JRC SCIENTIFIC INFORMATION SYSTEMS AND DATABASES REPORT

Quality Checks of meteorological data

QuackMe

Architectural and technical guide

2022

Joint
Research
Centre

Contact information
Name: Andrea Toreti
Address: Via Enrico Fermi 2749, I-21027, Ispra (VA), Italy
Email: andrea.toreti@ec.europa.eu

EU Science Hub
https://ec.europa.eu/jrc

JRC128152

PDF ISBN 978-92-76-46870-7 doi:10.2760/96559

Luxembourg: Publications Office of the European Union, 2022

How to cite this report: M. Bratu, A. Toreti, M. Zampieri, and A. Ceglar, *Quality Checks of meteorological data - QuackMe*, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-46870-7, doi:10.2760/96559, JRC128152.

Contents

Acknowledgements

We thank E. Müller, I. Oliveira, and G. van der Grijn for their support during the testing and refinement phase of the software.

Authors

Marian Bratu (Fincons Spa)
Andrea Toreti, Matteo Zampieri, and Andrej Ceglar (European Commission, Joint Research Centre)

Abstract

Reliable meteorological data from ground weather stations are an essential source of information for sectoral climate services. Checking their quality in near-real-time is a key step to detect and, whenever possible, correct errors. Here, we describe the technical elements of a new software developed to perform such a task : QuackMe. The software, free and open-source, can be retrieved in the dedicated JRC github page.

1 Introduction

The main purpose of this report is to describe the development and the maintenance of a new software *Quality Checks of Meteorological Data* QuackMe. This software is designed to obtain high-quality daily meteorological data. It processes meteorological observations from ground weather stations, identifies and removes erroneous values, identifies and flags suspicious values, and fills sub-daily gaps whenever possible. QuackMe replaces an old quality control software (developed in the '90s) in place at the JRC: AMDAC.

QuackMe satisfies some key essential criteria:

- open source;
- flexible to modify and easy to configure;
- clear and easy to modify code;
- able to use external data whenever necessary or required;
- easily adaptable to different input data formats;
- able to identify errors and suspicious values for fast and easy problem solving;
- adaptable to different platforms and operating systems;
- integrated with a user-friendly interface;
- able to report wrong or suspicious observations in a clear way.

