

DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

Oliver BOCHNÍČEK

Slovak hydrometeorological institute



Budapest, 13.5.2014

DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

The **aim** of the work is to:

- Demonstrate **steps, forego mathematical processing** (e.g. homogenization)
- Create **graphical form of historic and present data overview** to get immediate **access to all time series** duration **en bloc**
- **Quantify unknown data history** on daily and hourly measurements
- Demonstrate **forms of rmetadata registration**
- Describe **up to date possibilities of historic (paper) data QC**
- Point out **problems of future input** of the **historical** (and **present voluntary**) climatological data

DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

Presentation is a mosaic of items dealing with:

1. Station History
2. Data Flow
3. Data Input
4. QC
5. Metadata catalogue
6. Metadata SW
7. Station Relocation & possible prolongation example
8. Time Series (selected)
9. What's next

All these point are connected with one idea DARE.

DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

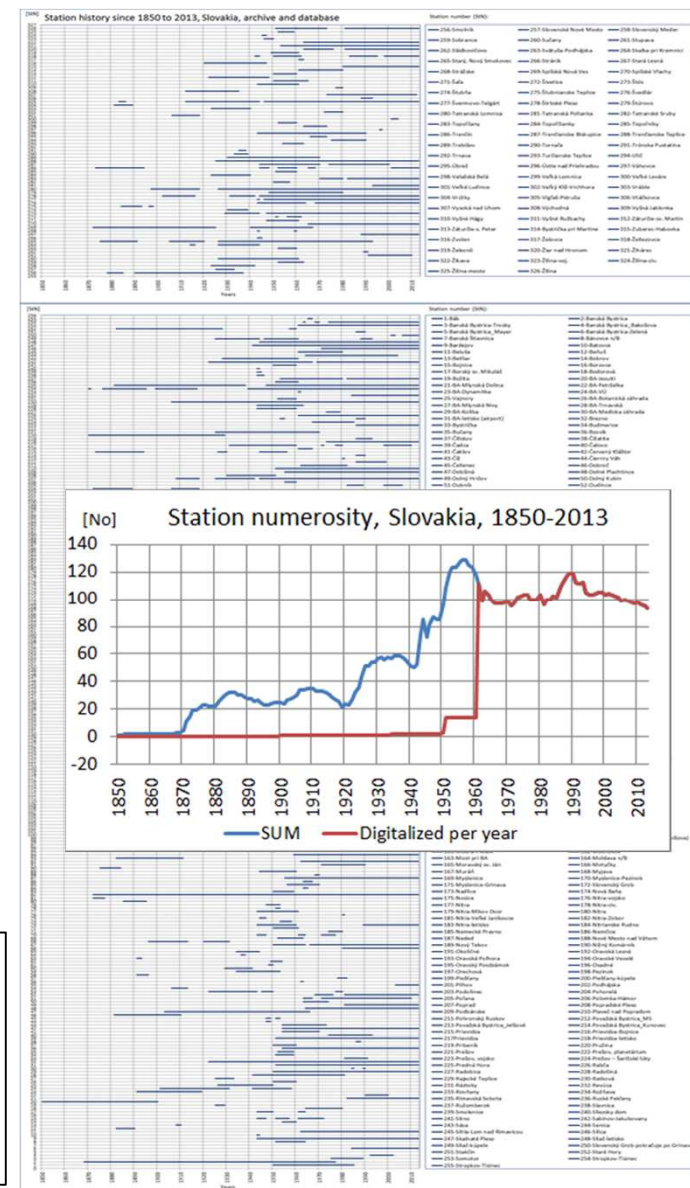
Object description:

1. Station History

- Archive catalogue is organized **by the name of the station**.
- In according with that (the past registration of the station and present name and indicative marking) we decided to prepare a **complex list of all stations** (with or without indicative) by the name of the station **in one line graph**. Station history visualization is appropriate especially for **finding intersection** of those stations (time series) at concrete period of interest that can be **useful for data gaps filling and prolongation** of corresponding time series as a first step of composition time series of homogenization process.

