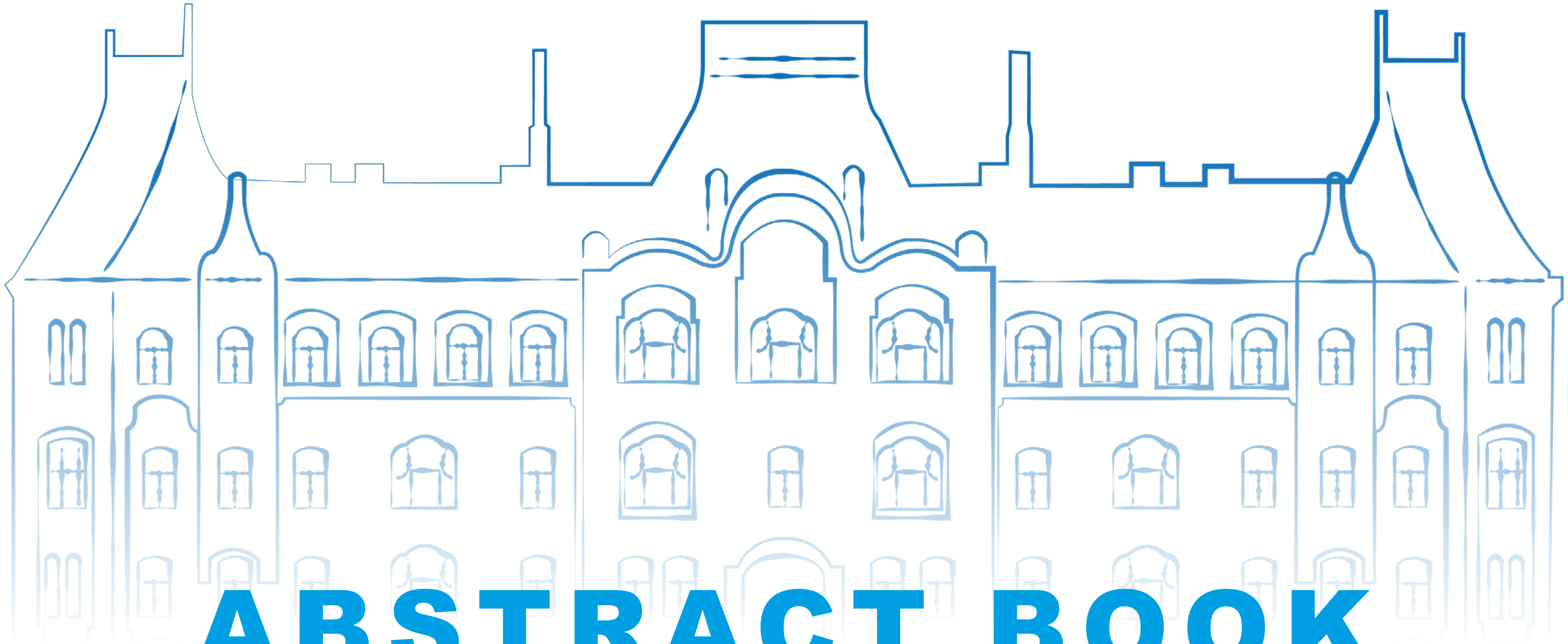

ABSTRACT BOOK

11TH SEMINAR FOR HOMOGENIZATION AND QUALITY CONTROL IN CLIMATOLOGICAL DATABASES
AND
6TH INTERPOLATION CONFERENCE JOINTLY ORGANIZED WITH THE 14TH EUMETNET DATA MANAGEMENT WORKSHOP

9–11. MAY, 2023
HEADQUARTERS OF THE HUNGARIAN METEOROLOGICAL SERVICE, BUDAPEST, HUNGARY

2023

11TH SEMINAR FOR HOMOGENIZATION AND QUALITY CONTROL IN CLIMATOLOGICAL DATABASES AND 6TH INTERPOLATION CONFERENCE JOINTLY ORGANIZED WITH THE 14TH EUMETNET DATA MANAGEMENT WORKSHOP

Budapest, Hungary 9 - 11 May 2023

Organized by Hungarian Meteorological Service (OMSZ)

Supported by EUMETNET Climate Program, OMSZ and WMO

Edited by

Mónika LAKATOS
Tamás SZENTIMREY
Dorottya SZABÓ

Published by

Hungarian Meteorological Service

www.met.hu

DOI:10.21404/11.SemHQC6.ConfSI.11.DMWS2023

PROGRAMME

TUESDAY – 09.05.2023.

ORALS: 20 min (+5 min. discussion)

8:30-9.20 Registration

9:20-9:40 Opening and Commemorate Sándor Szalai and Victor Venema

9:40-10:30 Homogenization

Tamás Szentimrey
Development of new version MASHv4.01 for homogenization of standard deviation

Peter Domonkos
ANOVA correction model in relative homogenization: Optional or obligatory?

10:30-11:00 Coffee break

11:00-12:40 Homogenization (cont.)

Xiaolan L. Wang, Yang Feng, Vincent Cheng, and Hong Xu
On the development of the first homogenized monthly precipitation dataset for Canada

Jose A. Guijarro
Climatol v. 4: New functionalities and tools (online)

Belinda Lorenzo, José A. Guijarro, Andrés Chazarra, César Rodríguez-Ballesteros, José V. Moreno, Ramiro Romero-Fresneda, Maite Huarte, Ana Morata
Operational homogenization of daily climate series in Spain: experiences with different variables

Barbara Chimani, Oliver Bochniček, Michele Brunetti, Manfred Ganekind, Juraj Holec, Beatrix Izsák, Mónika Lakatos, Melita Perčec Tadić, Veronica Manara, Maurizio Maugeri, Pavel Stastny, Olivér Szentes, Dino Zardi
Revisiting HISTALP Precipitation dataset

12:40-14:00 Lunch

14:00-15:15 Homogenization (cont.)

Olivér Szentes, Mónika Lakatos, Rita Pongrácz
Historical precipitation data sets in Hungary

Gregor Vertačnik
Homogenisation and interpolation of Slovenian temperature series in period 1950-2020

Kristína Szabóová, Katarína Mikulová, Peter Kajaba, Maroš Turňa
Homogenization of moisture characteristics data in Slovakia

15:15-15:45 Coffee break

15:45-17:00 Homogenization (cont.)

Emilie Lebarbier, Annarosa Quarello, Olivier Bock, Ninh Khanh Nguyen
GNSSseg, a Statistical Method for the Segmentation of Daily GNSS IWV Time Series (online)

Ninh Khanh Nguyen, Olivier Bock, Emilie Lebarbier, Annarosa Quarello
Sensitivity of Change-Point Detection and Trend Estimates to GNSS IWV Time Series Properties

Khanh Ninh Nguyen, Olivier Bock, Emilie Lebarbier
A statistical method for the attribution of change-points in segmentation of IWV difference time series

17:30-19:30 Ice breaking

WEDNESDAY – 10.05.2023.

9:00-9:50 Homogenization (cont.)

Elin Lundstad
Homogenization of the global historical database (HCLIM) (online)

Zbigniew Ustrnul, Bogdan Bochenek, Agnieszka Wypych
Homogeneity of long-term traditional circulation types series – problems and challenges

9:50-10:20 Poster pitches, 2 min per each

10:20-10:45 Coffee break

10:45-12:25 Data and quality control

Joelsson, L.M.T, Södling J., Engström, E
The 19th century precipitation measurement method, a field study

Cedric BERTRAND, Michel JOURNEE and Luis GONZALEZ SOTELINO
Introduction of low cost sensors (LCSs) in the Belgian climate observation network: a first assessment

Ingels, Romain and Journée, Michel
Data rescue by volunteers of precipitation and temperature extremes from centennial meteorological reports of the Royal Meteorological Institute (1881-1900): a first assessment

Tibor Rácz
Correction of processed historical records of siphoned rainfall recorders

12:25-14:00 Lunch

14:00-15:15 Data and quality control (cont.)

Giulio Bongiovanni, Michael Matiu, Alice Crespi, Anna Napoli, Bruno Majone, Dino Zardi
Preliminary results on a new dataset of daily observations from a dense network of weather stations covering the Extended Alpine Region (1991-2020) (online)

Hela Irha, Ana Weissenberger
Comparison of data measured at climatological and synoptic observation times

Petr Štěpánek, Gerard van der Schrier, Pavel Zahradníček
Data Quality Control applied on ECA&D

15:15-15:45 Coffee break

15:45-16:35 Data and quality control (cont.)

Niko Filipović
AQUAS – Data Quality Control at GeoSphere Austria

Marosz M.
Climate services products in Polish NMHS (online)

18:30-21:30 Conference dinner

THURSDAY – 11.05.2023

9:00-10:15 Spatial interpolation and gridded datasets

Tamás Szentimrey
Statistical modelling of the present climate by the interpolation method MISH - theoretical considerations

Beatrix Izsák, Tamás Szentimrey, Zita Bihari and Zsófia Barna
Hungarian development of observation based temperature dataset

Melita Perčec Tadić, Zoran Pasarić, José A. Guijarro
Homogenised data series, monthly temperature grids and climate monitoring products for Croatia

10:15-10:45 Coffee break

10:45-12:25 Spatial interpolation and gridded datasets (cont.)

Kian Abbasnezhadi and Xiaolan L. Wang
Comparison of gridding methods for precipitation over Canada and assessment of station/data density effects on gridding results

Brian O'Sullivan, Gabrielle Kelly
Spatio-temporal Imputation of Missing Rainfall Values to Establish Climate Normals

Mark McCarthy (Met Office), Ed Hawkins (University of Reading), Dan Hollis, Michael Kendon, Tim Legg
Extending the UK gridded rainfall dataset HadUK-Grid to 1836 with citizen science data recovery

Rustemeier, Elke; Finger, Peter; Schneider, Udo; Ziese, Markus; Hänsel, Stephanie
HOMPRA Europe 2 – An update of a gridded precipitation data set from European homogenized time series

12:25-14:00 Lunch

14:00-15:15 Spatial interpolation and gridded datasets (cont.)

Markus Ziese, Elke Rustemeier, Udo Schneider, Peter Finger, Astrid Heller, Raphaele Schulze, Magdalena Zepperitz, Siegfried Fränkling, Bruno Heller, Jan Nicolas Breidenbach, Stephanie Hänsel
Quality Control and Grid Creation at the Global Precipitation Climatology Centre (GPCC)

Markus Ziese, Zora L. Schirmeister, Carina-Denise Lemke, Jakub P. Walawender, Christoph Schweim, Damien Pichon, Stefania Grimaldi, Gonçalo Gomes, Peter Salamon
Quality Control and Creation of Grids of Meteorological Variables for the Copernicus Emergency Management Services

Anna-Maria TILG, Johann HIEBL, Angelika HÖFLER, Anna ROHRBÖCK, Christoph FREI
SPARTACUS: A km-scale daily dataset of the surface climate in Austria – overview and new developments (online)

15:15-15:45 Coffee break

15:45-17:00 Spatial interpolation and gridded datasets (cont.)

