13-02-2014	13-02-2014 23:50	17-02-2014	outlier	frr
305	606900	14-02-2014	01-01-3000	18-02-2014	for lav	frr
633	608000	14-02-2014	01-01-3000	17-02-2014	601 alt for lav	frr
601	608000	14-02-2014	01-01-3000	17-02-2014	Alt for lav	frr
302	606900	25-02-2014	25-02-2014 23:00	26-02-2014	outlier	frr
301	606900	25-02-2014	25-02-2014 23:00	26-02-2014	outlier	frr
112	614100	26-02-2014	26-02-2014 23:00	27-02-2014	outlier	frr
550	613200	26-02-2014	26-02-2014 23:00	27-02-2014	dataudfald	frr
504	613200	26-02-2014	26-02-2014 23:00	27-02-2014	dataudfald	frr
633	571520	26-02-2014	26-02-2014 23:50	27-02-2014	for høj	frr
601	571520	26-02-2014	26-02-2014 23:00	27-02-2014	for høj	frr
302	606900	26-02-2014	26-02-2014 23:00	27-02-2014	for lav (0)	frr

Lot's of metadata



Thermometer and raingauge arrangement July 1923 in Jacobshavn, Greenland

Observer Guldager by his instruments. Jacobshavn 1923



Different types of data

- Observation - "raw" observed data
- Basis climate data - 1 hourly (basis) quality checked data, derived values as daily, monthly, annual, reference etc., complete series for Denmark (missing and faulty data replaced by interpolated data). For Greenland and the Faroes faulty data are removed.
- Climate grid data - 1 hourly (basis) and derived in time and resoulution quality checked interpolated data 1x1 km

Overview: DMI Climate data availability

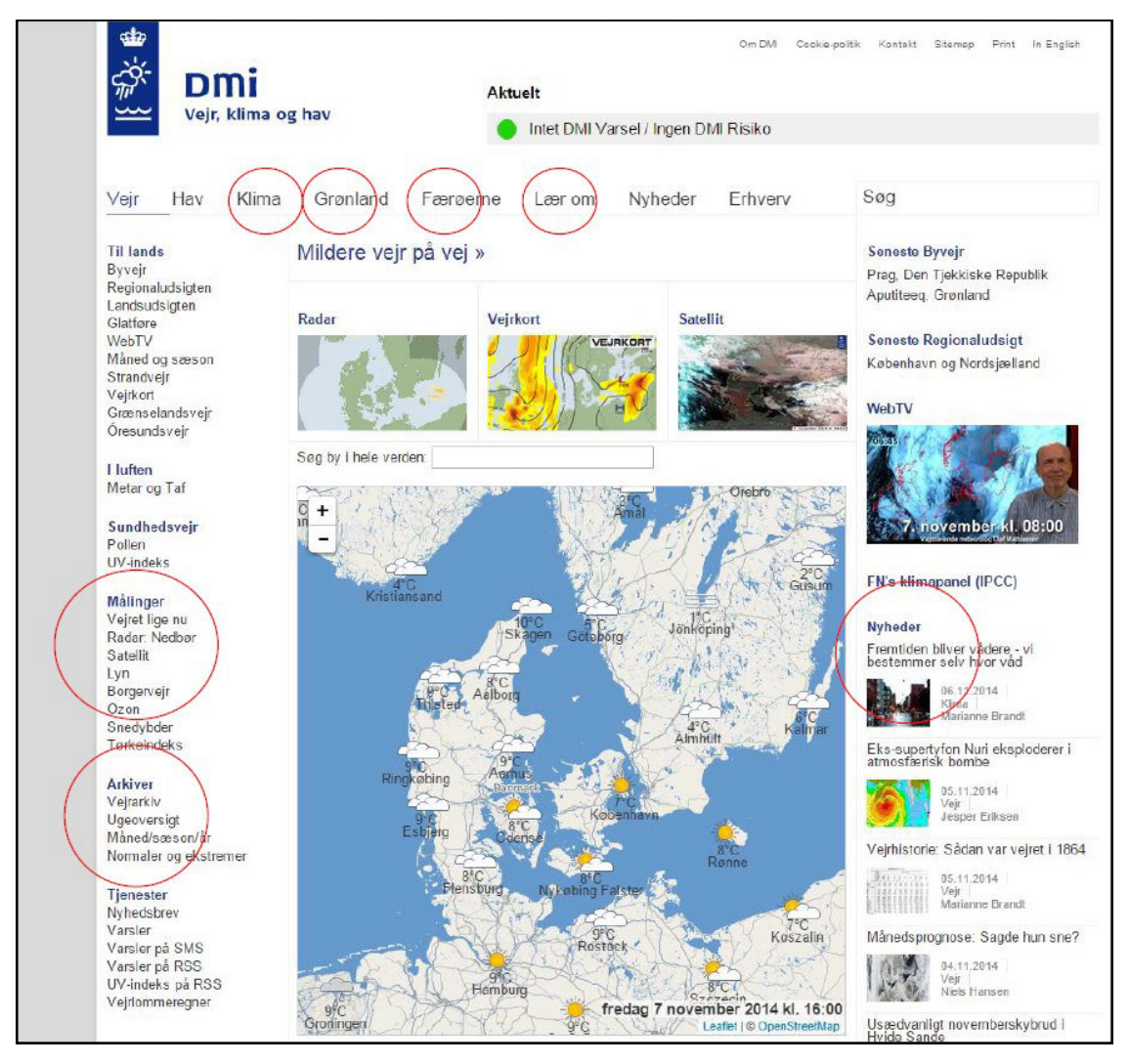
Before 1873: Few series: The Round Tower/Old Botanical Garden, Cph (from 1768) and composite SW Greenland monthly temp. series (from 1784)

From 1873:
- -Annual series, 11 parameters from several locations DK, FR, GR
- -Monthly series, 11 parameters from several locations DK, FR, GR
- -Daily series, 5 parameters from several locations DK, FR, GR
- -Observations of atmospheric air pressure (msl) (WASA data sets)
- -Country DK monthly and annual series of temperature, preciptitation, sunshine and cloud cover, extremes
- -Storm DK list and classification of storms

From 1953: Faroe Islands Dataset (Tórshavn) available for the public
From 1958: Greenland Dataset (all stations) available for the public
From 1958: Digitised observations in the climate databases at DMI
From 1961: The climate of Greenland, The Faroes Islands and Denmark in general with reference values 1961-90, monthly data app. 1961-1999 for many parameters, stations and country DK
From 1989: Climate grid values 10x10 and 20x20 km (selected values)
From 2001: Reference values 2001-2010
From 2001: Climate grid values 1x1 km (from 2001 selected values; from 2011 many values)
From 2001: Basis data DK in the climate databases at DMI (from 2001 selected values; from 2011 many values)
From 2014: Basis data NA (North Atlantic) in the climate databases at DMI

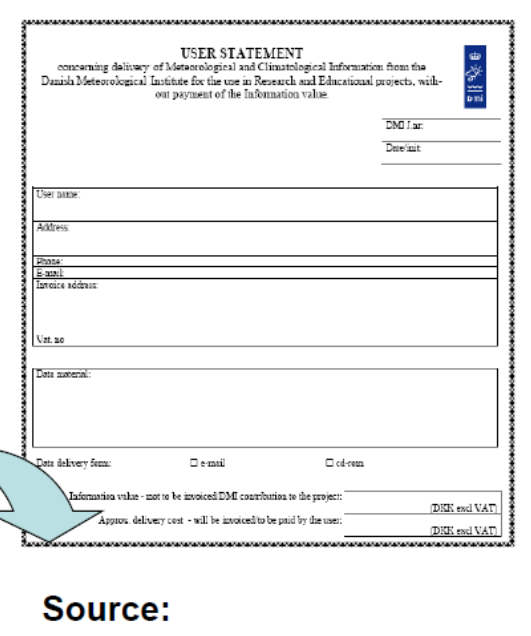
Climatic Data Guide and Data Policy exist.

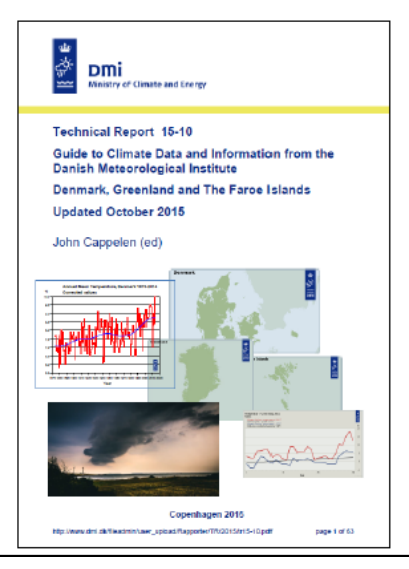
Climatic Data Guide, Data Policy and how to find climate information at dmi.dk



Besides data open to the public other data and services are available for a fee.

DMI provides free (besides delivery costs) and unrestricted access to meteorological data and products for educational and research purposes, if some conditions are met.





Technical Report 15-10
Guide to Climate Data and Information from the Danish Meteorological Institute
Denmark, Greenland and The Faroe Islands
Updated October 2015
John Cappelen (ed)

Source:
DMI Technical Report No. 15-10
John Cappelen (ed), 2015:
Guide to Climate Data and Information from the Danish Meteorological Institute, Updated October 2015

Link (latest version) :
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-10.pdf

In general: The Climate of Greenland, The Faroe Islands and Denmark

Monthly time series app. 1961-1999 and climatological standard normals 1961-90

- •Large number of stations
- •Many parameters
- •Descriptions of the Climate in text

http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/1999/tr99-5.pdf
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/1999/tr99-5.zip

Example:
John Cappelen et al. (2000):
The Observed Climate of Greenland, 1958-99 - with Climatological Standard Normals, 1961-90.
DMI Technical Report No. 00-18.

And many more….

http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/1998/tr98-14.pdf
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/1998/tr98-14.zip

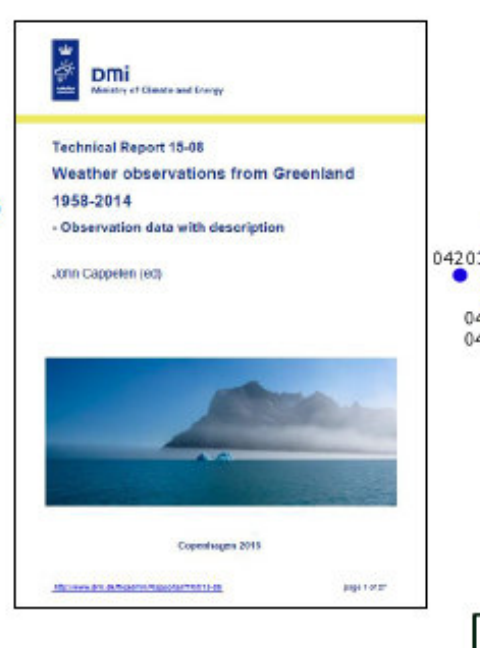
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2000/tr00-18.pdf
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2000/tr00-18-data_files.zip
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2000/tr00-18-wind_roses.zip

Greenland dataset

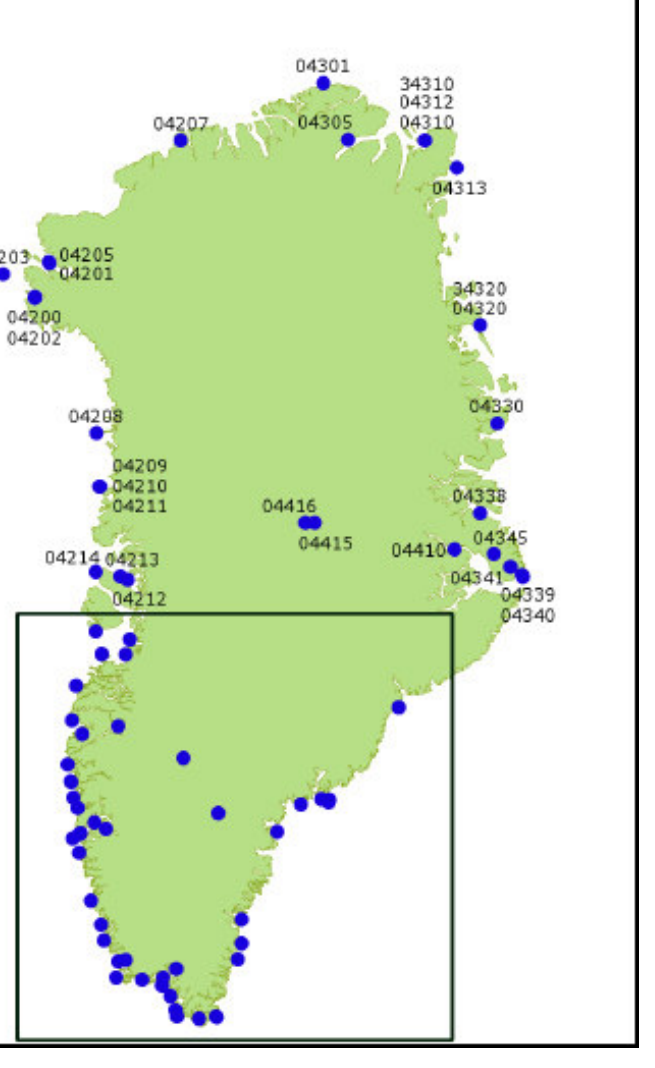
Greenland Dataset 1958-2014 available for the public, free of charge

Quality controlled observations of:

Dry bulb, max. and min. temperatures
Precipitation
Wind direction
Wind Speed
Cloud cover
Cloud Cover
Air pressure
Humidity
Depth of snow



Technical Report 15-08
Weather observations from Greenland 1958-2014
- Observation data with description
John Cappelen (ed)



Source:
DMI Technical Report No. 15-08
John Cappelen (ed), 2015:
Weather observations from Greenland 1958-2014, Observation data with description
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-08.pdf (report)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-08_new_dataformat_2014.zip (data1)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-08_old_dataformat_1958_2013.zip (data2)

Tórshavn, The Faroe Islands dataset

Tórshavn Dataset 1953-2014 available for the public, free of charge

Quality controlled observations of:

Dry bulb, max. and min. temperatures
Precipitation
Sunshine/Radiation
Wind direction
Wind Speed
Cloud cover
Air pressure
Humidity
Depth of snow

Source:
DMI Technical Report No. 15-09
John Cappelen, 2015:
Weather observations from Tórshavn, The Faroe Islands 1953-2014, Observation data with description
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-09.pdf (report)
http://www.dmi.dk/fileadmin/user_upload/Rapporter/TR/2015/tr15-09.zip (data)

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Office of Meteorology and Climatology MeteoSwiss

UNIVERSITÄT BERN
OESCHGER CENTRE
CLIMATE CHANGE RESEARCH

Senamhi

Homogenization of temperature and precipitation measurements in sparse station networks applying HOMER

S. Gubler[1], S. Hunziker[2], M. Begert[1], C. Oria[3], M. Croci-Maspoli[1], T. Konzelmann[1], S. Brönnimann[2]
[1]Swiss Federal Office of Meteorology and Climatology MeteoSwiss, Zurich, Switzerland
[2]Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Bern, Switzerland
[3]Meteorological and Hydrological Service of Peru (SENAMHI), Lima, Peru

1) Challenge: Homogenization of sparse station networks in the Peruvian Andes

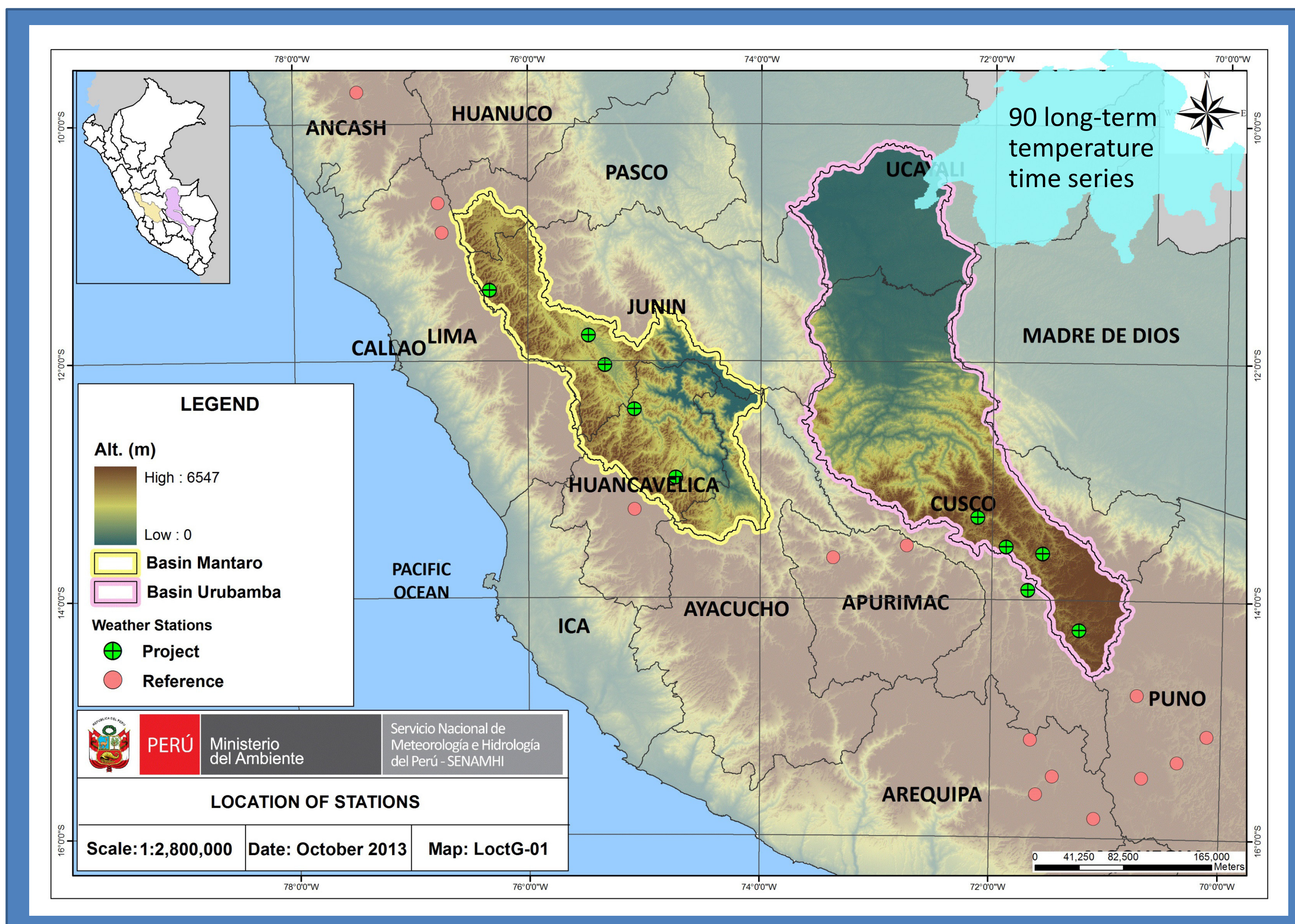


Fig. 1: Map of pilot regions of the CLIMANDES project (Urubamba and Mantaro river basins). Only few long-term (1964-today) high-quality stations exist in the pilot regions (green dots). To increase the station number for relative homogenization, the stations from the pilot regions are complemented with stations from surrounding regions (red dots).

