11th EUMETNET
Data Management Workshop

Placing climate data to social service: From observations to archives

ZAGREB, CROATIA
18 – 20 OCTOBER 2017

Programme

Wednesday, 18th October 2017

08:30 – 09:30 **registration**

09:30 – 10:30 Welcome addresses – Meteorological and Hydrological Service of Croatia and Croatian State Archive representatives and WMO talk (Peer Hechler)

10:30 – 11:00 **coffee break**

Session 1: Data Rescue
Chair: Peer Hechler

11:00 – 11:20 J. Sigró: The UERRA approach and results to enhance data availability and accessibility to support European regional reanalyses of high-resolution
11:20 – 11:40 L.Wern: Data rescue of Swedish data
11:40 – 12:00 Discussion

12:00 – 13:30 **lunch break**

Session 2: Data Quality Control
Chair: Jose Guijarro

13:30 – 13:50 S. Noone: The Copernicus Climate Change Service Global Land and Marine Observations Database
13:50 – 14:10 M.T. Burguera: Daily quality control of climatic variables involved in SPEI drought index calculation
14:10 – 14:30 C. Sigg: Probabilistic Plausibility of Surface Data
14:30 – 14:50 Short poster presentations

14:50 – 15:20 **coffee break**

15:20 – 16:00 **Poster Session and DARE Side Meeting**

17:00 – **Guided city tour**
19:00 – **Icebreaker**

Thursday, 19th October 2017

Session 2: Data Quality Control
Chair: Dubravka Rasol

09:00 – 09:20 Y. Brugnara: The EUSTACE daily LSAT dataset: A collection of 36000+ unique in-situ daily temperature series with inhomogeneities identified - CANCELLED

09:20 – 09:40 M. Musa: Next Generation of Quality Management Tools at MeteoSwiss

09:40 – 10:00 C. Hutin: New rainfall and climate quality control systems at the Met Office

10:00 – 10:20 L. Wern: Quality control, archiving and retrieving data Swedish data

10:20 – 10:50 **coffee break**

Session 2: Data Quality Control
Chair: Ole Einar Tveito

10:50 – 11:10 T. Szentimrey: Automated statistical quality control (QC) procedures in MASH-MISH systems

11:10 – 11:30 A. Paul: Digitalization and Data Quality Control of Historical Data at ZAMG

11:30 – 11:50 P. Petrovic: The Validity of Snow Density and Snow Water Equivalent Measurements In Serbia

11:50 – 12:10 G. Vertacnik: Detection of temperature inhomogeneities using observations of snow and rain

12:10 – 13:30 **lunch break**

Session 3: Homogenisation
Chair: Ingeborg Auer

13:30 – 13:50 A. Wypych: The longest meteorological measurement series in southern Poland

13:50 – 14:10 V. Venema: The error worlds of the global benchmarks for the International Surface Temperature Initiative (ISTI)

14:10 – 14:30 J. A. Guijarro: Data Rescue of two centennial Spanish series (Maó and Oviedo)

14:30 – 14:50 J. A. Guijarro: Series homogenization, missing data filling and gridded products with Climatol

14:50 – 15:20 **coffee break**

15:20 – 16:00 **Poster Session and HISTALP Side Meeting**

19:00 – **Workshop Dinner**

Friday, 20th October 2017

Session 3: Homogenisation
Chair: Mark McCarthy / Dan Hollis

09:00 – 09:20 P. Domonkos: Temperature homogenisation with ACMANTv3: Explanatory factors of high accuracy
09:20 – 09:40 E. Lundstad: Increasing quality of long-term time series by combining new homogenization routines and siting classification
09:40 – 10:00 A. Argiriou: Homogenization of the Hellenic cloudiness time series
10:00 – 10:20 Discussion

10:20 – 10:50 **coffee break**

Session 4: Data Bases
Chair: Anita Paul

10:50 – 11:10 M. Prohom: ARTYDOC, the digital archive of the Meteorological Service of Catalonia
11:10 – 11:30 A. Squintu: Quality Check and Homogenization of ECA&D temperature dataset
11:30 – 11:50 G. Van der Schrier: An Ensemble Version of the Daily E-OBS Dataset
11:50 – 12:10 O.E. Tveito: Preparing for the 1991-2020 standard climate normals

12:10 – 13:30 **lunch break**

Session 5: Data Analysis
Chair: Victor Venema

13:30 – 13:50 M. McCarthy: Management and development of a suite of national climate monitoring products for the UK
13:50 – 14:10 A. Baronetti: Assessment of daily rainfall data recorded by two different networks in Piedmont (North-West Italy)

14:10 – 14:50 **Final Discussion**

14:50 – 15:20 **coffee/Closing of DMW**

SESSION 1:
DATA RESCUE

The UERRA approach and results to enhance data availability and accessibility to support European regional reanalyses of high-resolution

M. Brunet(1,2), L. Ashcroft(1), J.R. Coll(1), A. Gilabert(1), **J. Sigró**(1), P. Jones(2), P. Unden(3)

(1) Centre for Climate Change, University Rovira i Virgili, Tarragona, Spain
(2) Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom
(3) Sveriges Meteorologiska och Hydrologiska Institut (SMHI), Sweden
e-mail: manola.brunet@urv.cat

Under the EU funding support, the UERRA (Uncertainties in Ensembles of Regional Reanalysis) ongoing project (2014-2017) was aimed at producing ensembles of regional reanalyses for several decades and estimating the associated uncertainties in the derived data sets. This project has a core component of data rescue (DARE) of meteorological observations at the synoptic scale over European data-sparse sub-regions and periods.

In this contribution, we will give detailed information about approaches and outcomes of the DARE activities carried out, what will include information about where the recovered datasets are publicly available. Under the UERRA-DARE component, about 9 million (~142 million) of station values were recovered (accessed from open data sources in Catalonia, Norway and Sweden through involving relevant staff in these meteorological services) over European datasparse sub-regions for the pre- and post-1950 periods.

DARE effort focused on rescuing synoptic observations for meteorological stations in western and eastern Mediterranean, North Africa, the Middle East, the Balkans, central and eastern Europe and Scandinavia mostly for the post-1950 period. These targets were adopted after exploring data availability in one of the most used climate data archives in Europe: the Meteorological Archival and Retrieval System (MARS) at the European Centre for Medium Range Weather Forecast (ECMWF). The targeted variables have been air pressure, air temperature, dew-point, relative humidity, cloud cover and wind speed and wind direction (precipitation, snow-depth, fresh snow) at the hourly (daily and sub-daily) scales.

Data rescue of Swedish data

Sverker Hellström, **Lenart Wern**

Swedish Meteorological and Hydrological Institute, Sweden
lennart.wern@smhi.se

- Digitization of meteorological data in Sweden – experiences and challenges
- How much is done, how much data is still only in paper journals
- Digitization tool – data is immediately stored in our database
- Tool for correcting already existing data in our data base
- Meta data – station numbers, geographical location, observational times
- Challenges and problems – Observations at non-standard observing times, mixed up stations series and so on.

SESSION 2:
DATA QUALITY CONTROL

The Copernicus Climate Change Service Global Land and Marine Observations Database

Peter Thorne (1), Corinne Voces (1), Matthew Menne (2), Eric Freeman (2), Robert Dunn (3), David Berry (4), Ag Stephens (5), Liz Kent (4), John Kennedy (3), Chris Atkinson (3), Kate Willett (3), Jay Lawrimore (2), Huai-Min Zhang (2), **Simon Noone** (1), and Anthony Kettle (1)

(1) Maynooth University, Irish Climate Analysis and Research Units, Geography, Maynooth, Ireland
(2) NOAA's National Centers for Environmental Information, Asheville, NC, USA
(3) Met Office, Exeter, UK
(4) National Oceanography Centre, Southampton, UK
(5) Science and technology Facilities Council, Swindon, UK
e-mail: peter.thorne@nuim.ie

This presentation shall outline the planned service provision for a new Copernicus Climate Change Service concerning the availability of in-situ fundamental climate data records. The service brings together a number of European parties working in tandem with NOAA NCEI to provide via the C3S Data Store improved access to land and marine surface meteorological records for climate research. This presentation shall provide a high level overview of service aims and timeline. On the marine side the service shall aim to improve the existing ICOADS holdings with improved quality flagging, duplicate removal etc. On the land side a set of integrated holdings across Essential Climate Variables and timescales is envisaged. Data shall be made available via the C3S data store under a common data model. The Service shall interact with sister lots concerned with data rescue, provision of baseline /reference network data, and provision of in-situ data products and the broader Copernicus Climate Change Service and Copernicus services.