Jouke de Baar, Gerard van der Schrier, Theo Brandsma
Community climatology: Combining official observations, crowd-sourced observations and covariates to obtain high-resolution gridded climate data sets (online)

Jouke de Baar, Gerard van der Schrier
Multi-fidelity adaptive sampling: Towards optimal station location choice of combined official and crowd-sourced weather observation networks (online)

POSTERS (size: A0, portrait)

Miroslav Trnka, Petr Štěpánek, Zdeněk Žalud, Pavel Zahradníček, Jan Balek, Daniela Semerádová, Lucie Kudláčková, Martin Možný
Climate services in the Czech Republic

Zsófia Barna, Beatrix Izsák, Ildikó Pieczka
Comparison of different interpolation methods for six-hour temperature data series

Kinga Bokros, Mónika Lakatos, Olivér Szentes
Long-term changes of cold indices in Hungary using homogenized and raw data series

Joelsson, L.M.T, Schimanke, S. Engström, E
Homogenisation of Swedish monthly precipitation from 1880

Vanda Pires, Paula Drumond, Teresa Ferreira, José Godinho, João Ferreira
Validation and homogenization of hybrid temperature and precipitation series in mainland Portugal (online)

Vanessa Seitner, Barbara Chimani, Leopold Haimberger
Comparison of homogenization methods for global radiation data in Austria

Maura Lousada, Carlos Pereira, Rui Cavaleiro, José Godinho, Tânia Cota, Ricardo Deus
Comparison of multi-satellite GPM IMERG and ERA5 reanalyses, cross-evaluation with Ordinary Cokriging and ground-based data in South West Africa

Zita Bihari, Kinga Bokros, Beatrix Izsák, Attila Kovács, Mónika Lakatos, Annamária Marton, Zsófia Molnár, Olivér Szentes, Bernadett Szolnoki-Tótiván
Climate services at the Hungarian Meteorological Service based on gridded data series

Helena Lebo Andreis
The historical phenological data migration at DHMZ

Peter Fleischer, Peter Kajaba, Martin Kubov, Jozef Rozkošný, Jaroslava Slavková
Use of homogenized data on spring phenology phases of oaks and linden

Vladimir Ribičić
Using ERA5-Land Hourly - ECMWF Climate Reanalysis in controlling ground measurements

ABSTRACTS

HOMOGENIZATION

Development of new version MASHv4.01 for homogenization of standard deviation

Tamás Szentimrey
Varimax Limited Partnership
szentimrey.t@gmail.com

The earlier versions of our method MASH (Multiple Analysis of Series for Homogenization; Szentimrey) were developed for homogenization of the daily and monthly data series in the mean i.e. the first order moment. The software MASH was developed as an interactive automatic, artificial intelligence (AI) system that simulates the human intelligence and mimics the human analysis on the basis of advanced mathematics. This year we finished the new version MASHv4.01 that is able to homogenize also the standard deviation i.e. the second order moment. The problem of standard deviation is related to the monthly and daily data series homogenization.

In essence the theme of homogenization can be divided into two subgroups, such as monthly and daily data series homogenization. These subjects are in strong connection with each other of course, for example the monthly results can be used for the homogenization of daily data. In the practice the monthly series are homogenized in the mean only, while there exist some trials to homogenize the daily series also in higher order moments. These procedures are based on a popular assumption that is the adjustment of mean is sufficient for monthly series, and the adjustment of higher order moments is necessary only in the case of daily data series. In general, it is tacitly assumed that the averaging is capable to filter out the inhomogeneity in the higher order moments. However, this assumption is false, since it can be proved if there is a common inhomogeneity in the standard deviation of daily data then we have the same inhomogeneity in monthly data. Therefore we developed a mathematical procedure for the homogenization of mean and standard deviation together. The presentation will focus on the mathematical questions raised and the applied solutions, tricks.

This developed procedure was incorporated into our new software version MASHv4.01, which is based on the examination of different type of monthly series and the monthly results are applied for the homogenization of daily series. We remark if the data are normally distributed (e.g. mean temperature) then the homogenization of mean and standard deviation is sufficient since in case of normal distribution if the first two moments are homogenous then the higher order moments are also homogeneous.

ANOVA correction model in relative homogenization: Optional or obligatory?

Peter Domonkos
dpeterfree@gmail.com

Recently a book has been published by Elsevier (Domonkos, Tóth and Nyitrai, 2022) "Climate observations: data quality control and time series homogenization". One main goal of writing that book was to create a thematically ordered document of the actual knowledge of time series homogenization. Homogenization tool can be sorted to the following categories: i) sufficiently studied and clear advantage of one or a few of distinct methods; ii) sufficiently studied, but a large variety of methods offer similar homogenization accuracy; iii) an exact evaluation would need further studies; iv) tools for special, less frequent homogenization tasks. The ANOVA correction model is an example of sufficiently studied homogenization tools which have clear and significant advantage on homogenization accuracy. The original model, introduced by Mestre in 1996, proposes the calculation of correction terms from a simplified equation system of the relative homogenization model. In the common ANOVA model, a simplification is that the climate is spatially constant. In 2008, Szentimrey presented a more correct model of relative homogenization whose variables can also be calculated by the solution of an equation system once the break detection phase has been completed. Finally, Domonkos (2022) showed that the most effective use of station specific metadata is their inclusion as break dates in the ANOVA model, disregarding the statistical significance of breaks. In spite of the long history of ANOVA correction models and their known advantage, practical homogenization often avoids their application with various argumentations. In the presentation some episodes of the methodological development will be recalled, from which a few steps favored the spread of ANOVA correction application, while others hindered it. The message of this presentation is clear: the use of an ANOVA correction model is obligatory under frequently existing conditions of relative homogenization.

On the development of the first homogenized monthly precipitation dataset for Canada

Xiaolan L. Wang, Yang Feng, Vincent Cheng, and Hong Xu
Climate Research Division, Science and Technology Branch, Environment and Climate Change Canada
xiaolan.wang@ec.gc.ca

This study developed a comprehensive semi-automatic data homogenization procedure to produce gap-infilled and homogenized monthly precipitation data series for 426 long-term/critical stations in Canada, which was then used to assess Canadian historical precipitation trends. Data gaps in the 426 series were infilled by advanced spatial interpolation of a much larger dataset. The homogenization procedure used repeatedly multiple homogeneity tests without and with reference series to identify changepoints/inhomogeneities, the results from which were finalized by manual analysis using metadata and visually inspection of the multiphase regression fits. Up to six reference series were used, including a gridded dataset based on in situ observations and a reanalysis dataset. As a result, 299 out of the 426 data series were found to be inhomogeneous. These series were homogenized using quantile matching adjustments. The homogenized dataset shows better spatial consistency of trends than does the raw dataset.

ONLINE

Climatol v. 4: New functionalities and tools

Jose A. Guijarro
Retired from the State Meteorological Agency (AEMET), Spain
jaguijarro21@gmail.com

The R package climatol has been updated to version 4.0, adding new functionalities and tools:

- Functions to help in preparing the input data in the required format: csv2climatol, sef2climatol and xls2csv are added to the existing db2dat, daily2climatol and rclimdex2climatol.
- The call to the homogenization function is now very simple, since only three parameters are needed: homogen(Variable, FirstYear, LastYear). The rest of parameters are guessed or have safe default values. However, users should revise the results and optionally run homogen again if they want to modify any parameter.
- A new initial quality control has been implemented in the homogen function, showing, for every series, boxplots of values (detecting and deleting clear errors), differences between consecutive values and lengths of sequences with the same value.
- Alternatively to SNHT the user can choose the Cucconi test to simultaneously detect changes in the mean and the variance, but only if reference series do not change along time, because that would introduce variability in the spatial anomalies.
- New tools: (1) fix.sunshine prunes any excess in sunshine hours produced in adjusted daily series. (2) QCthresholds obtains monthly thresholds for every series to implement Quality Control alerts in Climate Data Management Systems.
- New graphical products: Former versions produced wind roses and Walter & Lieth climograms. Version 4 adds 2D density plots, months:hours isopleths, running trends of different periods, meteograms and Intensity-Duration-Frequency curves.

These new functionalities, added to the existing functions dahstat (various statistics) and dahgrid (grids of normalized values) for generating products from the homogenized series, make climatol a useful package for many climatological practices.

Operational homogenization of daily climate series in Spain: experiences with different variables

Belinda Lorenzo, José A. Guijarro, Andrés Chazarra, César Rodríguez-Ballesteros, José V. Moreno, Ramiro Romero-Fresneda, Maite Huarte, Ana Morata
Spanish State Meteorological Agency (AEMET), Leonardo Prieto Castro 8, 28071 Madrid, Spain
blorenzom@aemet.es

The calculation of the new climatological standard normals for the period 1991–2020 was a motivation to carry out the homogenization of the required climatic variables in the Spanish Meteorological Agency (AEMET).

The national observation network has undergone changes along its history that often introduce non-climatic interferences to the series. On the other hand, for the calculation of various parameters and climatic indices it is essential to have complete daily series. With this in mind, the homogenization of the daily series of precipitation, maximum and minimum temperature, sunshine hours, relative humidity, station level pressure, mean wind speed and maximum wind gust was carried out.

This paper shows how the homogenization process was performed, covering the period 1975–2020 with carefully selected daily data sets from the national climatological database. The homogenization software Climatol v.4.0 was used for this process, and derived variables such as average temperature, sea level pressure and vapor pressure were calculated from their related homogenized series.

The peculiarities and issues of each variable are explored and, finally, the homogenization results were used to readily calculate the 1991–2020 climatological standard normals with the dedicated software CLINO_tool v.1.5.

REVISITING HISTALP PRECIPITATION DATASET

Barbara Chimani(1), Oliver Bochniček(2), Michele Brunetti(3), Manfred Ganekind(1), Juraj Holec(2), Beatrix Izsák(4), Mónika Lakatos(4), Melita Perčec Tadić(5), Veronica Manara(6), Maurizio Maugeri(6), Pavel Stastny(2), Olivér Szentes(4), Dino Zardi(7)

(1) GeoSphere Austria, Vienna, Austria (2) Slovak Hydrometeorological Institute (SMHU), Bratislava, Slovak Republik
(3) Consiglio Nazionale delle Ricerche, Istituto di Scienze dell'Atmosfera e del Clima (CNR-ISAC), Bologna, Italy
(4) Országos Meteorológiai Szolgálat (OMSZ), Budapest, Hungary (5) Croatian Meteorological and Hydrological Service (DHMZ), Zagreb, Croatia
(6) Universitá degli Studi di Milano, Department of Environmental Science and Policy, Milano, Italy
(7) University of Trento, Department of Civil, Environmental and Mechanical Engineering (DICAM), Trento, Italy
barbara.chimani@zamg.ac.at

HISTALP is a long-term climatological database for the Greater Alpine Region, including monthly data of precipitation, temperature, sunshine duration and air pressure. The activities leading to this dataset started in 1997 with the final database being in place in 2003. Since then annual updates of the data are done. The dataset is freely available for research and education and frequently used in different climate related studies.

