The EuMetNet Data Rescue portal

Ingeborg Auer (ZAMG), Barbara Chimani (ZAMG),
and members of the EUMETNET Expert Team on data rescue*)

Introduction

Long time series are essential for climate research. They allow for analyses of former climate, calculating reanalyses and model evaluations. Time series reaching back to the end of the 18th / begin of the 19th century are of great importance as connection between the instrumental period and paleoclimatology. For the best possible resolution in time and space it is essential to use all the available data. For reliable information on climate variability and extreme events long time series are essential.
But for analyses with high spatial resolution for Europe the available data is only sufficient since about 1960. Although different "Data Rescue Programs" (e.g. http://www.climatol.eu/DARE) have been exercised , there is still a lot of data stored in paper-based archives and therefore not available for climate research but exposed to destruction and oblivion. The poster summarizes the results of a EUMETNET Data Rescue Workshop held in Budapest during the 8th Seminar for Homogenization and Quality Control In Climatological Databases in May 2014 (http://www.met.hu/en/omsz/rendezvenyek/homogenization_and_interpolation/programme/).

Aims:

EUMETNET-Dare is complementing world wide running programs and does not intend to duplicate existing work. Therefore the following milestones were defined for EUMETNET-DARE:
1) data inventory of already digitised data and data still only available in paper archives
2) enhancing the data on digitalised data sets

For a follow-up project the following tasks are planned:
3) homogenisation of the datasets
4) publishing the data in relevant data bases

Activities:

To fulfil task 1 of the current project a questionnaire was sent to all the meteorological services of the WMO region VI. (Fig.1). The questionnaire contained questions on the digitised and non-digitised data with a length of ~100 years (~50years for mountain stations), but also questions concerning the time and amount of money needed for digitising the data and the willingness to provide the data to a international database.
Additionally to the information given by the meteorological services the information displayed on the MEDARE-webpage (http://www.omm.urv.cat/MEDARE/) were included into the available information sheets on the EUMETNET-DARE webpage (https://www.zamg.ac.at/dare/)
Detailed information on the available data can be found in Fig 1. As indicated not all of the countries have returned the questionnaire and not for all of the countries information can be found in the MEDARE-Database, as this project was focusing on the Mediterranean area.
An example of the information given on this website is shown in Fig.2.

The spread in the progress in regard to data rescue in the different countries is quite big. While some of the countries have already digitised all or most of their data and are keeping them in their own database or even make them available via their webpage, other countries are lacking the capacity to start a data rescue progress at the moment.
For some of the participating countries their progress in data rescue is displayed on this poster.

Problems:

- Financial support for data rescue activities is hard to find
- Data rescue is seen as important but can only been done in small amount of spare time between other important operational activities.
- Free distribution of the data can pose a problem for some of the participating countries, due to legal reasons and/or insufficient financial support by the government
- Uncertainties where data of some stations, that belonged to different countries in course of time, is now stored and if it still exists
- Internal structure of meteorological services in different countries (distribution of responsibilities,...)

Austria:
Nearly half of the Austrian data is already digitised and fully quality controlled. But still about 20 % of the data has to be digitised. The missing about 31% of data are during different states of quality control.

If you know about some long-term time series available in your country or stored in your archive and you can't find this information on the DARE-Homepage, please let us know about it (dare@zamg.ac.at).
If you would like to fill in the questionnaire, please feel free to download it from the EUMETNET-DARE-Homepage.

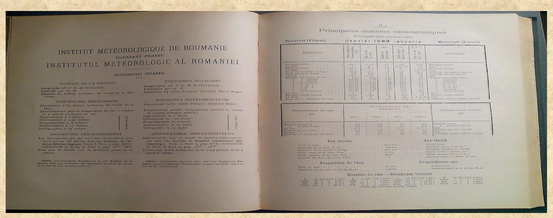
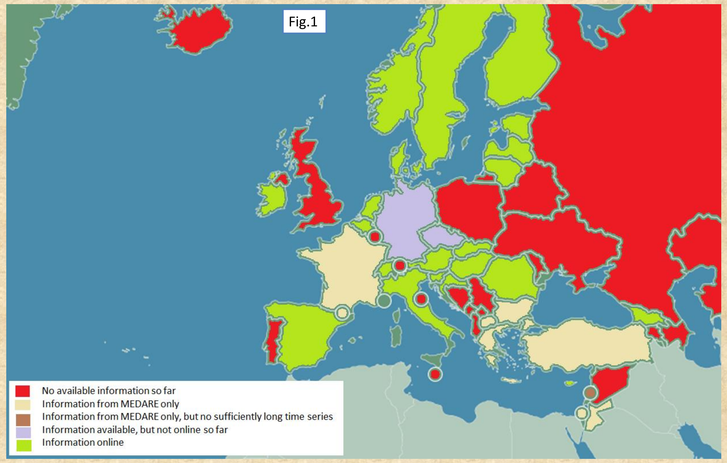
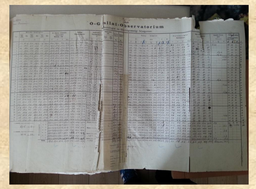