Daily quality control of climatic variables involved in SPEI drought index calculation

Miquel Tomas Burguera(1), Sergio M. Vicente Serrano(2), Santiago Beguería(1), Maria Yolanda Luna(3), Ana Morata(3), José Carlos Gonzalez-Hidalgo(4), Azucena Jiménez4

(1) Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas, Avda, Spain
(2) Instituto Pirenaico de Ecología, Consejo Superior de Investigaciones Científicas, Zaragoza, Spain
(3) Agencia Estatal de Meteorología (AEMET), Spain
(4) Departamento de Geografía y Ordenación del Territorio, Universidad de Zaragoza, Zaragoza,Spain
e-mail: mtomas@eead.csic.es

A quality control procedure has been implemented in R to detect suspicious daily data of 6 climatic variables (maximum temperature, minimum temperature, relative humidity, wind speed, sunshine duration and rainfall) from the national database of AEMET. The variables were selected due to their importance in the SPEI calculation.
We divided the quality control in the detection of two types of suspicious data: i) codification errors and ii) abnormal values. A final step of spatial coherence was tested but not implemented due to its high percentage of false negatives. The coherence between variables is only tested for maximum and minimum temperature. While the general scheme of the implemented controls is the same for all the variables, some specific controls and thresholds exist for each variable.
As codification errors we classified those errors related with detection of: i) exact values appearing at distinct points of the database during a minimum number of days duplicated data, ii) the same value appearing in ‘n’ consecutive days, depending the threshold ‘n’ on the variable (consecutive values) and iii) values codified using non-standard units for the variable according to AEMET metadata (erroneous units).
As abnormal values controls we classified those errors related with: i) values out of the range of the variable, ii) extreme values considered not feasible in Spain and iii) extreme values considered as suspicious values.
Results obtained shows that a procedure of detection of codification errors has to be implemented in the whole climate database, due to the detection of duplicated data between distant points, temporally and spatially speaking.
While the temporal analysis of detected errors clearly show some specific moments when errors are more present in the database (apparently related with digitalization processes), the spatial analysis does not show a clear pattern.

Improving marine quality control processes for data from automatic systems
CANCELLED

Fraser Cunningham

Marine Data Specialist, Met Office, Edinburgh, UK
e-mail: fraser.cunningham@metoffice.gov.uk

The current suite of marine quality control (QC) checks used by the Observations Quality Management (OBQM) team at the Met Office were designed specifically for manual observing ships reporting in delayed-mode on the synoptic hours. With the continued drive towards automation, the number of hourly observations received from automatic networks, for example the Met Office's ship borne Autonomous Marine Observing System (AMOS) network, is only going to increase. This represents a significant problem as many of the present QC checks applied are not appropriate for hourly data from automatic systems, or simply do not work on data received in real-time.

To address this, the OBQM team embarked on a project to modernise QC processes and software in order to improve the quality of all marine observations and create efficiencies. The vision for the project was:

"To improve the existing marine QC checks, QC processes and database infrastructure to better accommodate automatic observations in addition to the existing delayed mode data in order to make QC processes more efficient and effective"

This paper will outline the goals of the project, the approach taken and report on successful outcomes.

Probabilistic Plausibility of Surface Data

Christian Sigg

Federal Office of Meteorology and Climatology MeteoSwiss, Swiss
e-mail: christian.sigg@meteoswiss.ch

MeteoSwiss is upgrading its meteorological data processing system, with the data ware-house at its center, surrounded by closely integrated applications for the aggregation, com-putation, interpolation and quality control (QC) of measurements. The upgrade consists of a technology refresh, a new architecture for the processing applications and an improved metadata model. This submission focuses on the novel representation and computation of probabilistic quality information (QI). An overview of the whole effort is provided in “Next Generation of Quality Management Tools at MeteoSwiss” submitted by Marc Musa.

In the old system, the QI consisted of discrete flags that were stored with each measure-ment. QC test results were grouped in four classes: physical limits, climatological limits, interparameter consistency and spatial consistency. If a measurement failed a test, the flag of the corresponding class was set. Although the flag based QI seems straightforward, it is difficult to interpret by the user. The strength of the test evidence, i.e. whether a measurement value is plausible or implausible given a certain flag, varies significantly between different test classes and measurement parameters. Thus, QI flags beyond physical limits found rare use in practice.

In the new system, the QI is the probabilistic plausibility of the measurement, represented on a continuous scale between 0 (implausible) and 1 (plausible). The probabilistic plausibility is calculated from the a-priori plausibility of the measurement and all available test results, following the well-known statistical procedure called Naďve Bayes. Using this procedure, each test result influences the probabilistic plausibility exactly according to its evidence. Summarizing all the available test evidence in a single number makes it easier for our users to interpret the QI and to set an appropriate threshold on the minimal data quality that they need in their application.

I will give a concise introduction to the Naďve Bayes method and its application to the QC of meteorological data. Using concrete examples, I will illustrate how automated testing and expert inspection combine to produce the probabilistic plausibility of a measurement value. Finally, I will present solutions to important practical issues, e.g. how to minimize the amount of metadata that needs to be maintained and how to add tests without extensive recalculations.

The EUSTACE daily LSAT dataset: A collection of 36000+ unique in-situ daily temperature series with inhomogeneities identified

CANCELLED

Yuri Brugnara(1,2) and Stefan Brönnimann(1,2)

(1) Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland
(2) Institute of Geography, University of Bern, Bern, Switzerland
e-mail: yuri.brugnara@giub.unibe.ch

EUSTACE is a EU Horizon 2020 project whose goal is to produce 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. As part of this, tens of thousands of in-situ data series were collected from various providers.

Surface in-situ observations are affected by several problems, in particular by inconsistencies in space and time in the measurement procedures, and by the lack of information concerning these procedures. Moreover, human errors often contaminate the quality of the data in different ways. The absence of an official global repository for meteorological observations means that often data undergo numerous modifications by different users and different versions of the same series can be found even within the same collection, with usually no information on the modifications applied.

Within EUSTACE, public collections (mainly GHCN-D, ISTI, and ECA&D) of daily maximum and minimum temperature observations, together with some additional dataset with restricted data policy, were merged. Duplicates were removed by comparing the observations in each individual year of each series with the data of the stations located within a 200 km radius. Moreover, an automatic quality control algorithm was applied to all series. We also carried out an homogeneity assessment, by using four different breakpoint detection algorithms (three relative tests and one absolute test), and provide breakpoint locations for each series.

The final result is a quality-controlled dataset of ca. 750 million unique observations over the period 1850-2015 that will be assimilated into other EUSTACE products. The dataset will be publicly available in NetCDF format, with the exception of a small fraction of the data (due to the policy of data providers); nonetheless the information derived from this fraction (e.g., for the homogeneity tests) will be public.

Next Generation of Quality Management Tools at MeteoSwiss

Marc Musa, Markus Abbt, Deborah van Geijtenbeek, Christian Sigg

Swiss Federal Office of Meteorology and Climatology MeteoSwiss, Zurich Airport, Swiss
e-mail: marc.musa@meteoswiss.ch

MeteoSwiss is upgrading its meteorological data processing system, with the Data Warehouse (DWH) at its central integrated data platform at MeteoSwiss, surrounded by closely integrated applications for the aggregation, computation, interpolation and quality control of measurements.

The existing quality management tools were developed 25 years ago and later integrated with the DWH. These tools cover the aspects of real time quality control and calculation, interpolation and enhanced quality control plus manual data correction. Over time the requirements regarding quality control have continually increased.

In order to meet these new requirement due to functional restrictions and technical reasons, MeteoSwiss decided to renew its concept for the whole quality control system for surface weather stations until the end 2018. The new concept comprises the replacement of the existing core applications in the DWH with a set of core quality tests, which will allow to use also other data sources like model or radar data besides the used surface data. Furthermore, the new system will provide standard interfaces for quality control applications outside of the core DWH system that provide more complex algorithms such as spatial consistency tests.

As a second project activity, the metadata models for aggregation and plausibility testing were re-designed to be simpler and more general, with the goal of reducing errors and effort in configuration data management.

The presentation gives an overview of the concept and targets of the project and illustrates the advantages of the new system compared to the old one with its limitations. As part of the project, MeteoSwiss also adapts its way to describe the quality information of measured or processed values and the way it will be integrated into the data delivery. The concept of the new quality information will be presented in “Probabilistic Plausibility of Surface Data” submitted by Christian Sigg.

New rainfall and climate quality control systems at the Met Office

Martyn Sunter

Observations Quality Manager, Met Office, Edinburgh, UK
e-mail: martyn.sunter@metoffice.gov.uk

Good quality daily observations of rainfall and temperature are vital in monitoring our climate and placing extreme events in a historical context. Rainfall observations are used widely within meteorology and hydrology for things like drought and flood management and to determine return periods for high intensity events. High quality temperature observations help verify weather and climate prediction models and so are crucial for climate change predictions and adaptation advice.
Over the past two years the Met Office has developed new rainfall and temperature quality control software. This has provided resilience in our systems, improved functionality and created efficiencies in the processing.
This paper will provide an overview of our new quality control software used for daily rainfall and temperature observations. The new system for rainfall, known as SODA, provides a map based user interface which compares observations spatially. Outliers are flagged by the system and highlighted to quality control staff for further investigation. The system may indicate the likely cause of the error and is able to make automatic corrections by reapportioning multi-day accumulations, for example. Quality control staff overlay rainfall radar accumulations and lightning strike maps to help assess the validity of the observations and make changes as required. A screen shot of the new system is shown below. The traditional technique for the quality control of daily climate observations (e.g. daily maximum and minimum temperature and daily sunshine) was a manual process which involved the inspection of large amounts of data by quality control staff. The new software has automated many of the processes and created significant efficiencies. The paper will explain the processes that have been developed which include the spatial comparison of temperatures to identify potentially suspect data. The can then be investigated further by the data analyst. Finally, we will present our future plans for improvements to our quality control systems.

Quality control, archiving and retrieving data Swedish data

Lennart Wern

Swedish Meteorological and Hydrological Institute, Sweden
lennart.wern@smhi.se

After data has been digitized and imported into the data base, it is very important to check data in different ways, step check, geographical check etc. When errors in data are found, data must be corrected in the data base as soon as possible. If data is not quality controlled by a skilled person wrong decisions can be taken by a stakeholder.
In Sweden meteorological observations are open and free for everyone to retrieve and use for any purposes. Tools for retrieving data are available and easy to use.