Due to the long existence of the HISTALP dataset, a new homogenisation activity was due in order to ensure the homogeneity of the updated time series. This was done in the course of the creation of a new version of the gridded precipitation dataset of the Greater Alpine region (LAPrec, https://surfobs.climate.copernicus.eu/dataaccess/access_laprec.php) within an international Copernicus project. Before starting the homogenisation activity, the existing data was revisited and an exchange with the data owners on the original data took place. This lead to corrected, historical original data as well as to the replacement of some stations used in HISTALP. Homogenisation was mainly done within the four climate regions of HISTALP (www.zamg.ac.at/histalp), with some special networks e.g. for especially long time series. The results were compared to the former version of homogenised HISTALP-precipitation data as well as to national homogenised datasets.

The analyses of the resulting dataset on trends and data range support the idea of a generally good quality of the homogenisation. For all comparisons with other homogenised datasets the timing of the homogenisation had to be taken in to account, additional differences were to be expected due to the availability and selection of reference series, choice of homogenisation method and availability of metadata (especially for interactive methods). The results of the comparison show that in the national homogenisations more breaks were detected than in the HISTALP one. This was to be expected due to the higher number of highly correlated reference series. Overall, the comparison gives confidence in the HISTALP dataset and its homogenisation. Nevertheless, for a small number of stations strong differences between the different homogenisations have been detected. Those will be assessed by a future step.

Historical precipitation data sets in Hungary

Olivér Szentes(1, 2), Mónika Lakatos(1), Rita Pongrácz(3)

(1) Climate Department, Hungarian Meteorological Service, Kitaibel Pál u.1, Budapest 1024, Hungary
(2) ELTE Faculty of Science, Doctoral School of Earth Sciences, Budapest, Hungary (3) ELTE Department of Meteorology, Pázmány Péter st. 1/A, Budapest 1117, Hungary
szentes.o@met.hu

To better understand the climate system and its changes, we need to analyze long data series with high-quality. However, measurement conditions often change: station relocations, instrument changes, changes of measuring time, changes in environmental conditions and even changes in the observer can all cause breakpoints, so-called inhomogeneities, in the climate data series that are not because of climate change. Homogenization is therefore needed to detect the trends that are strictly due to the impact of climate change.

For homogenization of data series, quality control and filling in the missing values we use the MASH (Multiple Analysis of Series for Homogenization) procedure at the Climate Department of the Hungarian Meteorological Service (OMSZ). By applying the MASHv3.03 software we have homogenized and quality-controlled data series available without missing data for further analysis. The MASH method is based on hypothesis testing. To homogenize the precipitation series, we used a multiplicative model with a significance level of 0.01. Inhomogeneities are estimated from the monthly data series. The monthly, seasonal and annual inhomogeneities are harmonized in all MASH systems (considering different station networks).

The Hungarian precipitation climate database is currently based on two MASH systems: 131 stations from 1901, and 500 stations from 1951. The large jump in the number of data series from 1951 is explained by the fact that the majority of data series in the database were digitized from the mid-20th century.

Extensive precipitation measurements in Hungary began in the 1850s. So recently, all the monthly precipitation data have been collected from the beginning of measurements to 1950, which have not yet been digitized. This allowed a significant expansion of the station systems used for homogenization of data from the first half of the 20th century and the second half of the 19th century.

In this presentation, all the new station systems used in the homogenization process of precipitation data series will be presented. The most important result is that we have first insight into the precipitation conditions in Hungary from the beginning of the measurements up to the present with more spatial details than before.

Acknowledgements: The research presented was carried out within the framework of the Széchenyi Plan Plus program with the support of the RRF-2.3.1-21-2022-00014 National Multidisciplinary Laboratory for Climate Change project.

Homogenisation and interpolation of Slovenian temperature series in period 1950-2020

Gregor Vertačnik
Slovenian Environment Agency
gregor.vertacnik@gov.si

Slovenian Environment Agency has prepared climate normals for the latest reference period, 1991-2020, and renewed analysis of climate change in Slovenia. Air temperature was the first and most important climate variable that has been analysed. First, long time series of five air temperature variables within the period of 1950-2020 were thoroughly checked by various quality control methods. Daily data was aggregated into time series of monthly data. Stations were grouped in five overlapping climate regions and the corresponding time series were homogenised by HOMER. Homogenised time series were merged via linear regression into final homogenised time series of monthly data. This data was then used to create homogenised daily time series.

Daily values for months having complete original data were adjusted to match monthly homogenised series. These adjusted values were used for spatial interpolation of daily data for months with missing original data. Interpolation was based on a subset of days with temperature conditions at predictor stations being similar to the ones at »interpolated« day. Variance of interpolated data was partly adjusted to the variance of original data. Resulting time series were thus a compromise between having proper variance and least interpolation error.

Time series of homogenised and interpolated values have been compared to the results of previous analysis and used for the calculation of climate data normals and climate trends as well as for the analysis of extreme events.

Homogenization of moisture characteristics data in Slovakia

Kristína Szabóová, Katarína Mikulová, Peter Kajaba, Maroš Turňa
Slovak Hydrometeorological Institute
kristina.szaboova@shmu.sk

We would like to investigate changes in humidity characteristics at the most homogeneously considered stations. We use data on relative humidity, soil moisture, potential evapotranspiration and Tomlain's irrigation indicator to investigate moisture characteristics. Twenty-three stations will be selected. It was chosen to represent the northern, south-western and eastern parts of Slovakia, as this state represents a very diverse orography. Those stations have the longest series of observations and have undergone homogenization in the MASH environment. The homogeneity is important for all climatic variables, in all spatial scales, and for the data of all kinds of temporal resolution. There are several methods and software for the homogenization of climate data series. For the purposes of the theme, we used the MASH (Multiple Analysis of Series for Homogenization - T. Szentimrey) program to homogenize the data different variables. This program enables to homogenize monthly as well as daily data series and also to fulfill limited data gaps. The measurements from measuring stations and next chosen elements had to be homogenized this software, which required adequate software for an automatic implementation of the needed steps. Three normal periods will be compared, namely the normal 1961–1990, 1981–2010 and the most recent normal, 1991–2020. The main goal of the work will be to investigate how the individual moisture characteristics change, to what extent the land dries out. If we managed to prove it, appropriate adaptation measures will be necessary.

ONLINE

GNSSseg, a Statistical Method for the Segmentation of Daily GNSS IWV Time Series

Emilie Lebarbier, Annarosa Quarello, Olivier Bock, Ninh Khanh Nguyen
Universite Paris Nanterre
emilie.lebarbier@parisnanterre.fr

Homogenization is an important and crucial step to improve the usage of observational data for climate analysis. This work is motivated by the analysis of long series of GNSS Integrated Water Vapour (IWV) data, which have not yet been used in this context. This paper proposes a novel segmentation method called segfunc that integrates a periodic bias and a heterogeneous, monthly varying, variance. The method consists in estimating first the variance using a robust estimator and then estimating the segmentation and periodic bias iteratively. This strategy allows for the use of the dynamic programming algorithm, which is the most efficient exact algorithm to estimate the change point positions. The performance of the method is assessed through numerical simulation experiments. It is implemented in the R package GNSSseg, which is available on the CRAN. This paper presents the application of the method to a real data set from a global network of 120 GNSS stations. A hit rate of 32% is achieved with respect to available metadata. The final segmentation is made in a semi-automatic way, where the change points detected by three different penalty criteria are manually selected. In this case, the hit rate reaches 60% with respect to the metadata.

Sensitivity of Change-Point Detection and Trend Estimates to GNSS IWV Time Series Properties

Ninh Khanh Nguyen, Olivier Bock, Emilie Lebarbier, Annarosa Quarello
Institut de Physique du Globe de Paris
knguyen@ipgp.fr

This study investigates the sensitivity of the GNSSseg segmentation method to change in: GNSS data processing method, length of time series (17 to 25 years), auxiliary data used in the integrated water vapor (IWV) conversion, and reference time series used in the segmentation (ERA-Interim versus ERA5). Two GNSS data sets (IGS repro1 and CODE REPRO2015), representative of the first and second IGS reprocessing, were compared. Significant differences were found in the number and positions of detected change-points due to different a priori ZHD models, antenna/radome calibrations, and mapping functions. The more recent models used in the CODE solution have reduced noise and allow the segmentation to detect smaller offsets. Similarly, the more recent reanalysis ERA5 has reduced representativeness errors, improved quality compared to ERA-Interim, and achieves higher sensitivity of the segmentation. Only 45–50% of the detected change-points are similar between the two GNSS data sets or between the two reanalyses, compared to 70–80% when the length of the time series or the auxiliary data are changed. About 35% of the change-points are validated with respect to metadata. The uncertainty in the homogenized trends is estimated to be around 0.01–0.02 kg/m^2/year.

A STATISTICAL METHOD FOR THE ATTRIBUTION OF CHANGE-POINTS IN SEGMENTATION OF IWV DIFFERENCE TIME SERIES

Khanh Ninh Nguyen(1,2), Olivier Bock(1,2), Emilie Lebarbier(3)

(1)Institut de Physique du Globe de Paris (IPGP), Centre National de la Recherche Scientifique (CNRS), Institut national de l'information géographique et forestière (IGN), Université de Paris

(2)Ecole Nationale des Sciences Géographiques (ENSG), Institut national de l'information géographique et forestière (IGN)

(3)Laboratoire Modal'X, UPL, Université Paris Nanterre

knguyen@ipgp.fr

Many techniques employed for homogenizing climate time-series from station data rely on comparing the data to a reference series. This approach helps eliminate the common climate signals and improves the accuracy of change-point detection methods. However, it can be challenging to determine if the detected change-point originates from the main series or the reference series. This study presents a statistical method to aid in making this decision. It works by combining the data from the main station with the data from at least one nearby station, where the data from each station is actually composed of two series: a target series and a reference series. In our study, we consider daily GNSS Integrated Water Vapor (IWV) measurements as the target series and the Fifth ECMWF reanalysis (ERA5) as the reference series. From these four base series, six series of differences were created, and a statistical test was performed to identify if there was a significant change in mean before and after the tested change-point in each of the six series. Finally, to determine which of the four base series was impacted by the change-point(s), a predictive rule was used. The statistical analysis employed in this study is a generalized linear regression approach taking into account both heteroscedasticity and autocorrelation. The predictive rule is constructed on a dataset built from the test results obtained on the real data using a resampling strategy. Cross-validation was used to evaluate the performance of four popular machine learning methods such as Linear Discriminant Analysis, decision tree (CART), random forest and the k nearest Neighbor. Out of those, the random forest was found to be the most effective. The results of the proposed method applied to a real dataset were consistent and plausible with respect to GNSS metadata and our knowledge of the data. The study concluded that 41% of the change-points are caused by GNSS data, 15% by ERA5 data, and 25% are the result of coincident detections.