Fig.1

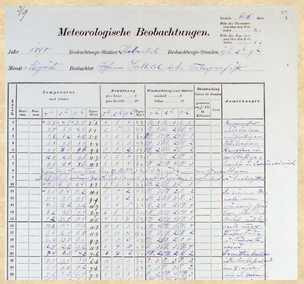




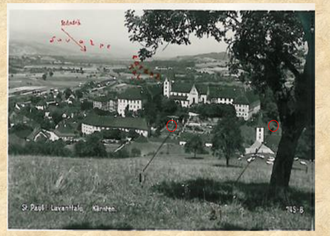
Italy:
Two organisations collect meteorological/climatological data since the second half of the 19th/ the first half of the 20th century on a national level (including parts of time series from former Yugoslavia and northern Africa). The data of 20 stations reach back more than one century. All the Italian data back to 1950 is already digitised. Between the two organisations working on national level a cooperation is going on leading to a common database of data and metadata and exchange of experience concerning data rescue. Taking into account the whole paper archive and all parameters and resolutions stored thereabout 35% of the data have already been digitised. For Florence and Naples hourly data is digitised back to 1925. As not all of the additional data given on the data sheets can be included into the database an effort is done for scanning all of them.

Slovakia:
This initiative includes rescue of metadata and data, a graphical data overview for time series duration and quantification of unknown data history on daily and hourly measurements. Quality control is done with a "Graphical interface for controlling and revision of climatological data". Data series are prolonged combining the data from different stations locations (e.g. caused by relocation).
Still the bigger part of the historical data is in paper form, but first attempts using a data recognition software have been done.

Hungary:
Majority of the data from climate stations from 1961 is already digitized. Earlier measurements are stored on microfilms, microfiches and in paper format. 15 long-term stations for temperature and additional more than 100 for precipitation (from 1901 and 1 mountainous station from 1952) are digitised, homogenized, completed and quality controlled. The digitization is a permanent process besides the registering of precipitation on regular basis. Data rescue took place partly within different projects, recently about 1 522 780 records of sunshine duration and climate data were digitized in the CarpatClim-Project.



Czech Republic:
Data rescue is done since 10 years with different progress in different regions. All data since 1961 is already digitised. Estimated about 95% of the long-term series have already been completely digitised. But data quality control is still ongoing and will take several years.
Problems concerning data distribution might be solved with a new law perhaps till 2020 .



Great Britain:
Data rescue is still at it's beginning and the different methods of automated versus manual are weighed. Additionally a catalouging project for getting an overview over the available data is done. The main effort is in retrieving enough digitised data so create gridded monthly datasets for the area.

Romania:
Data and metadata are addressed in the effort of data rescue. Sources are the national meteorological archives but also other documents from the library.
The data is archived in a web based (internal) database and the digitalisation of the meta data is also done via a web interface. For some of the stations , for which it is known that long time series should exists, it was not possible to recover the whole datasets.

Finland:
Rescue of meta data is an important issue. Nearly all of the data since 1960 is already digitised and SYNOP and CLIMATE data since the beginning of the time series. Digitising of data of the precipitation stations into a temporary database is underway (if digitising progress goes on with the same speed as until now this will take more than 20 years from now).
Quality control is done by a manual control during the digitizing process and will undergo automatic quality control afterwards.
There is so far no intention of digitising hourly data from aviation and co-operated stations.