Automated statistical quality control (QC) procedures in MASH-MISH systems

Tamás Szentimrey, Zita Bihari, Mónika Lakatos

Hungarian Meteorological Service, Hungary
e-mail:szentimrey.t@met.hu

The automated quality control (QC) procedures in order to indicate or correct the wrong data are indispensable in the databases. However the development of such software systems is impossible without adequate and advanced mathematical fundament. During the years we developed some software systems that are MASH (Multiple Analysis of Series for Homogenization) and MISH (Meteorological Interpolation based on Surface Homogenized Data Basis), with special attention for the strong mathematical basis.
The MASH system can be used for homogenization of daily and monthly series, furthermore in the system also an automated quality control (QC) unit has been built for the daily and monthly data. The principle of the QC procedure is certain multiple comparisons of the data by spatial interpolation technique and the analysis of interpolation errors for detection of wrong data is based on confidence intervals. In MASH system the necessary climate statistical parameters, such as expected values, spatial and temporal covariance structure, are estimated on the basis of examined data series by classic statistical way.
The MISH system can be used for spatial interpolation. The main difference between MISH and the geostatistical interpolation methods can be found in the amount of information used for modelling the necessary statistical parameters. In general at the geostatistical methods built in GIS the sample for modelling is only the predictors, which is a single realization in time. At MISH method we use the spatiotemporal data for modelling since the long data series form a sample in time and space as well. The long data series is such a specialty of the meteorology that makes possible to model efficiently the statistical parameters in question.
The new development of MISH is an automated real time quality control (QC) procedure for observed daily and monthly data. The mathematical principle of this procedure is the same as the QC procedure built in MASH. The main difference between the MASH and MISH QC procedures is, while at MASH it is developed for time series and the statistical parameters are estimated from the series in classic statistical way, at MISH it is a real time test and modelled statistical parameters are used.

Digitalization and Data Quality Control of Historical Data at ZAMG

Anita Paul, Silke Adler, Joachim Albenberger, Hermann Galavics, Wolfgang Lechner, Wolfgang Lipa, Angelika Manhard, Alexander Mandl, Daniela Teuschler

ZAMG-Zentralanstalt für Meteorologie und Geodynamik, Vienna, Austria
e-mail: anita.paul@zamg.ac.at

Extreme events, like long or heavy periods of precipitation, drought periods or late frostcombined with masses of warm air brought on by an early spring – as this year - affects both people and nature. Negative effects of this climate change not only provoke debates but lead to crop failure (due to damage caused by hail or flash flooding) and generate financial losses throughout the global community. So –

- What is in range and what is left outstanding?
- Has this historically been a trend and can we determine trends based on this data?
- Is the climate, in general, getting more extreme?

Such questions could be only answered in the framework of climate change studies and analysis, which need long term data in high temporal resolution in order to be valid. This is the primary purpose of data rescue projects and our starting point for this type of analysis through the digitalization of historical data.

Since 2008 we have been digitalizing historical climate sheets that are stored in our ZAMGArchive. Simultaneously, we worked on completing a full data inventory of the historical (1851-1948) ZAMG yearbooks.

A further and deeper step in the complete digitalization of the historical yearbooks is planned. First we will scan the yearbooks using 'Bookeye' (specialized image access scanner) and subsequently digitalize the pdf-files with OCR (Optical Character Recognition). The final step will be to initiate the data quality control with our in-house developed software application DCT (Data Correction Tool). This tool will check the plausibility of the inputted data using a Multi-stage-quality-control so we are confident in the analysis we are providing.

This presentation shows the importance of recovering and maintaining historical data for climate change research and gives a general overview of the quality control measures applied at our institute.

The Validity of Snow Density and Snow Water Equivalent Measurements In Serbia

Predrag Petrović

Republic Hydrometeorological Service of Serbia, Belgrade, Serbia
email: predrag.petrovic@hidmet.gov.rs

Although snow density and snow water equivalent are out of focus in climate surveys, these data are still very important. The utilization of these data is more oriented to dealing with some practical problems, including water management or civil engineering rather than to science itself. Since good results always come from high quality resources, it is necessary to examine validity of these data.

A simple method for snow density / snow water equivalent data quality control is designed on a basis of two main conditions: a) snow water equivalent should not be larger than total precipitation since formation of snow cover, and b) bulk snow density cannot be lower without new snowfall.

This method creates upper and lower limit of expected bulk snow density / snow water equivalent. Thus, measured snow water equivalent might be smaller than the lower limit, which indicates errors in measurements and / or reports. On the other hand, measured snow water equivalent might be even significantly higher than the upper limit. Such records are featured especially in a mountainous regions rather than in lowlands. These are not necessarily errors, since snow samples are most likely taken from thicker layers of snow packed from wind gusts and blizzards.

The utilization of these data might decide about their validity. While snow water equivalent higher than the model's upper limit might be false for water management or hydrological cycle, these data might be quite useful for calculating maximum snow loads in civil engineering.

Detection of temperature inhomogeneities using observations of snow and rain

Gregor Vretačnik

Slovenian Environment Agency, Slovenia
e-mail: gregor.vertacnik@gov.si

Modern homogenisation methods are usually based on comparison of station time series within the same climate region. Succesfulness of such approach is highly dependant on station density and thus less appropriate for early period of measurements. Old temperature data are often supplemented by precipitation measurements and observations of weather phenomena.
It may be possible to extract aditional information from suplemental data for homogenisation of temperature series. In a moderate climate with a mix of rainfall and snowfall in winter one can compare temperature data with precipitation type observations in order to assess changes in temperature-precipitation type relation. As snow-to-rain transition near 0 °C is only weakly dependant on climate, changes in the relation could be attributed to inappropriate microlocation of thermometer screen or thermometer bias. Temperature shift can be used for crude estimate of inhomogeneity size in winter mean temperature series. We show application of this proposed methods to some very long Slovenian temperature time series.

SESSION 3:
HOMOGENISATION

The longest meteorological measurement series in southern Poland

Agnieszka Wypych(1,2), Petr Stepanek(3), Zbigniew Ustrnul(1,2), Pavel Zahradnicek(3)

(1) Institute of Meteorology and Water Management – National Research Institute, Krakow, Poland
(2) Jagiellonian University, Krakow, Poland
(3) Global Change Research Centre, Czech Academy of Sciences, Brno, Czech Republic

Southern Poland is a region where meteorological observations have been conducted since the end of 18th century. The oldest station, located in Krakow, was established in and has been operating constantly for over 225 years. Before 1850s measurements were obtained at 17 stations (working during different periods and according to different manuals). The network was later extended and at the end of 19th century the number of stations located in the south of Poland was the largest comparing to other regions of the country.
The aim of the work is to prevent the oldest handwritten manuscripts from destruction. Data rescue programme has already been planned by Polish national meteorological service. 19 stations with measurements and visual observations dated back to the beginning of 20th century or earlier were designated for the first step of the project. The workflow schedules maintenance activities, digitalization as images to save the original handwriting or as numbers to make the data useful for further studies.
Research phase of the programme contains quality control and homogenization procedures. Historical station in Krakow, having the longest and unbroken records (already digitized in numeric format), was established as a main point and homogenized at first. Monthly air temperature data from the years 1792-2010 were processed using already homogenized secular series in the Czech Lands as reference stations. The results of homogenization procedures of ProClimDB software demonstrated significant impact of urbanization on air temperature measurements in Krakow with unchanged location of the station. Breaks detected in Krakow data series confirmed the necessity of detailed quality control and conscientious homogenization procedures for other stations after digitalization. Further analysis will be also conducted for other meteorological elements, especially air pressure, at selected stations.

The error worlds of the global benchmarks for the International Surface Temperature Initiative (ISTI)

Victor Venema (1), Katherine Willett (2), Renate Auchmann (3), Enric Aguilar (4), Peter Thorne (5), Claude Williams (6), Matt Menne (6), Lucie Vincent (7), Rachel Killick (2), Stefan Brönnimann (3), Zeke Hausfather (8), Ian Jolliffe (9), Thordis Thorarinsdotir (10), Steve Easterbrook (11), Robert Lund (12), Colin Gallagher (12), Giuseppina Lopardo (13), David Berry (14), and Lisa Alexander (15)

(1) University of Bonn, Meteorological Institute, Bonn, Germany;
(2) Met Office Hadley Centre, United Kingdom;
(3) Oeschger Center for Climate Change Research & Institute of Geography, University of Bern, Switzerland;
(4) Center for Climate Change, C3, University Rovira I Virgili, Tarragona, Spain;
(5) Geography Department, Maynooth University, Ireland;
(6) National Climatic Data Center, Asheville, North Carolina, USA;
(7) Climate Research Division, Environment Canada, Canada;
(8) Berkeley Earth, Berkeley, CA, USA;
(9) Exeter Climate Systems, University of Exeter, United Kingdom;
(10) Statistical Analysis, Pattern Recognition and Image Analysis (SAMBA), Norwegian Computing Centre, Oslo, Norway;
(11) Department of Computer Science, University of Toronto, Canada;
(12) Department of Mathematical Sciences, Clemson University, USA;
(13) Instituto Nazional di Recerca Metrologica (INRiM), Italy;
(14) National Oceanographic Centre, Southampton, UK;
(15) Climate Change Research Centre, University of New South Wales, Australia
e-mail: victor.venema@uni-bonn.de

Our surface temperature data are good enough to give us confidence that the world has warmed since 1880. However, the data are not perfect and the main source of uncertainty for secular trends is changes in the way temperature was observed and how well these inhomogeneities can be removed.