ONLINE

Homogenization of the global historical database (HCLIM)

Elin Lundstad
Norwegian Meteorological Institute
elinl@met.no

Instrumental meteorological observations are crucial for the analysis of climate backwards in time to reconstruct climate variations. However, the collection of instrumental data dating back to 1658 allows many of the long climate series to have often been affected by inhomogeneities (artificial shifts) due to changes in measurement conditions (relocations, instrumentation, change in environment, etc.). To deal with this problem, modern homogenization procedures have been used and developed to detect and adjust inhomogeneities. Homogenization in climate research means the removal of non-climatic changes. This presentation describes the homogenization of the early instrumental dataset (HCLIM; https://doi.pangaea.de/10.1594/PANGAEA.940724) of monthly mean temperature time series. New homogenization algorithm validation methodology is assessed here on early instrumental data, and its use to assess the skill of three different algorithms, when applied to early instrumental data. The methods tested were PHA, BART and CLIMATOL. Results and challenges of using these methods on early instrumental data will be shown. From the database that has been created (published January 2023), it has been necessary to clean up the data and at the same time use the data to see what strengths and weaknesses exist with the database. The clean-up has consisted of changing the data format for use for homogenisation, and testing latitude, longitude and meters above sea level and the possibility of obtaining verified data for comparison. You can also see that there is some data that cannot be used for this purpose. Something can be used for completely different purposes, e.g., finding extremes. It has also been a learning process to use these homogenization programs.

Homogeneity of long-term traditional circulation types series – problems and challenges

Zbigniew Ustrnul(1,2), Bogdan Bochenek(2), Agnieszka Wypych(1)
(1) Jagiellonian University in Krakow, Poland
(2) Institute of Meteorology and Water Management – NRI, Poland
zbigniew.ustrnul@uj.edu.pl

Typically, the problem of the homogeneity of climatological series concerns meteorological data, most often air temperature and precipitation. Rarely is this issue considered in the context of other elements, such as atmospheric circulation. This analysis considers several manual classifications of circulation types in terms of their homogeneity. These were typologies for Hungary (by Peczely), for Southern Poland (by Niedzwiedz) as well as for Central Europe (well-known Grosswetterlagen). The variability of these types was analyzed for a period of over 70 years and their assessment was proposed using the Machine Learning (ML) method and ECMWF reanalyses. The period 1991–2020 was taken for training the system. The obtained results allow to conclude that ML can be helpful in detecting doubts that can be identified when considering particular types of circulation.

DATA AND QUALITY CONTROL

THE 19TH CENTURY PRECIPITATION MEASUREMENT METHOD, A FIELD STUDY

Joelsson, L.M.T., Södling J., Engström, E.
Swedish Meteorological and Hydrological Institute
magnus.joelsson@smhi.se

Systematical precipitation measurements has been conducted in Sweden since the end of the 19th century. The measurement equipment has undergone several changes over the years: Wind shields have been installed, the material and the size of the mouth and the body of the cans have been modified, and the measurements have been automated. These changes can possibly have caused homogeneity breaks in the data. Some changes were applied to the network almost simultaneously, which makes detection of potential homogeneity breaks challenging with relative homogenisation methods. To quantify the difference between the historical and modern precipitation measurement methods, as well as the impact of the installation of the wind shield, a field study was set up.

From November 2016 through May 2021 precipitation was measure at two sites (at SMHI, Norrköping and in Katterjåkk in the northernmost mountain regions of Sweden) using two precipitation cans alike those used for precipitation measurements in Sweden in the late 190th century. One of each pair of cans was equipped with a wind shield.

In Norrköping, where small shares of the annual precipitation falls as snow, the historical cans collects 9 % (with wind shield) and 7 % (without wind shield) more than the automatic gauge.

For Katterjåkk, where a larger share of the annual precipitation falls as snow, the quota of difference between the historical cans and the automatic gauges is 16 % and 5 % respectively.

There is a very low correlation between the departure of the measurements with the historical cans from the modern measurements and other weather observables such as wind speeds and temperature. There is however a clear shift between super- and subzero temperature for the measurements without wind shield, where the historical can compared to the modern instrument collects significantly less at subzero temperature than at superzero temperatures.

The historical cans on average collect more precipitation than the modern automatic gauge, which is consistent with results of previous comparison. Underestimation of measurements without wind shields are most severe in snowy conditions.

Introduction of low cost sensors (LCSs) in the Belgian climate observation network: a first assessment

Cedric Bertrand, Michel Journee, Luis Gonzalez Sotelino
Royal Meteorological Institute of Belgium
cedric.bertrand@meteo.be

There is a greater need for accurate and quantitative climate data based on meteorological observations. Weather observation data are currently used for the real-time preparation of weather analyses, forecasts and severe weather warnings, for the study of climate, for local weather-dependent operations, for hydrology and agricultural meteorology, and for research in meteorology and climatology. All these applications need the execution of high-quality measurements. Weather stations need to be equipped with well calibrated sensors in order to ensure the reliability of measurements. RMI currently operates roughly fifteen fully equipped synoptic AWSs performing a complete set of meteorological observations. RMI's AWS deliver high quality information (known uncertainty) and robust communications (little to no down-time) from which users can rely. Installing a sufficient number of AWSs to ensure an adequate spacing is however difficult to achieve because of economic and geographical limitations such as expensive installation and maintenance costs and difficulties in selecting installation sites.

Advances in sensor networks and Internet of Things (IoT) technologies have created a new epoch in environmental monitoring, facilitating the collection of high-resolution spatiotemporal dataset which could be used to supplement the existing dataset from the current fully equipped AWSs network operated by RMI. Although IoT technology presents plausible tools to expand current capacity in environmental monitoring, the use of low-cost sensors (LCSs) however has raised several concerns especially pertaining to their accuracy, reliability, in-field applicability and performance.

This contribution provides an overview of the RMI's experience of introducing LCSs sensors in its climate observation network. Advantages and drawbacks of the approach will be presented and discussed together with some recommendations.

Data rescue by volunteers of precipitation and temperature extremes from centennial meteorological reports of the Royal Meteorological Institute (1881–1900): a first assessment

Romain Ingels, Michel Journée
Royal Meteorological Institute of Belgium
romain.ingels@meteo.be

Data repositories and archives play a critical role as source for observational data used in the study of past weather events and climate. Therefore, data storage is the first and essential step in the process of making the data usable. The Royal Meteorological Institute (RMI) has been digitizing its archives in image scan format for several years now, thanks to BELSPO's national DIGIT programme, in order to save them. This represents a significant amount of data that can be made available as soon as it has been encoded in a way that can be easily used by everyone. In terms of human resources, it would take a lot of time and people to carry out such a task. This is why RMI is going to launch a first campaign of keying by volunteers of its archives of climate bulletins from 1881 to 1900 via the Zooniverse platform. For this campaign, volunteers will be asked to encode daily precipitation and extreme temperature data, including their metadata. Zooniverse provides a popular way of developing citizen science projects. Using their Project Builder interface, a custom website was created to enable volunteers to transcribe the data from the images.

This first campaign will assess the usefulness and reliability of the results collected through volunteer participation. The focus will be on the design of the encoding page, the feedback from volunteers on the tool provided, the quality of the results but also the popularity of such a process. An first analysis of the results will be carried out to assess whether RMI will carry out other campaigns in the future on a wider range of climate parameters.

CORRECTION OF PROCESSED HISTORICAL RECORDS OF SIPHONED RAINFALL RECORDERS

Tibor Rácz
MATE Hungarian University of Agriculture and Life Sciences
raczt167@gmail.com

In the first three-quart of the last century, the sub-daily rainfall data have been collected mainly with the siphoned rainfall recorders. The operation of these instruments demands a periodical drop-down of water by their siphon unit to ensure the registration of the rainfall on a limited-width paper, while the pen returns to the starting point of the paper. Although the break of the measurement takes only 10-30 seconds, the break in the registration causes surely a systematic error, an under-measurement. In the case when the registration ribbon is available, the correction can be performed. A part of the registration ribbons had been processed, and after it, only the maximum data for certain time units can be seen. In this situation, the inherited errors cannot be corrected using the mentioned procedures. For the processed data a new procedure has been developed, and the systematic error of the siphoned rainfall recorders can be revised even if we know only the extracted maximums of the rainfalls for the given time units. The presentation and the paper show the procedure and its use for data of Hellmann-Fuess rainfall recorders, exploring furthermore the effect of the data correction on the IDF curves.

ONLINE

Preliminary results on a new dataset of daily observations from a dense network of weather stations covering the Extended Alpine Region (1991–2020)

Giulio Bongiovanni (1,2), Michael Matiu (2), Alice Crespi (3), Anna Napoli (2), Bruno Majone (2), Dino Zardi (2)

(1) University School for Advanced Studies Pavia (IUSS), Pavia, Italy

(2) University of Trento, Trento, Italy

(3) Eurac Research, Bolzano, Italy

giulio.bongiovanni@iusspavia.it

The Alpine area is one of the most vulnerable and sensitive mountain regions of the Mediterranean area to the continuous warming of climate and it is considered an important hotspot of climate change. In particular, climate change is expected to have strong impacts on all components of the hydrological cycle, including river regimes, with consequent effects on the services offered by the freshwater ecosystem, as well as on water availability for users, thus affecting several socio-economic sectors.

Several observational products of key climate variables are available to the scientific community and have been widely used to evaluate the extent of the ongoing effects of climate change in mountain regions. However, most products suffer from some limitations: for example, some have not been updated to recent decades, or their spatial coverage is quite limited, especially in view of reliably assessing trends at higher elevations. Therefore, more efforts must be dedicated to produce new harmonised and high-resolution products able to permit more robust assessments of climate change in mountain regions. Even though the Alpine region is densely covered with a high number of in-situ weather stations, the collection and management of such data for the whole Alpine area is a challenging task due to strong fragmentation and diversity of data sources.

We present here preliminary outcomes from a project aiming at preparing a new extensive observational dataset for the Extended Alpine Region (EAR). The first step of the project consisted in gathering in-situ daily measurements of meteo-climatic variables from a variety of meteorological and hydrological services in the area covering all the available period up to 2020. The observational network includes more than 9000 in-situ weather stations. Such high density allows for an extended and uniform coverage both in space and elevation. After a preliminary phase of pre-processing to unify formats, data have been subjected to a quality control aimed at checking internal, temporal, and spatial consistency of time series, and outlier identification. In addition, different techniques were exploited to assess data homogeneity.

COMPARISON OF DATA MEASURED AT CLIMATOLOGICAL AND SYNOPTIC OBSERVATION TIMES

Hela Irha, Ana Weissenberger

Croatian Meteorological and Hydrological Service
hela.irha@cirus.dhz.hr

Climatological observations in Croatia are taken according to local mean time. For meteorological stations that lie on the 15th meridian, local mean time and standard (Central European) time are the same. If a main meteorological station is farther than 2,5° from the 15th meridian, and the difference between local mean time and standard time is more than 10 minutes, meteorological data are collected both at standard time for synoptic purposes and at local mean time for climatological purposes.

This study discusses whether the difference in data measured at so slightly different observation times is significant enough to maintain this practice at the main meteorological stations operated by Croatian Meteorological and Hydrological Service. Statistical analysis was carried out to quantify the difference in monthly mean temperature calculated from the data from the two observation times. Air temperature data measured manually and by automatic sensors were both used. Results show no significant change in temperature values nor the difference in monthly mean temperature measured at synoptic hour and at the climatological observation time only 15 minutes earlier. The possible causes for small discrepancies between the data series are also discussed.