Motivation

The aim of the project CLIMANDES is to homogenize stations in two pilot regions in the Peruvian Andes, namely the Mantaro and Urubamba basins (Fig. 1). The low stations density in and surrounding the regions however challenges relative homogenization. In this study, a "thinned" Swiss network (Fig. 3) was used to mimic conditions in Peru to investigate the performance of homogenization in low station density networks using HOMER.
Trends of the raw data in and surrounding the pilot regions are depicted in Fig. 2.

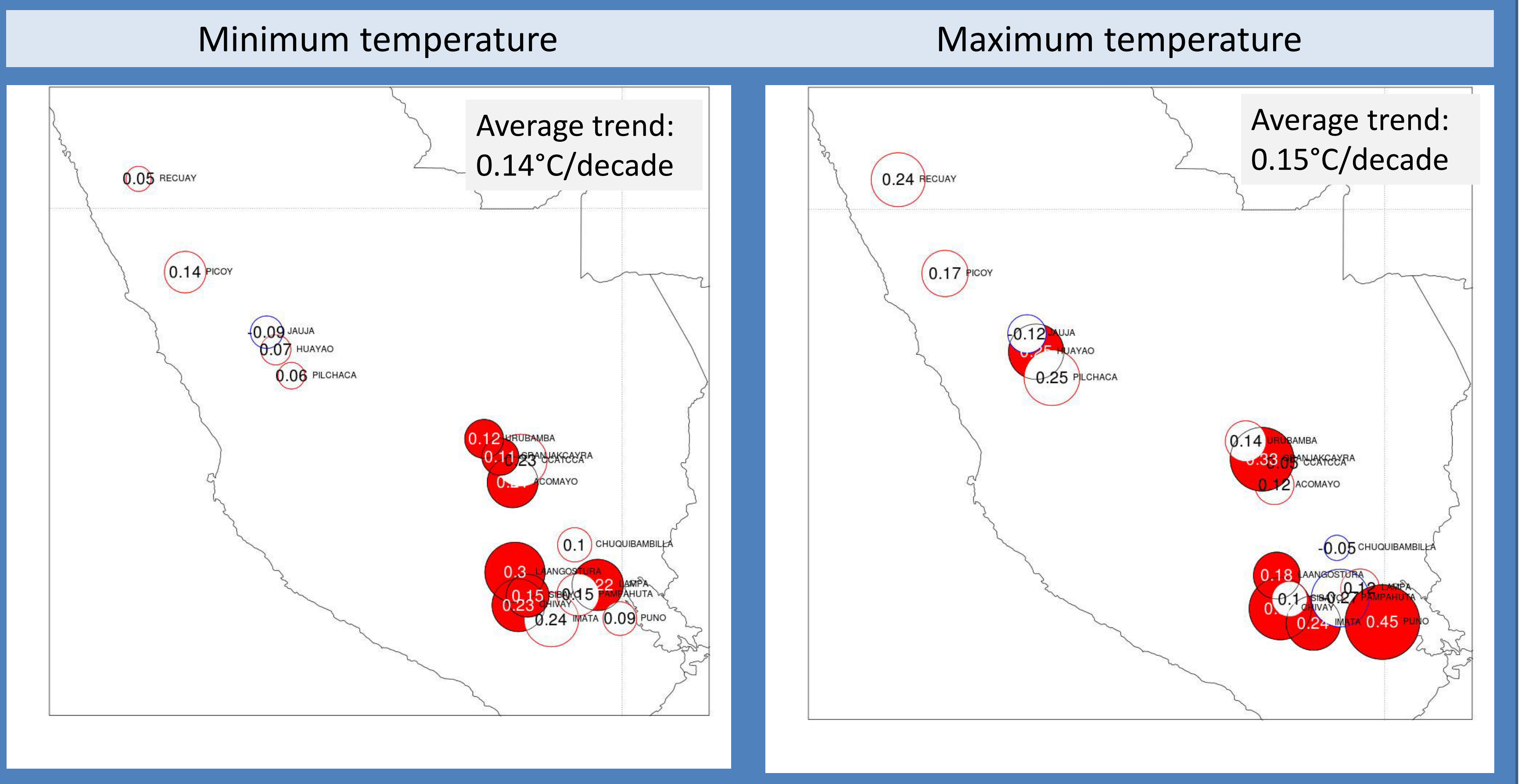


Fig. 2: Trends [°C/decade] of the raw data in and surrounding the pilot regions. The size of the circle indicate the magnitude of the trend. Filled circles are significant on the 5%-level.

2) Approach: Simulating the performance of homogenization on an artificially thinned Swiss network

Homogenization of sparse networks using HOMER

Comparison of homogenization in dense and sparse networks simulated on an artificially thinned Swiss network (Fig. 3) using HOMER. The study was done for 4 different settings to investigate the influence of human interaction and the use of metadata on homogenization:

1. Automatic run (auto)
2. Manual run without metadata (manu)
3. Inclusion of metadata a-priori regardless of statistical break detection (meta-pre)
4. Inclusion of metadata to confirm or reject potential breakpoints a posteriori (meta-post)

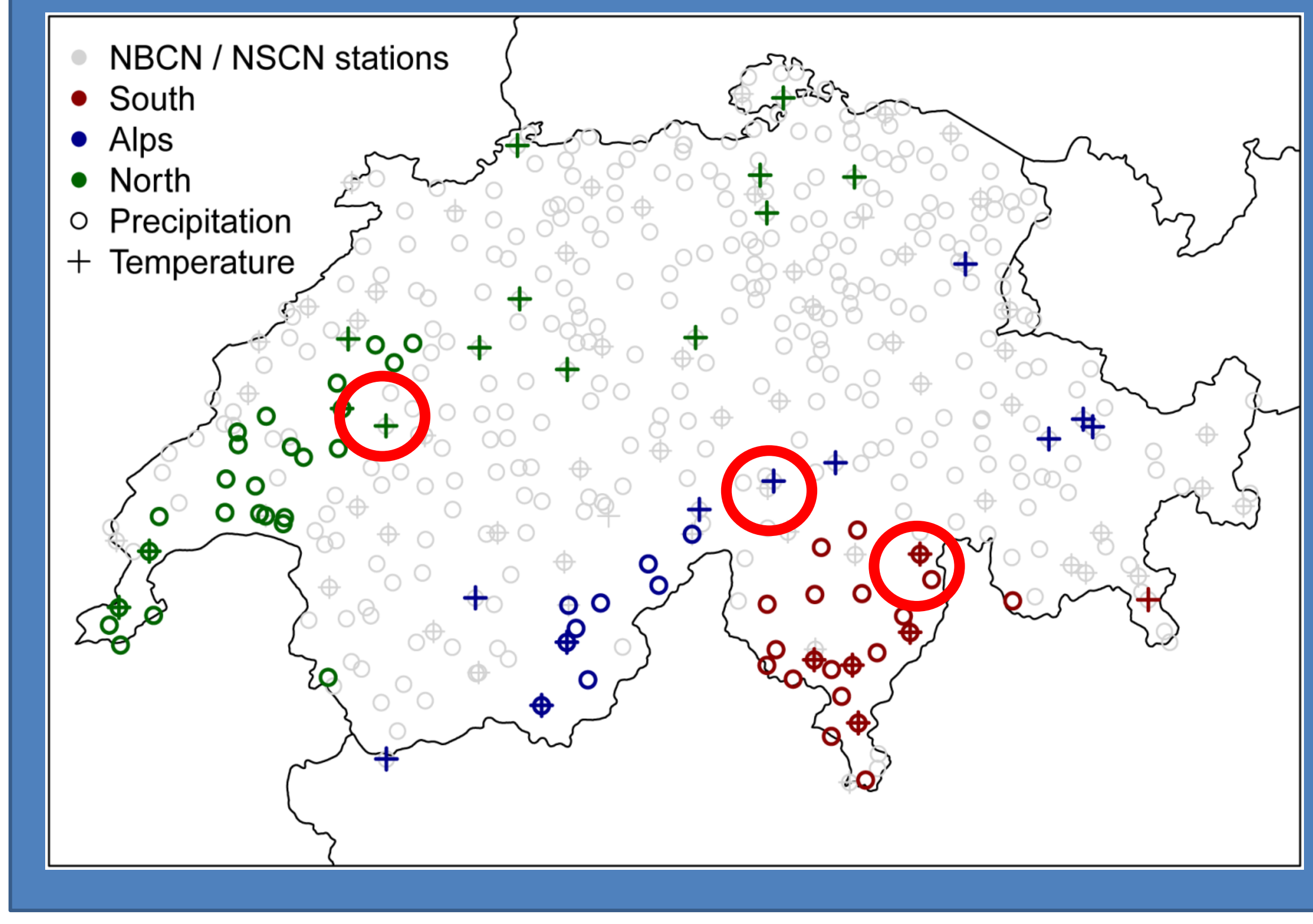


Fig. 3: To compare homogenization in dense and sparse networks, 3 dense and 30 sparse networks were subset from the Swiss network according to correlation structures encountered in Peru (Fig. 4):

- dense ($R^2 \geq 0.85$) and
- sparse ($0.6 \leq R^2 \leq 0.8$ (T) and $0.45 \leq R^2 \leq 0.6$ (P)) networks.

Sparse networks were complemented with stations from surrounding regions (e.g. ECA&D for temperature).

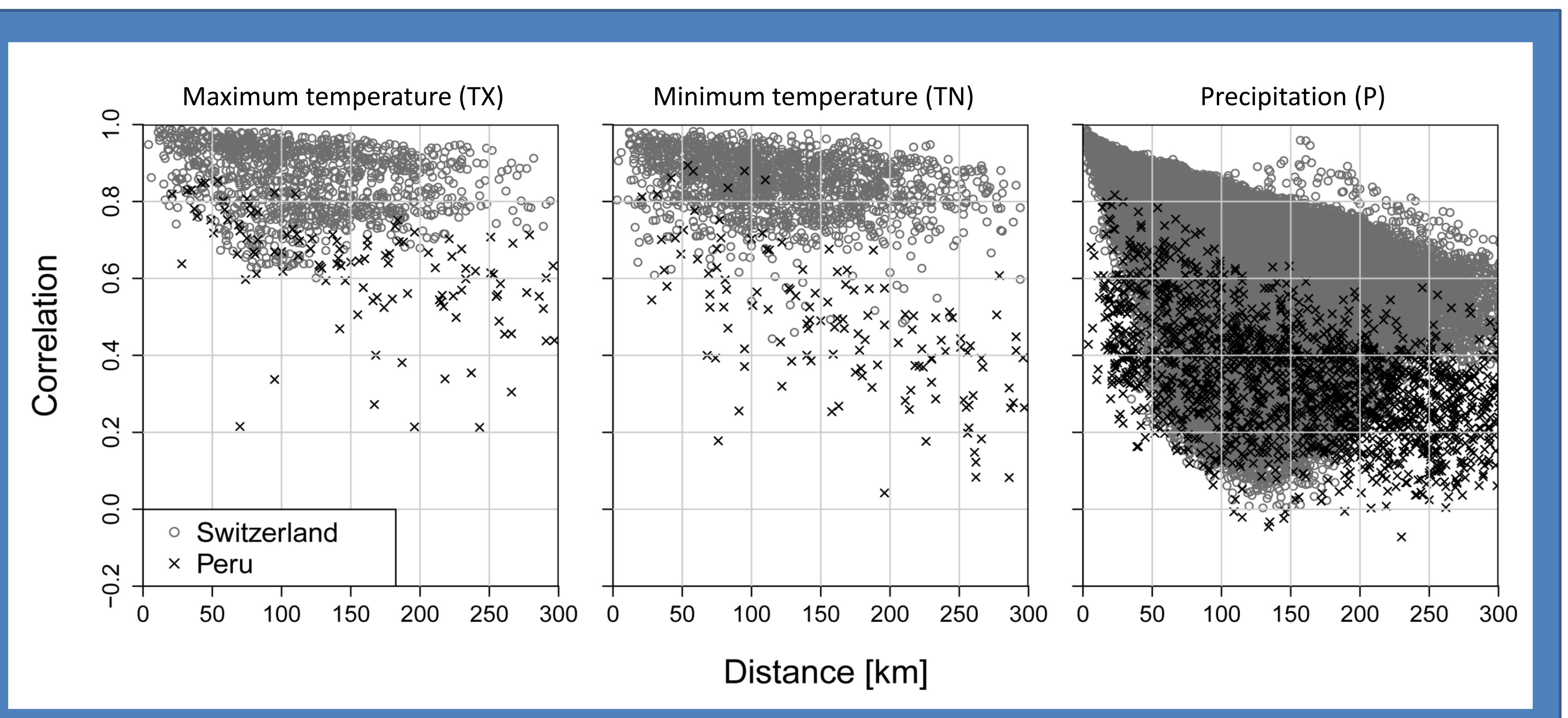


Fig. 4: Normalized Spearman correlation of Peruvian (black crosses) and Swiss (grey circles) data as a function of the distance between two stations. Correlations are lower in Peru than in Switzerland for all investigated parameters (e.g., maximum temperature (TX), minimum temperature (TN), and precipitation (P)).

3) Results: A) Performance of homogenization in sparse networks and B) application of HOMER on Peruvian network

A)

- Auto provides unreliable results for temperature and precipitation in dense networks
- Sparse networks: Improved data quality for both manu and meta-post
  - ➢ HOMER (with human interaction) improves data quality in sparse networks
  - ➢ For temperature, the variability and the bias of the trends in sparse networks is decreased after homogenization.

B)

Homogenization increases average trend in TX (0.15 to 0.29°C/decade) and decreases average trend in TN (0.14°C to 0.11°C) (Fig. 2 and 6). The daily temperature range hence increased in the last decades.

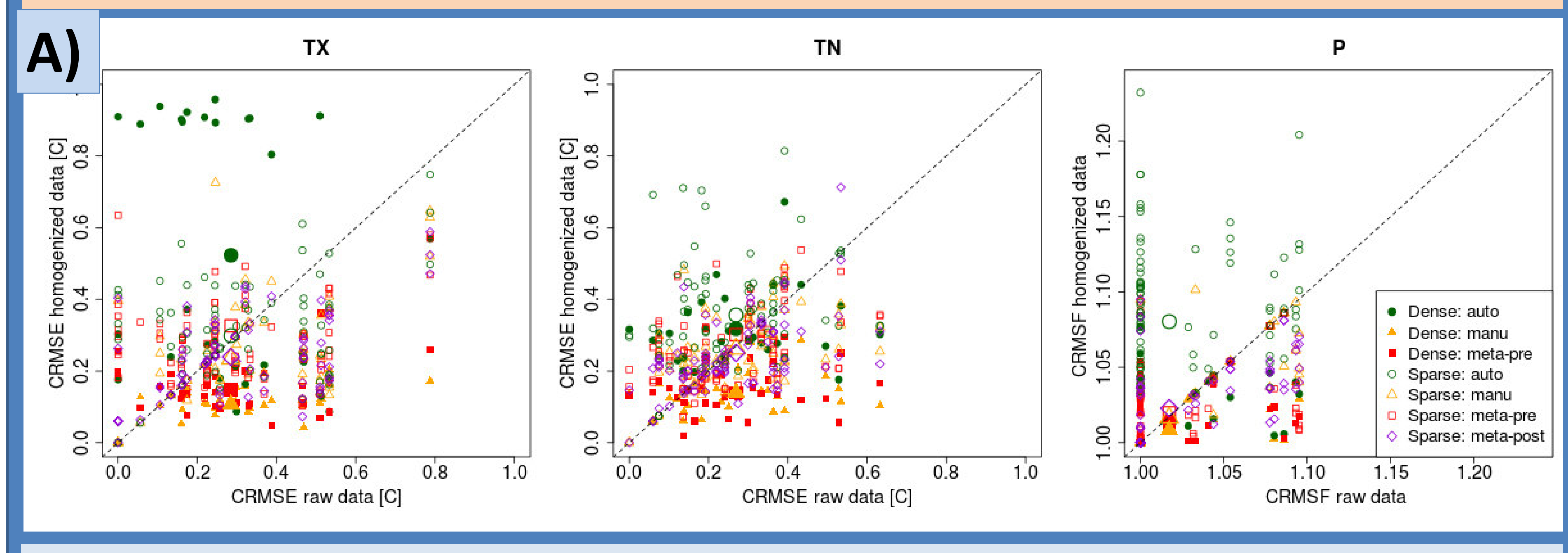


Fig. 5: CRMSE (CRMSF for P) of dense (filled symbols) and sparse (open symbols) networks for all experiments (colors). Points below the 1-1 axis indicate stations that improve after homogenization.

