Experience of Climate Data rescue at Météo-France

Sylvie Jourdain
Météo-France, Direction de la Climatologie et des Services Climatiques, France

sylvie.jourdain@meteo.fr

Sea Level Data Archaeology Workshop, 11 mars 2020
Outlines

• **Context**
  - Météo-France
  - French climate archives

• **Climate Data rescue at Météo-France**
  - Material of Climate Data Rescue
  - Preserving, organizing and cataloguing
  - Original sources imaging
  - Data and metadata digitizing
  - Most important difficulties
  - Strong points of Climate Data rescue

• **References**
Historical weather observations are vital to climate research.

It is only through looking at long-term trends and comparing historic climate with the present day that scientists can build a bigger picture of what is happening across the planet.

**France has a rich heritage of surface and upper-air meteorological observations.**

But a wealth of old climate records still only exist in original form, such as paper, microfiche or microfilm.

**Météo-France holds data either in digital form or in original paper form.**

Météo-France's climate database contains French digital data (subdaily and daily instrumental data) back to 1688.

But most of observations from the 19th century and early 20th century exist only in hard-copy forms and need to be recovered.

**Fast all the data post-1960 are stored in the climate database.**
Context: French Climate Archives

Météo-France, as National Meteorological Service (NMS), is in charge of preserving the memory of the French climate.

Météo-France's holds historical meteorological records collected by the successive French national meteorological services (7.5 linear kilometers). These public and historical records are stored at 90 locations in France mainland and French overseas territories.

Meteorological records start from 1780 for France and from 1920 for French former colonies.

French National Archives (FRAN)

In addition, from 1971 to 1991, the previous French NMS, transferred 2 linear kilometers of physical meteorological records to the French National Archives (Archives nationales).


About holdings of historical climate data at French archives: Brunet et al. (2014)
Climate Data Rescue (DARE)

DARE is an ongoing process and involves rescuing the data and metadata. Metadata are of crucial importance to interpreting measurements and observations.

All the entire process of DARE is a long chain of actions:

- Archiving, searching and locating climate records (microfiche or paper media)
- Imaging documents, storaging of image files on the Météo-France storage server and on disks
- Digitizing (keying) metadata and data values
- Quality-controlling digitized data and metadata and data already in the database
- Integrating digitized data and metadata into the French National Climatological Database BDCLIM

National coordination is made by the Direction of Climatology and Climate Services
Material of Climate Data Rescue

Original sources of data and metadata
Instrumental observations were usually recorded on various meteorological forms depending on responsible organization, network, type of the station and date.

Big challenge for data recovering due to the wide variety of types of form sheet with different parameters, units, instructions, etc.

Example on the right
Weather and meteorological data on sheet
Observed Parameters: visibility, wind direction, wind strength, cloudiness
Measured Parameters: daily rainfall and snow depth
Metadata available on sheet
Producer, Title of the source, Departement, Date, Station, Hour of observations (reference ?), Form reference, Name of observer

Digital exhibition on French meteorological sources:
http://archivesduclimat.meteofrance.fr

Météo-France
Material of Climate Data Rescue

Stripchart for temperature, humidity, wind, rain, sunshine and pressure since the end of the 19th century
Unfortunately, a vast collection of diagrams has been destroyed in the 1990's

Important metadata are available on the stripchart
Responsible organization, Title of the source, Departement, Date, Station, Dimensions of instrument (pluviograph) and Reference of stripchart

Pluviogram Bastia-Poretta (Corsica) 18-19 August 1957
Material of Climate Data Rescue

Original Source of metadata
Instructions for observer
Observing practices depend on the network and are very important

Example on the right
Observations should be made every morning at 7h, winter time

Rain: value measured at 7h (total from the previous day 7h). This value will be attributed to the previous day in the database.

Wind direction: 3 letters (windrose in 16 directions)

Wind strength scale: 0 to 6

Monthly precipitation: Calculation without the value of the first day. The value of the first line should be attributed to the last day of the previous month. (Source of errors)
Material of Climate Data Rescue

Original Source of metadata
Information sheets about the station with geographical localization, local environment, instruments etc.

Unfortunately, information sheets about the station and meteorological records are rarely stored at the same repository.

FRAN stores 2 linear km of Meteorological records without any information sheet about the station

French climate database stores metadata and allows digitization of metadata

Difficulty: hand-written and textual metadata

Météo-France
As recovering data is time-consuming, prioritization is important. Priority to the GCOS Essential Climate Variables (ECVs) and to numerical information

- daily temperature and precipitation
- subdaily surface humidity, pressure, temperature, wind speed and wind direction and cloudiness
- Upper air humidity, temperature, wind speed and direction
- Snow depth

The Lightning ECV has not been yet recovered. Inventorying of Thunderstorm sheets is in process. Recovery of lightning in historical landforms is particularly difficult: textual information.
Preserving, organizing and cataloguing

**Preserving** Meteorological records are unique historical sources, but they are continuously at risk of being damaged, lost or destroyed (endless battle because of lack of resources and space). They should be preserved.