DATA QUALITY CONTROL APPLIED ON ECA&D

Petr Štěpánek(1,3), Gerard van der Schrier(2), Pavel Zahradníček(1,3)
(1) Global Change Research Institute (CzechGlobe), Brno, Czech Republic
(2) Royal Netherlands Meteorological Institute (KNMI), De Bilt, the Netherlands
(3) Czech Hydrometeorological Institute, regional office Brno, Czech Republic
stepanek.p@czechglobe.cz

The ECA&D data-set collects observations for the Essential Climate Variables of more than 22.400 stations from all countries in Europe. Large number of these series are affected by outliers, repeated values and other issues in the measurements. These quality issues need to be identified prior to further processing of the data into e.g. the production of the gridded E-OBS datasets as these issues may lead to erroneous estimates of climate impact indices and trends. The data quality control – called MetQC method and developed at GCRI, combines checks on duplicate series, repetitive values and outliers with an inter-stations comparison. Earlier work compared five different quality check methods on four benchmark data-sets covered by ECA&D. This work indicated that the MetQC approach performed well in comparison against the other methods and the new method replaces the more basic approach at ECA&D in which straightforward stand-alone tests were conducted based on physical limits and logical constraint testing. A strong aspect of the MetQC method is that it provides an estimate for an alternative value of a suspect value based on values of surrounding stations, with a quantification of the reliability of this alternative. The method is fully automated, and identified data issues are accompanied by further outputs like time series plots and maps, that help users in a final decision about how to proceed with the found suspicious values and outliers. The MetQC method has been further refined and tailored to the application at ECA&D to be capable of handling the vast dataset in a reasonable time. The presentation will focus on the improvement in quality for ECA&D and comparisons in terms of numbers of flagged data with focus on air temperature and precipitation time series.

AQUAS – Data Quality Control at GeoSphere Austria

Niko Filipović

Data Quality and Digitisation
GeoSphere Austria
Hohe Warte 38, 1190 Vienna
niko.filipovic@geosphere.at

Austria Quality Service - short AQUAS - is an automatic tool for quality control of meteorological observation data. The software was developed at ZAMG (now GeoSphere Austria) in Vienna as part of quality management in the field of real-time processing of near-surface observation data. The data is collected by the automatic weather station network TAWES, which is operated by the Austrian national weather service. Since the observation network is subject to constant change, AQUAS offers the possibility of quickly integrating new stations or stations from third-party networks into the quality control procedure.

The basis for quality control in AQUAS are the standard methods for checking meteorological and climatological data (e.g. plausibility check, temporal, spatial and internal consistency tests, etc.). These test procedures are continuously improved and further developed in AQUAS. Individual system components are designed for testing of incoming observation parameters in real-time. For the highest possible flexibility, each parameter can be processed autonomously, independent of the other measured variables of a weather station. In addition, data from other sources like radar and satellite data as well as Numerical Weather Prediction model data and data from hydrological monitoring network can also be used within AQUAS.

Examples for the operational use of AQUAS are presented as well as the current state of research on the quality control of 10-minute wind speed data. A method is depicted in which the daily sums of global radiation and sunshine duration are used to detect instrument malfunctions.

ONLINE

Climate services products in Polish NMHS

Michal Marosz
Institute of Meteorology and Water Management - National Research Institute
mmarosz@imgw.pl

Climate change and its consequences are a significant contemporary issue. Polish NMHS (IMGW-PIB) introduced a service of climate monitoring to its portfolio many years ago (in the early 2010s) in the form of a monthly bulletin. It has evolved to a bulleting presenting most of the aspects of climatological features variability (atmospheric circulation, solar radiation, temperature, pluvial characteristics recorded extremes from the preceding month.

However, its fixed form prevents the presentation of many aspects of climate variability due to technical issues and the restrictions posed by the level of expertise of the end users. Due to those constraints, IMGW-PIB developed its climate monitoring web page (klimat.imgw.pl), which is complementary to the monthly bulletin. We present the thermal and pluvial conditions variability from the current year with respect to multiannual conditions (1991–2020 normals) at all synoptic stations. Also, in the context of recorded extremes, the exceedance probabilities (temperature, precipitation totals, wind speed) are calculated and subsequently presented as maps. This is accompanied by annual reports. This allows us to inform the general public about the contemporary variability of climate in Poland and provide additional information on climate change.

SPATIAL INTERPOLATION AND GRIDDED DATASETS

STATISTICAL MODELLING OF THE PRESENT CLIMATE BY THE INTERPOLATION METHOD MISH - THEORETICAL CONSIDERATIONS

Tamás Szentimrey
Varimax Limited Partnership
szentimrey.t@gmail.com

In the statistical climatology the climate can be formulated as the probability distribution of the meteorological events or variables. The purpose of the statistical climatology is to estimate or model the climate probability distribution or equivalently the climate statistical parameters. Furthermore the meteorological data series make possible to estimate or model the climate statistical parameters in accordance with the establishments of statistical climatology principles.

Our method MISH (Meteorological Interpolation based on Surface Homogenized Data Basis; Szentimrey and Bihari) was developed for spatial interpolation of meteorological elements. According to the mathematical theorems the optimal interpolation parameters are known functions of certain climate statistical parameters, which fact means we could interpolate optimally if we knew the climate. Consequently according to the principles of climatology the modelling part of software MISH is based on long meteorological data series. The main difference between MISH and the geostatistical interpolation methods built in GIS is that the sample for modelling at GIS methods is only the predictors, which is a single realization in time, while at MISH method we use spatiotemporal data for modelling, which form a sample in time and space alike.

We focus the methodology of the modelling subsystem built in MISH. This subsystem was developed to model the following climate statistical parameters for half minutes grid: monthly, daily expected values, standard deviations and the spatial, temporal correlations. Consequently the modelling subsystem of MISH is completed for all the first two spatiotemporal moments. If the joint spatiotemporal probability distribution of a given meteorological element is normal (e.g. daily, monthly mean temperature) then the spatiotemporal moments above uniquely determine this distribution, which is the mathematical model of the present climate for this meteorological element.

Development of observation based temperature dataset

Beatrix Izsák(1), Tamás Szentimrey(2), Zita Bihari(1) and Zsófia Barna(1)

(1) Hungarian Meteorological Service

(2) Varimax Limited Partnership Budapest, Hungary

izsak.b@met.hu

As climate models have evolved, there has been an increasing need to study the intraday characteristics of temperature and precipitation, which has required refinement of the daily grid point observation database over time. In this presentation we will describe the process of creating a six-hourly temperature database and the main features of the databases.

For temperature we have four daily measurements for 58 stations since 1970. Using these, and by further development of the MASH and MISH software, a 6-hourly (0, 6, 12, 18 UTC) grid point observation database was produced for the area covering Hungary with 1233 grid points, for the period 1970–2022 for temperature.

During the homogenization of the six-hourly temperature data series with MASH, the breakpoints detected in the daily data series are automatically used as META data. The homogenized station data series are used for interpolation with the MISH software. The MISH interpolation system uses the results of the modelling subsystem. However, these systems were designed for daily, monthly data series, so improvements were needed when applying them to hourly temperature data. The methodological improvement is to use the modelled climate statistics parameters for daily values to interpolate hourly values. Spatial trend values were modified using regression coefficients between daily and hourly values.

Homogenised data series, monthly temperature grids and climate monitoring products for Croatia

Melita Perčec Tadić, Zoran Pasarić, José A. Guijarro
Croatian Meteorological and Hydrolological Service
melita.percec.tadic@cirus.dhz.hr

We are presenting homogenised mean monthly air temperature series, monthly temperature grids and derived climate monitoring products for assessing the air temperature conditions and the observed temperature change in Croatia. Monthly mean temperatures from 122 Croatian stations are homogenised and high resolution monthly gridded data are developed for the 1981–2018 period. The hierarchical clustering is introduced to define climate regions in Croatia needed for homogenisation. The breaks of homogeneity are detected by the standard normal homogeneity test. Further on, the regression kriging is applied to produce 1 km x 1 km monthly grids for each month in the analysed period. The quality of the interpolation was assessed by leave-one-out cross-validation and the root mean square error of 0.7°C. The quality of spatial interpolation is estimated with normalised error maps. Climate normals and trends are derived from homogenised station data and monthly grids. After 2000, average annual anomalies from the 30-years climate normal 1981–2010 were positive and up to 1.4 °C warmer than the average, and just occasionally negative. The significant strong warming is observed and mapped over the entire Croatian territory in April, June, July, August and November, being stronger inland than on the coast. Annual trends were significant and in the range from 0.3 °C/decade to 0.7 °C/decade. That suggests that our region could face consequences such as devastating heatwaves, water shortages, loss of biodiversity and risks to food production, especially as being part of the Mediterranean where it seems that the observed trends are 2-2.5 times stronger than the global mean. We can hope that some of the presented climate monitoring products can help in assessing the vulnerability and the risk from climate change and help with the mitigation of the potentially affected sectors like forestry, agronomy, tourism, water management, energy production or consumption, health or others.

Comparison of gridding methods for precipitation over Canada and assessment of station/data density effects on gridding results

Kian Abbasnezhadi, Xiaolan L. Wang
Climate Research Division, Environment and Climate Change Canada
xiaolan.wang@ec.gc.ca

This study compares, for the very first time, the Optimal Interpolation (OI) used to produce the Canadian Gridded (CanGRD) data against Thin-Plate Smoothing Spline and ordinary kriging for representation of precipitation climate and variability and regional mean precipitation trends. The comparisons also include comparing the results from applying the gridding methods to different precipitation variables, including total monthly amounts, the anomalies and relative anomalies, and the climate normals of monthly precipitation. Two observation-based gridded precipitation datasets were used as `truth', each sampled at full and subsets of a typical long-term precipitation station network in Canada to obtain surrogate monthly observation data for gridding. The gridded data were then evaluated against the corresponding truth. The results show that, among the three gridding methods, kriging is overall the best, while CanGRD\nobreakdash-OI is the worst in presenting the precipitation climate and variability, and the CanGRD-based trend estimation method notably overestimates the regional mean precipitation trend for the North, but underestimates it for the South. It is also shown that the best regional mean precipitation trend estimates are obtained by using kriging to grid climate normals and relative anomalies, separately. This approach was then used to assess the impact of different station/data densities on the gridded data. The results show that the lower the station/data density goes below its critical threshold, the less accurate is the gridded data, and that the station/data density of the traditionally used Canadian precipitation dataset is associated with large biases in both the regional mean precipitation amounts and trends.

Spatio-temporal Imputation of Missing Rainfall Values to Establish Climate Normals

Brian O'Sullivan, Gabrielle Kelly
University College Dublin
brian.osullivan3@ucdconnect.ie

When producing climate normals, the presence of missing data can severely limit the coverage of a climate monitoring network. The WMO recommends that only stations with at least 80% completeness during the averaging period should be considered when producing Long-Term Averages (LTAs, 30-years). This greatly reduces the availability of data collected by Met Éireann (The Irish meteorological office) to only 175 stations out of a potential >1100. In this study, monthly data is considered from 474 rain gauge stations that are 50% complete or greater from 1981–2010. The imputation of missing values for said stations enables their inclusion in the calculation of LTAs for Irish rainfall, greatly improving overall spatial coverage of the island.

