Building a Metadata Global Registry for Recovered Meteorological Data

Maria Antónia Valente[1] (mavalente@fc.ul.pt), Thomas Cropper[1] (tecropper@fc.ul.pt),
Alexander Stickler[2] (alexander.stickler@giub.unibe.ch) and
Rob Allan[3] (rob.allan@metoffice.gov.uk)

[1]Instituto Dom Luiz – FFCUL , Lisbon – Portugal;
[2]Oeschger Centre – University of Bern, Switzerland; [3]Meteorological Office, Exeter, UK
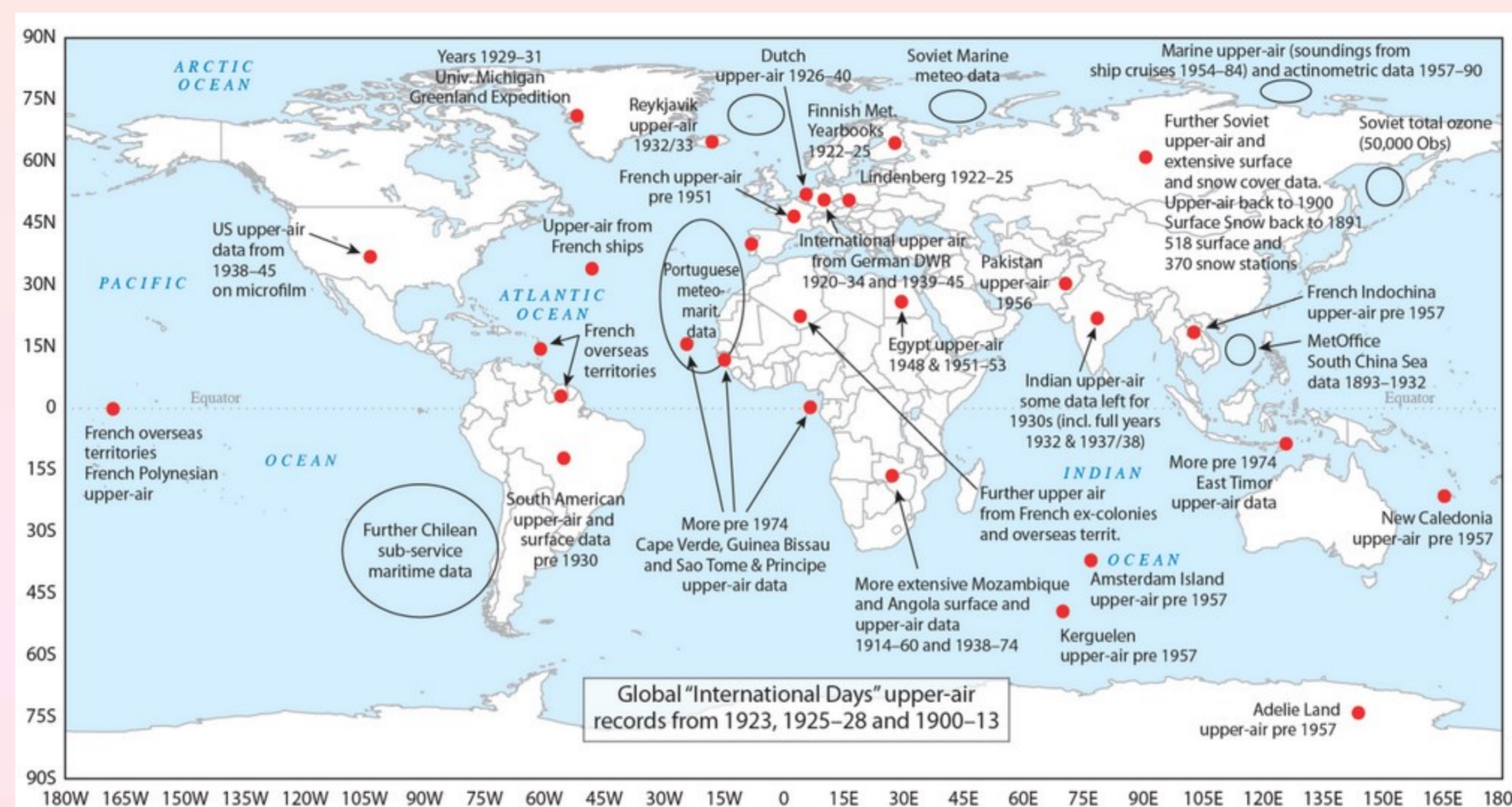
1 - Project ERA-CLIM2 is developing an interactive online Global Registry for Historical Meteorological Metadata information on data that has so far been recovered and is in the process of recovery. It includes separate inventories for surface, upper air and maritime metadata. The objective is to gather in one place the maximum possible number of data records information available to the science community at large. The tool aims to supply lists of datasets with its exact geographical location and time span, as well as the existing ECVs (Essential Climate Variables).

