

ICOS ETC DATA PRODUCT DESCRIPTION

Level2 Meteo Sensor Half-hourly Data

Data Type Label: ETC L2 Meteosens

Data Level: 2

File name structure: ICOSSETC_CC-###_METEOSENS_L2 (with CC-### = ICOS site code, see also the "Update and versions" paragraph below).

Stations: Class 1, Class 2, Associated

Format: csv

Description: half-hourly meteorological data and all the related quality flags and connected variables measured by each single sensor used in the station. Data are processed and quality checked by the ICOS ETC starting from the raw data at higher time resolution (Class 1 and 2 stations). Data not gap-filled. The code used is available in the ICOS ETC GitHub repository. In the Associates stations the data are calculated by the station teams.

Update and versions: released once or twice per year. When released twice per year (only Class1 and 2 stations), the middle-year version (*interim*) is created by appending the new processed data to the last official release. This version is identifiable by the file named ICOSSETC_CC-###_METEOSENS_INTERIM_L2

Variables and units: variable names are composed by the variable code and three numeric indexes or positional qualifiers, used to indicate relative positions of observations at the site (e.g. different points in space, along a vertical profile). The information related to sensor model, position and calibration for each variable are available in the metadata. Being aggregated from data at 1 to 60 seconds resolution (in general), the standard error (identified by the suffix _SE) and the number of single measurements used in the calculation (identified by the suffix _N) are reported for each variable and half-hour. For Class 1 and Class 2 stations, each variable is also provided with a quality check flag (suffix _QC) informing the user of the presence of format issues, quality, disturbances (as reported in the metadata) or other reasons for rejection of records and half hours, and the amount of data discarded for each reason. The flag is made of a numeric string using one cipher (0-9) for each potential data rejection source. The first cipher is fixed to 8, and 8 is also used in all the other

positions when the corresponding condition cannot be checked. The second and third position (values 0-1) indicates if the corresponding file was present in the database, and if its format was compliant, respectively. The fourth and fifth positions (values 0-5) are relative to the quality tests applied to the data, namely an out-of-range and a step tests: the value of the ciphers depends on the amount of records discarded (from 0: no data rejected, 1: less than or equal to 10% data rejected, up to 5: $\geq 40\%$ data rejected). The sixth position indicates how many records were missing in the file (values 0-5 with the same criteria explained above). The seventh position is relative to the presence of any type of disturbances reported by the station team via the BADM system (values 0-5). The eighth position indicates whether the mapping of the corresponding variable was missing in the BADM system (values 0-1). Ninth and tenth positions summarise on the presence of the half hour after all the tests and checks applied, respectively by indicating whether the corresponding half hour had greater than or equal to 30% of records rejected/missing, and greater than or equal to 6 consecutive minutes missing, respectively (the two conditions applied to discard or retain a half hour). The cipher in the eleventh position (0-1) is simply indicating whether the corresponding half hour was missing in the raw files, and the twelve and last position is dedicated to reporting on the automatic calibration of the sensor measuring that variable (currently only for G).

Variables codes, units and descriptions are available here:
<https://hdl.handle.net/11676/rIIMSsmAoDi2W3W44rGzQI3X>

Metadata: variable codes and units available in the Carbon Portal. Full metadata in the Archive product, including sensor model, position, history.