Two spatio-temporal imputation methods are investigated. The first proposed method is regression-kriging with a spatio-temporal variogram. Following initial trend removal by Ordinary Least Squares (OLS) from the data, an empirical variogram is produced from the resulting residuals. A spatio-temporal correlation structure is considered consisting of spatial, temporal and joint Matérn covariance models. Known as a sum-metric spatio-temporal covariance, this model is fit to the empirical variogram by surface-fitting through the Limited-memory BFGS algorithm with bound constraints on the model parameters (L-BFGS-B). After the initial variogram fit, the resulting covariance matrix is used to repeat trend removal on the original data by Generalised Least Squares (GLS). New residuals are obtained from this GLS output and are again used to produce a fitted sum-metric variogram. Residuals corresponding to missing values are interpolated by kriging at each step, with kriging weights obtained from the fitted variogram. The final imputed values are obtained by adding these interpolated residuals to the estimated trend for each missing data entry.

Ordinary Kriging with trend removal by Elastic-Net Regularisation is the second imputation technique investigated. Elastic-Net regression is used to detrend the original data, and a sum-metric spatio-temporal variogram is produced and fit to residuals similar to the first method. In this case, the initial fitted variogram is used to calculate kriging weights. Missing values are imputed as the sum of their interpolated residuals from kriging with the estimated trend from Elastic-Net regression. While regression-kriging displays a stronger representation of the underlying spatio-temporal dependence structure and subsequently better performance, this second method boasts much lower computational costs for only a small trade-off in prediction accuracy.

Both methods have been evaluated by ten-fold cross-validation and root mean squared error computed, showing a good improvement compared to their purely spatial counterparts. These methods present themselves as strong options for imputing missing precipitation data for moderately-sized datasets with adequate spatial and temporal dependency.

The work presented in this study is a joint-research between Brian O'Sullivan and Prof. Gabrielle Kelly, University College Dublin, Ireland.

Extending the UK gridded rainfall dataset HadUK-Grid to 1836 with citizen science data recovery

Mark McCarthy(1), Ed Hawkins(2), Dan Hollis(1), Michael Kendon(1), Tim Legg(1)

(1) Met Office

(2) University of Reading

mark.mccarthy@metoffice.gov.uk

The UK Met Office produce a maintain a gridded dataset of key climatological parameters spanning the late 19th Century to present. However the extent of paper records of observations still exceeds that of those available digitally, particularly for dates prior to 1960. One collection held by the Met Office national archives are known as the “Ten year rainfall sheets” a collection of loose-leaf pages providing monthly and annual rainfall totals for decadal periods for thousands of locations across the UK and Ireland. A recent citizen science project led by colleagues at the University of Reading in collaboration with the Met Office has successfully digitised 5.2 million historical monthly rainfall observations from this collection, spanning the period 1677 to 1960. This rich data source has been used to extend the Haduk-grid gridded monthly rainfall product for the UK to 1836. In this presentation I will summarise the steps from the archives to the gridded dataset, the improvements to our national climate monitoring this provides, limitations of the new data, and highlight its value for identifying certain hydrological extremes.

HOMPRA Europe 2 – An update of a gridded precipitation data set from European homogenized time series

Elke Rustemeier, Peter Finger, Udo Schneider, Markus Ziese, Stephanie Hänsel
GPCC, Deutscher Wetterdienst
elke.rustemeier@dwd.de

Reliable observational data are essential for robust climate analyses, especially long-term trends. HOMPRA Europe 2 (HOMogenized PRecipitation Analysis of European in-situ data) is a gridded monthly precipitation dataset based on homogenized time-series. The carefully selected database of more than 5500 stations is a subset of the time series collected by the Global Precipitation Climatology Centre (GPCC). The subset is characterized by few missing values < 20% and intensive quality control.

Compared to its predecessor, the new HOMPRA Europe 2 product covers a longer period from 1951-2015 and is syncronized with the more recent data of the GPCC Full Data Monthly product, so that the data can be extended to the present without a break, although a homogeneous time series cannot be guaranteed for these more recent data.

The actual homogenization process consists of multiple steps: In the first step, for each station, comparable time series are selected. The decision is based on the correlation and Ward's method of minimum variance performed on the deterministic first derivative. For the artificial breakpoint detection, natural variability and natural trends are temporarily removed using the previously selected neighboring time series. This ensures that only artificial changes can be detected. The actual detection applied to annual values is based on the segmentation algorithm of Caussinus and Mestre (2004). In the final step, the breaks detected in the monthly data are corrected using multiple linear regression (Mestre, 2003).

An additional random variation of the neighboring stations shows the robustness of the correction of the individual time series. Finally, the actual HOMPRA Europe 2 product is created by interpolating the homogenized series onto a 1° grid using the modified spheremap interpolation schemes in operation at the GPCC (Becker et al., 2013 and Schamm et al., 2014).

References:

Caussinus, H., und O. Mestre, 2004: Detection and correction of artificial shifts in climate series, Journal of the Royal, Statistical Society. Series C (Applied Statistics), 53(3), 405-425. Mestre, O., 2003: Correcting climate series using ANOVA technique, Proceedings of the fourth seminar

Willmott, C.; Rowe, C. & Philpot, W., 1985: Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring The American Carthographer, 12, 5-16 Becker, A.; Finger, P.; Meyer-Christoffer, A.; Rudolf, B.; Schamm, K.;

Schneider, U. & Ziese, M., 2013: A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present Earth System Science Data, 5, 71-99 Schamm, K.;

Ziese, M.; Becker, A.; Finger, P.; Meyer-Christoffer, A.; Schneider, U.; Schröder, M. & Stender, P., 2014: Global gridded precipitation over land: a description of the new GPCC First Guess Daily product, Earth System Science Data, 6, 49-60

Quality Control and Grid Creation at the Global Precipitation Climatology Centre (GPCC)

Markus Ziese, Elke Rustemeier, Udo Schneider, Peter Finger, Astrid Heller, Raphaele Schulze, Magdalena Zepperitz, Siegfried Fränkling, Bruno Heller, Jan Nicolas Breidenbach, Stephanie Hänsel

Deutscher Weterdienst

markus.ziese@dwd.de

The Global Precipitation Climatology Centre (GPCC) was established in 1989 on request of WMO at Deutscher Wetterdienst (DWD). Under WMO auspice, it collects in situ precipitation data. Those data are quality controlled and stored in a relational data bank system. Gridded precipitation analyses are produced based on these data by means of a modified SPHEREMAP scheme.

Data sources for the GPCC are near real-time SYNOP and CLIMAT data, global and regional data collections and as backbone the data provided by the national meteorological and hydrological services (NMHS). Different quality control procedures are applied to the data depending on the source and created analysis. For the near real-time analysis based only on SYNOP data, a fully automatic quality control is performed. The later available CLIMAT data pass also an automatic quality control, followed by a visual inspection for spatial consistency together with the SYNOP data. Statistical checks were applied to data from all the other sources and flagged data were manually confirmed, corrected or rejected. This comprises also a comparison with existing data in the data collection.

The data are stored in a relational data bank system. Each type of data source has its own slot and flags document the performed quality control and changes. For each record, two values can be stored – the original value and a corrected one. This is a roll-back option, if additional information is later available and the decision about a correction or deletion of a value could be revised. Due to the source specific slots, a cross-comparison between the sources is possible and gaps from one source could be filled by data from another source.

The gridded data should represent the mean area precipitation amount in each grid cell. To create those grids, a modified SPHEREMAP scheme is applied. The modifications concern the treatment of observations near the grid point to avoid the utilization of single observations as representative for the grid cell. Other modifications increase the number of applied stations in station-dense regions. For all grid creations, the first step is the interpolation to the corners of a sub-grid and averaging of those four values to the grid cell mean of the sub-grid. Depending on the final grid resolution, in a second step a certain number of sub-grid cells are averaged, taken care of the area of each sub-grid. Knowing that the spatial correlation of precipitation anomalies is much larger than that of precipitation totals, absolute (monthly) and relative (daily) anomalies are interpolated and superimposed with gridded long-term means (monthly) or monthly totals (daily). Absolute values are only interpolated for the creation of gridded long-term means.

We will present an overview of the GPCC quality control and issues detected within the data. Furthermore, the gridding procedure will be presented.

Quality Control and Creation of Grids of Meteorological Variables for the Copernicus Emergency Management Services

Markus Ziese(1), Zora L. Schirmeister(1), Carina-Denise Lemke(1), Jakub P. Walawender(1), Christoph Schweim(2), Damien Pichon(3), Stefania Grimaldi(4), Gonçalo Gomes(4), Peter Salamon(4)

(1) Deutscher Wetterdienst (2) KISTERS AG

(3) KISTERS France (4) European Commission Joint Research Centre

markus.ziese@dwd.de

The Meteorological Data Collection Centre (MDCC) is operated by a consortium of KISTERS AG and Deutscher Wetterdienst (DWD) on behalf of the European Commission to support the Copernicus Emergency Management Services (CEMS) with near-real time in-situ meteorological data. CEMS services using MDCC data are the European Flood Awareness System (EFAS), the European Drought Observatory (EDO) and the European Forest Fire Information System (EFFIS). MDCC collects near-real time data from currently 35 data providers, processing about 14,000,000 data records per day. Apart from precipitation and air temperature, which are of key importance for the CEMS services, also wind speed, water vapour pressure, relative air humidity and solar radiation data are collected, amongst other variables.

All the collected data are imported into a data management system (WISKI by KISTERS), which applies automatic quality checks during the import of the data. Threshold and rate-of-change validations are applied as well as cross-validation, statistical checks and temporal consistency analyses. The latter are used to identify gaps in the time-series, from single missing observations to longer failures. Thresholds depending on climate zone and season are among the recent implementations aimed at improving the quality control process. Currently, an additional data validation framework is being implemented allowing even more complex analysis of time series data including spatial validations. Quality-controlled observations are used to compute aggregated values (min, mean, max, or totals) over several different aggregation periods (six hours, one day, one month, one year).

Depending on the CEMS service, availability of gridded data is required. Gridded data products are created using the validated station data with additional post-processing. Therefore, the data undergo height correction (if necessary), followed by a spatial interpolation using a modified SPHEREMAP scheme. Several optimizations are added to SPHEREMAP in order to create 70 grids for 6-hourly and daily aggregation levels at 1 arc minute spatial resolution for an extended European domain (13.454.100 grid cells) every day. In order to interpolate temperature (min, max, average) and vapor pressure, a height correction is also applied to the gridded data.

The presentation will provide an overview of the used and planned quality control techniques with real data examples focusing on false acceptances and rejections. A brief overview of the implemented optimizations to SPHEREMAP will be also presented.