2 - The Global Registry comes in the sequence of the initially developed ERA-CLIM UBERN Metadatabase tool

3 - The UBERN Metadata Base tool contains inventories of surface, upper air and transmission records and will also incorporate ships logbooks in ERA-CLIM2 (Chile and Portugal at least). It is on-line at (www.oeschger-data.unibe.ch/metads) is editable and new records can be added by registered users (see A. Stickler, S. Brönnimann, M. A. Valente, J. Bethke, A. Sterin, S. Jourdain, E. Roucaute, M. V. Vasquez, D. A. Reyes, R. Allan, and D. Dee, 2014: ERA-CLIM: Historical Surface and Upper-Air Data for Future Reanalyses. *Bull. Amer. Meteor. Soc.*, **95**, 1419–1430. doi: http://dx.doi.org/10.1175/BAMS-D-13-00147.1).

ERA-CLIM2 data recovery map

4 - The new registry adapts the UBERN tool in order to be able to assimilate and merge the inventories already existing at global scale DARE initiatives, such as ERA-CLIM, UERRA, ECA&D, ISTI, ISPD, CHUAN, IGRA, IEDRO, CRU. It also allows the possibility to integrate regional and NMS inventories.

It will be possible to list stations and search the inventories, edit individual records, add new station entries or new Inventories (under site administration supervision). Editing will only be allowed to registered users.

The site is being built using php language to ensure a high level of security to the database. A javascript package (dhtmlxGrids) allows us to construct the editable inventory tables.

The ERA-CLIM2 inventories were the first to be uploaded and are being used to test the tool. The ERA-CLIM2 inventories cover the 5 continents during the 1863-2011 period, with particular emphasis on the regions shown on the ERA-CLIM2 map on the right. Rescue efforts have been concentrated primarily on upper air data all over the world, but also on surface data in the tropical regions and in Chile.

5 – In ERA-CLIM2 imaging and digitisation of upper air and surface data are proceeding apace. Maritime data has been imaged and is still awaiting digitisation. QC and homogenisation are also being applied to upper air data and QC only (for the moment) to surface data.

Many ECV variables are being digitised and time series are being constructed, The current situation of the data processing is signalled on the inventories by 6 levels: 0 – data in hardcopy ; 1 –imaging completed; 2 – digitised raw data format; 3 – basic QC applied; 4 – Compreensive QC applied and formatted data; 5 – Homogenised data.

The global registry will provide a unique_identity_number for each station, the WMO number, name, location (country, longitude, latitude, altitude), time resolution, starting and ending dates, the existing variables, comments, data processing level the original source and a link to data holders and/or contacting persons or institutions.

It is possible to perform searches by the several parameters listed in the Metadatabase, ex. search by station name, country, period, variable,etc. The tool supplies a list with the available entries for each search. It will also be possible to chose a lat/lon limited area and get the list of stations included in this region. Crossed searches will also be allowed in the future

ERA-CLIM2 Global Registry Home Page

6 - One of the main forthcoming developments will be the ability to plot on a lat/lon map the chosen stations (see the maps presented here for the surface and upper air ERA-CLIM2 stations). So, if the user wants for instance to plot all stations available for a certain year, this tool will have this capability.

The next phase in terms of uploading inventories, will be to seek the other global DARE metadatabases. They will be adapted to the ERA-CLIM2 format and a unique identity number given to each station. The ISPD inventory database is especially important to include, as the ERA-CLIM2 reanalyses are using version 3 of ISPD. Many ERA-CLIM2 stations have already been included in ISPD3 and the Global Registry will certainly detect duplicated records of surface or msl pressure.

One of the advantages of the Global Registry will therefore be the detection of possibly duplicated efforts in rescuing data. It will also help to prevent unnecessary re-digitisation or re-imaging. Nevertheless sometimes it is useful to crosscheck data that has been digitised twice by different groups, as it helps to eliminate some pending errors and fill small gaps.