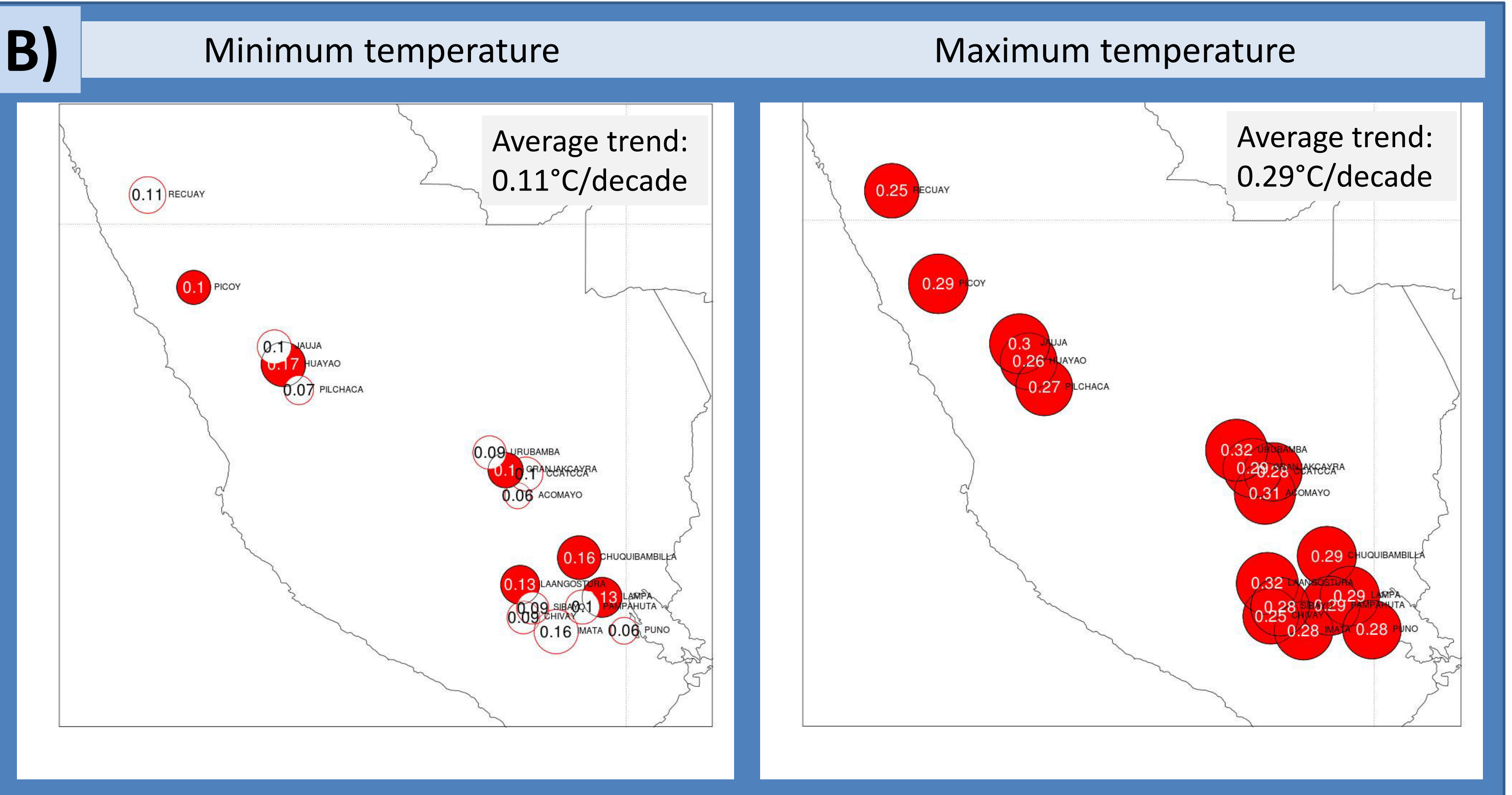


Fig. 6: Trends [°C/decade] of the homogenized data using HOMER. Metadata was available for the region of Cuzco. Average trends decrease from 0.14 to 0.11°C/decade for TN and increase from 0.15 to 0.29°C/decade for TX (e.g., Fig. 2). After homogenization, the spatial variability of the trends is reduced.

Contact: stefanie.gubler@meteoswiss.ch

LINEAR TREND OF SEASONAL SUM AND MAXIMUM DAILY PRECIPITATION IN SLOVENIA FROM 1961 TO 2011

Gregor Vertačnik, Mojca Dolinar, Renato Bertalanič,
Damjan Dvoršek, Matija Klančar, Mateja Nadbath

Long-term precipitation trends represents one of the main aspects of climate change. Besides the commonly assessed change in the mean precipitation sum also the change in maximum values is important to comprehensively assess the impact of climate change. Seasonal precipitation sum and daily maximum precipitation in Slovenia in period 1961–2011 were analysed to estimate the linear trend of both precipitation statistics.

Data

Measured and operationally quality-controlled daily precipitation data have been thoroughly checked again by experts to raise the quality of the whole dataset to the uniform level. Monthly precipitation series from 266 Slovenian stations have been homogenised and missing values interpolated. Resulting series were used to estimate the 51-year linear trend of seasonal precipitation sum. Linear trend of the seasonal daily precipitation maximum was estimated at a subset of 92 stations, where both the percentage of missing values in quality-control time series and the size of discovered inhomogeneities in the time series of annual precipitation were low.

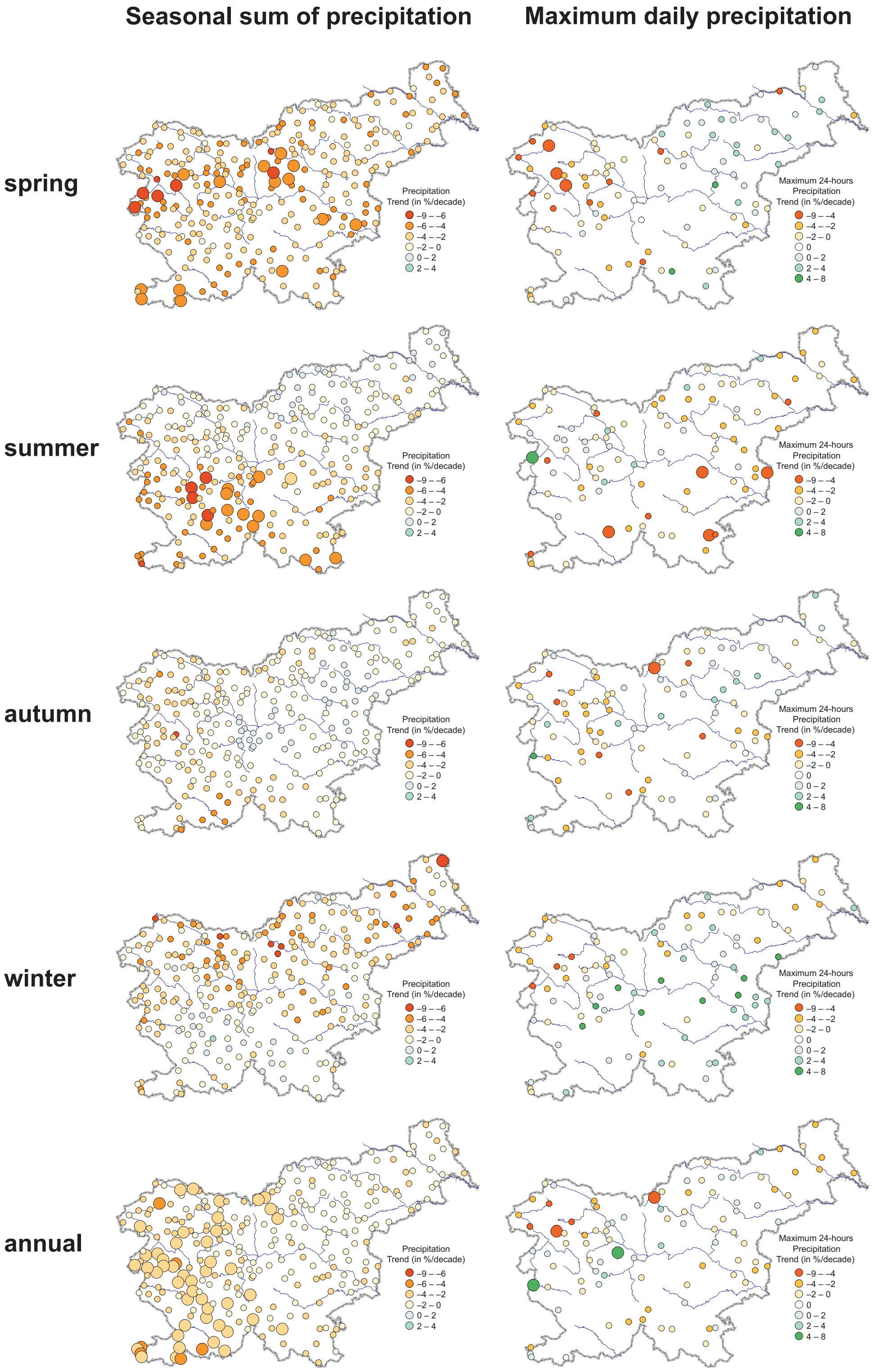
The seasonal and annual linear trend

Time series of seasonal precipitation sum and seasonal daily precipitation maximum are mostly weakly to moderately correlated. The correlation strength varies between seasons and climatic regions.

Linear trend of seasonal values using Theil-Sen method shows partly coherent results between the sums and the maximums. While some smaller regions exhibit statistically significant decrease in the sum in spring or summer, these mostly don't correspond with significant decreases in the daily extremes. Moreover, the state-level spring precipitation sum is clearly decreasing while the daily maximum shows no significant trend.

The autumn trends are weak at all stations for both statistics. In winter there is a north-to-south pattern, more inclined to the decrease in seasonal sum than in daily maximum. On the other hand, there are almost no stations where the ratio of the two statistics shows significant trend or the trend in normalised series of the two statistics is significantly different.

We can conclude that the change in the precipitation sum and the daily precipitation maximum in our case is either similar or the period of investigation is too short regarding the interannual variation and long-term trend of precipitation.



Linear trend in the period 1961–2011. Larger circles mark statisticaly significant trends at the level of 5 %.

Slovenian Environment Agency, Meteorological Office
Vojkova cesta 1b, Ljubljana, Slovenia
e-mail: gregor.vertacnik@gov.si

10th EUMETNET Data Management Workshop
St. Gallen, Switzerland
28–30 Oktober 2015

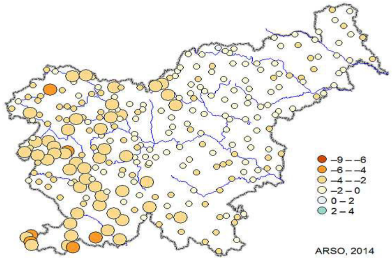
10th EUMETNET Data Management Workshop
St. Gallen, Switzerland, 28–30 October 2015

Climate of Slovenia on homogenized data 1961–2014

REPUBLIC OF SLOVENIA
MINISTRY OF THE ENVIRONMENT AND SPATIAL PLANNING
SLOVENIAN ENVIRONMENT AGENCY

M. Dolinar, G. Vertačnik, R. Bertalanič, D. Dvoršek,
M. Klančar, D. Gartner and M. Nadbath (E-mail: mateja.nadbath@gov.si)

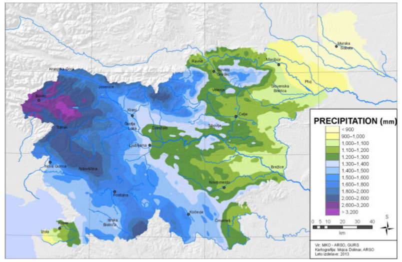
Precipitation



ARSO, 2014

Trend of annual precipitation in % per decade

Precipitation trend varies seasonally and is mostly statistically insignificant. The trend of annual sum is negative and it ranges from –4 to –2 % per decade, it is statistically significant at the 5 % level only in western Slovenia. The spatial pattern of annual precipitation variability is not uniform. In some years there has been precipitation deficiency in one part of the country whereas the other part has experienced the precipitation surplus according to 1981–2010 mean value.



Annual mean precipitation in mm 1981-2010

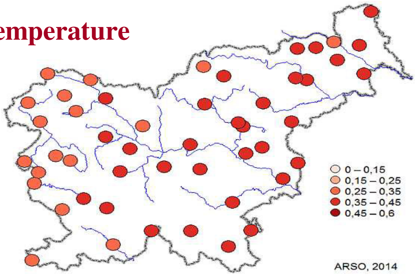
It took us six years to quality control, homogenize and analyze time series of air temperature, precipitation, snow depth, sunshine duration, air pressure and reference evapotranspiration for the period 1961–2011 in Slovenia.
The task was accomplished in 2014.

Climate of Slovenia is illustrated as annual trends and variability of **mean air temperature** and **precipitation sum**.
Trends are calculated for the period 1961–2011 but variability is shown for the period 1961–2014.

In Slovenia the year 2014 was unique: it was the warmest and the wettest year of all since 1961 on.

Annual variability is expressed as the deviation from the 1981–2010 mean value. Precipitation in %; temperature in °C.

Mean air temperature

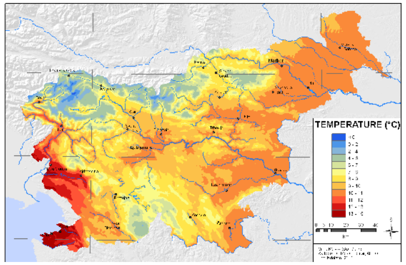


ARSO, 2014

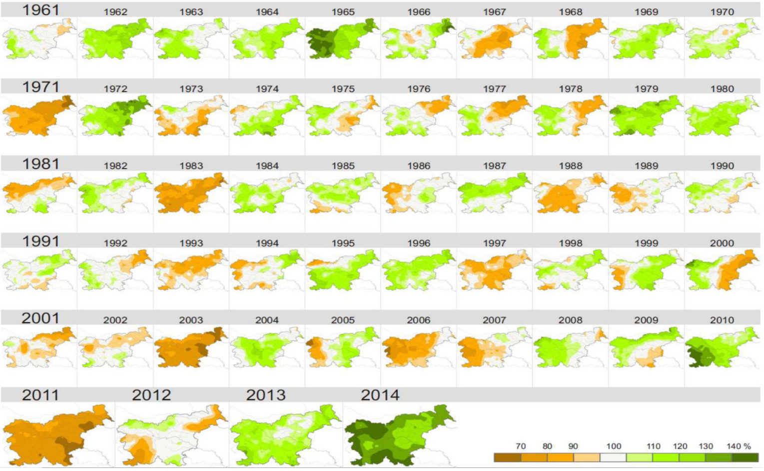
Trend of annual mean air temperature in °C per decade

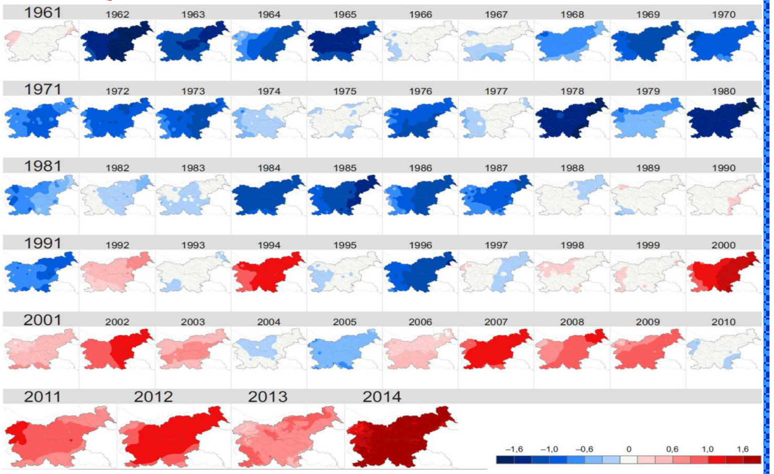
The most uniform trend is shown at mean air temperature; it is positive and statistically significant at the 5 % level in all seasons, except in autumn. On annual level the trend is around 0,33 °C per decade.