Previous assessments of homogenisation methods were based on regional and national networks. The Benchmarking and Assessment Working Group (BAWG) of the International Surface Temperature Initiative (ISTI; www.surfacetemperatures.org) will perform a global assessment replicating the network of the global ISTI raw data collection. The aim is to quantify the skill of homogenisation algorithms on a global scale against realistic benchmarks.

The benchmarking involves the creation of homogeneous synthetic worlds of station temperature data, the deliberate contamination of these with known inhomogeneities and the assessment of the ability of homogenisation algorithms to detect and remove these inhomogeneities. The ultimate aim is threefold: quantifying uncertainties in surface temperature data; enabling more meaningful product intercomparison; and improving homogenisation methods.

This presentation will detail how the inhomogeneities are generated. We will generate ten “blind” error worlds (errors are unknown to user) with different assumptions about the inhomogeneities and several “open” error worlds (errors are known). The aim is that all blind worlds are realistic, but we also aim to ensure that enough diversity is included so that reality likely falls within the spread. The open error worlds will be partially more idealised and can be useful for understanding the performance of the homogenisation methods.

The error worlds differ in the size distributions and frequencies of (large-scale trend bias producing) inhomogeneities. We will make several assumptions about the autocorrelations of the inhomogeneities (as a random walk and as a noise from baseline) and how the errors are correlated between stations and networks to produce global average trend errors.

Assessment of parallel temperature measurements networks

F. Acquaotta(2), A. Baronettia, S. Fratiannia(2), G. Fortinc, D. Garzenaa, D. Guenzia

(1) Dipartimento di Scienze della Terra, Universitŕ di Torino, Italy
(2) Centro Interdipartimentale sui Rischi Naturali in Ambiente Montano e Collinare, Universitŕ di Torino, Italy
(3) Département d'histoire et de géographie, Université de Moncton, Moncton, Canada
e-mail: fiorella.acquaotta@unito.it

Long historical climate records usually contain non-climatic changes that can influence the observed behaviour of meteorological variables. The availability of parallel measurements offers an ideal occasion to study these discontinuities as they record the same climate. The transition from manual to automatic measurements has been analysed in this study. The dataset has been obtained from two independent climate networks in the Piedmont region, in Northwest Italy and in Québec, Canada.

The first selection of pairs of stations was based on three parameters, that is, the overlapping period, the difference in elevation and the distance. The second step was to evaluate the exposure of the selected stations and their characteristics, the type of instrumentation, the neighbours and the general conditions. Therefore, the dataset contain 16 pairs of stations with up to 15 years of overlapping. On average, the overlapping period is 12 years with 8760 daily data matched for a pair of stations.

The dataset was divided in three groups by the instruments utilized. The first the pairs of stations have a manual thermometer and a thermograph, the second two thermograms, old type and new type, and the third a termistance and a temperature probe.

In order to be able to make a direct comparison between the daily temperature series, any values that were missing in one series were also set to be missing in its counterpart before the monthly statistics were computed. Non-parametric tests were applied to the daily values to evaluate the preliminary relationships between the pairs of series. The root mean square error (RMSE) was used to identify the mean difference between the two series, while the Spearman method was used to evaluate the correlation coefficient. The Kolmogorov–Smirnov test (KS) was applied to determine whether two datasets could have come from the same distribution, while the Wilcoxon rank sum test (W) was considered to establish whether two samples had identical population medians. A $p = 5\%$ significance level was used for all the tests. The thresholds by percentile were calculated on a daily scale to identify the different temperature types. Five classes of temperature were established: extremely cold, cold, medium, heat and extremely heat. The number of events and the mean values of temperature were calculated for each class and for each pair of series.

The transition between the networks has highlighted important differences in the temperature values. In the first and in the third group there are not a clear relationship between the instruments. The behaviour between the pairs of stations depend to the peculiarity of the locations. In the second group in the extreme classes, extreme cold and extreme heat, the new type of thermograms recorded a greater number of events. These differences produce a spurious change in the temperature of the analysed area, thus showing the importance of having a homogeneous dataset to identify real climate variations.

Series homogenization, missing data filling and gridded products with *Climatol*

J.A. Guijarro

State Meteorological Agency, Spain
e-mail: jguijarrop@aemet.es

Observational climatic series are very important to study climate variability and trends, both in time and space. Yet most often raw series bear missing data and inhomogeneities due to changes in the observation conditions (relocations, changes of instrumentation or in the environment, etc). Statistical methods are then needed to homogenize the series and fill their missing data in. Moreover, as new observations are continuously been added to the database, these statistical procedures need to be applied repeatedly, suggesting that they should be as automatic as possible.

The R package *Climatol* provides tools to easily obtain products derived from homogenized and completed climate series, implemented in functions such as:

- *db2dat*: to get data from a data-base and generate input files for *Climatol*.
- *homogen*: to homogenize series and fill all their missing data, even in very short lived series. The number of series is limited only by the available RAM in the computer. Can be applied to any climatic variable, at the daily or monthly scale. Detected break-point dates may be adjusted by known reliable metadata. Lots of diagnostic graphics allow quality control of the input data and the homogenization process, enabling the user to re-run the process with better adjusted parameters.
- *dahstat*: to list means (climate normals), percentiles, trends and their p-values, etc.
- *dahgrid*: to interpolate homogenized and normalized values in a user-defined grid and save them in NetCDF format.
- *rosavent*: to plot wind-roses.
- *diagwl*: to plot Walter&Lieth climate diagrams.

Recent experiences with this package include the homogenization of monthly precipitations and temperatures in Spain (and the calculation of 1981-2010 normals), monthly mean wind speeds in Spain, daily mean and maximum peak gusts in Australia, Denmark, Finland, Norway, Spain and Sweden, and daily temperature, precipitation, relative humidity and solar radiation series from the two Spanish Antarctic bases. (In data sparse areas, series from nearby reanalysis grid-points have effectively acted as suitable references.)

Temperature homogenisation with ACMANTv3: Explanatory factors of high accuracy

Peter Domonkos

University Rovira i Virgili, Spain
e-mail: peter.domonkos@urv.cat

In testing the efficiencies of homogenisation methods, ACMANT often produces the lowest residual root mean squared error and lowest residual trend bias. The success of ACMANT is based on three principles, i) appropriate model selection, ii) state-of-art statistical tools, iii) empirically developed algorithm and parameter setting with the use of an own developed test dataset.

The own test dataset consists of 20 subsets for the examination of 20 kinds of inhomogeneity problems. Each subset contains 200-500 networks with 5-20 time series in each. Inhomogeneities are randomly positioned and of random size. The total number of test series is 61000.

The explanatory factors of high efficiency in details:

A) In separating inhomogeneities from climate signal: i) Deviations from neighbouring series are considered with unifying the information by composite reference series, ii) Gaps of reference composites are filled with interpolation before the use of the series.

B) In the detection of long-term biases: i) The model of optimal segmentation approaches well real problems of relative homogenisation, ii) Optimal segmentation provides the best estimation of break positions for a given model setup, iii) Caussinus-Lyazrhi criterion is a good estimator for the number of breaks.

C) In the detection of short-term biases (up to 2 years), each short section of the series is compared with its temporal neighbourhood to detect possible outlier sections.

D) For adjustments: i) ANOVA correction model both for the corrections in annual mean and for finding the seasonality of biases, ii) Spatial interpolation for very short biases (< 6 months)

E) Three-step iteration for improving accuracy: i) Each step repeatedly contains all of the following operations: interpolation for gap-filling, construction of composite reference series, filtering of short-term biases, detection of long-term biases and the calculation of adjustment terms with ANOVA model. ii) In the first two steps ensemble homogenisation with the exclusion of one potential reference composite, then applying the minimum adjustment size of the ensemble members.

In addition, further favourable characteristics are included in ACMANTv3 in the detection of bias seasonality and daily data homogenisation.

Increasing quality of long-term time series by combining new homogenization routines and siting classification

Elin Lundstad(1), Mareile Wolff(1), Gabriel Kielland(1), Nina Elisabeth LarsgÍrd(1), Aslaug van Nes(1), Hildegunn NygÍrd(1), Ted Torfoss(1), Petr Stepanek(2)

(1) Norwegian Meteorological Institute, Observation and Climate Department, Norwegian Meteorological institute, Oslo, Norway
(2) Global Change Research Institute, Czech Academy of Sciences, Brno, Czech Republic
e-mail: elinl@met.no

Air temperature is one of the meteorological quantities whose measurements are particularly sensitive to exposure. For climate studies in particular, temperature measurements are affected by the state of the surroundings, by vegetation, by the presence of buildings and other objects, by ground cover, by the condition of, and changes in, the design of the radiation shield or screen, and by other changes in equipment (WMO, 2011). Consequently, WMO's commission for instruments and methods of observation (WMO-CIMO) has recently established a siting classification scheme to help determine the given site's exposure to possible artificial influences in the immediate vicinity. WMO-CIMO's siting classification scheme evaluates slope, vegetation, distance to heat sources and water bodies and shadow. The classification process helps network and station owners to consider the exposure rules in an objective manner, and also identifying means to improve the exposure of the the siting. Additionally, the siting environment can be documented in a comparable form. It is, however, still under discussion how correct the siting classification evaluates the siting exposure and several groups are working on quantification of the influences and improvement of the scheme.

The decision if a detected break in the times series of a site is caused by a change in the climate element or by an external change, depends highly on the recorded metadata for that site. For this study, the air temperature time series for selected sites in Norway were homogenized. Homogenization was performed for time series of the mean, maximum and minimum daily temperatures. Breaks which could be related to exposure changes are especially evaluated. Based on the historical metadata information available, the authors have applied the CIMO siting classification to all considered sites. Significant changes in the environment are reflected by changes in the siting class. The historical siting classes allowed a systematic analysis of which kind of environment changes translate to detectable breaks in a time series of air temperatures. It was found that certain changes were only visible in one or both of the extreme value time series while the time series of the mean temperature seemed unaffected. Also the order of magnitude and the "direction" were different for different changes.