7 - As data is being rescued, quality controled and homogenised, registered users that are responsible for these tasks will be able to update their respective entries themselves, under the guidance of the Global Registry administration. For every updating task there will be a message sent to the administrators with information on who and what was changed. Alternatively new updated inventories can be submitted to the site administration
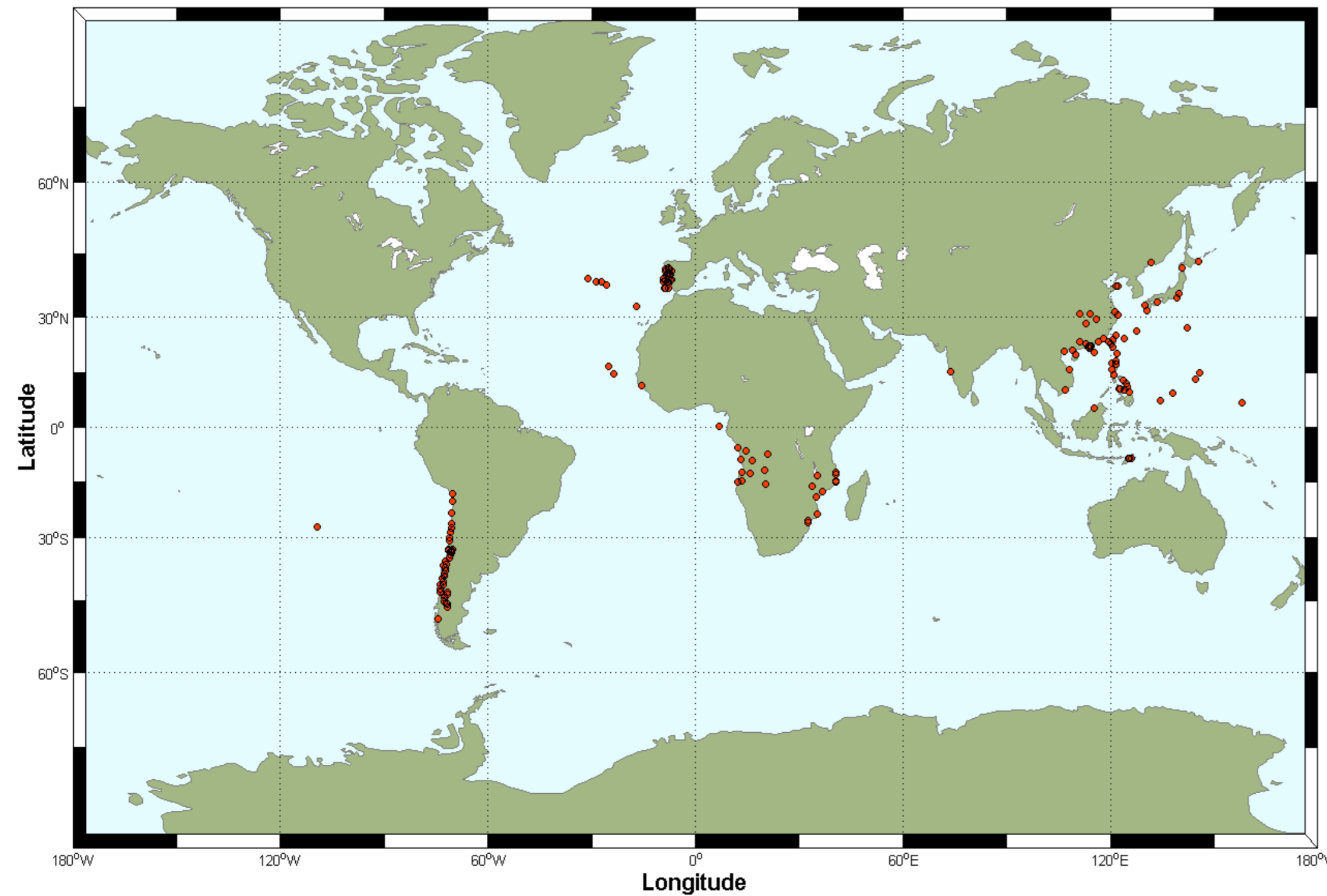
It is highly recommended that this Global Registry is built in cooperation with the www.IDARE-Portal.org initiative taking place under WMO auspicies at KNMI (see Peter Siegmund's talk), which is the site for listing all DARE initiatives taking place at the moment.

The 2 full sets of surface and upper air data being subjected to rescue, quality control and homogenization processes in ERA-CLIM2 can be seen in the bottom plots on this poster.

All this work is being performed under the Atmospheric Circulation Reconstructions over the Earth initiative.

ERA-CLIM2 Global Registry - Surface Data Inventory

ERACLIM2: Surface Inventory 1863-2011

ERACLIM2: Upperair Inventory 1891-1972