History of climatological station network in Slovakia:

- 326 climatological stations of various length
- Data availability from 9815 yrs.
- Only archive (paper form) data 4275 yrs.
- Digitalized data 5540 yrs.
- Ratio (only archive/all) = **unknown history = 44%**



DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

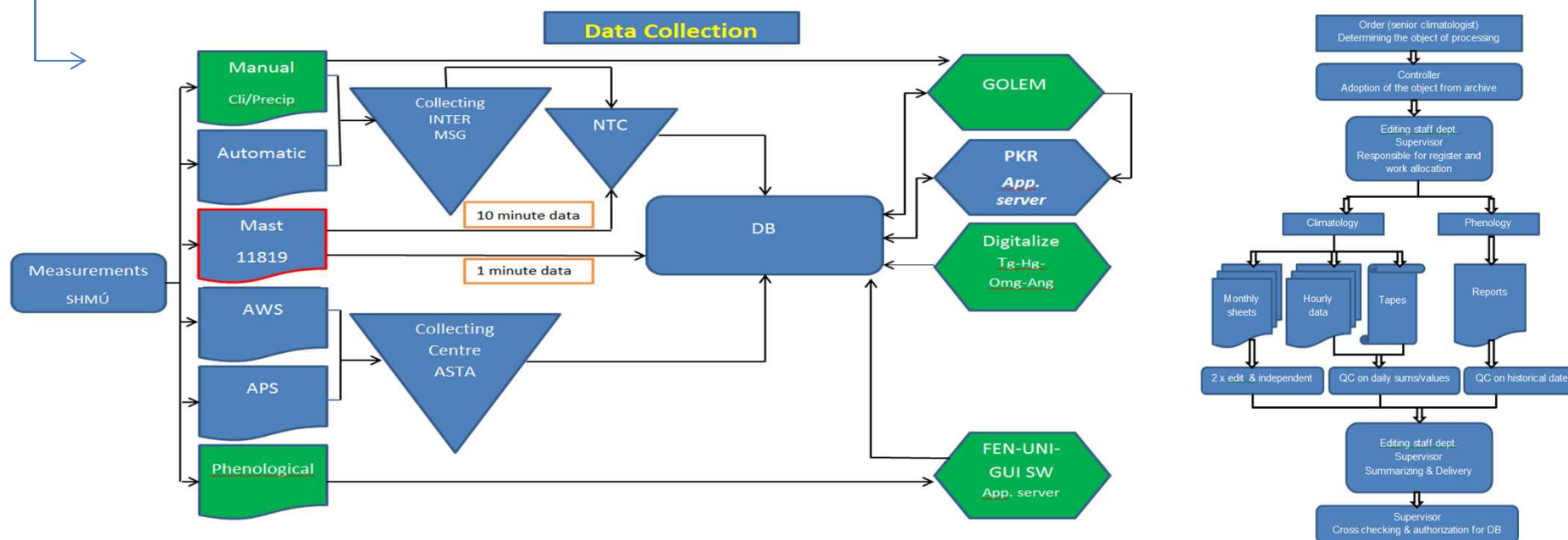
Object description:

2. Data Flow

- Data flow scheme is focused on data **source**, **frequency** and **object** of measurement, way of data **collecting**, **places** and **processes** that are executed and final data **placement**.

3. Data Input

- Data input scheme explains the **rule** and **competence** in data management from archive through senior climatologist, archive person, controller, edit staff, supervisor and database administrator.



DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

Object description:

4. QC

- QC is carried out by PKR SW as a “**Graphical interface for controlling and revision of climatological data**” useable for **present** and **historical** revision and control as well, because of the same way of input data into the DB (see Data Flow and Data Input scheme).
- based on **Flag Marking**
- **Flag** – means data quality result according to the results of particular control algorithms

5. Metadata catalogue

- Established 20 years ago
- Enable to **write down & search** metadata information in word document.
- Consider to be the first complex form of metadata register.
- Still in filling and accessible.

