PaleoProxy Data Base (PPDB): A comprehensive geodatabase to archive and manage paleoproxies data

D.Strigaro, M.Mattavelli, I.Frigerio, M.DeAmicis

Geomatic Laboratory - Department of Earth and Environmental Sciences – University of Milano Bicocca

Introduction

Glaciers are sentinels of climate and environmental change, many marine regions pollen and other biological and geological entities can provide information on past climate variations. The Project of Interest NextData will favour the implementation of measurement networks in remote mountain and marine areas and will develop efficient web portals to access meteo-climatic and atmospheric composition data, past climate information from ice and sediment cores, biodiversity and ecosystem data, measurements of the hydrological cycle, marine reanalyses and climate projections at global and regional scale. In this work a methodology to recover, store, access and disseminate paleoproxies data is presented. Studied proxies are ice and marine cores, sediment cores, pollen, tree rings, speleothems (Bradley R. S. 1999).

Data Retrieving

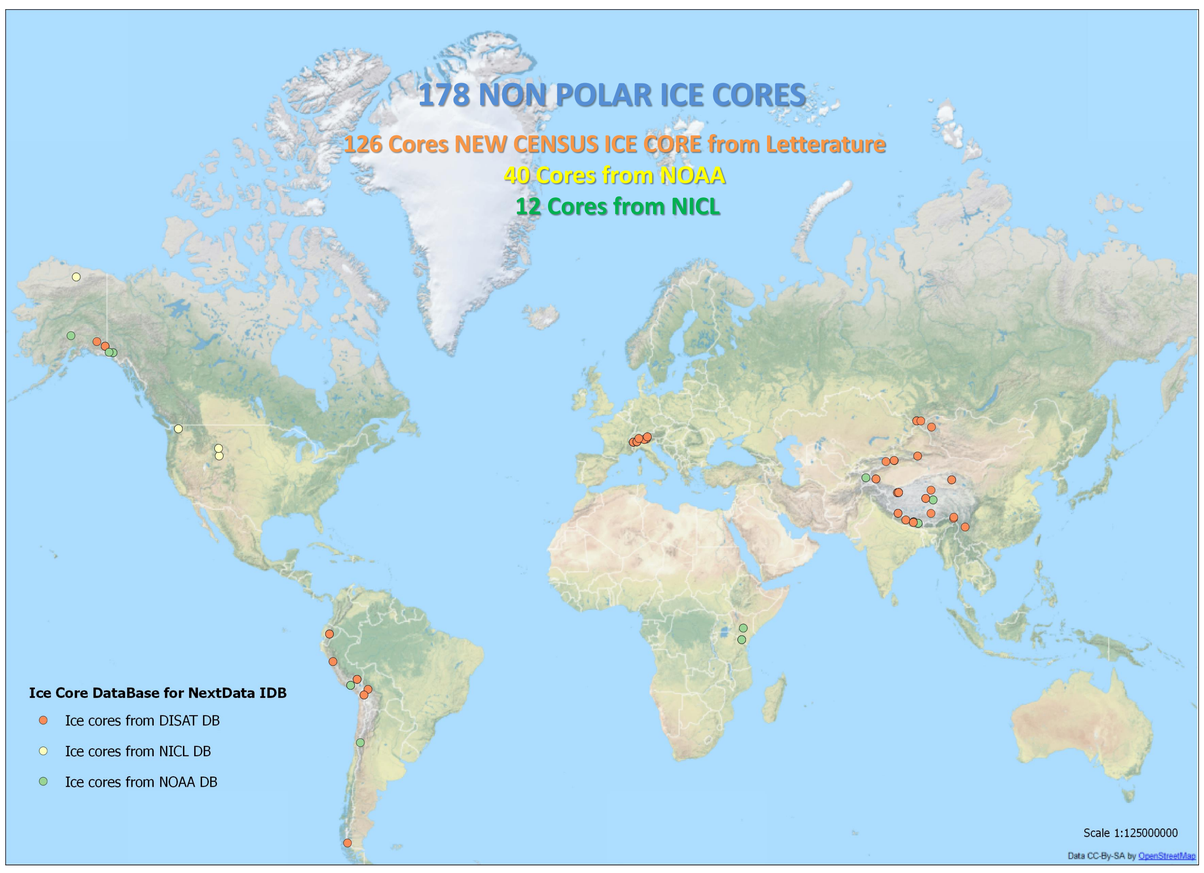
A total of 178 different non polar ice cores has been collected coming from 4 different sources.

- 52 comes from NOAA and NICL database;
- <u>126 has been collected from literature and so georeferenced and stored for the first time in an ice core Db;</u>

<u>A total of 86 pollen core has been collected coming from the Laboratory of Palinology and Paleoecology of CNR-IPDA</u>

To standardize the data, each parameter has been defined by a IUPAC name for chemical value and SI (International System of Units) for units of measurements.