ONLINE

SPARTACUS: A KM-SCALE DAILY DATASET OF THE SURFACE CLIMATE IN AUSTRIA – OVERVIEW AND NEW DEVELOPMENTS

Anna-Maria Tilg(1), Johann Hiebl(2), Angelika Höfler(1), Anna Rohrböck(1), Christoph Frei(3)

(1) GeoSphere Austria, Department Climate-Impact-Research, Vienna, Austria

(2) Consultant of GeoSphere Austria, Aschbach-Markt, Austria (3) MeteoSchweiz, Zurich, Switzerland

anna-maria.tilg@geosphere.at

SPARTACUS is a high-resolution spatial analysis of daily surface climate variables for the territory of Austria. It is an operational climate-monitoring dataset produced by the national weather service of Austria. This dataset is used for climate monitoring and is applied in various applications, which require spatiotemporal information about the past like drought or snow cover modelling. It has a spatial resolution of 1 km to 1 km and provides data since January 1, 1961 with a temporal resolution of one day and further temporal aggregates (month, season and year).

In the current operational version the surface in-situ observations of the parameters air temperature, precipitation amount and sunshine duration are interpolated with parameter-specific methods (Frei et al. (2015); Hiebl and Frei (2016); Hiebl and Frei (2018)). The number of considered stations depends on the parameter itself, but the station networks used for the analyses are chosen to be relatively constant over time, in order to reduce temporal inconsistencies. Despite this, local inhomogeneities may still be present, because not all the station series were homogenized.

A gridded dataset of parameters describing surface air humidity is currently under development. Such a dataset is highly relevant for example for drought monitoring. The actual spatial analysis is performed for the parameter of dew point depression, because this allowed adapting the interpolation method previously used for air temperature. To enable a conversion from dew point depression into other air-humidity variables like relative air humidity, a gridded dataset of the mean 24-hour air temperature was developed in parallel. This newly developed dataset overcomes the issue of the shift of the observation time from 21 o'clock to 19 o'clock in 1971 in Austria and takes the advantage of having increasing lengths of automatic station observations. After analysing the pending evaluation of the humidity dataset, it is planned to operationalise the daily interpolation of dew point depression.

The presentation will provide an overview of the currently available components of SPARTACUS with illustrative examples, key-results from the evaluation, and references from applications. Moreover, preliminary findings will be shown from the current developments that may be relevant for similar attempts in other regions with complex topography.

References

Frei C, Willi M, Stöckli R, Dürr B (2015) Spatial analysis of sunshine duration in complex terrain by non-contemporaneous combination of station and satellite data. Int J Climatol 35/15:4771–4790. https://doi.org/10.1002/joc.4322

Hiebl J, Frei C (2016) Daily temperature grids for Austria since 1961 – concept, creation and applicability. Theor Appl Climatol 124:161–178. https://doi.org/10.1007/s00704-015-1411-4

Hiebl J, Frei C (2018) Daily precipitation grids for Austria since 1961 – development and evaluation of a spatial dataset for hydroclimatic monitoring and modelling. Theor Appl Climatol 132:327–345. https://doi.org/10.1007/s00704-017-2093-x

ONLINE

Community climatology: Combining official observations, crowd-sourced observations and covariates to obtain high-resolution gridded climate data sets

Jouke de Baar, Gerard van der Schrier, Theo Brandsma

KNMI, De Bilt, The Netherlands

jouke.de.baar@knmi.nl

Problem. When talking to end-users about gridded climate data sets for the Netherlands, we received the question if we can provide hourly gridded maps of past and current weather, representative for a 1 km X 1 km resolution. Examples of areas of impact, where such high-resolution data could prove useful, are the spread of infectious airborne diseases, wind damage to trees, etc. The effective resolution of the maps KNMI currently provides is considerably lower, so this would involve innovation.

Approach. Two possible approaches to increase spatial resolution are (a) the use of alternative data sources like crowd-sourced observations, and (b) the use of covariate information (like land use) for downscaling. However, neither of these approaches provides sufficient improvement in the effective spatial resolution. In the present research, we include both solutions at the same time, thus combining official observations with crowd-sourced observations and covariates. This may significantly improve the effective spatial resolution of gridded data sets.

Results. We are currently in the process of providing multi-year high-resolution gridded data sets for the Netherlands, for air temperature, precipitation, relative humidity, wind speed, wind gust and wind direction. We will present the first results and insights obtained from this process.

ONLINE

Multi-fidelity adaptive sampling: Towards optimal station location choice of combined official and crowd-sourced weather observation networks

Jouke de Baar, Gerard van der Schrier

KNMI, De Bilt, The Netherlands

jouke.de.baar@knmi.nl

Problem. In climate research as well as severe weather early warning services, end-users are asking for high-resolution gridded data sets. Recently we have seen several promising results in fusion of official and crowd-sourced weather observation networks, particularly in combination with covariates (like land use). Considering that we might want to expand our weather observation network, by adding more official and/or crowd-sourced stations, what would be the optimal way of doing this? Where in a country should new stations be located, should they be official or crowd-sourced stations, and how much improvement can we expect from adding these stations?

Approach. It is interesting how much we can learn from collaborating between different scientific disciplines. For example, in computational fluid dynamics (CFD) based vehicle design, various lessons have been learned in the area of multi-fidelity adaptive sampling (MF-AS). It turns out that, with some adjustments, we can apply the ideas and techniques of MF-AS to design a strategy for efficient growth of combined official and crowd-sourced observation networks.

Results. First study, we identify typical statistics of air temperature fields in the Netherlands by analyzing a year of hourly data. Then, based on these statistics, we illustrate how MF-AS can be modified and applied to optimally expand and shape a multi-fidelity weather observation network.

POSTERS

Climate services in the Czech Republic

Miroslav Trnka(1), Petr Štěpánek(1,2), Zdeněk Žalud(1), Pavel Zahradníček(1,2), Jan Balek(1), Daniela Semerádová(1), Lucie Kudláčková(1), Martin Možný(1,2)

(1)Climate services in the Czech Republic Global Change Research Institute CAS

(2)Czech Hydrometeorological Institute

zahradnicek.p@czechglobe.cz

Global Change Research Institute (GCRI) is a public research institution, European center of excellence investigating the ongoing global change and its impact on the atmosphere, biosphere and human society through the use of the latest techniques and instrumentation. The research focuses primarily on the development of climate and its future scenarios, on the carbon cycle and the effects of changing conditions on the production and biodiversity of ecosystems and on the impacts on the future development and behavior of our society.

Our mission is provides for the public, state authorities, politicians and experts the climate information (free of charge). This is accomplished through special web portals. The most important is drought monitoring system (www.intersucho.cz), that's focused on monitoring agricultural drought and their forecast for 10 days ahead by numerical weather prediction models. We have expanded this drought monitoring to a global level thanks to the cooperation with the platform www.windy.com in the last two years. Climate change also brings an increased risk of wildfires. Therefore, it is necessary to have fire risk monitoring based on fire weather including its forecast (www.firerisk.cz). Another important tool is portal focused on monitoring and forecasting on-line system - focused on selected biotic and abiotic noxious factors (www.agrorisk.cz). These include early warning of late spring frosts, strong winds, high temperatures, drought or pets and diseases. Most of the products are created in broad cooperation between professional and state organizations such as the Czech Hydrometeorological Institute, Mendel University or the State Land Office.

All these publicly available products are based on precisely prepared meteorological and climatic data. Measured data are after quality control and homogenization. Numerical weather prediction models or climate model outputs are bias corrected. The results are also interpolated in the 500×500 m resolution by again own methods, which are developed for individuals purposed of the products (from hourly data to the long-term averages, various elements etc.)

Comparison of different interpolation methods for six-hour temperature data series

Zsófia Barna(1), Beatrix Izsák(1), Ildikó Pieczka(2)
(1) Hungarian Meteorological Service,Kitaibel Pál Street 1, H-1024, Budapest, Hungary
(2) ELTE Eötvös Loránd University, Institute of Geography and Earth Sciences, Department of Meteorology, Pázmány Péter sétány 1/A, Budapest H-1117, Hungary
zsofibarna0528@gmail.com

The application of the MISH (Meteorological Interpolation based on Surface Homogenized Data Basis) interpolation system in meteorological practice and its effectiveness are outstanding. Interpolation with the MISH system compared to ordinary kriging and inverse distance methods used in climate studies has a much smaller error value in the data series of all databases, so its use is much more recommended compared to other interpolation methods. Its reason is the fact that in meteorology we cannot assume that there is no spatial trend since the expected value differs in different landscape units due to the country's complex orography. A major advantage of the MISH system is that it uses homogenized station data series to model statistical climate parameters. When applying the MISH system to hourly values we used the daily model results and modified the spatial trend values using the regression coefficients between daily and hourly values.

Cross validation technique was used to compare different interpolation methods which is available in the Geostatistical Analyst toolbox of the ArcGIS software. This software randomly divides the station system into training and test datasets. For each subset and each hourly value 50 interpolations were performed. After that we calculated the RMSE (root mean square error) values for all three methods. The results are shown in this poster.

Long-term changes of cold indices in Hungary using homogenized and raw data series

Kinga Bokros, Mónika Lakatos, Olivér Szentes
Hungarian Meteorological Service
bokros.k@met.hu

Several indices are used to describe cold weather and cold extremes for instance ice days, frost days, cold and sever days or the cold-spell-duration index (CSDI). In order to determine the changes of the cold extremes caused by climate change, completed, quality-controlled and homogenized minimum and maximum temperature data series have to analyse. In our study we calculated the annual values of these indices for the whole area of Hungary from 1901 to 2022. We used raw data series first to derive the cold indices and then homogenized data series created by Multiple Analysis of Series of Homogenization (MASH) (Szentimrey) method and compared the results. Linear trend model was used to estimate the changes, and to test its significance. The station data series (raw and homogenized) were interpolated to a 10km resolution grid with Meteorological Interpolation based on Surface Homogenized Data Basis (MISH) (Szentimrey & Bihari) interpolation method to ensure evenly the territorial coverage. The countrywide averages of the specific indices are considered as the average of the grid points covering Hungary. Above the countrywide averaged values, indices series and trends for Budapest station were examined more detailed since the beginning of the 20th century until 2022.

HOMOGENISATION OF SWEDISH MONTHLY PRECIPITATION FROM 1880

Joelsson, L.M.T., Schimanke, S., Engström, E.
Swedish Meteorological and Hydrological Institute
magnus.joelsson@smhi.se

An ongoing project at the Swedish Hydrological and Meteorological Institute aims to homogenise the Swedish monthly precipitation data set. Systematical precipitation measurements has been conducted in Sweden since the end of the 19th century and the entire data set covers more than 2 000 individual time series.

The first approach is to follow the method of the homogenisation of the monthly temperature data set (Joelsson et al. 2023), where time series initially are merged with an automatic merging script. Preliminary results indicate the use of data from over 1 800 stations in over 1 200 merged time series. The merging shifts the distribution of length of time series towards longer time series: Time series with at least 60 years of data is increased by about 25%. Merging also restricts the need of interpolation of missing data: The data coverage in the merged data set is about 25% higher compared to the original data set.

Following Joelsson et al. (2023), the data will be homogenised with the automated version of HOMER. Monthly precipitation data set of the Irish meteorological service (Met Éireann) has previously been homogenised using HOMER (Coll et al. 2018).