The charts of annual variability of mean air temperature show characteristic pattern: air temperature before 1991 was mostly colder while after 1992 it was mainly warmer in comparison to the 1981–2010 mean value.



Annual mean air temperature in °C 1981-2010





CLIMATE DATABASE IN CROATIA

Helena Lebo Andreis
Meteorological and Hydrological Service of Croatia (MHSC)

CONFIGURATION

PostgreSQL:

- free
- cross-platform
- open source

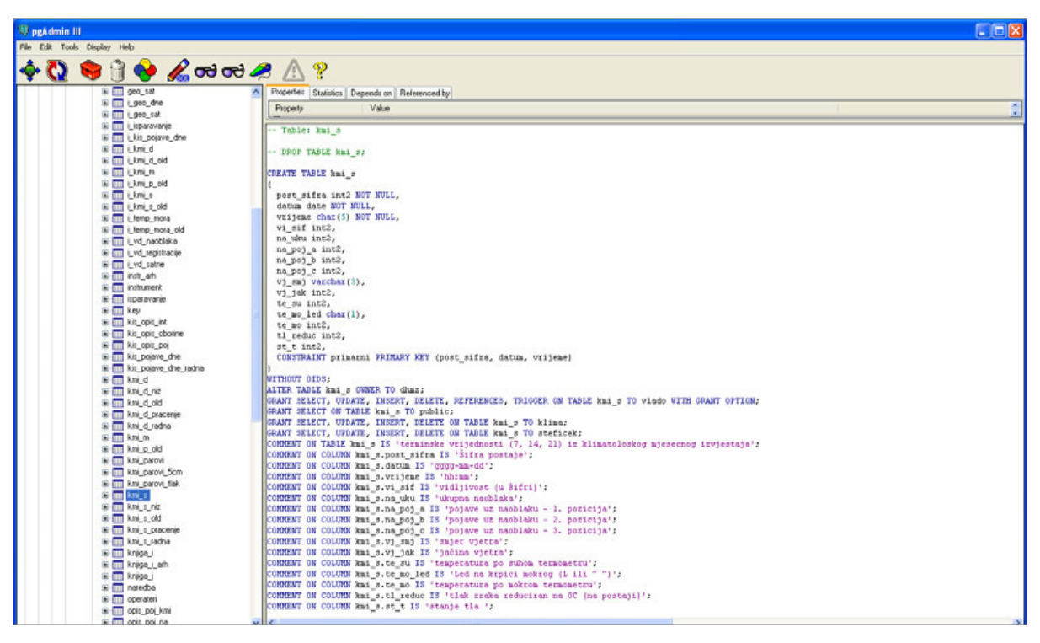
About climate database

At MHSC we have been using relational database since 2007. The climate database has been developed in IT Department of MHSC.

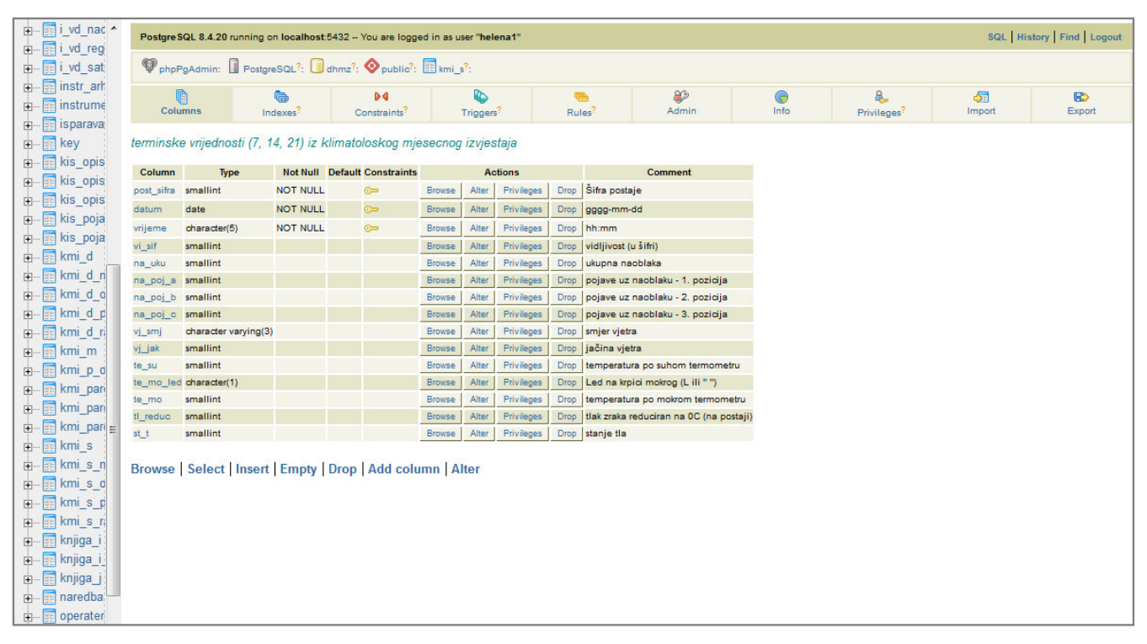
Technical data

- Database:**PostgreSQL** 8.4.20
- Platform: **Linux** CentOS 6.4
- HTTP Server: Apache 2

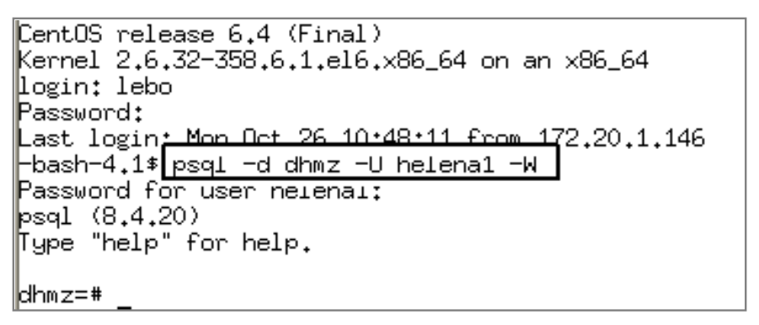
MAINTENANCE



Graphical interface of pgAdmin tool (locally installed)



Graphical interface of phpPgAdmin tool (web-oriented tool)



```
CentOS release 6.4 (Final)
Kernel 2.6.32-358.6.1.el6.x86_64 on an x86_64
login: lebo
Password:
Last login: Mon Oct 26 10:48:11 from 172.20.1.146
-bash-4.1$ psql -d dhmz -U helena1 -W
Password for user helena1:
psql (8.4.20)
Type "help" for help.

dhmz=# _
```

Example of PostgreSQL command line: login

Backup

Entire database is regularly backed up on the remote location.

Responsibility

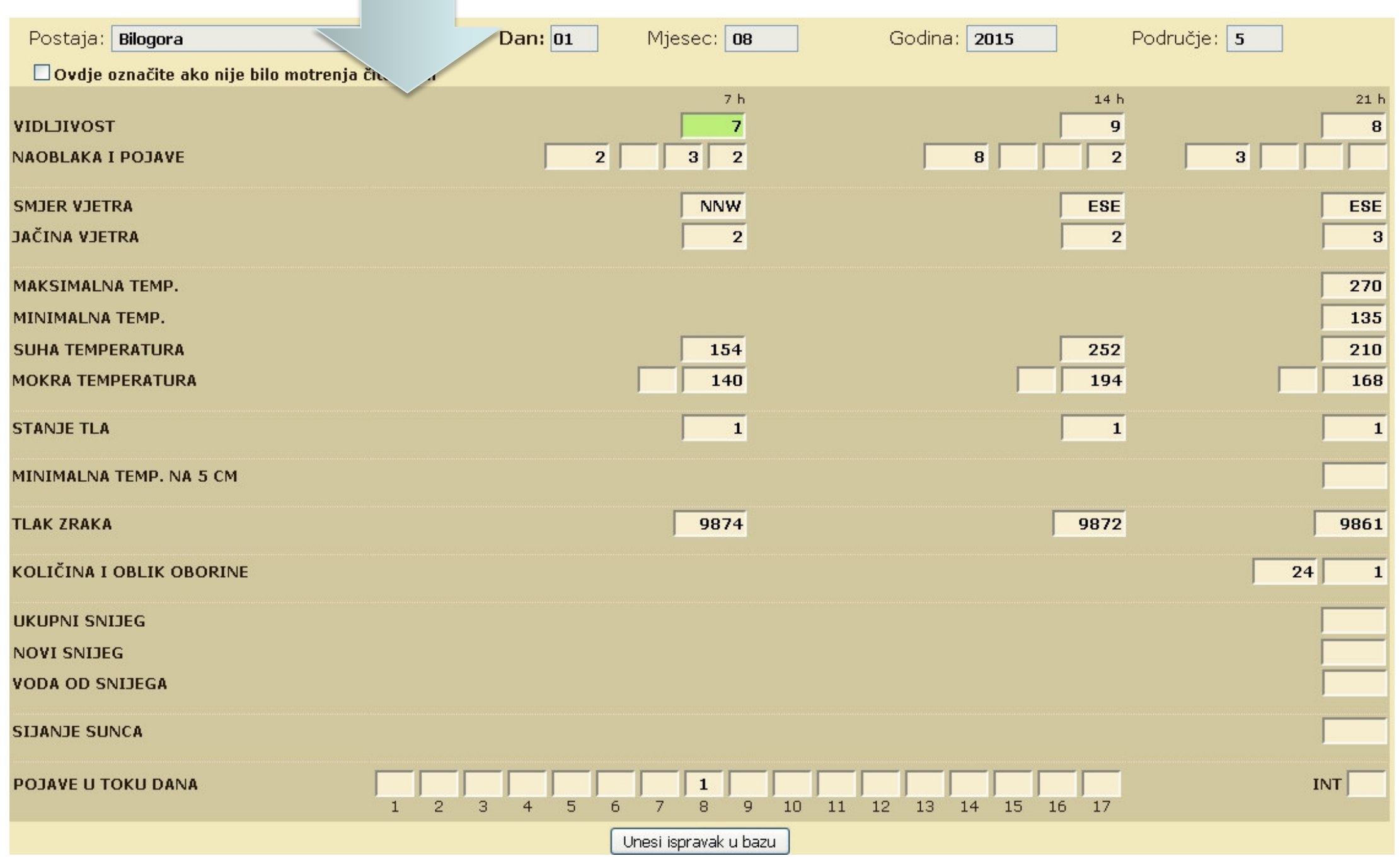
Database administration, creating queries, monitoring, backing up is a responsibility of IT Department of MHSC.

Tools

For designing tables, creating users, privileges, etc. we use graphical tools **pgAdmin** and **phpPgAdmin.**

Those features are also available in command line.

USER INTERFACE



Postaja: Bilogora Dan: 01 Mjesec: 08 Godina: 2015 Područje: 5

☐ Ovdje označite ako nije bilo motrenja či...

	7 h	14 h	21 h
VIDLJIVOST	7	9	8
NAOBLAKA I POJAVE	2 3 2	8 2	3
SMJER VJETRA	NNW	ESE	ESE
JAČINA VJETRA	2	2	3
MAKSIMALNA TEMP.			270
MINIMALNA TEMP.			135
SUHA TEMPERATURA	154	252	210
MOKRA TEMPERATURA	140	194	168
STANJE TLA	1	1	1
MINIMALNA TEMP. NA 5 CM			
TLAK ZRAKA	9874	9872	9861
KOLIČINA I OBLIK OBORINE			24 1
UKUPNI SNIJEG			
NOVI SNIJEG			
VODA OD SNIJEGA			
SIJANJE SUNCA			

POJAVE U TOKU DANA: 8: 1; INT

Input form for climatological station, accessed through a web browser.
Data are controlled on screen before being stored in the database.

Developers

- IT Department of MHSC
- programming languages and tools: PHP, JavaScript, HTML, CSS, SQL, Google maps, ...

Users

- employees of MHSC
- data input, providing quality controls, corrections, running reports, exchanging data etc.

User interface

- web-oriented, in a browser
- user-friendly
- easy user-developer communication

DATA STORAGE

DATA INPUT

2 copies

DATASET 1

- **KEEPING ORIGINAL DATA**
- **NO CORRECTIONS ALLOWED ***
- **USED FOR COMPARISON WITH CORRECTED DATA ****

DATASET 2

- **QUALITY CONTROL**
- **CORRECTIONS**
- **VERIFICATION**
- **REPORTS *****

* except for typing mistakes

*** see section **"Reports"**

** see poster **"The influence of subjective corrections on climate data"**

Input control

- Control of users' privileges
- On-screen control of data ("sanity check")

Data storage

After input, data are automatically stored in two copies.

- One copy is kept unchanged (see **dataset 1**),
- Other copy is controlled and corrected if necessary (see **dataset 2**). This set of data is used for reports, etc.

Verification

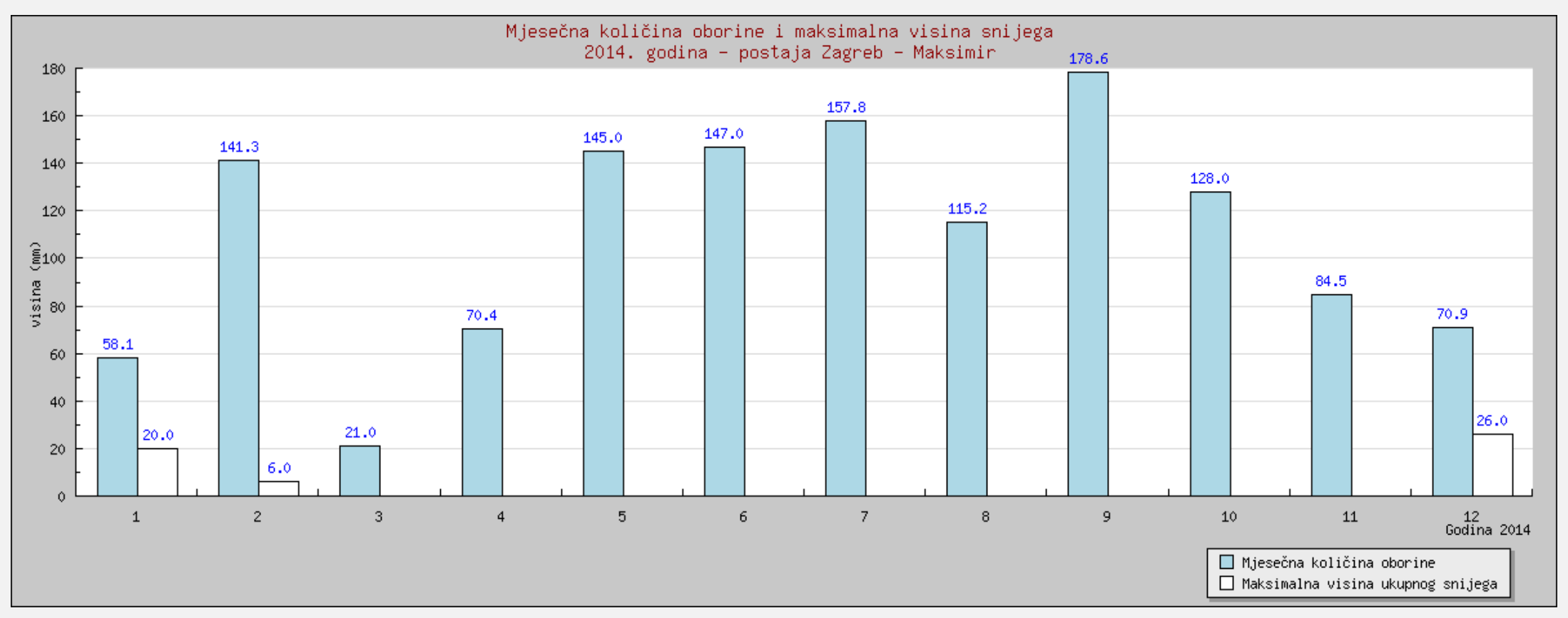
Verification indicators are stored in a separate table, for later use.