In this paper the results of this study are presented. 44 meteorological stations in Norway are evaluated, homogenized and classified applying a common metadata scheme. The four criteria of the WMO CIMO siting classification (slope, vegetation, distance to heat sources and water bodies and shading) were analyzed separately for those stations. The analysis gives us results that evaluate patterns where we can match breaks and do the classification. We can fine-tune the classification and it gives us a better decision base for the breaks in the homogenization. For example, we can see from the analysis that what happens with temperature measurements if a car park is nearby or whether a house is being built or a greenhouse is being built. This combination with site classification gives an increased value in homogenization and helps to systematize the breaks. Also the site classification can be improved, as it can be checked for the different influences of the different exposure criteria (vegetation, heat sources, slope and shadow).

Homogenization of the Hellenic cloudiness time series

Athanassios A. Argiriou(1), Anna Mamara(2), Elias Dimadis(1)

(1) Laboratory of Atmospheric Physics, Department of Physics, University of Patras, Greece
(2) Hellenic National Meteorological Service, Athens, Greece
e-mail: athanarg@upatras.gr

Cloudiness is an important meteorological factor not only because it affects the radiative energy balance of Earth, but also becaues it influences other weather parameters, like precipitation. This first influence is twofold since cloudines vaires the albedo of the Earth and because, it modifies the atmospheric long wave radiative exchange. The determination of cloudiness became easier due to the advent of modern remote sensing space technology and other automated sky monitoring techniques. However cloudiness time series produced by these methods are available only for the recent years, limiting their use for climatological studies and especially for studies related to climate change. Long records of cloudiness are the result of human estimations made during the synoptic observation hours at each weather station. All weather time series records, may suffer from inhomogeneities. This is even more true for cloudiness time series, since in addition to the usual causes of inhomogeneities, a value attributed to cloudiness is not an instrumental measurement but the - subjective – result of an evaluation of the observer. In this work we attempt for the first time to analyze and homogenize all cloudiness time series coming from the weather station network of the Hellenic National Weather Service (HNMS). Data come from 36 WMO stations in total, and cover the period from 1975 to 2004. Raw data comprise (synoptic) hourly observations. Hourly data have been subjected to quality control, after which, time series of daily and monthly averages have been produced. The methods - software tools used for the homogenization are ACMANT, CLIMATOL, HOMER and MASH.

SESSION 4:
DATA BASSES

ARTYDOC, the digital archive of the Meteorological Service of Catalonia

Marc Prohom, Mňnica Herrero and Montserrat Cańas

Area of Climatology, Meteorological Service of Catalonia (SMC)
e-mail: mprohom@meteo.cat

In recent years the Meteorological Service of Catalonia (SMC) has promoted a long process of identification, cataloguing and imaging of documentary sources containing meteorological and climate information. Currently (May 2017) the volume of digitized images exceeds 6 Tb, and every year this number increases. The challenge once this information is preserved is providing a good cataloguing system, to be able to look at the material efficiently.
ARTYDOC software is the tool chosen by the SMC for the management, classification, description and dissemination of the digital archive of the institution. As for management the tool lets you to define the user's profile, to decide who participates and how, as well as maintaining the traceability of the system. The classification allows access to documents through various levels related hierarchically: observation sheets, publications, correspondence, photographs ... The system permits precise description of the documents configurable by the manager, giving to digital objects background information, organization and content. Finally, ARTYDOC also allows the dissemination of the digital heritage of the SMC, deciding how much you want to open up the network to the general public.

Quality Check and Homogenization of ECA&D temperature data-set

Antonello Squintu(1), Gerard van der Schrier(1), Petr Štepanek(2), Pavel Zahradniček(2),
Tamas Szentimrey(3), Yuri Brugnara(4), Albert Klein Tank(1)

(1)KNMI, De Bilt, Netherlands
(2)Czech Glob, Brno, Czech Republic
(3) Hungarian Meteorological Service (OMS), Budapest, Hungary
(4)University of Bern - Department of Geography, Bern, Switzerland
e-mail: antonello.squintu@knmi.nl

ECA&D data-set collects over 4800 temperature series from all countries of Europe and Mediterranean countries. Large number of these series are affected by outliers, repeated values and other mistakes in the measurements. These quality issues, coupled with the inhomogeneities due to relocations or changes of the instruments may lead to erroneous estimates of climate impact indices and trends. In the context of Copernicus project C3S 311a Lot 4 and Horizon 2020 EUSTACE Project, KNMI is taking part in the selection and the development of Quality Check and Homogenization procedures. In collaboration with Czech Globe and OMSZ three Quality Check procedures have been tested on four sample regions covered by ECA&D (Catalonia, Sweden, Bavaria and Slovenia), choosing the most suitable one to be applied to the entire European Dataset.
In cooperation with University of Bern a procedure of Break Detection has been developed and applied to ECA&D data-set. It consists in an agreement-system of three different break detection methods having different statistical basis (Kuglitsch et al., 2012). Results of the break detection have been used for the calculation of the amplitude of the breaks, based on a quantile matching process based on the method by Trewin (2012). In order to perform a more effective homogenization, the results of this procedure have been used as input and reference set for a second iteration of the same process, taking advantage of the higher signal-to-noise ratio and of the better quality of the series themselves after the first iteration.

The homogenization process has been analyzed and checked on some critical points such as: limitation due to self induced signals during the second iteration, role of geographical distribution of the reference series and validation of the homogenized series comparing them with manually homogenized series provided by national meteorological services and universities which are partners of ECA&D. This method will be then compared to other homogenization procedures on the same tested regions used for QC evaluation.

An Ensemble Version of the Daily E-OBS Dataset

Richard Cornes, **Gerard van der Schrier**, Else van den Besselaar, Phil Jones

Royal Netherlands Meteorological Institute (KNMI), Netherlands
e-mail: schrier@knmi.nl

Daily, gridded datasets formed solely from station data provide an important tool for the evaluation of reanalysis datasets. However, uncertainty in the gridded datasets is rarely taken into consideration in such comparisons, primarily because few gridded datasets provide estimates of uncertainty. In this study we describe a new version of the daily, Europe-wide E-OBS gridded temperature and rainfall dataset that contains a number of significant developments over earlier releases. Notably, uncertainty estimates are generated from a 100- member ensemble of realizations of each daily gridded field. In order to produce a consistent ensemble of realizations, the original interpolation method used in E-OBS has been overhauled. In contrast to the kriging/spline method used in earlier versions, we have made use of Generalized Additive Modelling. Topographic effects are integrated through the modelling of gridded background fields that are calculated using month-by-month climatological averages calculated over the period 1961-90. In addition to altitude, these background fields also incorporate additional environmental factors, such as coastal proximity and slope/aspect. To assess the success of the new interpolation method, the gridded data are compared against several high-resolution gridded datasets produced by National Meteorological Services (NMS) across Europe. These datasets generally contain many more station data and are therefore expected to replicate the true daily fields of temperature or precipitation more accurately than E-OBS, but do so at a scale that is comparable to E-OBS. On the whole the new version of E-OBS has reduced error relative to these NMS datasets compared to the current operational version of the dataset (version 14.0). The new version of E-OBS is expected to aid evaluation of global-scale reanalyses across Europe, but particularly European regional reanalyses.

Preparing for the 1991-2020 standard climate normals

Ole Einar Tveito, Herdis Motrřen Gjelten, Elin Lundstad, Eirik J. Fřrland

Norwegian Meteorological Institute, Norway
e-mail: ole.einar.tveito@met.no

Climatological normals have traditionally been applied to describe expected weather and climate conditions at given locations, and to provide a reference for climate indicator products. WMO guidelines recommend that these reference values should be based on thirty year periods of observations. The current standard normal period is 1961-1990, and is used for the production of the climate anomalies for Norway. These normals are also applied in several national technical guidelines and regulations, so even if it has been recommended to apply intermediate normal periods such as 1971-2000 or 1981-2010 the standard normal period has been applied in order to keep the climate anomaly products consistent with these guidelines. The next standard normal period is 1991-2020, and the Norwegian Meteorological Institute has recently started to prepare the basis for calculating new standard station normals. The observation network has changed considerably during the last 20-30 years. Automatization, station relocations and reduction of the network implies large challenges in order to achieve complete high quality climatological time series. The automatization process has introduced new sensors and instruments, and also in many cases relocation of stations. Homogenisation of the series covering the entire period is therefore necessary to provide a consistent basis for the new normals.

In addition are many series discontinued within this period, mainly as a result of a reduction of the manual observation network. There is therefore a need to inter-/extrapolate series that does not cover the entire period. Several multivariate statistical analysis tools such as multiple regression and regional principal component analysis (PCA) will be considered for filling data gaps. Also spatial interpolation methods applied for observation gridding might be applied. The homogenization and gap filling analyses will be carried out as an iterative process until a sufficient number of complete and homogenised high-quality time series are achieved. In order to have sufficient training period for the statistical analyses the time series considered in this analysis will cover at least the period 1961-2020 (60 years).

We will present the outline of this activity, and show examples and preliminary results of the evaluation of analysis tools that will be implemented for estimating the new standard 1991-2020 climate normals.