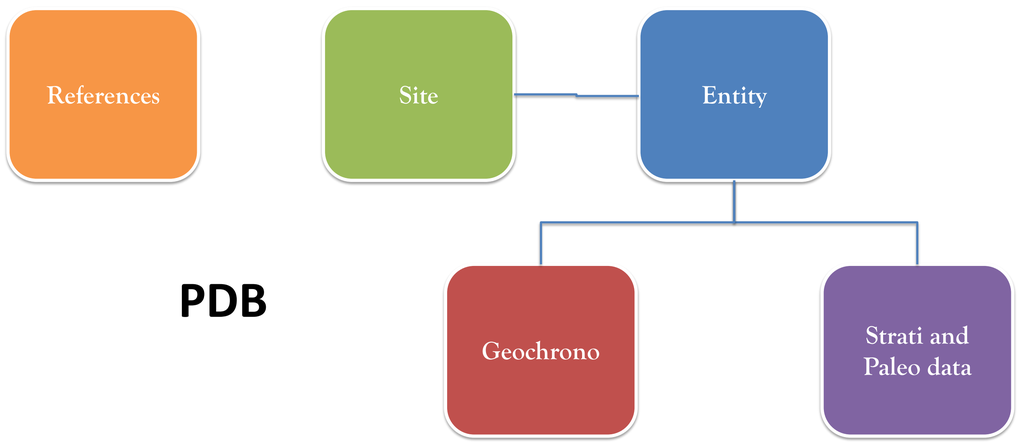
The NOAA and NICL databases, related to ice core, are not structured to archive each singles numeric value coming from ice analysis so they didn't offer the possibility to search data with queries based for instance on the data provider, on the parameter of interest and or the ice core name.

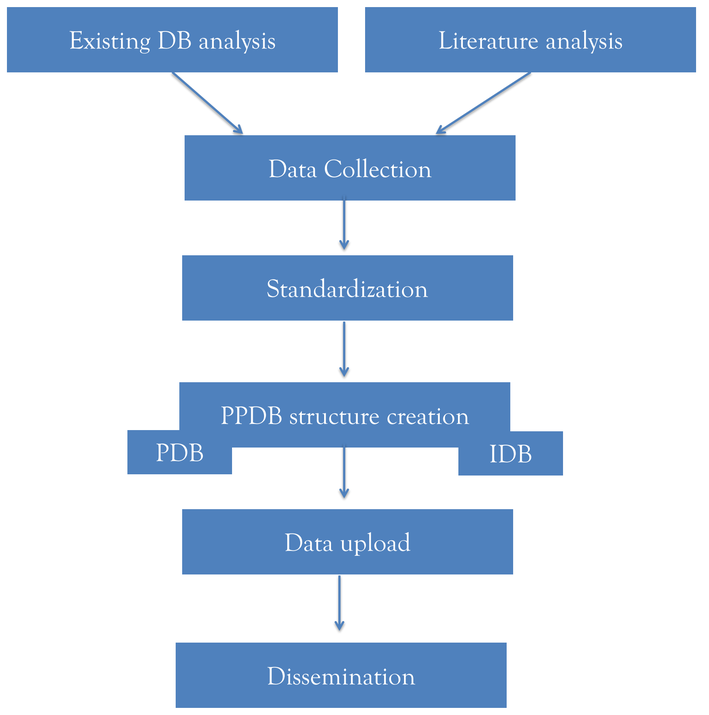


Pollen DataBase (PDB)

From the individuation of the principal entities, a new suitable tables schema has been drawn to permit their archivation:

- **site**: location of recovery of one or more core;
- **entity**: in this table can be found all entities that we have selectionated like core, section and superficial sample;
- **stratigraphic and paleoenvironmental data**: all data about pollen count or percent and stratigraphic section data about cores;
- **references**: in this section, are archived dataprovider data and references about articles where the information and data have been found;
- **geochrono**: a set of tables that register each chronology developed for an entity. In addition in this section are stored the name of the model used to derive the sample ages, the name of the person who prepared the chronology, and the date of the last modified chronology (in development).





Structure of the Paleo Proxies DataBase (PPDB)

The structure of the database has two main quality:

- Maintain the interoperability of different type of data,
- Find a way to standardise the paleoclimatic data.

Using Postgresql, an unique geodatabase called PPDB is initialized. Inside it, are built different schemas for each typology of data used in this context.