Fig. 2

name	temporal resolution	temperature	start	end	data gaps (T) (years)	fillable (T)	humidity	start	end	data gaps (humidity) (years)	cloudiness	start	end	data gaps (cloudiness) (years)	precipitation	start	end	data gaps (precipitation) (years)	wind velocity	start	end	data gaps (wind velocity) (years)	wind direction	start	end	data gaps (wind direction) (years)	air pressure	start	end	data gaps (air pressure) (years)	sunshine	start	end	data gaps (sunshine) (years)	fillable (sunshine)	snow	start	end	data gaps (snow) (years)	mountain
Innsbruck-Univ.	daily, monthly	X	1877	active	yes	no	X	1882	active	yes	X	1877	active	yes	X	1877	active	yes	X	1877	active	yes	X	1877	active	yes	X	1877	active	yes	X	1922	active	yes	yes	X	1916	active	yes	
Patscherkofel	daily, monthly	X	1940	active	no	no	X	1940	active	no	X	1940	2010	no	X	1903	active	yes	X	1940	active	no	X	1940	active	no	X	1946	active	yes	X	1961	active	no	yes	X	1919	active	yes	x
Salzburg	daily, monthly	X	1874	active	yes	no	X	1874	active	yes	X	1874	active	yes	X	1874	active	yes	X	1880	active	yes	X	1874	active	yes	X	1874	active	yes	X	1940	active	no	yes	X	1936	active	yes	
Sonnblick	daily, monthly	X	1886	active	no		X	1886	active	no	X	1886	active	no	X	1890	active	no	X	1886	active	no	X	1886	active	no	X	1886	active	no	X	1886	active	no		X	1938	active	no	x
Feuerkogel	daily, monthly	X	1930	active	no		X	1930	active	no	X	1930	active	no	X	1930	active	no	X	1930	active	no	X	1930	active	no	X	1933	active	no	X	1930	active	no		X	1930	active	no	x
Kremsmünster	daily, monthly	X	1876	active	no		X	1874	active	no	X	1874	active	no	X	1874	active	no	X	1874	active	no	X	1874	active	no	X	1874	active	no	X	1887	active	no	yes	X	1917	active	no	
Wien-Hohe Warte	daily, monthly	X	1872	active	no		X	1872	active	no	X	1872	active	no	X	1872	active	no	X	1872	active	no	X	1872	active	no	X	1872	active	no	X	1880	active	no	yes	X	1916	active	yes	
Graz-Univ.	daily, monthly	X	1894	active	no		X	1894	active	no	X	1894	active	no	X	1894	active	no	X	1894	active	no	X	1894	active	no	X	1894	active	no	X	1911	active	yes	yes	X	1896	active	yes	
Hochobir	daily, monthly	X	1881	1944	no		X	1880	1944	no	X	1880	1944	no	X	1880	1944	no	X	1880	1944	no	X	1880	1944	no	X	1880	1944	no	X	1883	1944	yes	yes	X	1924	1944	no	x
Kanzelhöhe	daily, monthly	X	1938	active	yes	partly	X	1938	active	yes	X	1938	active	yes	X	1938	active	yes	X	1938	active	yes	X	1938	active	yes	X	1938	active	yes	X	1938	active	yes	yes	X	1938	active	yes	x
Villacher Alpe	daily, monthly	X	1921	active	yes	no	X	1921	active	yes	X	1921	active	yes	X	1921	active	yes	X	1921	active	yes	X	1921	active	yes	X	1933	active	no	X	1933	active	yes	yes	X	1925	active	no	x
Feldkirch	daily, monthly	X	1936	active	no	no	X	1936	active	yes	X	1936	active	no	X	1895	active	yes	X	1936	active	no	X	1936	active	no	X	1936	active	no	X	1958	active	no	yes	X	1895	active	yes	
Langen am Arlberg	daily, monthly	X	1952	active	no	no	X	1952	active	no	X	1952	active	no	X	1895	active	yes	X	1952	active	yes	X	1952	active	no	X	2007	active	no	X	2007	active	no	yes	X	1895	active	yes	
Schröcken	daily, monthly	X	1948	active	yes	no	X	1947	active	yes	X	1947	active	no	X	1947	active	no	X	1947	active	no	X	1947	active	no	X	2007	active	no	X	2007	active	no	yes	X	1895	active	yes	
Galtür	daily, monthly	X	1956	active	no	no	X	1938	active	yes	X	1938	active	no	X	1938	active	no	X	1938	active	no	X	1938	active	no	X	1996	active	no	X	1996	active	no	yes	X	1901	active	yes	x
Kufstein	daily, monthly	X	1937	active	yes		X	1946	active	no	X	1936	active	yes	X	1905	active	yes	X	1936	active	yes	X	1936	active	yes	X	1963	active	no	X	1992	active	yes	yes	X	1923	active	yes	
Landeck	daily, monthly	X	1946	active	no		X	1946	active	no	X	1946	active	no	X	1895	active	no	X	1946	active	no	X	1946	active	no	X	1986	active	no	X	1986	active	no	yes	X	1897	active	no	

Spain:
All data from about 1961 are already digitised. Scanning of the longest manuscript data reports (~1860-1960) and digitalisation of monthly precipitation and temperature data for this period is also finished. Current efforts, limited by staff shortage, include imaging of radiosonde data from before 1970 and paper strips records as well as the digitalisation of additionally found data.
The next open tasks will be imaging and cataloguing all other data documents and analogue recordings, and their digitalisation.

*) Members of the EUMETNET Expert Team on Data Rescue: Anita Paul (Austria), Miroslav Řepka (Czech Republic), Anna Frey (Finland), M. Lakatos (Hungary), Seamus Walsh, Mary Curley, Mairead Treanor (Irland), Maria Carmen Beltrano, Gianpaolo Mordacchini (Italy), Ancuta Manea, Elena Mateescu (Romania), Oliver Bochníček (Slovakia), José A. Guijarro (Spain), Tim Legg, M. McCarthy (United Kingdom),

Contact Information:
Ingeborg Auer
Abteilung Klimaforschung, Zentralanstalt für Meteorologie und Geodynamik
ingeborg.auer@zamg.ac.at, Tel. 0043 1 360 26 / 2202, www.zamg.ac.at

Acknowledgement:
This project is funded by EUMETNET