SESSION 5:
DATA ANALYSIS

Management and development of a suite of national climate monitoring products forthe UK

Mark McCarthy, Dan Hollis, Michael Kendon, Tim Legg, Ian Simpson

Met Office, UK
e-mail: mark.mccarthy@metoffice.gov.uk

Here we present the latest developments in the production and management of historical climate datasets for the UK derived from the national in-situ observing network. A hierarchy of climate data sets will be presented along with the data management and post-processing that is undertaken to produce them. This activity covers data recovery to real time station observations to gridded datasets and national climate series and associated products. The presentation will outline the technical data management activities being adopted to improve the traceability, reproducibility and accessibility of these climate records, and how they underpin the development of robust climate services for public, government and private sector. We also outline the main research gaps and plans for future development including homogenization and improved geo-spatial interpolation.

Assessment of daily rainfall data recorded by two different networks in Piedmont (North-West Italy)

A.Baronetti(1), F. Acquaotta(2), S. Falzoi(1), D. Garzena(1), D. Guenzi(1), F. Spanna(3), S. Fratianni(1,2)

(1) Earth Sciences Department, University of Turin, Turin, Italy
(2) Centro Interdipartimentale sui Rischi Naturali in Ambiente Montano e Collinare, University ofTurin, Turin, Italy
(3) Piedmont Region, Phytosanitary Sector, Turin, Italy
e-mail: alice.baronetti@unito.it

Rainfall data are essential for various hydrological applications related to water resource management, power production, irrigation, flood control, forecasting and validation of remotely sensed data from space platforms. Studying and analysing extreme rain events, dry and wet periods and trends can help in planning and manage the effects of the climate change.
The availability of daily precipitation data is necessary to make good climatic analysis and to better understand extreme events. One of the main problem of rainfall measurement is his quite small representativeness for large area. For this reason, is necessary have a high density of rain gauges spread in whole the study area.
In Piedmont are available several networks managed by public and local authorities. In this study, we will analyse the correlation of two independent automatic climate networks, to give a novel contribution to the analysis of rainfall daily data in the Piedmont region.
The dataset comes from the meteorological stations of the Regional Agency for Environmental Protection (ARPA) and the Agro-Meteorological network (RAM).
The comparison was performed using couple of daily pluviometric parallel measurements, of ARPA (reference serie) and RAM serie. The selection was based on five parameters: the overlapping period, the difference in elevation, the distance, the aspect and their characteristic. Once defined them for each couple the free and open source script CoRain written in R language was used. This script uses an innovative analysis approach to compare two parallel rain series (with an overlapping period). CoRain combines a set of well-known statistical tools and highlight the overestimations and underestimations due to rain gauges.
The precipitation were split in different classes of intensity calculated on the ARPA series, and also the number of events for each class have been observed.
This methodology allowed to highlight the percentage of precipitation events that can be considered equal between the pairs of series and, at the same time, to underline the type of events that can induce the greater difference between the two stations.
The comparison between the two networks, highlight that the major difference between the precipitation records is related to the type of instrument. In fact, was found a systematic inhomogeneity in the number of rainy days, defined as days withmore than 1mm precipitation. About the extreme events, an overestimation of the ARPA stations is observed, while they tend to underestimate weak events of rainfall.
This project shows the importance of comparing different meteorological networks in the study area to better understand extreme events of rainfall.

POSTERS

1

Current status of European Data Rescue Activities

Ingeborg Auer, Barbara Chimani and members of the EUMETNET data rescue team

Zentralanstalt für Meteorologie und Geodynamik, Austria
e-mail:ingeborg.auer@zamg.ac.at

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 Europe there is a quite good data coverage since the 1960ies, however to capture the full climate variability including extremes the time series are often too short and its spatial density does not allow to capture small scale events. Although a considerable part of long-term series has already been digitized and made available, there are still millions of data to be recovered and rescued. Due to a number of completed or running activities (http://www.climatol.eu/DARE) the number of digital available data has been increasing continuously, however an extended overview has not been made available so far. EU-DARE is contributing to the European climate services by providing an extended inventory of digitized and non-digitized climate data, focusing on centennial or even longer daily data, long-term mountain stations of at least 50 years and data in sparse regions.

The data inventory installed by the EUMETNET's Data Rescue Activity initiative (https://www.zamg.ac.at/dare/) offers information about already digitized data and data waiting to be rescued and is updated every year. The updates demonstrate that the amount of available digital data is growing from year to year, but differs from country to country considerably. Recently a "lost and found activity" has been introduced into the website facilitating retrieving data of stations which for political reasons in the past have been situated in another country.

2

Data Rescue Approach (visualisation, methodology, examples and homogenization scheme) in Slovakia

Oliver Bochniček, Pavel Faško, Ladislav Markovič, Zuzana Palušova, Peter Kajaba

Slovenský hydrometeorologický ústav, Bratislava, Slovakia
e-mail: oliver.bohnicek@shmu.sk

Digitalization is essential for the history evaluation and future expectation. Main step is (except having accessible archive) to have precise data field overview (archive & database) as well as metadata. Digitalization should be guaranteed by the professional staff as well as implemented standard procedures. Different data input (manual, automatic) needs different quality control and way of processing. Station network automation brings a kind of inhomogeneity as well as relocation stations. There exist a lot of stations in archive which do not have an indicative, despite they are very important for the time series data construction (sometimes very near stations have totally different names). Map of the municipality of Slovakia in connection with overview of the length of time series (climatological and precipitation station network separately) sorted by name helps to determine the next usability for extension of the time series. Sometimes parallel climatological observation exists within one city (e.g. Bratislava time series (since 1850) including 12 locations with time overlay of several years, sometimes decades).

In the less populated mountain regions of Slovakia, precipitation stations have been in operation in the past having completed measurements and observations meantime. At present, due to the lack of rainfall stations in medium and high mountain climates, the perceived results of measuring from the old rainfall stations are very valuable. Because they do not exist in digital form, digitization will have to be done after selection and review of data from these stations. When selecting these stations, hydrological catchments were respected so that the new data were well applicable in hydrological practice. Newly created time series (alongside existing ones mostly at lower altitudes) of atmospheric precipitation and snow cover should enrich source of information for mountain area. Example of possible review and creation of precipitation station network for hydrological practice in selected river basin is shown.

Surrounding change causes gradual or abrupt change in the time series. Appropriate methods could help make them useful for climate change detection and assessment. A simple schematic diagram of the homogenization package use is shown.
People should be aware of having treasure in climatological archives (sometimes even 50 percent is paper-based only) and try to digitalize them as soon as possible through coordination with international authority in respect to the experience exchange and help. Local experts play irreplaceable role in data release for the scientific purposes.

Data Rescue of two centennial Spanish series (Maó and Oviedo)
Moved to oral presentations - Session 3: Homogenisation

J.A. Guijarro(1), A. Jansŕ(2), M.A. Mora(1) and V.M. González(1)

(1) State Meteorological Agency, Spain
(2)University of the Balearic Islands, Spain
e-mail: jguijarrop@aemet.es

Long instrumental climatic series are of paramount importance to study climate variability and trends, but there are a limited number of observatories with series spanning more than 100 years. However, many old meteorological observations remain in paper archives of observations and strip recording charts, and rescuing these old records may help extending many series backwards.

This communication presents two recent projects that have enabled the incorporation of daily data to two long Spanish series from which only monthly summaries were available in their older parts: Maó (Minorca, Balearic islands) and Oviedo (Asturias, northern Spain).

Although there are some previous meteorological records, the Maó series is considered initiated by J. Carreras in 1863 and continued by M. Hernández from 1885 to 1926. The original annotations are located in the Maó municipality archives and the cultural association Ateneu de Maó respectively, and were studied by P. Carreras with a fund by the Institut Menorquí d'Estudis. She provided a detailed history of the observations and digitized the daily precipitation data, thereby extending this series backwards more than 70 years. However, imaging of the original records and digitizing the other climatic observations is still a pending task.

Observations in Oviedo began at its University in 1851 under the direction of L. Salmeán and were interrupted in 1936 due to the civil war. Resumed in 1946, they definitively ceased in 1958. All data available in the observation logs have been digitized under the direction of M.A. Mora and V.M. González in spread-sheets (1,147,832 records), and current efforts are devoted to control their quality and consistency with the monthly summaries held in the National Climatic Data Base. As overlapping observations are available in the area for the period 1941-1985 (La Cadellada station) and from 1973 (when the current Oviedo observatory was established), statistical homogenization techniques allow the reconstruction of this latter series back to 1851.

The presentation includes the discussion of the problems encountered along the development of these projects and basic statistical descriptions of the digitized series and their homogenization.

4

Data rescue of upper wind data obtained from pilot balloon observations in Spain: First works

Manuel Lara

AEMET, Delegación Territorial Extremadura, Spain
e-mail: mlaraj@aemet.es

In 1913, a network of upper wind observation started to work in Spain. Network activity suffered from ups and downs due to both of world wars and to the Spanish civil war. It is known that in 1925, the network had 16 observatories which launched 3335 balloons. In 1944, the number of observatories reached 26, which launched 5669 balloons. The method of observation has been documented in detail.

Observation workbooks from years between 1924 and 1953 have been located for the Observatory of Badajoz. Using a software which incorporates different quality controls, although little efficient for recording, data from a whole year have been digitalized. Some qualitative comparisons have been made between observed winds and winds obtained from the ERA-20C reanalysis, these comparisons sometimes show significant differences.

Data provided for the observation of pilot balloons during the first half of 20th century is estimated to be of great utility in future reanalysis.