Ice core DataBase (IDB)

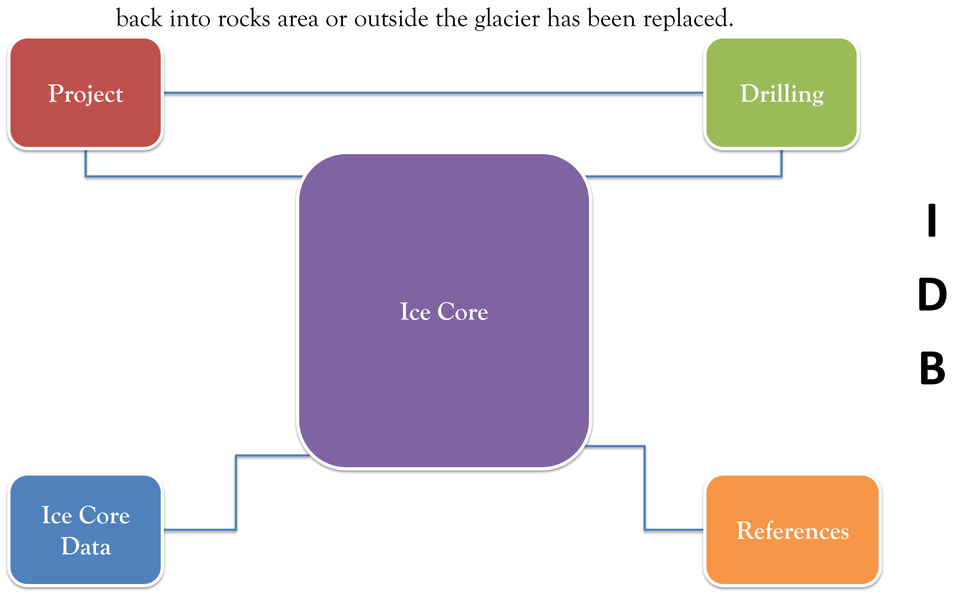
- IDB has been be structured as a geographical database where the spatial information has been defined by a couple of coordinates that identify a point inside a non-polar glaciers, where the ice cores has been drilled.

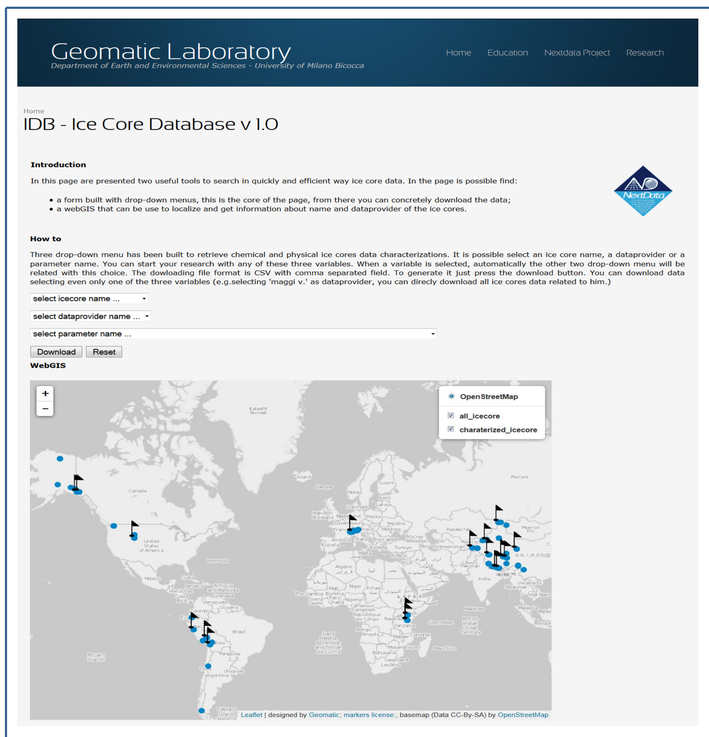
The schema can be divided by the following sections:

- **Project:** section that contain the principal information about the perforation project such as the year of perforation and the principal investigator;
- **Drilling:** table that contain the information about the drilling site and the year of drilling.
- **Ice core:** section that stored the principal parameters that can define the IC.
- **Ice core data:** table that provide a definitive description of what the data represents, including chemical and physical properties, identifies the characteristic or the measurable factors of the values being parameterized. n this section, are archived dataprovider data and references about articles where the information and data have been found.
- **Reference:** section in that there are archived the numeric value of the analysis.

In most cases, the spatial information is obtained from the literature and reported as one coordinate for the different ice cores drilled in the same glacier.

A quality control of coordinates was made on the basis of information retrieved from maps and imagine found in literature and the ice core or the perforation site fall back into rocks area or outside the glacier has been replaced.





Retrieving data

A geoportal has been built to provide a simple and useful mask where users can download all data. Only a web interface for ice core is developed with three key words: dataprovider, parameter and ice core name. The web interface can be accessible at this link: http://geomatic.disat.unimib.it/idb. Thanks to OGC standards a webgis is built to visualize the geospatial distribution of the cores.

Conclusion

Currently this is the first and unique geodatabase where the raw numeric value, derived from the measurement of different paleoproxy, have been stored. Unlike other databases, PPDB allow users to search a specific chemical or physical values stating from a particular dataprovider name or parameter. This is essential to enable a rapid data search and to quickly compare different proxy belong to the same category.

The importance of this project is firstly the intersection between different paleoproxy that are archived in the same database, secondly the interaction between data characterization, their geographic information and their metadata. Through the development of web GIS application it is possible to share environmental datasets and provide easy access for users lacking GIS knowledge.

Milano 10-12 Settembre 2014
the future of THE ITALIAN GEOSCIENCES of the future