Additional preliminary will be presented.

Joelsson, L. M. T., Engström, E., and Kjellström, E. (2023). Homogenization of Swedish mean monthly temperature series 1860–2021. International Journal of Climatology. 43(2), 1079–1093 https://doi.org/10.1002/joc.7881

Coll, J., Curley, M., Walsh, S., and Sweeney, J. (2018). HOMERUN: relative homogenisation of the Irish precipitation network. Wexford: EPA. Report, (242), 43.

ONLINE

VALIDATION AND HOMOGENIZATION OF HYBRID TEMPERATURE AND PRECIPITATION SERIES IN MAINLAND PORTUGAL

Vanda Pires, Paula Drumond, Teresa Ferreira, José Godinho, João Ferreira
Portuguese Institute for Sea and Atmosphere, I.P
vanda.cabrinha@ipma.pt

Knowledge of the climate is fundamental in the planning and management of socio-economic activities and essential to mitigate the consequences and the risks associated with climate change.

The analysis of climate data provides support for the evaluation of historical data and for decision-making, especially in studies on climate change. Ensuring its quality is essential. The data collection process, through meteorological stations, can present problems, where inconsistencies can occur.

The Portuguese Sea and Atmosphere Institute, I.P (IPMA) has meteorological data series, whose first observations date back to 1865. Maintaining their quality and homogeneity is fundamental.

For the calculation of the 1981-2010 and 1991-2020 climate normal, a rigorous treatment of temperature and precipitation climate data series from the IPMA network was carried out, using validation and homogenization processes.

For each meteorological station, series of daily air temperature and precipitation data were considered for the period of 1980 to 2020. The first phase of the process consisted of building daily data series obtained from various types of records: classic meteorological stations, automatic weather stations (AWS), recording instruments and values obtained from SYNOP.

A validation of the in situ observation data was carried out and the periods of missing data were identified. The missing daily values were complemented with data from the ERA5 model (reanalysis dataset of the ECMWF model) available on the Copernicus Climate Change Service platform, through the Climate Data Store (CDS) service. After filling in the gaps, a homogenization process was performed using the RClimDex package software.

The IPMA observation network has changed considerably during the last 20–30 years, introducing non-climatic changes such as automation and relocation. Homogenization was therefore necessary to provide a consistent basis for the new normal.

In a first stage, the homogenization process was made without filling the gaps with data from ERA 5 model and, in a second stage, it was carried out with the introduction of that data. Generally, the introduction of model data didn't originate significant variations in the homogenization. Some breakpoints were identified, having been validated or rejected (significant variations of data values) according to the metadata of the series.

After this process and considering hybrid series of temperature and precipitation climate data, the 1981–2010 and 1991–2020 climate normal were calculated and will be used not only in climate monitoring, but also in long climate analysis studies.

Comparison of homogenization methods for global radiation data in Austria

Vanessa Seitner(1), Barbara Chimani(1), Leopold Haimberger(2)

(1) Department Climate-Impact-Research, GeoSphere Austria, Vienna, Austria

(2) Department of Meteorology and Geophysics, University of Vienna, Vienna, Austria

vanessa.seitner@geosphere.at

In contrast to other meteorological parameters like temperature or precipitation, less attention was given to the homogenization of radiation data. Radiation processes are of great interest for both natural and anthropogenic contributions to the climate system and its changes. The global solar radiation as that part of the solar radiation that hits the Earth's surface is the basic source of energy of the climate system.

Less attention was paid to the homogenization of radiation data, because other parameters (temperature, precipitation) are more important for most applications. For global solar radiation there are long measurement time series in Austria, going back to 1983. To use them for detection of recent climate change and subsequent identification of atmospheric drivers, these time series have to be homogenized first. Possible sources for inhomogenities in the time series are for example relocations of stations, changes in calibration or tilt angle, as well as instrument replacements.

There are many methods for the detection of inhomogeneities in time series. Different homogenization methods (SNHT, ACMANT and python ruptures) have been applied to selected stations. The detected breaks of all methods are compared with each other and with metadata, in order to decide which breaks need to corrected.

The next step will be the application and test of correction methods, which will be rechecked to make sure that the correction of the break points has led to a homogenous time series and the analyses of long-term developments like trends. The method which shows the best agreement between the detected and real breaks will be used in the future to homogenize the Austrian global radiation data.

The poster will provide information on the Austrian long-term radiation network and illustrate the results of the different break detection methods.

Comparison of multi-satellite GPM IMERG and ERA5 reanalyses, cross-evaluation with Ordinary Cokriging and ground-based data in South West Africa

Maura Lousada, Carlos Pereira, Rui Cavaleiro, José Godinho, Tânia Cota, Ricardo Deus

Portuguese Institute for the Ocean and Atmosphere (IPMA)

maura.lousada@gmail.com

In recent decades, southern Africa has become increasingly vulnerable to weather variability, with large areas affected by both dry season droughts and wet season floods. In Angola, more than 1 million people were affected by these droughts in 2016, with an estimated economic impact of more than US$749 million. In the 2009 floods, which also affected the borders with Namibia and Zambia, 30000 people were left homeless in the same country. Recently, from December 2022 to January 2023, heavy rains again caused severe flooding in various regions in southern Angola and in the bordering areas of Namibia.

Given the sparse and mostly uncoordinated weather monitoring networks and the fact that most countries' economies rely on rudimentary rainfed agriculture systems, access to reliable data is of paramount importance.

This work is integrated into an ongoing project funded by the European Union (EU), FRESAN (Strengthening Resilience and Food Security in Angola). Preliminary results of the adequacy from IMERG/GPM-era data, and ERA5 reanalysis datasets for a region that spans from central Angola to the Namibia border are presented. The accuracy, of hourly and daily precipitation for different seasons, from both, IMERG and ERA5 were evaluated trough a cross validation of stochastic estimates with weather stations from (SASSCAL) The Southern African Science Service Centre for Climate Change and Adaptive Land Management. The same set of weather stations was used as primary data in different ordinary cokriging models, over an area of circa 600 000 Km^2 . Grids derived from IMRG and ERA5 were used as secondary data for different models, both, as secondary data available at all locations where the primary variable is to be estimated. The leave out one cross validation provided an estimate of the most suitable model.

Climate services at the Hungarian Meteorological Service based on gridded data series

Zita Bihari, Kinga Bokros, Beatrix Izsák, Attila Kovács, Mónika Lakatos, Annamária Marton, Zsófia Molnár, Olivér Szentes, Bernadett Szolnoki-Tótiván
Hungarian Meteorological Service
bihari.z@met.hu

At the Unit of Climatology in OMSZ we update the Hungarian homogenized and interpolated climate data series at the beginning of each year (HuClim data). With the MASH and MISH software, we produce the following daily data sets:

- maximum minimum and mean temperature (1901-)
- precipitation (1901-)
- air pressure (1961-)
- relative humidity (1951-)
- wind speed (mean and max) and direction (1997-)
- global radiation (2000-)

A number of climate services are based on these data series, as well as on the raw, but completed not homogenized but interpolated data of the current year:

- Freely accessible Meteorological Database (odp.met.hu)
- General climate analyzes on the OMSZ website
- Climate reports
- The KlimAdat geo-information system supporting climate change impact studies
- Agrometeorological analyses: soil moisture, drought and growing degree days maps
- Meteorological support of the Hungarian Agricultural Risk Management System
- Data provision for the ESA Danube Data Cube project
- Individual services

THE HISTORICAL PHENOLOGICAL DATA MIGRATION AT DHMZ

Helena Lebo Andreis
Croatian Meteorological and Hydrological Service
helena.lebo@cirus.dhz.hr

At DHMZ, we have a long series of phenological data, starting from the 1950s. The format of these data has changed through the past. Several years ago we launched a new phenological relational database. It was rather easy to store the data from 1995 to present. However, it was a challenge to migrate the historical data from more than 100 stations, due to their different format and storage system. The goal of this presentation is to show progress, challenges and results of this process.

USE OF HOMOGENIZED DATA ON SPRING PHENOLOGY PHASES OF OAKS AND LINDEN

Peter Fleischer (2,3,5), Peter Kajaba (1), Martin Kubov (2,3), Jozef Rozkošný (1), Jaroslava Slavková (1,4)

(1) Slovak Hydrometeorological Institute

(2) Technical University in Zvolen, Faculty of Forestry, Department of Integrated Forest and Landscape Protection

(3) Slovak Academy of Sciences, Institute of Forest Ecology

(4) Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Astronomy, Physics of the Earth, and Meteorology

(5) Administration of Tatra National Park

Jaroslava.Slavkova@shmu.sk;

Jozef.Rozkosny@shmu.sk

The homogeneity is important for all climatic variables, in all spatial scales, and for the data of all kinds of temporal resolution. There are several methods and software for the homogenization of climate data series. For the purposes of the article, we used the MASH (T.Szentimrey) program to homogenize the data. This program enables homogenizing monthly as well as daily data series and also fulfils limited data gaps. The measurements of temperature from selected 34 climatological stations were homogenized with this software, which required adequate software for the automatic implementation of the needed steps.

We used homogenized climate data to the analysis of the relationship between the leaf unfolding of two deciduous species (Quercus sp. and Tilia sp.) and temperature over the period of 25 years (1996–2020) in Slovakia. Specifically, we calculated the cumulated positive average monthly air temperature (CPAMAT III.-IV.) for a period of two months preceding the onset of the selected phenological phase (Braslavská et Borsányi, 1996). The results documented interannual variability in the dating of phenological phases within the species, while the differences among the species were also revealed. Significant correlations ($p<0.05$) were detected between the dating of leaf unfolding and air temperature. The trend analysis showed that the onset of leaf unfolding was slightly shifted to the earlier dates during the period of 25 years. However, not all trends were statistically significant.

USING ERA5-LAND HOURLY - ECMWF CLIMATE REANALYSIS IN CONTROLLING GROUND MEASUREMENTS

Vladimir Ribičić
Croatian Meteorological and Hydrological Service
vribicic@cirus.dhz.hr

Using ERA5-Land Hourly - ECMWF Climate Reanalysis in controlling ground measurements is the focus of this poster. The poster compares the advantages and disadvantages of ground and reanalysis data and explores their correlation. Reanalysis data was used as a way to overcome the challenges associated with using raw satellite data and insufficient ground data. Google Earth Engine was utilized to access the reanalysis data, and the data was then used in a desktop app in an attempt to visually control ground measurements such as temperature, precipitation amount, and atmospheric pressure on Croatian weather stations.

The reason for this type of analysis is the fact that while ground-based measurements provide mostly accurate and high-resolution data, they have limited spatial coverage. Graphical analysis and comparison are commonly used methods for quality control of ground data, but they have limitations since neighboring stations may be significantly far away or located in a different environment. On the other hand, satellite data provides global coverage and high spatial resolution, but may be less accurate in certain regions or under certain weather conditions. Therefore, comparison of ground data with reanalysis data can be a viable solution to overcome ground data limitations, but the user cannot take reanalysis data for granted and additional comparison to other data sources is also recommended. The study utilized Python as a tool to retrieve, process, store and plot data in a database for later analysis.