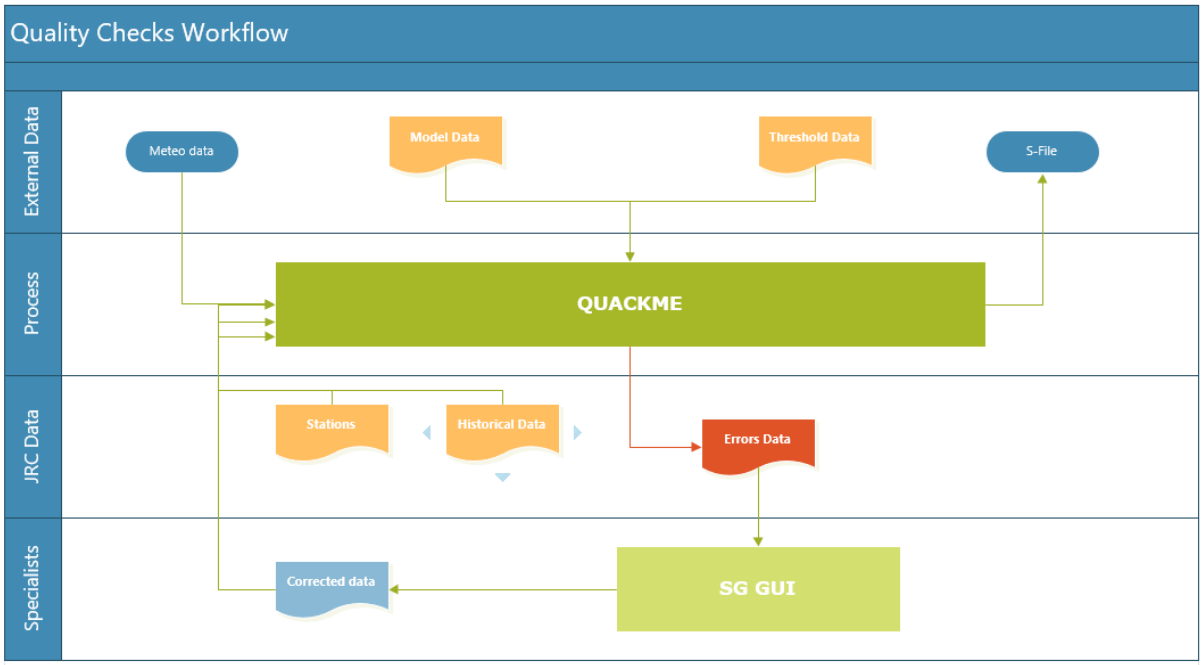
Table 1 Sub-daily input data

Description	***Note***	***Unit of Measure***
Station number	*Numerical value*	
Day Time	*Numerical value (YYYYMMDDHH24)*	*Including hour with 24H format*
Air temperature		*Celsius degrees*
Dew point temperature		*Celsius degrees*
Wind direction		*Degrees*
Wind speed		*Knots*
Solar radiation		J/cm^2
Sunshine duration		*Minutes*
Precipitation		*mm*
Relative humidity	*Percentage between 0 and 100*	
Atmospheric pressure		*hPa*
Cloud cover		*Octas*
Visibility		*km*
Snow depth		*cm*

2 Architecture

The architectural design of QuackMe is shown in Figure 1.

Figure 1. Schematic QuackMe workflow



The quality checks component is implemented as a workflow where each step produces a part of the input for the next step.

At each step, there are:

- Input:
    - output with checked/corrected data from a previous step (if it exists)
    - historical data, if necessary
    - external data from suppliers, if necessary
    - threshold data

- Output :
    - file with checked/corrected data (internal data format)
    - file with quality checks flag
    - file with errors/problems (xml format)
    - historical file, if necessary
    - data for suppliers (S-File)

In details, the workflow contains the following steps:

- *Input Converter*. It converts data from the meteorological data provider format to an internal data format.

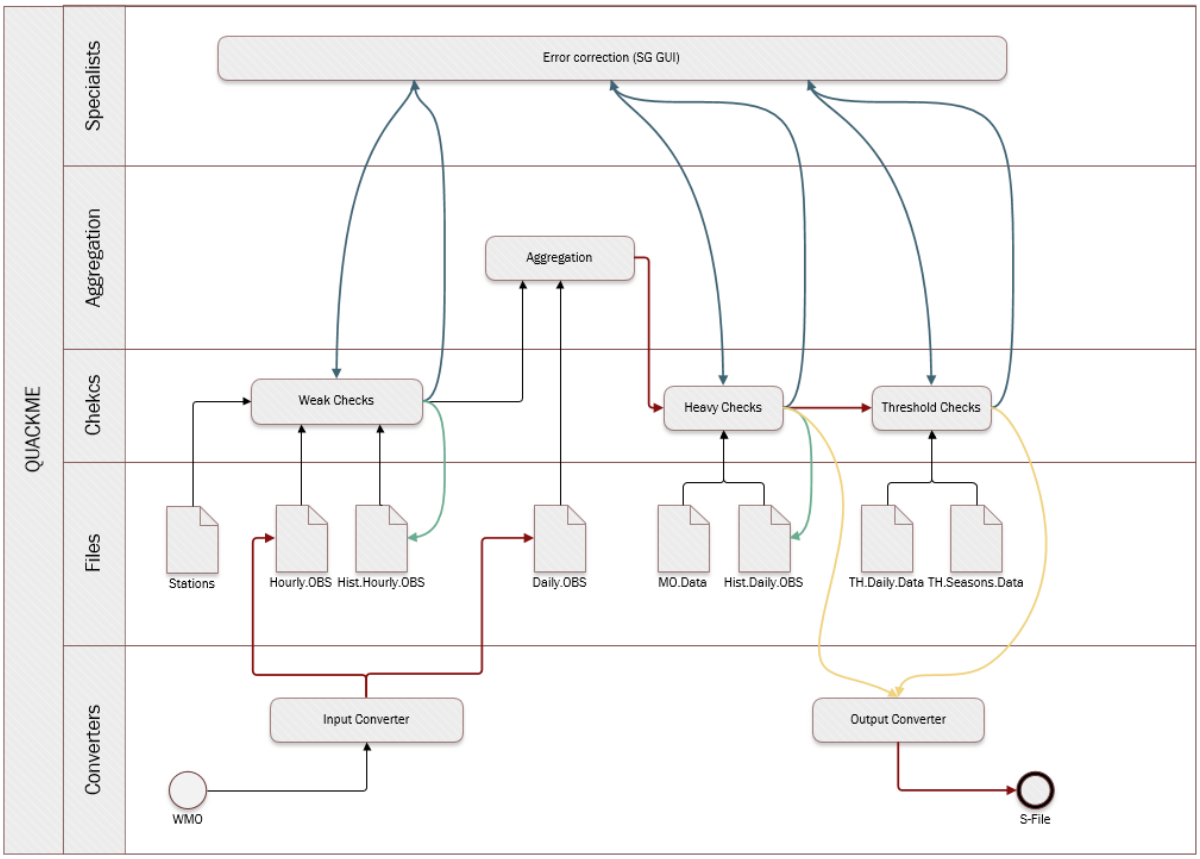
- *Weak Checks*. It performs the weak checks.
    - Input:
        - Data from meteorological data providers or checked/corrected data. It uses observations of the day under analysis and the ones from the previous day (when needed).
    - Output:
        - File with checked/corrected observational data (internal data format)
        - File with observations to be used for the controls of the subsequent day
        - File with errors (xml format)

