NOI A.G. / S.p.A.
Roberto Cavaliere
r.cavaliere@noi.bz.it
T +39 0471 066 676

EURAC Weather Data Collector

v1.2, 01.07.2022

Preliminary notes

The Center for Sensing Solutions of EURAC manages valuable research datasets withing Environmental Data Platform (EDP). The willingness is to make these datasets also available through the Open Data Hub and its channels.

The first dataset imported is related to historical meteorological data measured by a wide network of weather stations in the Autonomous Provinces of Bolzano and Trento. The dataset is quite relevant for climatological analysis, since the history of available measurements is quite long, and covers in some cases many decades. The data is made available through three different methods:

- **metadata**: https://edp-portal.eurac.edu/envdb/metadata
- **historical measurements**: https://edp-portal.eurac.edu/envdb/climate_daily [1]
- **climatological analysis**: i.e. the averages per month of the weather parameters measured along the whole time series available: https://edp-portal.eurac.edu/envdb/climatologies

More information about the data shared is available at: https://edp-portal.eurac.edu/cdb_doc/

Field Code Changed
Field Code Changed
Field Code Changed
Field Code Changed

[1] Heavy call, better to use this method by using the name of the station of interest, e.g. https://edp-portal.eurac.edu/envdb/climate_daily?station=eq.Cles

Field Code Changed

General design considerations for the integration in the ODH

The intrinsic data model is very in line with the simple model followed by the Open Data Hub – Mobility, based on the triple "station" – "type" – "measurement".

One challenge is related to the fact that most of the stations are probably already available in the Open Data Hub, being given by other Data Providers. The ideal approach is to link the measurements to existing stations and types, without introducing new ones. This is ensured by the fact that stations are characterized by the codes assigned by the reference public authorities owning them, e.g. the Autonomous Province of Bolzano and Trento. Since these codes already characterize these stations integrated in the Open Data Hub, they can be used as key to recognize stations that are already integrated in the Open Data Hub.

Commented [RCN1]: Waiting for feedback by EURAC in relation to the availability of the stationcode in their API, so that we can use this field to make the matching with the stations we have in the ODH

The other challenge is related to the integration of the climatology time series. Since they are near static data, an approach could be to consider them as additional metadata of a station. The choice is however to consider them as measurements as well, with an update time of one year.

Stations metadata

The mapping of the metadata follows the following specifications:

Web-service fields	ODH parameters
id	-
geom	-
lat, long	**pointprojection**
~~code~~id_source	**stationcode**
name	**name**

Table 1: Mapping between web-service and ODH fields (reference: "station" table).

Commented [RCN2]: If we don't received the stationcode (see above comment, to be mapped with our stationcode)

Commented [RCN3]: To be checked / added by EURAC

The following specifications have to be also considered:

- the ODH field **origin** is to set as **EURAC**.
- the ODH field **stationtype** is to set as **MeteoStation**.
- in case id_source = NULL a new station has to be created, otherwise the reference metadata is the one already available in the Open Data Hub (no metadata update)[2]
- the other fields provided by the web-service (including "id" which an internal EURAC code) have to be considered as **metadata** (i.e. will be stored in the "metadata" table of the ODH database)

Formatted: English (United States)

Formatted: English (United States)

[2] The assumption here is that the main Data Providers for these stations (i.e. the stations owners) are the reference source for the metadata.

Station types

Unfortunately, the types are retrieved directly with the measurements, and there is no specific method to get this information. This means that the Data Collector shall use the methods for the measurements and the climatologies also to manage the association between stations and types.

As already mentioned, it is very important to reuse the types that are already present in the ODH, where possible. The following mapping has to be considered:

Web-service fields	ODH parameters	New type?
tmin	**air-temperature-min**	**YES**
tmax	**air-temperature-max**	**YES**
tmean	**air-temperature**	**NO**
prec	**precipitation**	**NO**

Table 2: Mapping between web-service and ODH types (reference: "type" station)

Measurements

Data have to be considered as **measurement**, i.e. a field with double format.

The following specifications have to be also considered:

- the ODH field **period** is to set as:
  - **measurements**: 86400 (one day, in [s])
  - **climatological analysis**: 31536000 (one year, in [s])
- the ODH field **timestamp** is to set as:
  - **measurements**: mapping with web-service field "date"
  - **climatological analysis**: ~~same value as "created_on" (timestamp of the reading of the new data from the web-service)~~the timestamp must be configured following this specification
    - month = x (e.g. '5')
    - created_on (timestamp of call request) = dd.mm.yyyy (e.g. 01.07.2022)
    - timestamp = 01.x.yyyy-1 (e.g. 01.05.2021)

Formatted

Frequency in the API call

The data is not real-time, so it considered sufficient to interrogate the web-service once a day.

The Data Collector should not check if there are changes in the data, this will be controlled by the ODH core (writer component) before storing the data in the database.