**Organizing** Source should be organized in a logical manner, stored in acid-free archive boxes on sturdy shelves and inventoried.

**Météo-France's national program**

National program of preservation and inventorying Climate Archives has been managed and coordinated by an Archivist, expert in historical climate sources since 2014.

The same national template for inventorying sources and meteorological parameters is used by trained meteorologists and climatologists involved in this action in all of French territories.

Today 75% of the Météo-France's climate archives have been catalogued.

A record management software is used to manage Météo-France climate archives.

Digital catalogues are available on-line on the open platform for French public data [https://www.data.gouv.fr/en/](https://www.data.gouv.fr/en/) (key words: Météo-France)

**Difficulties**: Numerous repositories spread over the world, less and less human resources due to a new reorganization. Archives preservation is rarely the first priority!
Original sources imaging

DARE is a permanent and lengthy process that can be costly, labor-intensive and time-consuming. **It is necessary to prioritize the records to be imaged and digitized.**

**Advices:** All pages of original meteorological documents should be imaged to ensure that all available metadata are rescued. After historical records are imaged, the original documents must not be thrown away because digital images are also not a permanent archive.

**Access to the images**

Météo-France has an active digital program for climate documents since 2007, funded by Météo-France and several national and international projects.

Digital Climate publications (bulletins, yearbooks, etc.) are for free available on the Météo-France's digital library [http://bibliothque.meteo.fr](http://bibliothque.meteo.fr)

Regarding historical meteorological records, a public portal is being developed by Météo-France in order to provide the catalogue of all the meteorological records and selected collections in digital form (images) on-line. This portal will facilitate locating material held in the archives and will give access to hundreds of thousands pages of meteorological sheets. This portal should be published in few months.
There are so many French weather observations to be recovered!

The priority of data to be digitized is based on scientific considerations: lengthen, complete or create long-term series, contribute to research projects.

**Remarks and Advices**

Recovering data involves data keying and sources imaging for a full traceability.

The layout of the sheets changes frequently and observers can do mistakes. Therefore, before documents are distributed for keying, they are inspected in depth by experts in historical data. Based on this work, the station is identified, parameters are selected, units and time reference are noticed and noted down.

Good knowledge in the history of the meteorological observations (network, station, instrument, operational procedure, etc.) is required to write specification and to collect metadata.

Providing good instructions to keyers is fundamental to avoid quality problems.

Technique for manual keying data from handwritten records: Double keying and transcription data exactly as they are printed “key as you see”
Historical Pluviograms Digitizing

National action  from 2009 to 2019

- Use of Météo-France’s tool *Digitalise* with Calcomp digitizing tablets for digitizing pluviographic records

  Advantage of this tool: direct comparison in real time with daily data from accumulation raingauge already available in the climate database, but the software is now obsolete.

- EXTRAFLO project funded by ANR 2008-2013: Digitization of daily pluviograms  Mixed results: the comparison with daily observations from accumulation rain gauges has highlighted significant biases concerning the intensity ranges, well correlated with the type of equipment used. Impossible to get good quality infra-hourly data and too many missing data during extrem events (Soubeyroux et al.2011)

  Quality-checks: comparison with values from accumulation raingauge (daily values, yearly values and extreme data)

  Results: 14 series have been integrated into the climate database.

- Action is now closed: Météo-France has decided not to develop a new tool in light of the time-consuming task for charts digitization and the potential usage of this kind of series
Most important difficulties

- The list of repositories that hold either weather and climate observations or original documents is very long.

A large amount of material is not yet listed on-line, and many archives hold relevant material that is inadequately catalogued or not catalogued at all.

Records of one long-term series are stored in several repositories and sometimes partially outside Météo-France. In France, the access to the original records at public archives is difficult.

- The wide variety of type of sources makes the work more difficult and longer.
- The wide variety of the instructions for observers is a source of mistakes.
- The data and metadata already in the climate database can be incorrect because the required level of sources examination and traceability have changed.

- All the steps of DARE are time-consuming, labour-intensive and require rigor and traceability. It can be difficult to motivate agents for this tedious work.
Strong Points of Météo-France’s Climate Data Rescue

- A scientific research agreement between Archives nationales and Météo-France has been signed related to the Climate Data Rescue to affirm a commitment to collaborate together to enhance the meteorological collection.

- National coordination is made by two experts at the Direction of Climatology and Climate Services in Toulouse.

- Choosing the appropriate personnel is critical to the success of a data rescue project. Several climatologists expert in DARE are involved in the national action to guarantee high quality of sources preservation and cataloguing, and high quality of recovery.

- WMO Guidelines and C3S Data Rescue Service have been incorporated into the structure and all the colleagues involved in DARE activities have been trained. Operating mode, technical guide and standard forms are available in order to guarantee the homogeneity and tracability of the process.

- Météo-France is involved in ACRE initiative, is partner of the C3S Data Rescue Service consortium and informs about DARE projects on I-DARE portal https://www.idare-portal.org/. Effective support from the international collaboration.
References

- Publications


Thank you for attention

Questions about French climate archives
Contact: archives.climat@meteo.fr