- *Aggregation*. It transforms hourly observations into daily aggregated data and derives some specific agro-climate indicators

- *Daily/Heavy Checks*. It performs checks of the daily aggregated data:
    - Input:
        - Daily aggregated data
        - Daily data of the previous day
        - Model Output Statistics data or other external data sources to be used as reference
    - Output:
        - File with checked/corrected data
        - File with errors (xml format)

- *Threshold Checks*. It performs complex checks:
    - Input:
        - Data already processed by the Daily Checks
        - Daily data of the previous day
        - Daily and seasonal threshold-data
    - Output:
        - File with checked/corrected data
        - File with errors (xml format)
- *SConverter*. It converts the output of QuackMe to the S-File format.

Figure 2. Details of the QuackMe workflow and of its components



3 QuackMe components

3.1 Input converter

To have more flexibility on managing different formats of input data a converter has been developed. For each type of input format, it creates an interface (to be called from the main workflow) to convert the input data into the internal data format. This module gives the possibility to use more suppliers of meteorological data without changing the main workflow. The converter is also designed to check for duplicated lines in correspondence of <Station, Day&Time>. It gives an error message when one is found. It also replaces values equal to 9 in the N and L columns with 8.

3.1.1 Input data format

The input data format depends on the data provider.

3.1.2 Internal data format

The internal data format is a text file organized by columns. Its structure contains:

- First line, the header, with the name of the columns.
- Lines with values (associated to the available observations) separated by a single space character.
- No empty lines.

There are two formats: one for sub-daily observations and another one for daily observations.

Table 2 Conversion details from the standard input format (dat) to the internal one (csv)

Property description	DAT columns	Note	CSV Columns	Unit of Measure
Station number	Station	Numerical value	INDEX	
Day Time	DayTime	Numerical value (YYYYMMDDHH24)	YYYY, MM, DD, HHmm	
Air Temperature	TT	Numerical value	TL	Celsius Degree
Dew point temperature	TD	Numerical value	TD	Celsius Degree
Hourly maximum temperature	TX1	Numerical value	TX1	Celsius Degree
Hourly minimum temperature	TN1	Numerical value	TN1	Celsius Degree
6-hourly maximum temperature	TX6	Numerical value	TX6	Celsius Degree
6-hourly minimum temperature	TN6	Numerical value	TN6	Celsius Degree
12-hourly maximum temperature	TX12	Numerical value	TX12	Celsius Degree
12-hourly minimum temperature	TN12	Numerical value	TN12	Celsius Degree
Hourly precipitation	PREC	Numerical value	RR1h	mm

Property description	DAT columns	Note	CSV Columns	Unit of Measure
6-hourly precipitation	PR06	Numerical value	RR06	mm
24-hourly precipitation	PR24	Numerical value	RR24	mm
Precipitation	RR	Numerical value	RR	mm
Precipitation interval	TR	Numerical value (Hour number)	TR	
Snow depth	Snow	Numerical value	SNO	cm
Wind direction at 10m	DIR	Numerical value	DIR	Degree
Wind speed at 10 m	FF	Numerical value	FF	m/s
Cloud cover	N	Numerical value	N	Octas
Low cloud cover	L	Numerical value	L	Octas
Hourly radiation	RD	Numerical value	GL1h	J/m^2
24-hourly Radiation	RD24	Numerical value	GL24	J/m^2
Atmospheric pressure at station level	AP	Numerical value	QFE	hPa
Sea-Level reduced atmospheric pressure	QFF	Numerical value	QFF	hPa
Hourly sunshine	SH	Numerical value	Sh	minutes
24-hourly sunshine	SH24	Numerical value	SS24	minutes
Visibility	VIS	Numerical value	VIS	km
Relative humidity	RH	Numerical value	(added by code)	%
Soil State	SOIL	Numerical value	GS	

3.1.3 MOS/external data source file format

The MOS/external data source file (daily) contains hourly values for the interval: 00 of the day under analysis – 06 next day.