REPORTS

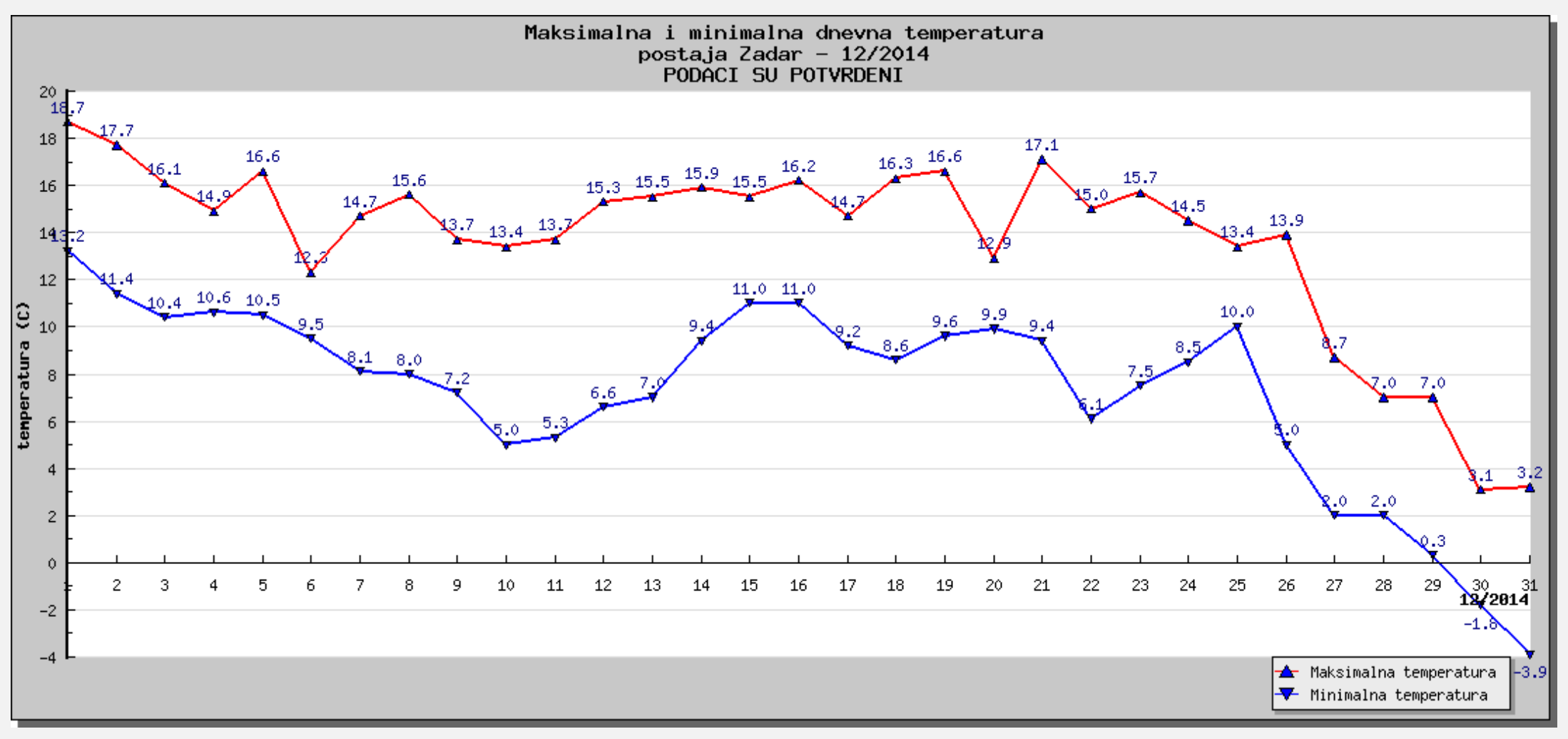
Once the data have been verified, users can view tables, charts, maps, etc. They are accessed through a web browser.

Programs generate reports on users' demand.

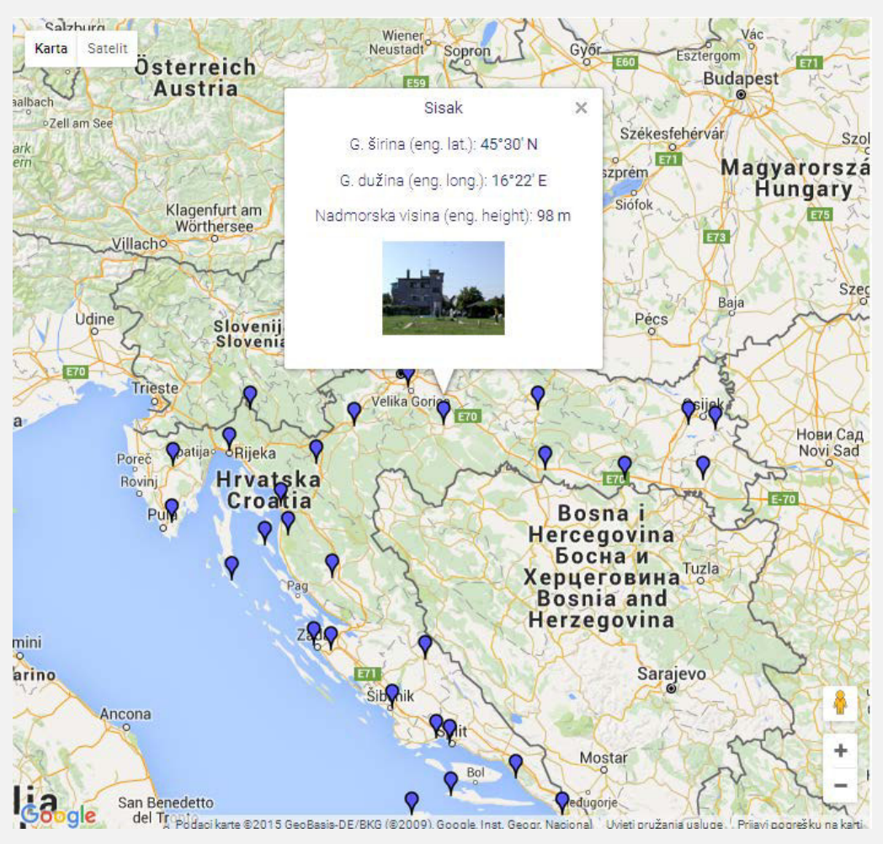
Some of the products are shown below.



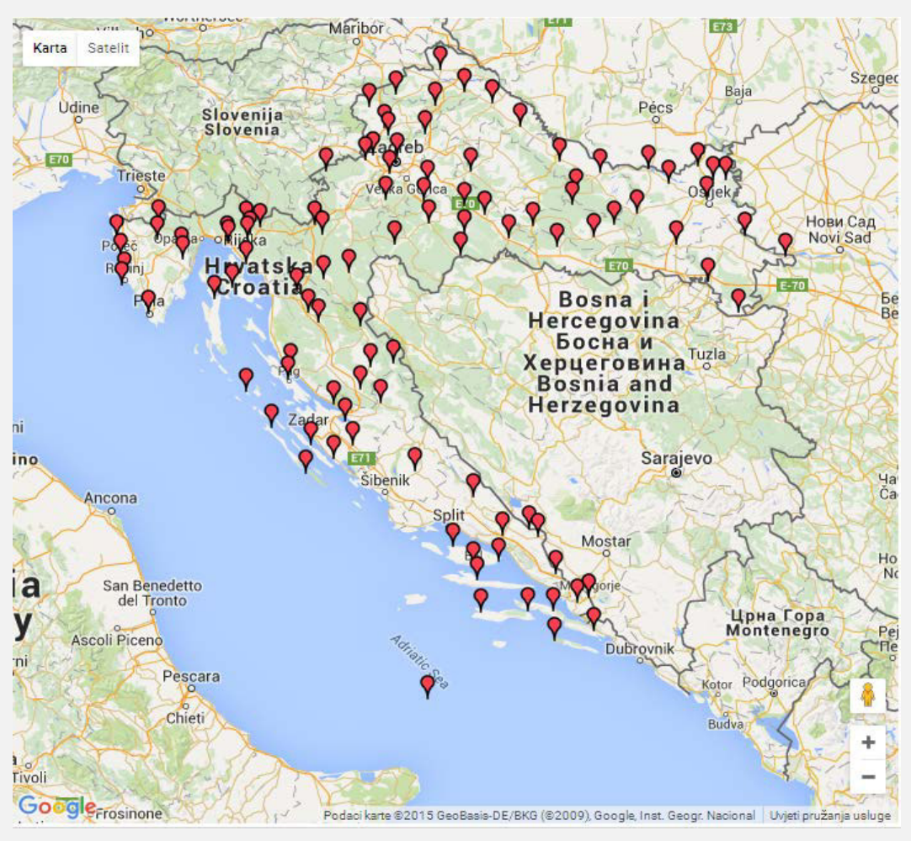
Monthly precipitation and maximum snow depth at station Zagreb Maksimir for year 2014



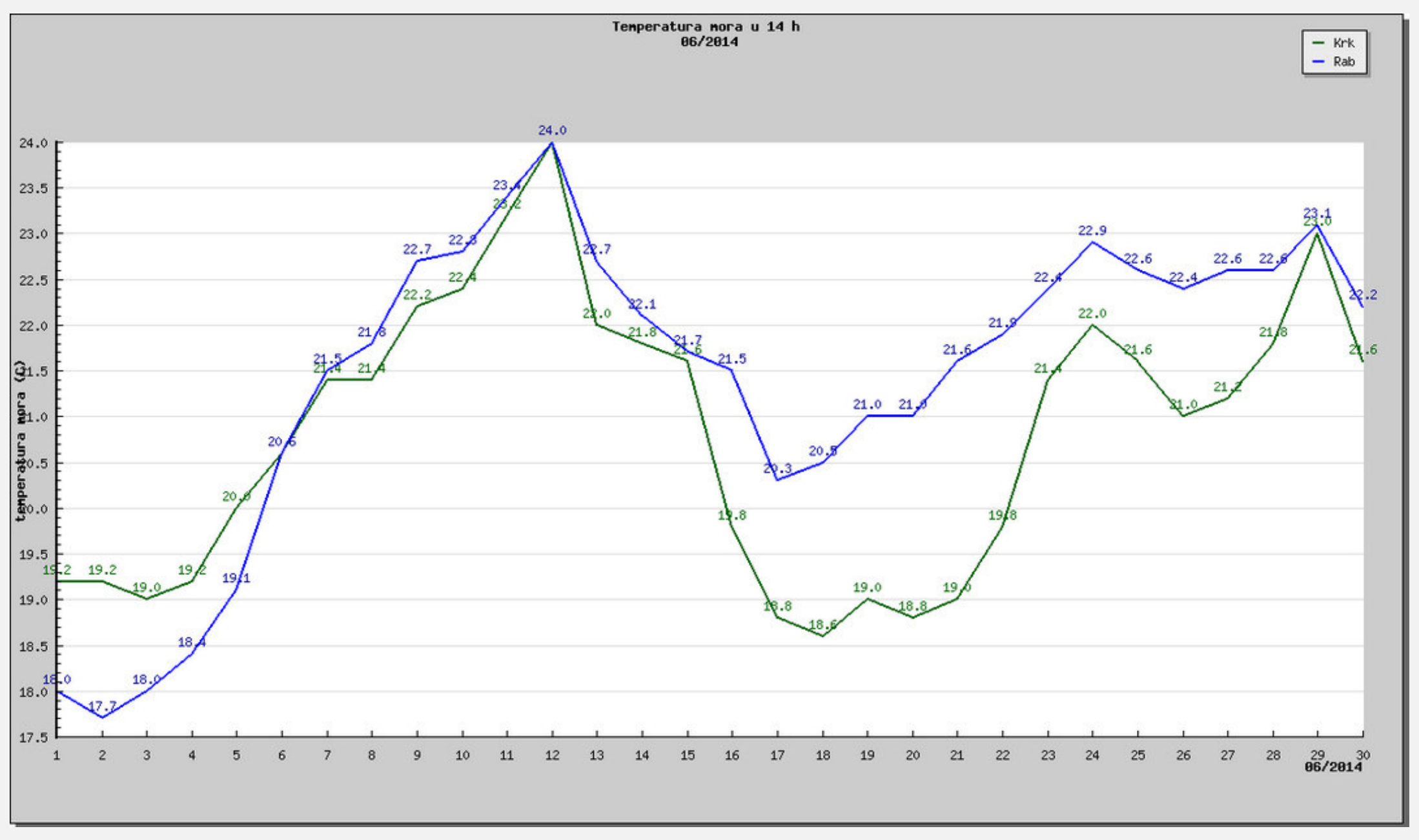
Daily maximum and minimum air temperature at station Zadar for December 2014



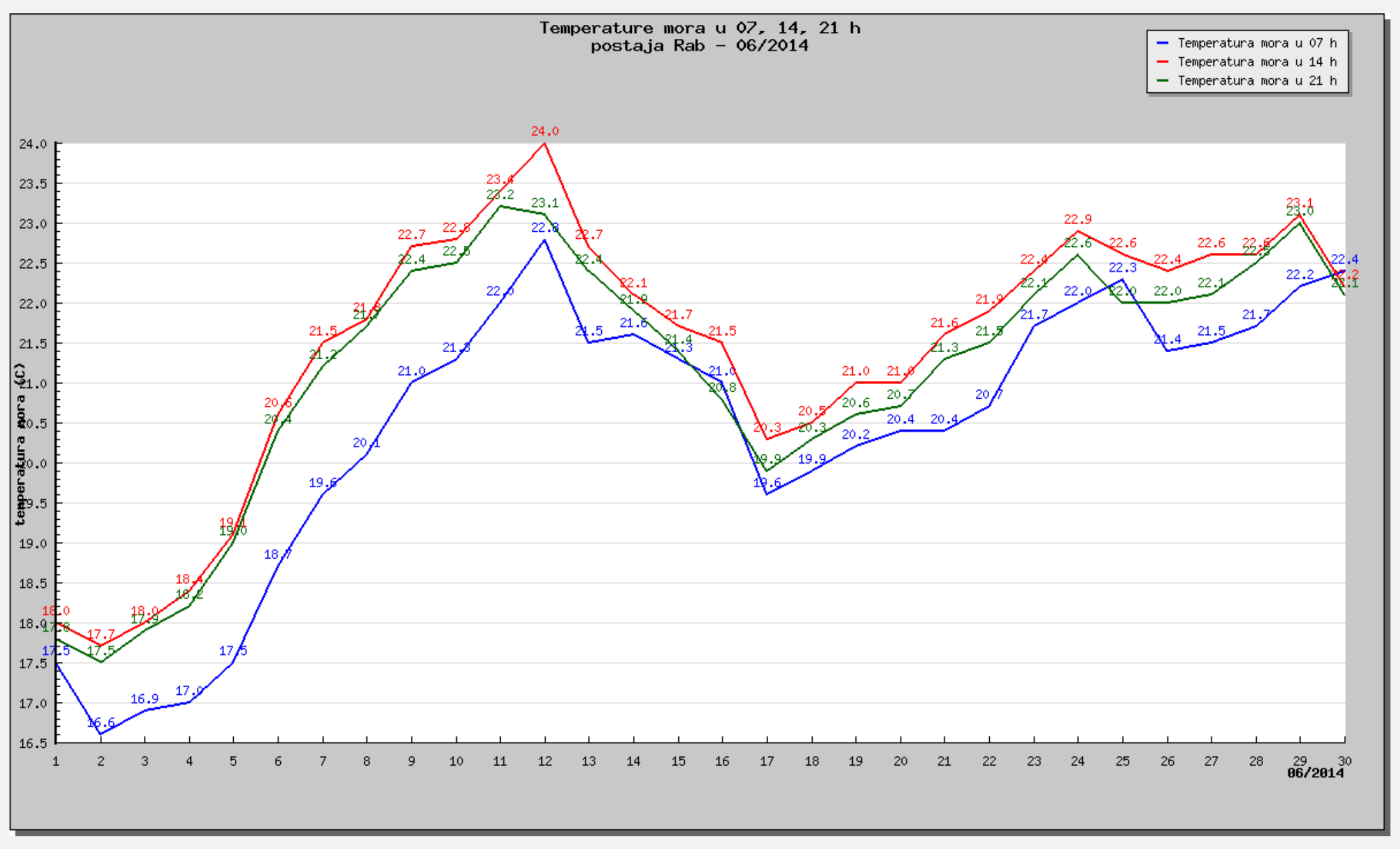
Main meteorological stations in Croatia on interactive map



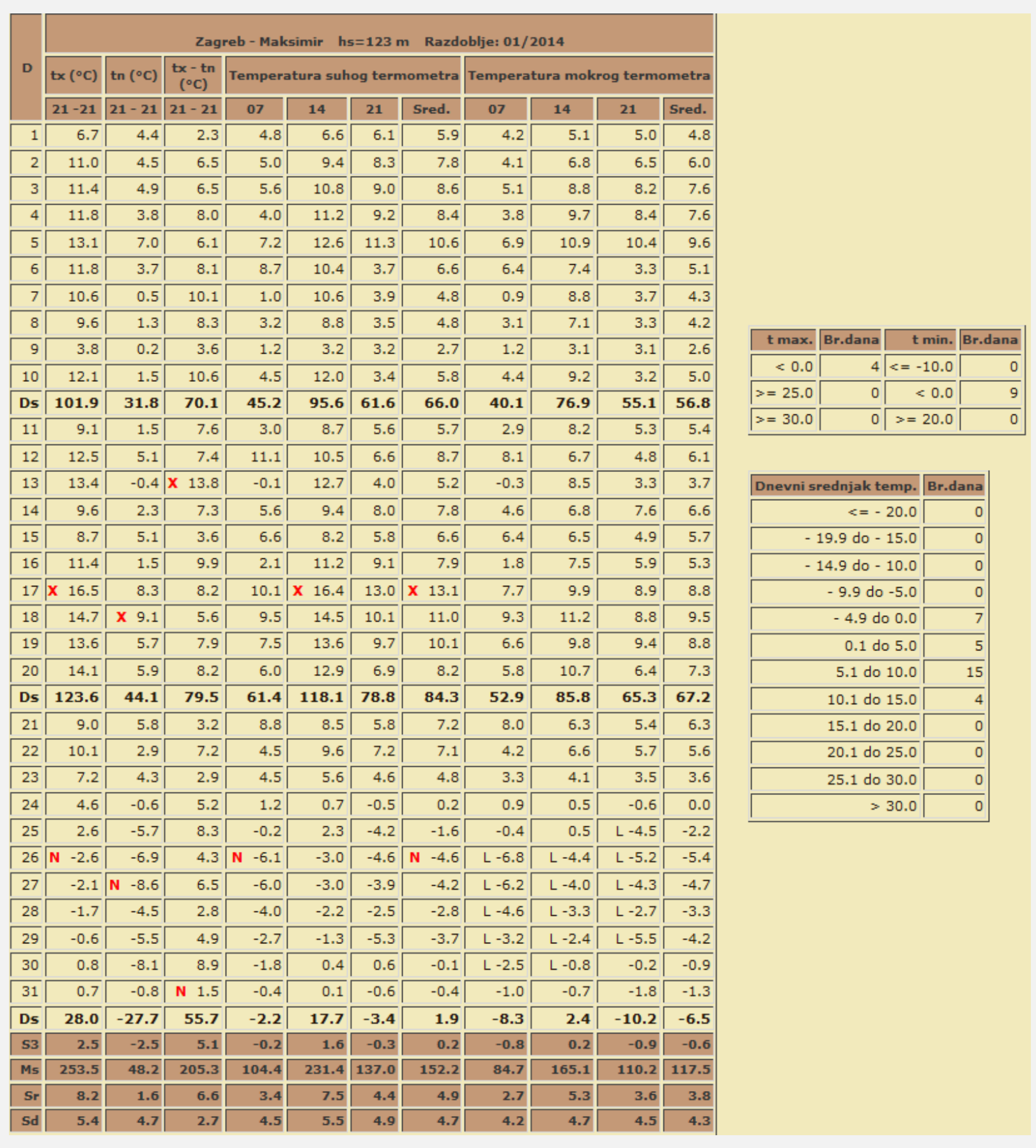
Climatological stations in Croatia on interactive map