5

Review of the historical data series in the Slovak Republic

Katarína Mikulová, Pavel Štastny

Slovak Hydrometeorological institute, Bratislava, Slovakia
e.mail: katarina.mikulova@shmu.sk

The first information about weather (mainly adverse) in Slovakia we can find in the old chronicles from towns or villages in our territory. But the first systematic meteorological observations was started in the half of 19th century, especially after the establishment of the Central Institution for Meteorology and Earth Magnetism in Vienna in 1851. This Institution established the first state network of observational stations at the former territory of Austria- Hungary. In our metadata records there was written that one of the first meteorological station in the territory of Slovakia was station in Banská Štiavnica at the Forest Academy since 1848. The presence of this meteorological station confirms the publication "Climatology" written by Otto Schwartz from the year 1891, but in the archive of the Slovak Hydrometeorological Institute has no records of these observations. The most of these oldest observations in the territory of Slovakia have unknown place of depositing, or we do not know if they have been preserved. We know that some of them probably could be in Vienna or Budapest archives.
The most of old records in our archive begins since 1871 after Austro-Hungarian settlement, when the Royal Central Institution for Meteorology and Earth Magnetism in Budapest continued with the measurements. The existing meteorological network in Hungarian part of the Empire was managed and developed with this Institution and some of new stations - in our territory were established. In the archive of the Slovak Hydrometeorological Institute we have a number of the historical meteorological records from 19th century with various beginnings of their observations and breaks or finish in their observational activities. Only the number of 11 from this collection of stations are available for the old time series reconstruction. Their observations continue from the time of their establishment up to now with only non-important replacements and/or the short interruptions of observations. The data gaps were mainly during the First World War and World War II. The list of these stations starts with Hurbanovo (since 1871) – this station is located in the same places during the history of observation, Liptovský Hradok (1881), Košice (1876), Lučenec (1871), Viglaš (1883), Rožnava (1870), Rimavská Sobota (1883), Michalovce (1871), Bojnice (1873), Nitra (1872) and Bratislava (1852). In Košice and Bratislava we can find the parallel observations in the same time periods, so we have more possibilities for the long time series reconstruction for these important cities. Recently, we have digitalized all records of Hurbanovo and Liptovský Hrádok stations. We are finishing with the quality control of the Hurbanovo dataset and preparing the same procedure for Liptovský Hrádok station. We have chosen the next station for reconstruction from the Košice records. With such a set of three station old data series from western, central and eastern part of Slovakia we can find the difference of the climate regime between them in the past.

6

Discovering and Rescuing Climate Records in Republic of Slovenia

Mateja Nadbath

Slovenian Environment Agency (ARSO), Ljubljana, Slovenia
e-mail: mateja.nadbath@gov.si

In Slovenian Environmental Agency (ARSO) run some activities regarding discovering and rescuing climate records. The most important findings, results and troubles will be shown on the poster.

Slovenia is a member of initiatives WMO MEDARE (The MEditerranean climate DAta REscue under World Meteorological Organization) and EUMETNET DaRe (Data rescue under European meteorological services network).

The most important findings on the field of discovering and rescuing climate records are:

- all climate records and metadata in period 1961–nowadays are digitized,
- climate records from period 1945–1961 are in ARSO's archive, partly digitized, metadata completely digitized,
- climate records for period 1918–1945 for the stations from central and eastern parts of the country are in ARSO's archive, partly digitized, metadata completely digitized,
- the climate records and metadata that still needs to be discovered are:
  - for years before 1918 for the stations from eastern parts of the country and partly also from central Slovenia and
  - for years before 1945 for the stations from western parts of Slovenia,
- the digitization of climate records makes slow progress due to lack of employee and timeconsuming work.

Results achieved in discovering climate records are:

- the list of missing climate records, metadata and yearbooks is made and published on the web pages of EUMETNET - DaRe Lost&Found site and on the I-DARE International data rescue portal,
- some historical climate records, metadata, and yearbooks have been found on the internet (Digital Library dLib Digitalna knjižnica Slovenije, Oberösterreichisches Landesmuseum, Google Books and Yuri Brugnara-The Data rescue initiative for southern Alps),
- in cooperation with Austrian (ZAMG) and Hungarian (OMSZ) colleagues, some historical climate records, metadata, and yearbooks have been found and scanned.

Troubles in discovering and rescuing the lost Slovenian climate records before year 1945 are:

- scattered climate records in unknown foreign archives for the sake of Slovenian history,
- German, Hungarian and Italian language used in Climate logbooks and yearbooks,
- the usage of Suetterlin alphabet in some logbooks, letters and sketches,
- different historical names of villages and towns from today because of different official languages.

7

Climate data background in Croatia

Željka Pogačić, Ines Srzić

Meteorological and Hydrological Service of Croatia, Croatia
e-mail: zeljka.pogacic@cirus.dhz.hr

Although meteorological observations in Croatia have not changed much throughout the years, it is very important for data users to know the specifications of the equipment and methods of observations. Considering the mandatory practices and recommendations of the World Meteorological Organization, meteorological services are those who are preparing manuals and operating meteorological networks.

According to the introduction above, our poster intends to show the methods of collecting climate data, best practices and standards in Croatia over the last 160 years. In other words, to present the climate date background – the basis for quality data usage.

Belgian homogenized reference climate time series

C. Bertrand, C. Delvaux, R. Ingels, V. Vrabel and M. Journée

Royal Meteorological Institute of Belgium (IRM/KMI), Belgium
e-mail: cedric@meteo.be

Long-term, high-quality and reliable instrumental climate records are indispensable pieces of information required for undertaking robust and consistent studies to better understand, detect, predict and respond to global climate variability and change. Accurate and homogeneous climate data are also indispensable for the calculation of related statistics that are needed and used to define the state of climate and climate extremes.

Huge amounts of climate data have been recorded since the earliest observational days in Belgium (i.e. instrumental data extend back in time at least to the 19th century over most regions of the country).
However, the existing data heritage is largely under-exploited because historical information still remained in hard copy and fragile media (easily accessible digital climate data are mostly restricted to the second half of the 20th century). To overcome such a limitation, the Royal meteorological Institute of Belgium (RMI) has undertaken data rescue activities aiming at transferring historical climate records from paper forms to new media (i.e. digital forms).

However, it is well known that long climatological time series often contain variations that are not only due to the vagaries of the weather or climate. At the same time, wrong or aberrant observations are common in most observational systems. All these factors reduce the quality of original data and compromise their homogeneity. Therefore the identification and correction of these aberrant observations and non climatic factors is essential before any reliable climate study can be carried out for a meaningful assessment of change in climate. Towards this objective, reliable Belgian climatological time series are being produced at RMI by applying state-of-the art homogenization methods to long-term air temperature and precipitation time series after preliminary data quality control checks.

9

Recovered weather data of Ljubljana from 1818 to 1850: comparison with HISTALP and Berkeley Earth data

Damijana Gartner, Gregor Vertačnik

Slovenian Environment Agency, Slovenia
e-mail: gregor.vretacnik@gov.si

Very old daily weather data from Ljubljana has been recently discovered in newspaper Laibacher Zeitung. The data comprise observed weather phenomena and measurements of air temperature, humidity and pressure three times a day. The data spans the period from 1818 to 1850 but are incomplete and are missing metadata. However, internal consistency of data and statistical comparison with homogenised regional (HISTALP) and global datasets (Berkeley Earth) can reveal the quality and applicability of the data, ranging from daily weather to decadal climate variability.

10

News from HISTALP

Barbara Chimani, Manfred Ganekind, Christoph Matulla

Zentralanstalt für Meteorologie und Geodynamik, Wien, Austria
e-mail: barbara.chimani@zamg.ac.at

HISTALP is still ongoing work: the quality of the available data is increased whenever possible and additional datasets are provided. Two main tasked were undertaken within the last year:

- In cooperation with all HISTALP-partners the dataset of monthly temperature was revised, including changes due to national data quality control procedures and the use of additional station data and information.
- The amount of data available for Austria was increased including now homogenized daily temperature and precipitation data and reconstructed data of lake surface temperatures.

Information and results of those improvements and extensions will be provided.

11

First stage of new phenological database at Meteorological and Hydrological Service of Croatia

Helena Lebo Andreis

Meteorological and Hydrological Service of Croatia, Croatia
e-mail: lebo@cirus.dhz.hr

At Meteorological and Hydrological Service of Croatia, there have been several versions of phenological database systems, but none of them are used any more, for different reasons.
It was decided to develop a new database system that would overcome all previous drawbacks and integrate all the data from different sources.
First of all, previous documentation was collected and the basic rules were set according to requirements.
Using these guidelines, the new schema was created, and the software was developed.
The first stage of database system includes various web forms - for station information, plant information, development stages for plants, photo upload and photo catalogue, and forms for observation data.
The next task was to test the new database system, and afterward to create a catalogue of stations, plants and development stages, and to enter the observation data using web forms. Future development includes transferring data from old databases, creating software for quality control, views, reports, data exchange etc.

12

Metadata Digital Information System of Bulgarian National Institute of Meteorology and Hydrology

Stanislava Radeva

National institute of Meteorology and Hydrology, Bulgaria
e-mail: StanislavaRadeva@meteo.bg

In its role of a National Meteorological Service, NIMH has many operational duties and scientific activities from the far 1887 to the present. Metadata records of meteorological stations have been maintained on paper over the years. All changes in their work are recorded fairly accurately. With the development of science and technology, it has become necessary to digitize the metadata of the stations and to update them in due course (immediately after the change), as well as to allow a wide range of scientists dealing with climate change and homogenization of data, etc. to it.
In order to meet the need for a fast reliable and cost-effective method of access to information, an institutional web-based information system was built. The specialists responsible for the maintenance of the meteorological stations report about each change in the stations for which they are responsible. Climatologists and operational workers from other areas use the station files for their research. A description of system will be presented.