6. Metadata SW

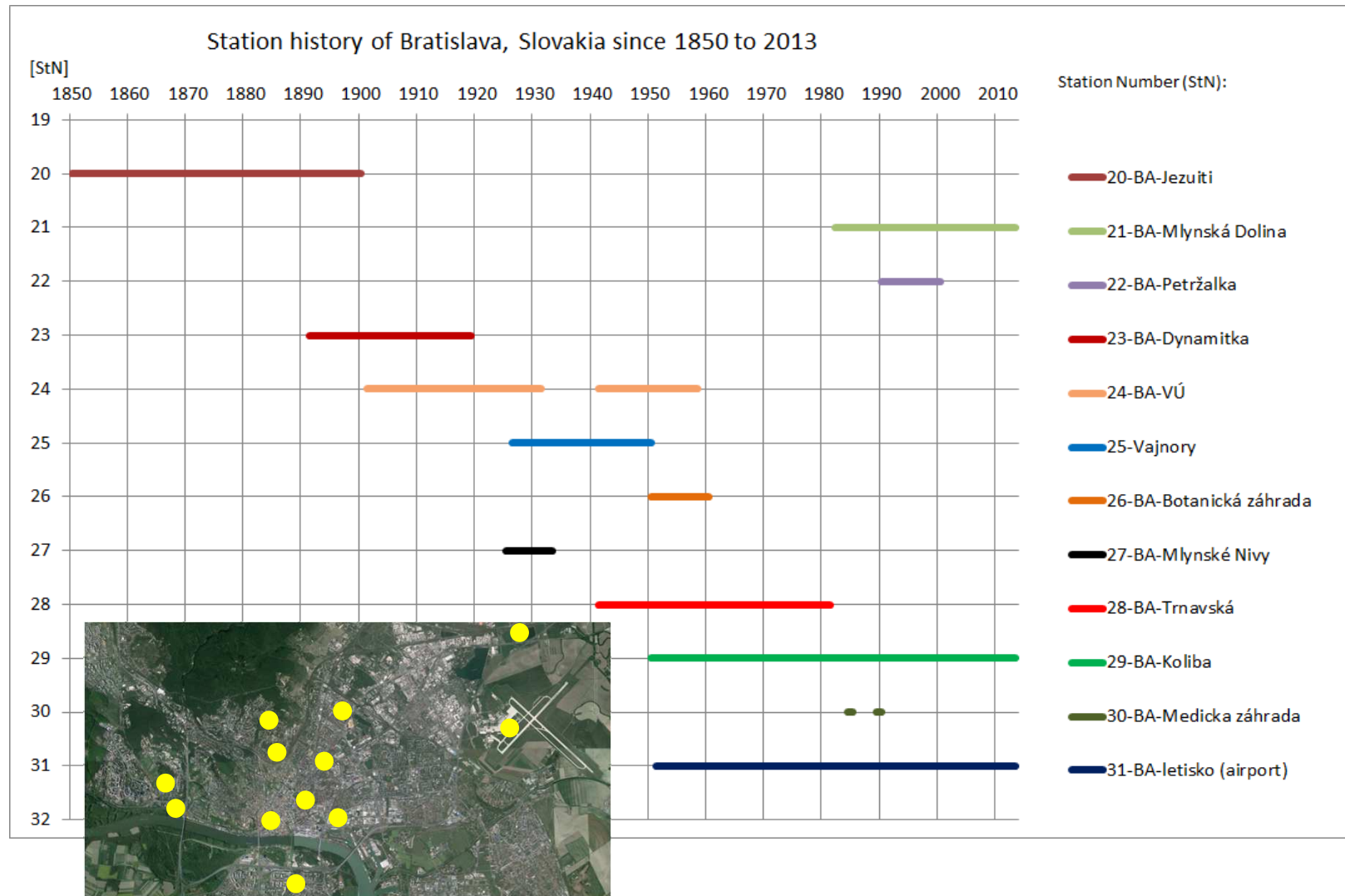
First **graphical interface of metadata register** more than 10 years in practice.