Sea temperature at stations Krk and Rab for June 2014, at 14:00 h



Sea temperature at station Rab for June 2014, measured at 07:00, 14:00 and 21:00 h



Zagreb - Maksimir hs=123 m Razdoblje: 01/2014

D	tx (°C)	tn (°C)	tx - tn (°C)	Temperatura suhog termometra				Temperatura mokrog termometra			
	21 - 21	21 - 21	21 - 21	07	14	21	Sred.	07	14	21	Sred.
1	6.7	4.4	2.3	4.8	6.6	6.1	5.9	4.2	5.1	5.0	4.8
2	11.0	4.5	6.5	5.0	9.4	8.3	7.8	4.1	6.8	6.5	6.0
3	11.4	4.9	6.5	5.6	10.8	9.0	8.6	5.1	8.8	8.2	7.6
4	11.8	3.8	8.0	4.0	11.2	9.2	8.4	3.8	9.7	8.4	7.6
5	13.1	7.0	6.1	7.2	12.6	11.3	10.6	6.9	10.9	10.4	9.6
6	11.8	3.7	8.1	8.7	10.4	3.7	6.6	6.4	7.4	3.3	5.1
7	10.6	0.5	10.1	1.0	10.6	3.9	4.8	0.9	8.8	3.7	4.3
8	9.6	1.3	8.3	3.2	8.8	3.5	4.8	3.1	7.1	3.3	4.2
9	3.8	0.2	3.6	1.2	3.2	3.2	2.7	1.2	3.1	3.1	2.6
10	12.1	1.5	10.6	4.5	12.0	3.4	5.8	4.4	9.2	3.2	5.0
Ds	**101.9**	**31.8**	**70.1**	**45.2**	**95.6**	**61.6**	**66.0**	**40.1**	**76.9**	**55.1**	**56.8**
11	9.1	1.5	7.6	3.0	8.7	5.6	5.7	2.9	8.2	5.3	5.4
12	12.5	5.1	7.4	11.1	10.5	6.6	8.7	8.1	6.7	4.8	6.1
13	13.4	-0.4	X 13.8	-0.1	12.7	4.0	5.2	-0.3	8.5	3.3	3.7
14	9.6	2.3	7.3	5.6	9.4	8.0	7.8	4.6	6.8	7.6	6.6
15	8.7	5.1	3.6	6.6	8.2	5.8	6.6	6.4	6.5	4.9	5.7
16	11.4	1.5	9.9	2.1	11.2	9.1	7.9	1.8	7.5	5.9	5.3
17	X 16.5	8.3	8.2	10.1	X 16.4	13.0	X 13.1	7.7	9.9	8.9	8.8
18	14.7	X 9.1	5.6	9.5	14.5	10.1	11.0	9.3	11.2	8.8	9.5
19	13.6	5.7	7.9	7.5	13.6	9.7	10.1	6.6	9.8	9.4	8.8
20	14.1	5.9	8.2	6.0	12.9	6.9	8.2	5.8	10.7	6.4	7.3
Ds	**123.6**	**44.1**	**79.5**	**61.4**	**118.1**	**78.8**	**84.3**	**52.9**	**85.8**	**65.3**	**67.2**
21	9.0	5.8	3.2	8.8	8.5	5.8	7.2	8.0	6.3	5.4	6.3
22	10.1	2.9	7.2	4.5	9.6	7.2	7.1	4.2	6.6	5.7	5.6
23	7.2	4.3	2.9	4.5	5.6	4.6	4.8	3.3	4.1	3.5	3.6
24	4.6	-0.6	5.2	1.2	0.7	-0.5	0.2	0.9	0.5	-0.6	0.0
25	2.6	-5.7	8.3	-0.2	2.3	-4.2	-1.6	-0.4	0.5	L -4.5	-2.2
26	N -2.6	-6.9	4.3	N -6.1	-3.0	-4.6	N -4.6	L -6.8	L -4.4	L -5.2	-5.4
27	-2.1	N -8.6	6.5	-6.0	-3.0	-3.9	-4.2	L -6.2	L -4.0	L -4.3	-4.7
28	-1.7	-4.5	2.8	-4.0	-2.2	-2.5	-2.8	L -4.6	L -3.3	L -2.7	-3.3
29	-0.6	-5.5	4.9	-2.7	-1.3	-5.3	-3.7	L -3.2	L -2.4	L -5.5	-4.2
30	0.8	-8.1	8.9	-1.8	0.4	0.6	-0.1	L -2.5	L -0.8	-0.2	-0.9
31	0.7	-0.8	N 1.5	-0.4	0.1	-0.6	-0.4	-1.0	-0.7	-1.8	-1.3
Ds	**28.0**	**-27.7**	**55.7**	**-2.2**	**17.7**	**-3.4**	**1.9**	**-8.3**	**2.4**	**-10.2**	**-6.5**
S3	**2.5**	**-2.5**	**5.1**	**-0.2**	**1.6**	**-0.3**	**0.2**	**-0.8**	**0.2**	**-0.9**	**-0.6**
Ms	**253.5**	**48.2**	**205.3**	**104.4**	**231.4**	**137.0**	**152.2**	**84.7**	**165.1**	**110.2**	**117.5**
Sr	**8.2**	**1.6**	**6.6**	**3.4**	**7.5**	**4.4**	**4.9**	**2.7**	**5.3**	**3.6**	**3.8**
Sd	**5.4**	**4.7**	**2.7**	**4.5**	**5.5**	**4.9**	**4.7**	**4.2**	**4.7**	**4.5**	**4.3**

t max.	Br.dana	t min.	Br.dana
< 0.0	4	<= -10.0	0
>= 25.0	0	< 0.0	9
>= 30.0	0	>= 20.0	0

Dnevni srednjak temp.	Br.dana
<= - 20.0	0
- 19.9 do - 15.0	0
- 14.9 do - 10.0	0
- 9.9 do -5.0	0
- 4.9 do 0.0	7
0.1 do 5.0	5
5.1 do 10.0	15
10.1 do 15.0	4
15.1 do 20.0	0
20.1 do 25.0	0
25.1 do 30.0	0
> 30.0	0

Monthly report for air temperature at station Zagreb Maksimir for January 2014

10th EUMETNET Data Management Workshop - "High quality climate data - the foundation of Climate Services", St. Gallen, Switzerland, 28th – 30th October 2015

10th EUMETNET Data Management Workshop. “High quality climate data – Knowledge base of Climate Services”
St. Gallen (Switzerland), 28th – 30th October, 2015

PRELIMINARY RESULTS OF A DISASTER MODEL OF EXTREME WIND GUST SITUATIONS IN PENINSULAR SPAIN

RODRIGO, M.*[1][2]***, LÓPEZ, J.A.***[2]

[1]CCS (Consorcio de Compensación de Seguros), SPAIN, [2]AEMET (Agencia Estatal de Meteorología), SPAIN

*e-mail: ccseguros@aemet.es

The CCS (Spanish Insurance Compensation Consortium) is the Spanish Agency that provides insurance coverage against weather events that involve an extraordinary risk. One of the extraordinary risks covered by the CCS refers to extraordinary wind, defined as those with wind gusts exceeding 120 km/h. The accurate delimitation of the areas in which this condition is met shows considerable difficulties in our country because of the lack of wind observations and the complexity of the terrain.

The operational procedure performed at AEMET (Spanish Meteorological Agency) for estimating the areas with maximum wind gusts applies a geostatistical technique, the universal kriging, based on the observational data of maximum wind gust and drawing also on external variables: terrain elevation, distance from the shore, and the HIRLAM (High Resolution Limited Area Model) forecasting model output of maximum wind gust field.

Extreme meteorological events sometimes go unnoticed but frequently become apparent due their impact on society with its corresponding consequences. Moreover, the catastrophic nature of a phenomenon depends not only on the extreme value that the climate element takes on, but also on other factors, such as population distribution or geomorphological features, among many others.

In this work we present some preliminary steps in the development of a disaster model of extreme wind gust situations in peninsular Spain. The main aim of this model is to provide significant information for covering risk against future extreme wind situations, as well as assessing the liability owing to this risk on a year time-frame.

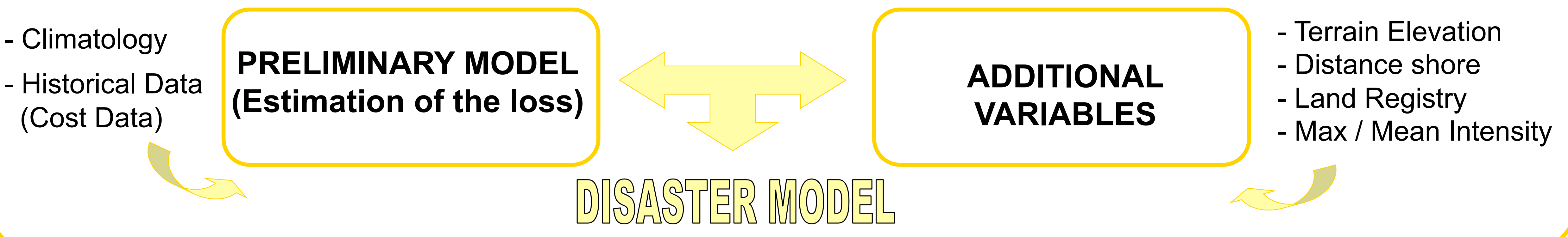
BACKGROUND

UNIVERSAL KRIGING INTERPOLATION:

- Observational Data
- External Variables

PLANNING

The first steps in the development of this model are:
a)the generation of climatology of the areas under extraordinary wind going back to 2009 and,
b) the estimation of incurred covered loss in these years based on data from the insurer.
The following step will be the estimation of a local (at municipality level at best) expected loss function from extraordinary winds depending on the intensity of the event.



- Climatology
- Historical Data (Cost Data)

PRELIMINARY MODEL (Estimation of the loss)

ADDITIONAL VARIABLES

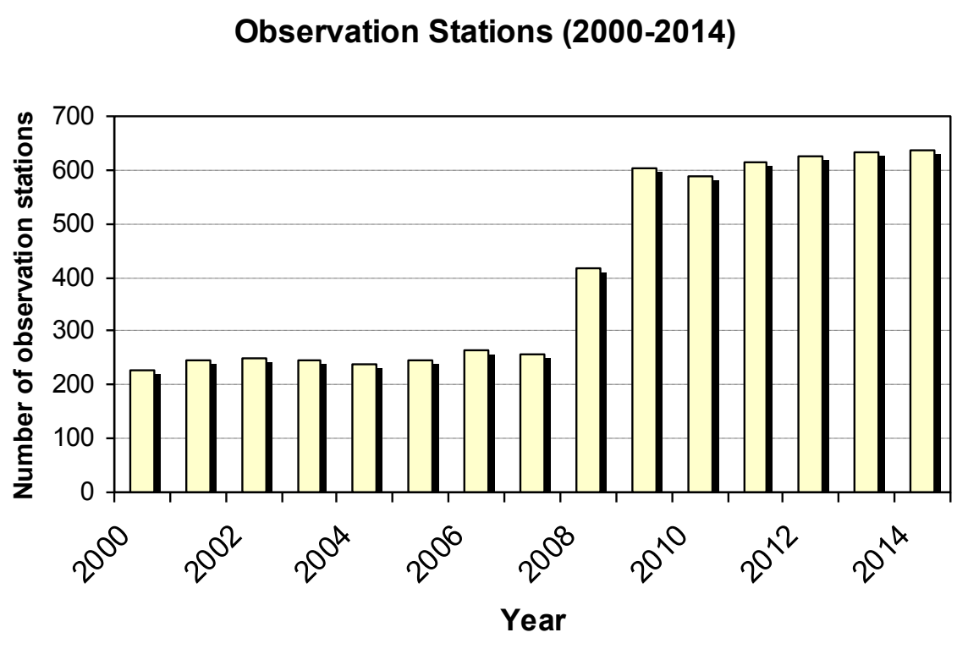
- Terrain Elevation
- Distance shore
- Land Registry
- Max / Mean Intensity

DISASTER MODEL

CLIMATOLOGY

OBSERVATION DATA

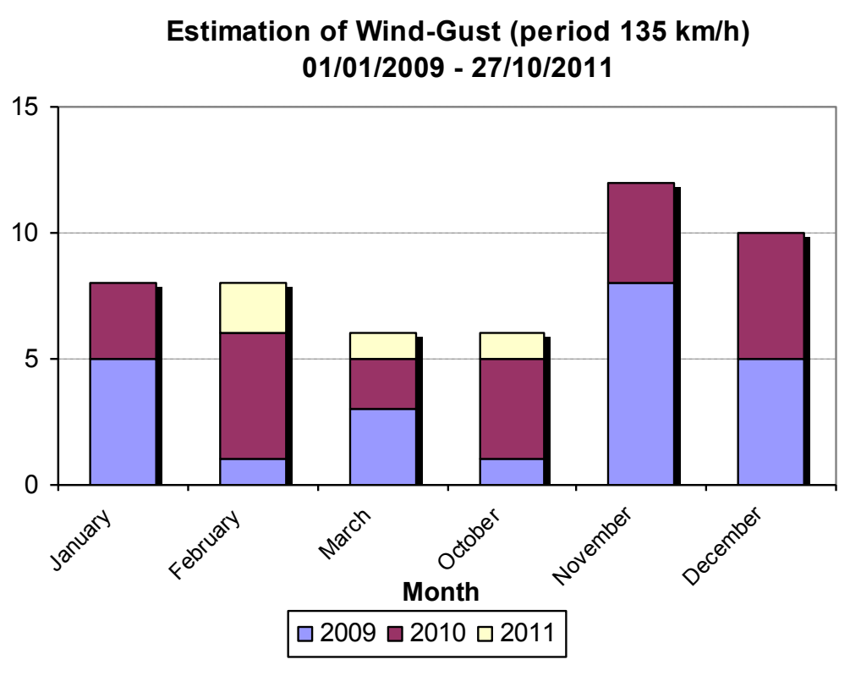
The climate study takes into account the necessary homogeneity in the number of stations with observational data annually.

