

# An observed database to characterize the weather conditions associated with subtropical cyclogenesis over southern-southeastern of Brazil



Rodrigo Yamamoto\*, Rosmerir P. Rocha

rodrigo.yamamoto@usp.br, Institute of Astronomy, Geophysics and Atmospheric Sciences (IAG) – University of Sao Paulo (USP) - Brazil



## Abstract

Due to the necessity of studying the climatic, dynamic and synoptic aspects of subtropical cyclones that occur in southern-southeastern coast of Brazil, the project “Ciclones Subtropicais no Atlântico Sul: Aspectos climáticos, sinóticos e dinâmicos” is being developed. The weather conditions associated with such cyclones is an important question that must be answered in this project. However, for such characterization it is necessary to use the local meteorological observations of wind, wind gust, rainfall, air temperature, etc.

The NCEP (National Center for Environmental Prediction) reanalysis have spatial and time resolutions that provide elements to study the synoptic and dynamics of meteorological events (cyclone, anticyclones, troughs, ridges, monsoons circulations, etc) and even the production of complex climatology. However, this analysis has coarse horizontal resolution (~250 Km) that often does not allow the identification of intense meteorological phenomena. A more precise characterization of location and intensity of weather conditions associated with subtropical cyclones would be performed using local observations.

Therefore, this work describes the methodology to construct a database of surface weather observations using a relational database management system (RDBMS) MySQL. The data source are SYNOP (Surface Synoptic Observations), METAR (Meteorological Aerodrome Report), NCDC (National Climatic Data Center) and CETESB (Environmental Agency of Sao Paulo State) that are available online through dynamic web pages.

## Method

**Implementation of the database (RDBMS):** A database was created in MySQL, to store and organize the meteorological surface data. The structure of the database was based on three table categories (detailed in Figure 2):

- a table with informations referring to the data source.
- a table with the weather stations definitions.
- two table sets (for hourly and daily data) corresponding to each weather variable.

**Data acquisition and preprocessing:** as the data sources, SYNOP, METAR, NCDC and CETESB were available in dynamic web pages, an iterative algorithm was developed to automate the web servers data requisition and to parse the raw data from server response.

**Post-processing of data and storage:** Most data sources are coded or in non-patterned formats, therefore algorithms were developed in C++ using the REGEX library (detailed in Figure 1). After data decoding and formatting, the valid gathered data were sent to the MySQL configured database.

## Conclusion

The work was extended to a data set that included the entire Latin America using the MySQL that give us a simple and versatile platform allowing the continued growth of the database. The next step is to establish a web interface to make the data available to general public.

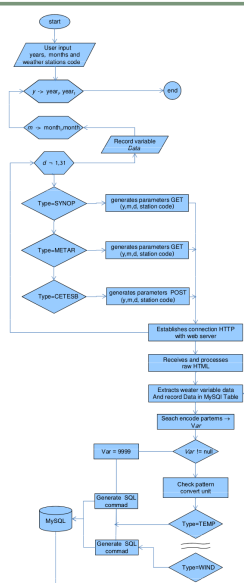


Figure 1 – flow diagram of algorithm

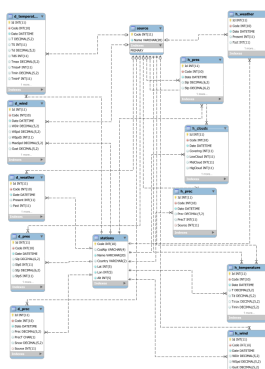


Figure 2 – schema of the database

## Data and results

The data in NCDC, SYNOP, METAR, and CETESB are available from 1942, 1999, 1997 and 1997, respectively.