- Enable **input** important station information into DB KMIS
- Enable **search** metadata information from DB KMIS

DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

7. Station Relocation & possible prolongation example

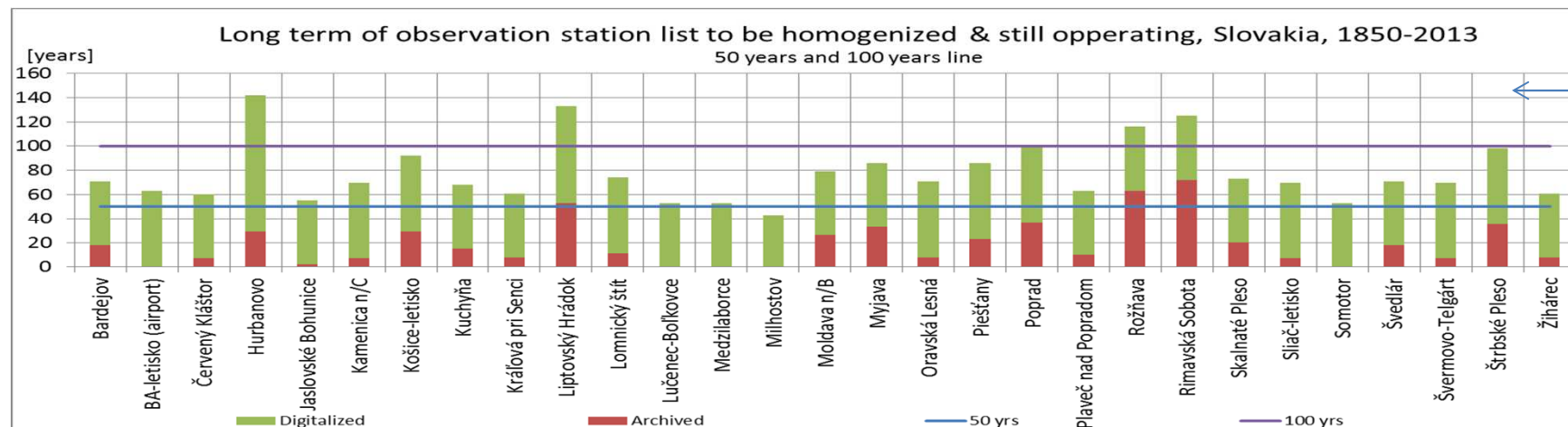
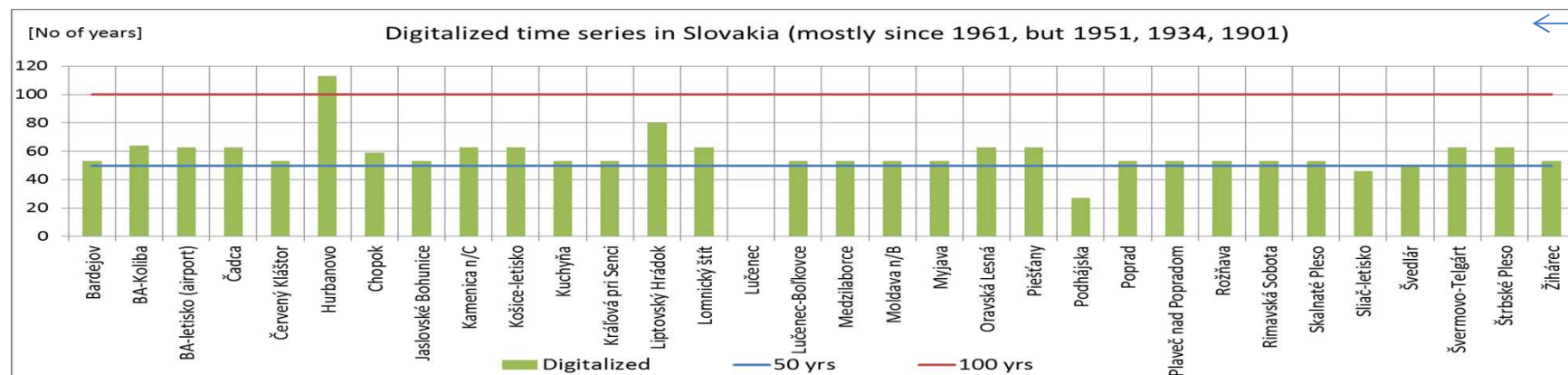
- Station relocation & possible prolongation is **based on the station history** [1] overview. Longest time series is explained via **part time series** of **Bratislava** which the oldest time series reaches up to **1850**.



DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

8. Time Series (selected)

- Aim of this item is to present result of “**EUMETNET Climate Program, Expert Team on Data Rescue and Recovery and Questionnaire - Data Rescue**”. Long-term series and mountain series digitized data that are available in DB are shown along with the table expression as well as data still to be digitized connected with continuation of present station network.

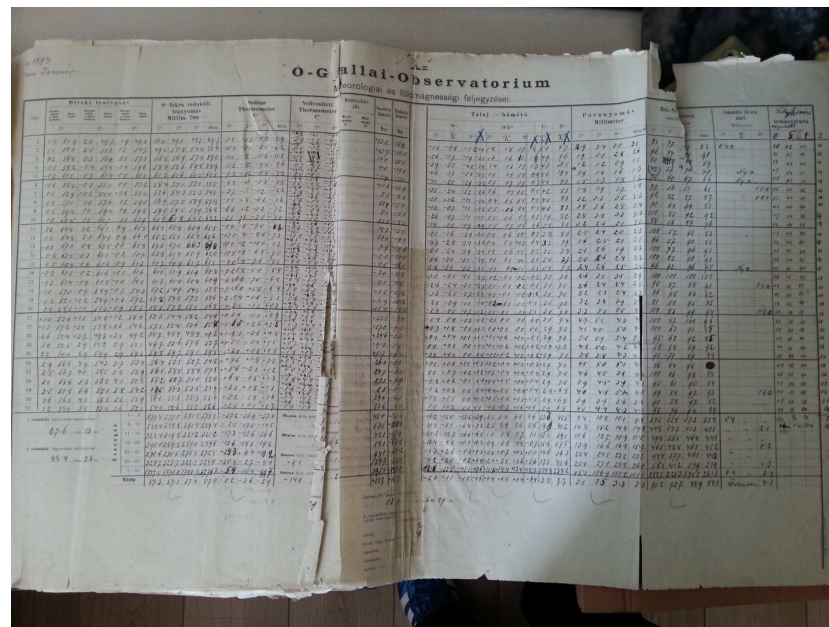


DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

9. What's next

Some remarks to the end are mentioned:

- **Historical data** play **important** role for recognition possible impacts of future weather.
- Most of our **data history** (because of paper form) is **unknown**.
- Not only daily or sub daily measurements but also **hourly data** are **important** for detecting possible mistakes (inhomogeneity) in daily or monthly evaluations especially extremes.
- Number of **professional staff capable to edit** historical measurements **goes down**. Students are not equivalent compensation.
- **Digitalization by recognition SW** (e.g. AbbyyFine Reader) from paper is quite **difficult** and long lasting especially due to frequently changing form, shape and angle of the hand made record. First attempts have been done.



The image shows a photograph of an open historical meteorological report from 1894, titled "O-Gallai-Observatorium". The pages are filled with handwritten data in a structured table format, showing various meteorological measurements over time. The report is written in Hungarian and includes columns for date, time, and various weather-related data points. The handwriting is in cursive, and the paper appears aged and slightly yellowed.

Fig. Monthly report, Hurbanovo (at present), 1894.

DARE-Technological sequence of (historical) data register (metadata), collection, transfer, digitalization, QC and homogenization in Slovakia

THANK YOU FOR YOUR ATTENTION