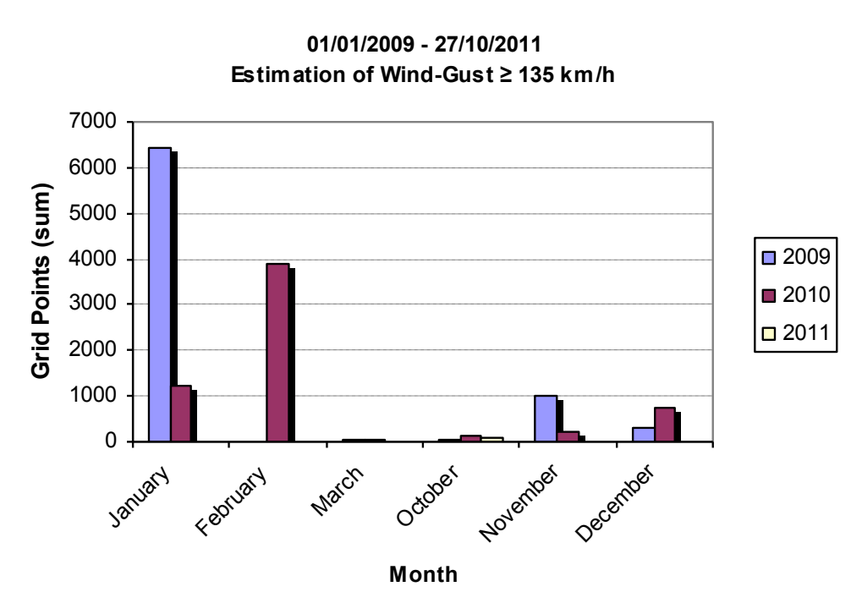


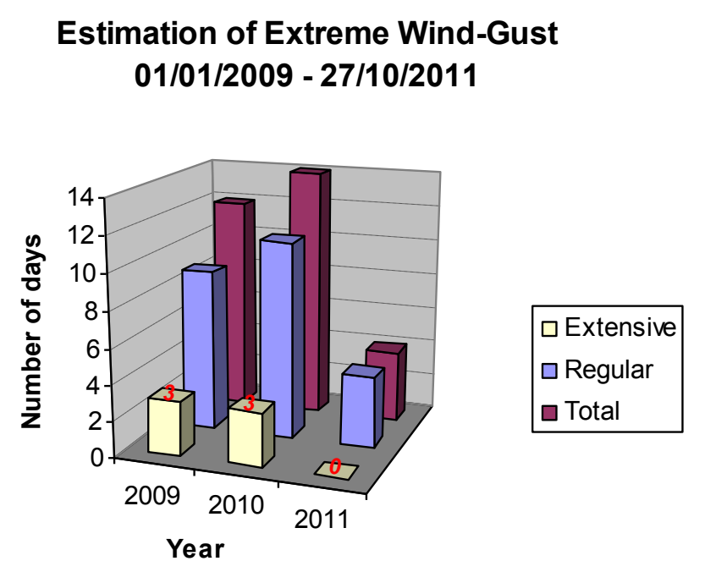
FRECUENCY

The selected days have at least one observation station with wind-gust exceeding 110 km/h (left). We present the climatology of the extension of the extreme wind-gust situations (centre and right). There are 2 periods with different wind gust limit.

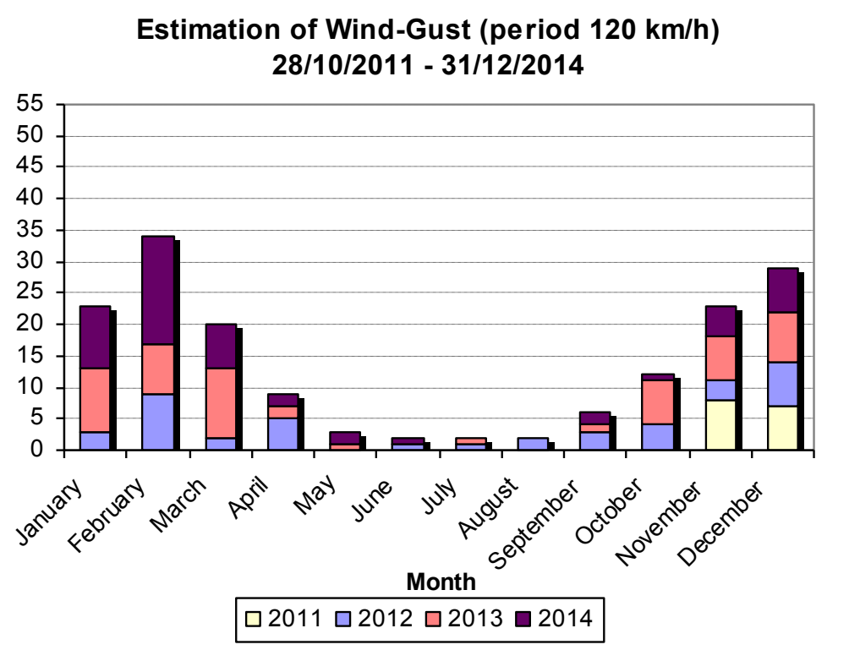
Period 135 km/h:
from 01/01/2009
to 27/10/2011

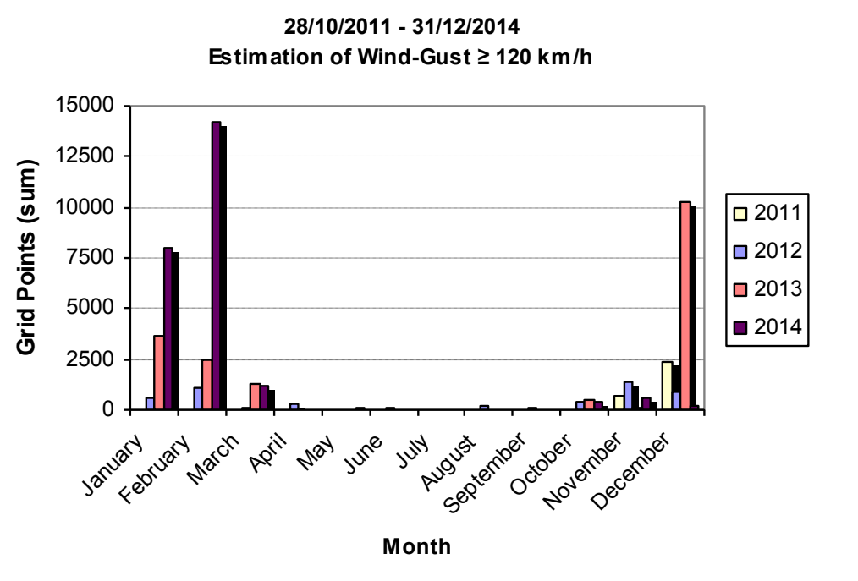


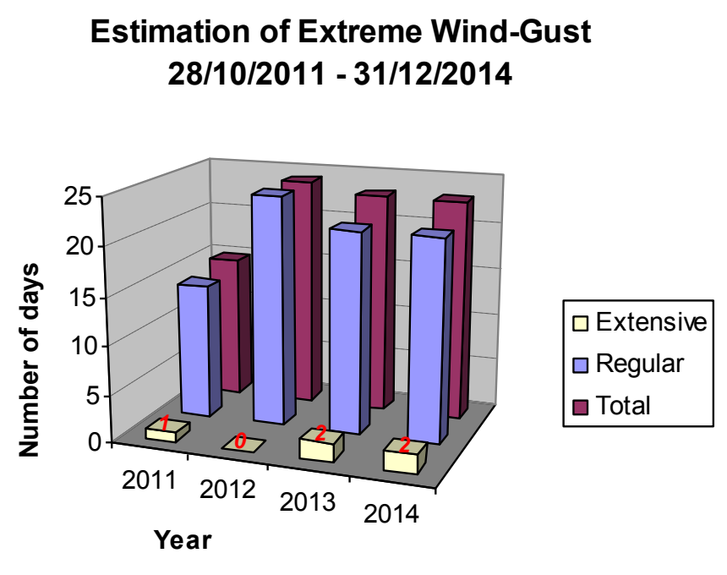




Period 120 km/h:
from 28/10/2011
to 31/12/2014



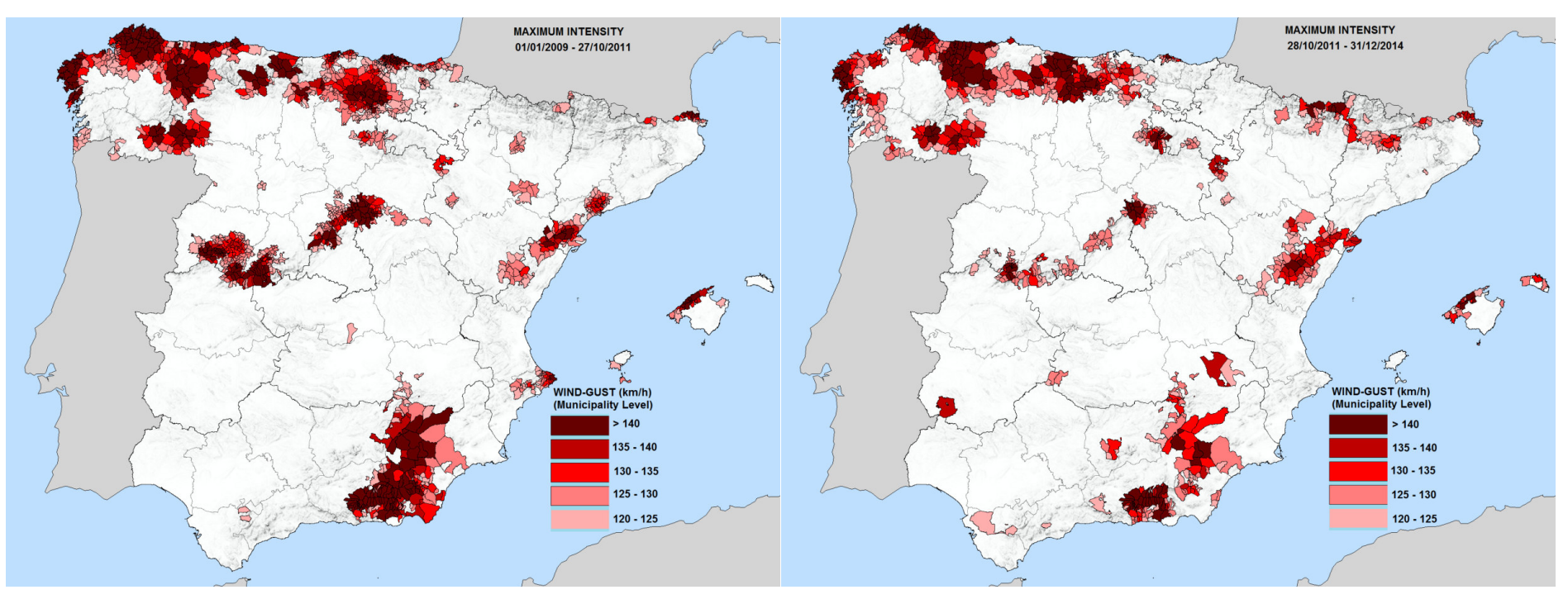




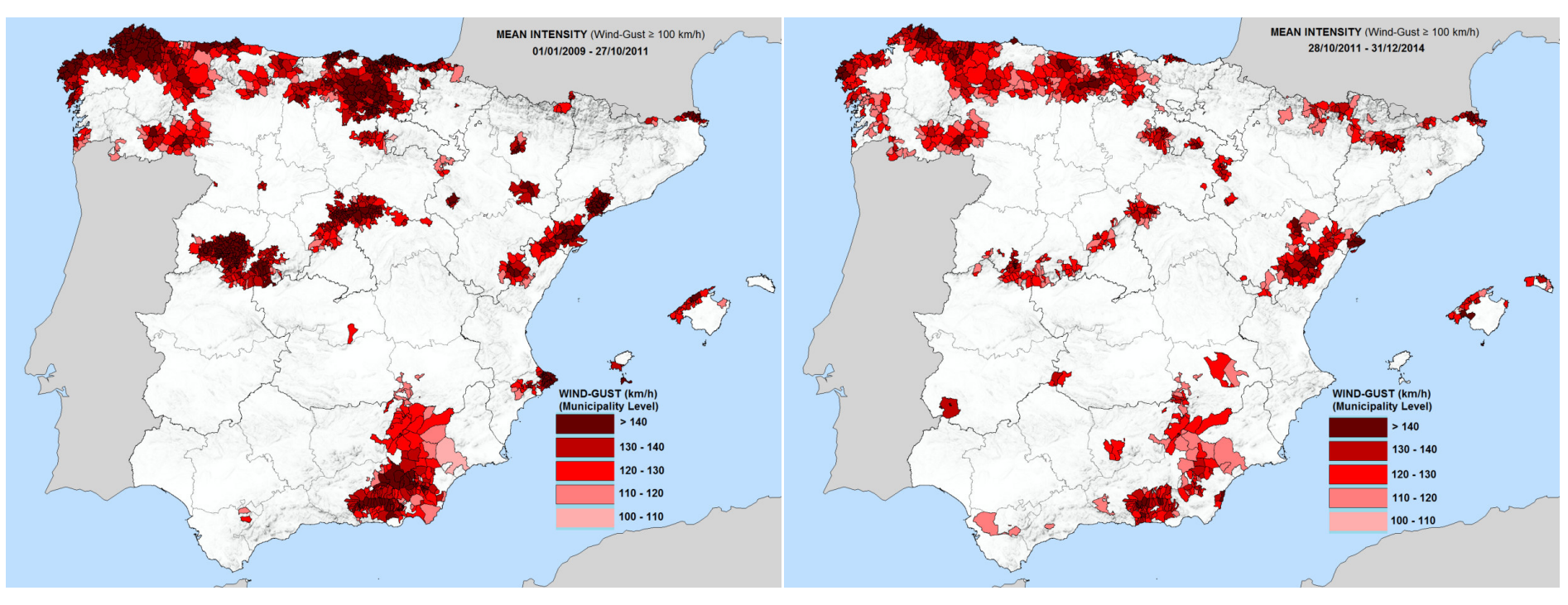
SPATIAL DISTRIBUTION

Intensity of the extreme wind-gust situations at municipality level.

MAXIMUM



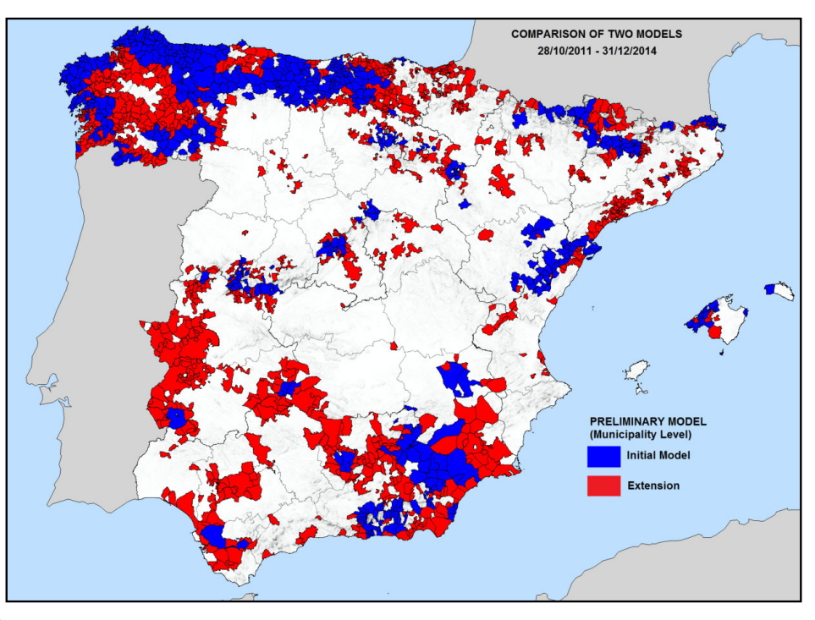
MEAN



SPATIAL DISTRIBUTION OF LOSSES

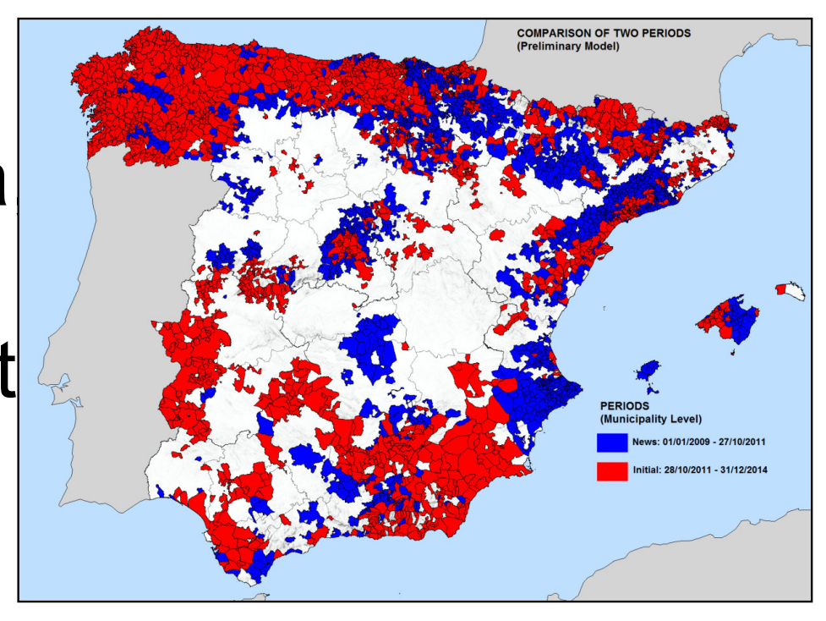
Estimation of expected loss cost from extraordinary winds, at municipality level, based on:
the climatology of the areas under extraordinary wind and the covered loss data from the insurer.