13

Recent developments in the European Climate Assessment & Dataset

G. van der Schrier, E. van den Besselaar, C. Photiadou, A. Squintu,
R. Cornes, A. Klein Tank, G. Verver

Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands
e-mail: schrier@knmi.nl

The European Climate Assessment & Dataset aims to provide daily meteorological stationbased data and analysis for scientific research in and outside the climate sciences. An important contribution to this activity is the E-OBS, with is the gridded data set with daily maps of temperature, precipitation and pressure for Europe, based on ECA&D.
In this presentation the latest changes to the ECA&D station coverage and products are shown. Emphasis will be on the newly introduced Quality Control procedures and homogenization of daily temperature data and the newly developed E-OBS which provides a new and more realistic extimate of uncertainty.
Finally, an overview is given of the development of new products for the Copernicus Climate Change Service.

14

Influence of human quality control on precipitation amounts
CANCELLED

Suzana Filić, Helena Lebo, Ana Weissenberger, Dubravka Rasol

Meteorological and Hydrological Service of Croatia, Croatia
e-mail: filic@cirus.dhz.hr

The network for precipitation measurements of the Meteorological and Hydrological Service of Croatia consists of 319 classical Hellman rain gauges and 67 automatic ombrographs, some of which are placed at the same location. Differences in measuring equipment can cause significant differences of measured precipitation amounts. In addition, data from conventional and automatic instruments go through the different quality control procedures. Data from automatic instruments seems to differ in both positive and negative amount according to the results of a quality control of automatic data, but influence of human quality control on conventionally measured precipitation amounts is unknown. Here we studied how much precipitation from conventional measurements is changed during the human quality control of precipitation amounts on monthly, seasonal and yearly scale in different climate regions of Croatia for the period of 2008. - 2016.

Data Rescue of the Long-term Daily Maximum and Minimum Air Temperature Records in Ireland

Carla Mateus(1), Mary Curley(2), Aaron Potito(1)

(1) Department of Geography, National University of Ireland Galway, Ireland,
(2) Met Éireann, Glasnevin Hill, Dublin, Ireland
e-mail: C.PEDROSOMATEUS2@nuigalway.ie

In Ireland, with the exception of the synoptic stations and a few other stations which are available from 1941, most of the daily air temperature records had not been digitised prior to the 1960s, and exists in the original observations logs, extremely fragile manuscripts and scattered publications stored in various archives, such as: Met Éireann, National Botanic Gardens, Royal Irish Academy, Royal Dublin Society, NUIG Library and County Council Libraries and Archives.

This research accomplishes data and metadata rescue, digitisation, quality control and homogenisation of available and accessible long-term instrumental daily maximum and minimum air temperature records since the early and mid-19th century to the present in Ireland: Phoenix Park (1831-2016), Trinity College Dublin (1840-2016), Botanic Gardens (1848-2016), NUI Galway (1861-2016), Fitzwilliam Square (1871-2016), Birr Observatory (1872-2016), Roches Point (1872-2016), Valentia Observatory (1872-2016), Markree Observatory (1875-2016), Blacksod Point/Belmullet (1885-2016) and Malin Head (1885-2016). Short-term daily air temperature series dated back to the 19th century and parallel measurements (e.g. in and outside observatory or the use of different screens) were also rescued.

The objectives of this contribution are to:

(1) Present the abundance and geographical distribution of early air temperature measurements in Ireland, mentioning the multiple archives and sources of data rescue.
(2) Enlighten the importance of data rescue of long-term and short-term air temperature records in Ireland.
(3) Introduce the methodology for data and metadata rescue in order to assure the completeness of the long-term climate stations.
(4) Expose the procedures for data and metadata digitisation.
(5) Display the rescued short and long-term air temperature series.
(6) Explain the importance of the most complete rescued metadata such as: observing time, thermometers type, thermometers level above ground, thermometers exposure and their location, instruments maintenance, station surroundings and enclosure, re-location of station, observing practices, observer’s comments, screen type and size, transition to Stevenson screens, conversion from manual to automatic station, number of gaps and its reason, previous corrections applied and why, parallel measurements, automatic sensor type and station inspections for the quality control process.

Following work will involve the automatic quality control and daily homogenisation of the short and long-term records of maximum and minimum air temperature using a number of methods, including the software MASH v3.03 (SZEMTIMREY, 2014) in conjugation with the station metadata in order to: fill missing data, perform additional quality control, establish composing reference series, accomplish homogeneity tests, achieve break-points detection in the series, effectuate the adjustment of non-climatic inhomogeneities and perform the validation process through a benchmarking of statistical methodologies.

16

The new Hungarian National Atlas

Zita Bihari, Mónika Lakatos, Péter Szabó, Tamás Szentimrey

Hungarian Meteorological Service, Hungary
e-mail: bihari.z@met.hu

A new national atlas is being published in Hungary, initiated by the Hungarian Academy of Sciences and coordinated by the Geographical Institute. This atlas will be available not only in a conventional (printed) form, but also in an electronic version (on DVD and on the web).
The printed version of the National Atlas of Hungary will be published in four volumes, with one volume being issued per year from 2017. The tentative titles and topics of the volumes are as follows: Natural Environment; Society; The State of Hungary and Its Place in the World; Economy.
This year, the "Natural Environment" volume will be available and will consist of the following chapters: geology, geophysics, relief, climate, waters, soils, flora, fauna, landscape regions, nature protection, environment protection, natural hazards and catastrophes.
The chapter on climate was developed by the Hungarian Meteorological Service. It provides an overview on the main characteristics of the climate of Hungary. In addition to the traditional maps, the chapter also depicts the climate change signal in the last 100 years and the expected changes in the middle and the end of 21st century.
The maps and graphs representing the recent and past climate are based on homogenized and quality controlled gridded data series produced by the MASH – MISH methods. The future projection maps show the results of the ALADIN-Climate and the RegCM regional climate models providing the simplest quantification of the simulation uncertainties.

17

Climate Maps for the Reference Period 1981-2010 in Spain

Andrés Chazarra, Elías Criado

Agencia Estatal de Meteorología, Madrid, Spain
e-mail: achazarrab@aemet.es

The Spanish Meteorological Agency (AEMET) has recently updated the climate maps in Spain in order to account for the reference period 1981-2010. These maps include information on the annual and monthly normal values of temperature and precipitation as well as on the mean number of precipitation days, snow days, fog days, hail days and storm days. The objective of this poster is to describe the data selection and the methods applied for obtaining the gridded fields for the climate maps. In many cases, additional variables such as the altitude, the latitude or the distance to the coast were used in the interpolation process. We will focus on the main issues that we have found in the spatial interpolation of some climatic variables such as snow days, fog days and snow days.

18

Operational suite for observational gridded data products in Belgium

M. Journée, C Delvaux and C. Bertrand

Royal meteorological Institute of Belgium (IRM/KMI), Belgium
e-mail: michelj@meteo.be

Observational gridded data provides an estimate of the spatial distribution of a meteorological variable based on observations. While instrumental measurements are taken at irregularly distributed stations, gridded data represents the meteorological variable on a predefined regular grid. Gridded climate data is of strong benefit in disciplines applying distributed quantitative models to examine the influence of weather and climate. Gridded data are also very convenient to provide estimations for any specific location of interest for the user.

An operational automated suite of gridded data covering Belgium at a spatial resolution of 5 km has been developed at the Royal Meteorological Institute of Belgium (RMI). These products encompass several variables (precipitations, air temperature, sunshine duration, solar radiation, relative humidity and wind speed) and are available in different temporal resolutions (daily, monthly, seasonal and annual values). These products are first generated in near realtime (i.e. the next day) with the observations already available and then updated for archiving as soon as the data quality control is completed.

The data processing includes automatic data filtering in order to eliminate gross errors, spatial interpolation procedures adapted to each parameter and temporal resolution and finally the generation of end-user products (such as areal statistics, climate maps, etc.).

19

Measuring practice and statistical analyses of short-term precipitation in Hungary

Monika Lakatos, Lilla Hoffmann, Zita Bihari, Tamás Szentimrey

Hungarian Meteorological Service (OMS), Hungary
e-mail: lakatos.m@met.hu

In the recent years the frequency and intensity of extreme events related with precipitation have increased in many regions of Europe. The flood risk increased in the recent decades in the region due to more extreme precipitation events. Extremes indices studies usually focus on daily precipitation sum, however there is a demand of sub-daily scale extreme analysis due to lack of the reliable data for understanding of the nature and drivers of global and regional precipitation extremes and change on different time scales which are relevant for the societies.

The measuring practice of short-term precipitation has completely changed with installing automatic climate stations. In the period preceding the automatization, mainly from 1967 to 1997 ombrographs registered the quantity of precipitation in Hungary. Evaluation of the rain register paper was carried out by selecting the largest precipitation amounts during 5, 10, 20, 30, 60, 180 min periods within a wet event, considering the slope of the curve of accumulated precipitation and registering these partial amounts. Automatic stations replaced the ombrographs in many places in national observing networks. At the Hungarian Meteorological Service the amount of precipitation is registered every ten minutes forming a data set for further analysis. One minute data are available only for few stations and for limited period.

The paper presents a statistical analysis of data registered by ombrographs and automatic measurements to describe the differences between the two types of sampling. Parallel measurements for a short period and one minute data are analysed to explore the relationship of the former and the recent measurements. The main aim is to make possible merging the data series for deriving long term series for climate change studies.