Study of models



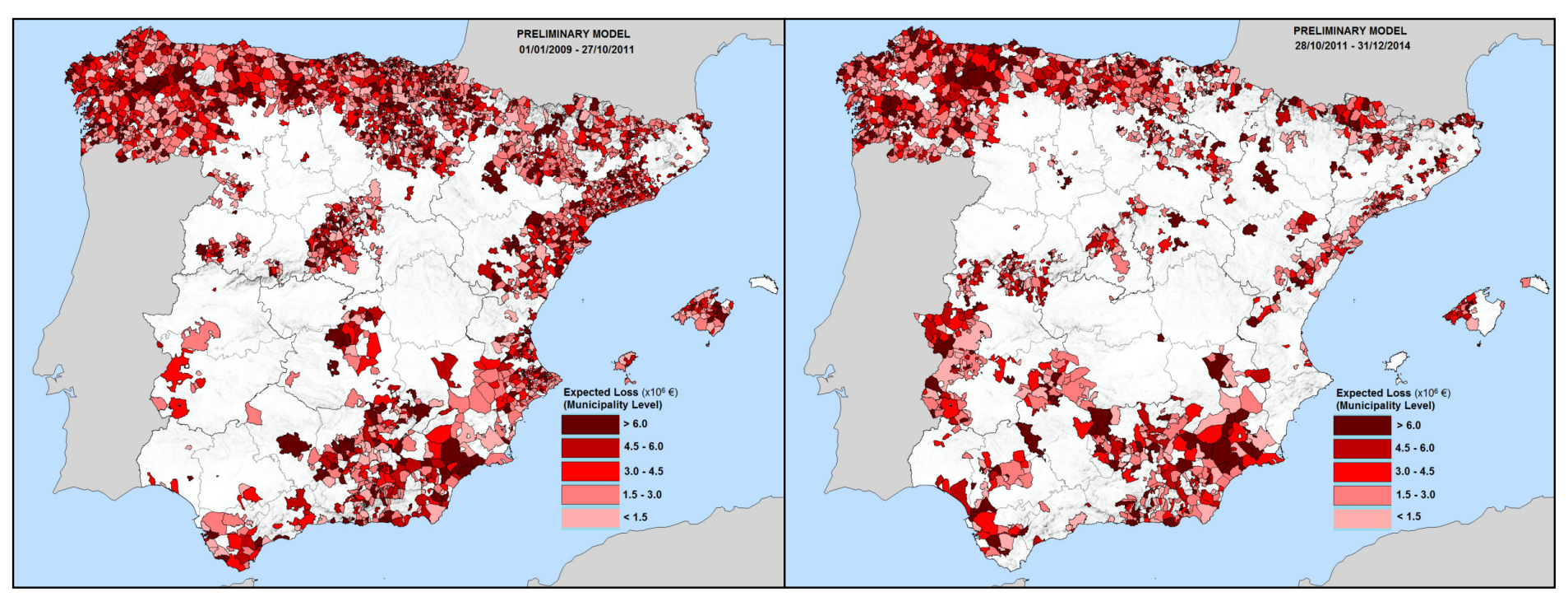
According to historical data the available information at municipality level.

Study of periods



PRELIMINARY MODEL

Expected Loss



Expected Loss Function
\+
Expected Intensity

DISASTER MODEL

The final methodology for the estimation of expected loss from extraordinary winds will apply a function based on the values of preliminary model and drawing also on physiographic variables, terrain elevation and distance from the shore, and population data or land registry. Furthermore, the catastrophic model can be improved by introducing other variables to measure the intensity of the situation (maximum intensity, mean intensity…)

SYNCHRONISATION

Climatology Date
Cost Date –from insurer-

DATA AND TOOLS

- AEMET and CCS Historical Data.
- R: A language and environment for statistical computing. URL: http://www.R-project.org/
- SAGA: System for Automated Geoscientific Analyses. URL: http://www.saga-gis.org/

ACKNOWLEDGEMENTS

This project is being sponsored by **CCS**, **AEMET** and **ATSISTEMAS**. The authors would like to recognize the support of many people belonging to these organizations for making this project possibe.

Adaptation to climate change in Hungary, the NAGIS project

Zita Bihari
bihari.z@met.hu
Hungarian Meteorological Service

The NAGiS project

The overall objective of the **National Adaptation Geo-information System (NAGiS)** project is to develop a multipurpose geo-information system that can facilitate the policy-making, strategy-building and decision-making process related to the impact assessment of climate change and the adoption of necessary adaptation measures in Hungary. Considerably NAGiS aims to play a key role in addressing the global challenge of climate change and developing smart adaptation measures tailored to regional, local needs.

Project promoter is the Geological and Geophysical Institute of Hungary.

Link: http://nater.mfgi.hu/en

Hungarian Meteorological Service (OMSZ) has been involved in the project from the beginning with producing climate data for the periods 1961-2010, 2021-2050 and 2071-2100.

- 0.1° spatial resolution
- Homogenized gridded data for the period 1961-2010 (extension of CARPATCLIM data for the territory of Hungary)
- Climate projections for 2 targets:
  1. 2021–2050: „short-term" planning
  2. 2071–2100: long-term strategy, robustness & significance

Impact studies in NAGiS based on meteorological data:

- Hydrology: ground water, drinking water
- Natural ecosystems
-Agriculture, forestry

The CRIGiS project

Recently OMSZ is supported in the frame of "EEA and Norway Financial Mechanism Hungary – Climate Change" for the extension of NAGiS to other sectors.

The main objective of the "Vulnerability/Impact Studies with a focus on Tourism and Critical Infrastructures (CRIGiS)" project is to prepare indicators based on existing dataset of NAGiS and on database to be established newly in the project. These indicators will be used to assess the vulnerability (due to climate change) which will foster the development of adaptation strategies and objective decision making.

The project is focusing on three important sectors within the tourism and critical infrastructure:

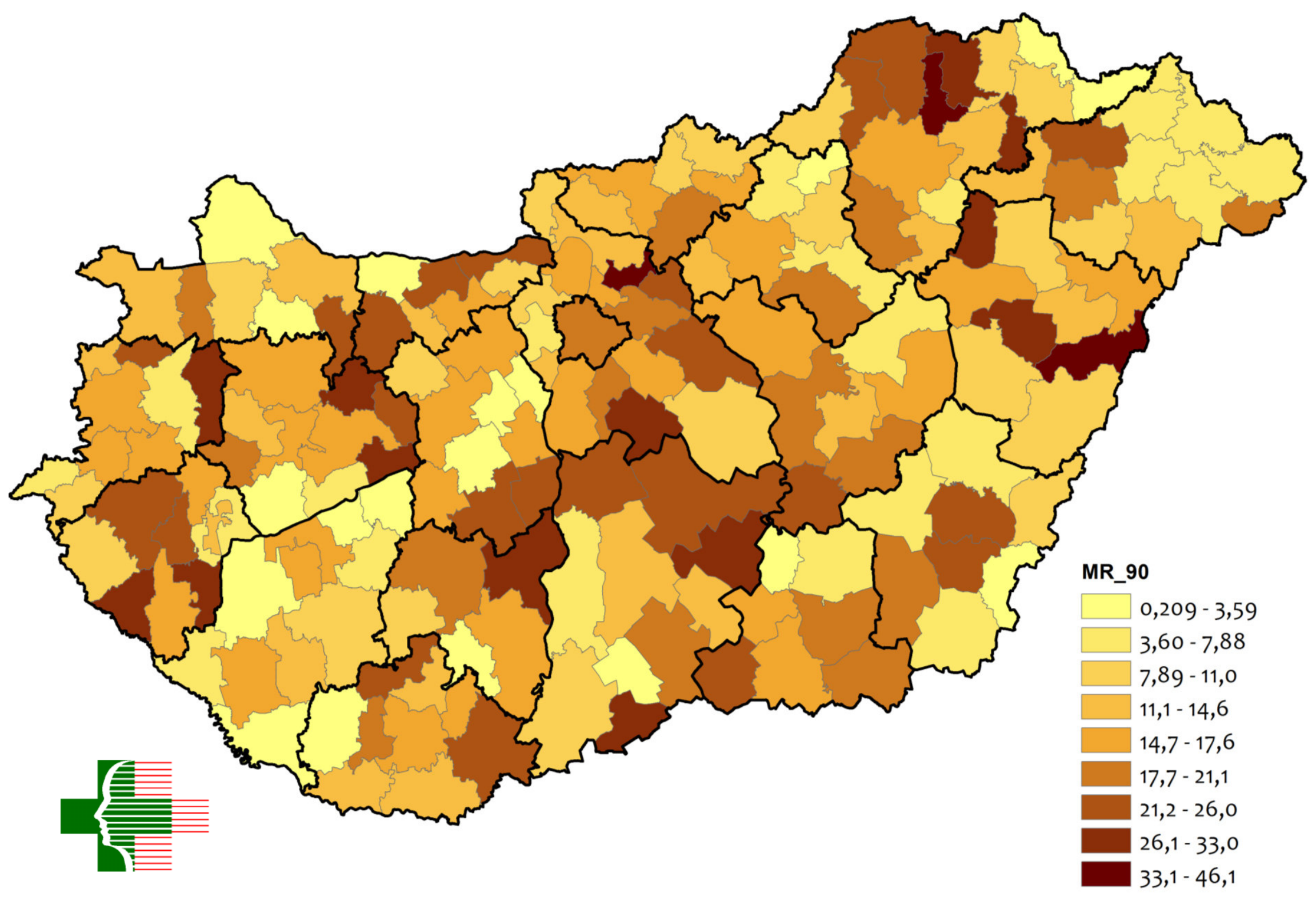
- Heatwave-induced excess mortality,
- Impacts of extreme weather events on road accidents,
- Impacts of climatic conditions on tourism.

The indicators will be quantified, on the one hand, for the past and present based on observational data, and on the other hand, for the future based on regional climate model outputs.

Link: kriter.met.hu

Preliminary results of the CRIGiS project

Heatwave-induced excess mortality

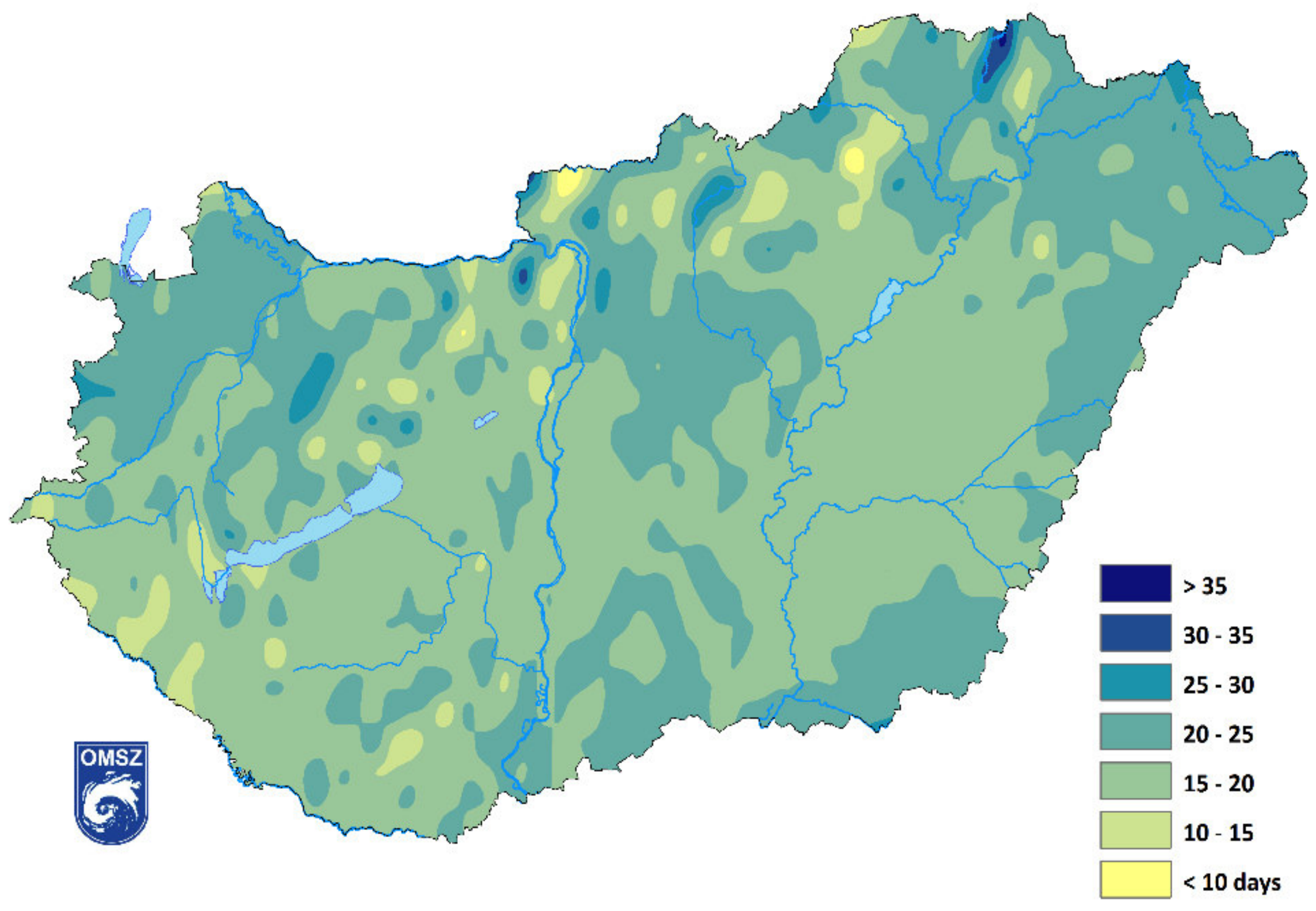


Excess mortality related to heatwaves (%) in summer (01.05-30.10), 2005-2013

Definition of heatwaves: days with temperature exceeding the 90% percentile of daily mean temperature

Páldy, A. , Bobvos, J.
National Centre of Environmental Health

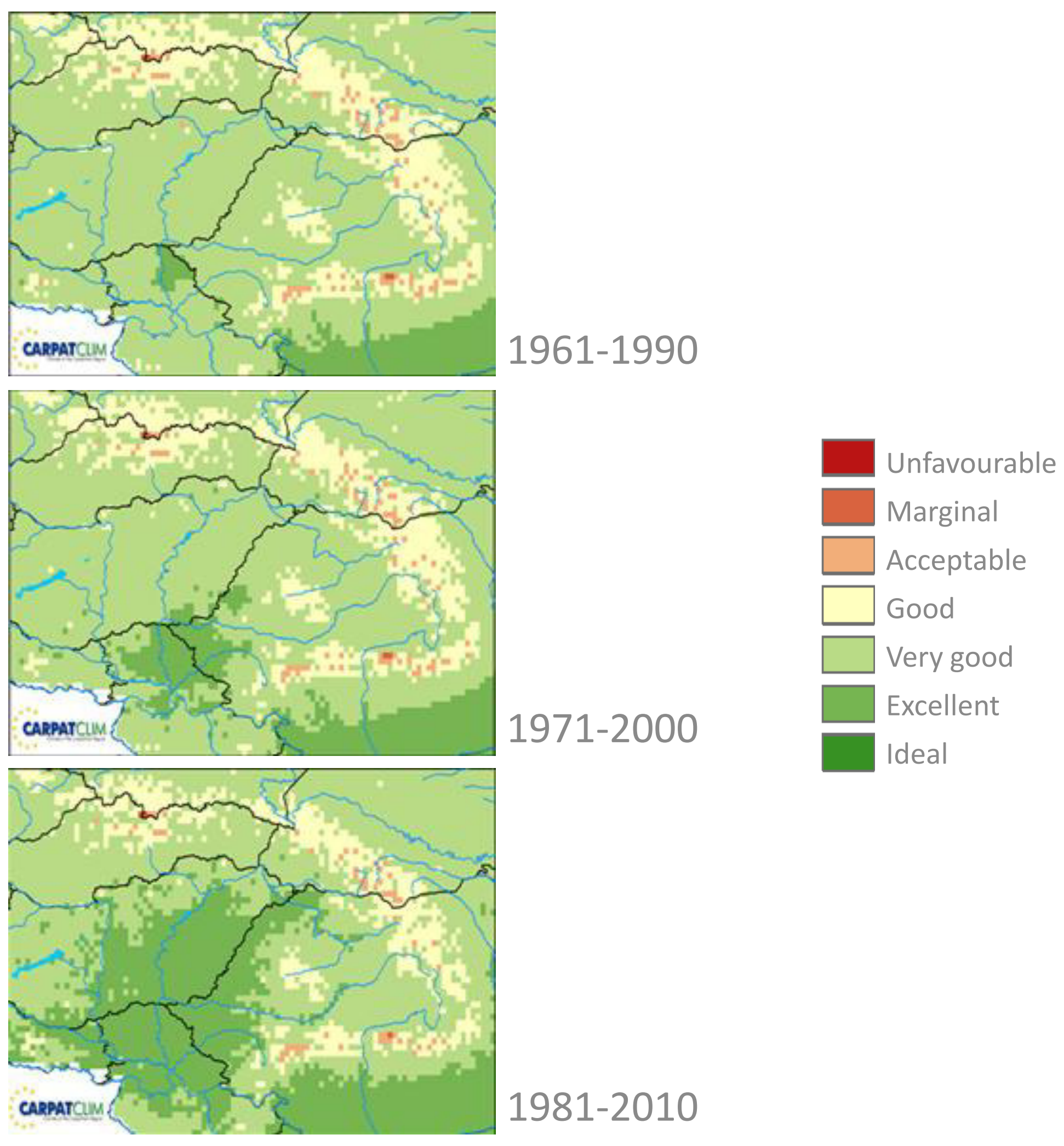
Impacts of extreme weather events on road accidents



Number of "zero crossing days with precipitation", 1981-2010

" Zero crossing days with precipitation" is one of the indicators that we apply in the project to describe the dangerous weather situations

Impacts of climatic conditions on tourism



Average of Tourism Climatic Index of May in three 30 years periods in the CARPATCLIM region

The size of "excellent" areas increased in the southern region

Consortium:
Lead partner: Hungarian Meteorological Service
National Centre of Environmental Health
National Directorate General for Disaster Management, Ministry of the Interior
University of Szeged, Department of Climatology and Landscape Ecology