Table 3 Conversion table from the external data source/MOS (CSV) to the DAT file

Property description	DAT columns	Note	CSV Columns	Unit of Measure
Station number	Station	Numerical value	INDEX	
Day Time	DayTime	Numerical value (YYYYMMDDHH24)	YYYY, MM, DD, HHmm	
Air Temperature	TT	Numerical value	TL	Celsius Degree
Dew point temperature	TD	Numerical value	TD	Celsius Degree

Hourly precipitation	PREC	Numerical value	RR1h	mm
6-hourly precipitation	PR06	Numerical value	RR06	mm
24-hourly precipitation	PR24	Numerical value	RR24	mm
Wind speed at 10 m	FF	Numerical value	FF	m/s
hourly maximum precipitation	RRRX1	Numerical value	LAYER1	mm
6-hourly maximum precipitation	RRRX6	Numerical value	LAYER2	mm
12-hourly maximum precipitation	RRRX12	Numerical value	LAYER3	mm
24-hourly maximum precipitation	RRRX24	Numerical value	LAYER4	mm

3.2 Weak checks component

It performs checks on sub-daily observations of the following parameters:

- 2-m air temperature
- Minimum and maximum 2-m air temperature
- Dew point temperature
- Cloud cover
- Precipitation
- Sea level atmospheric pressure
- Wind speed & direction
- Sunshine duration
- Relative humidity
- Atmospheric pressure
- Solar radiation

Currently, checks for visibility and state of the soil are not configured.

The following checks are implemented:

- 2-m air temperature
- Values outside -80 / +60 Celsius degrees are flagged as WRONG.
- Difference between two consecutive records (hourly observations) < 0.1 degree. The alert is raised after the error has been identified at least 15 times. Values are flagged as SUSPICIOUS. This control is not active for the observational network only reporting integer values.
- Difference between two consecutive records (either hourly or 3-hourly) > 15 degree. Values are flagged as SUSPICIOUS.
- Absolute difference between the observed 2-m air temperature and the associated MOS/External data source 2-m air temperature greater than 12 degrees. The value is flagged as WRONG.

- Absolute difference between the observed 2-m air temperature and the associated MOS/External data source air temperature greater than 8.5 degrees and less than/equal to 12 degrees. The value is flagged as SUSPICIOS.

- 2-m minimum air temperature
- Values outside -80 / +40 Celsius degrees are flagged as WRONG.
- Difference between two consecutive records (hourly observations) < 0.1 degree. Values are flagged as SUSPICIOUS and the error message given at the 16-th observation.
- Difference between two consecutive records (either hourly or 3-hourly) > 15 degree. Values are flagged as SUSPICIOUS.

- 2-m maximum air temperature
- Values outside -80 / +60 Celsius degrees are flagged as WRONG.
- Difference between two consecutive records (hourly observations) < 0.1 degree. Values are flagged as SUSPICIOUS and the error given at the 16-th observation.
- Difference between two consecutive records (either hourly or 3-hourly) > 15 degree. Values are flagged as SUSPICIOUS.

- Dew point temperature
- Values outside -80 / 35 Celsius degrees are flagged as WRONG.
- Dew point temperature > Air Temperature. If the difference between dew point temperature and air temperature is less than 1 degree then the dew point temperature is replaced with the air temperature, else the value is flagged as WRONG.
- Difference between two consecutive records (hourly observations) > 11 degree. Values are flagged as SUSPICIOUS.
- Difference between two consecutive records (hourly observations) < 0.05 degree. The alert is raised after the error has been identified at least 20 times. Values are flagged as SUSPICIOUS. This check is not active for the observational networks only reporting integer values.
- Absolute difference between the observed dew point temperature and the associated MOS/External data source dew point temperature greater than 15 degrees. The value is flagged as SUSPICIOUS.
- Absolute difference between the observed dew point temperature and the associated MOS/External data source dew point temperature greater than 11 degrees and less than or equal to 15 degrees. The value is flagged as SUSPICIOUS.

- Relative humidity
- Values outside 0 - 100% are flagged as WRONG.
- Difference between two consecutive records (hourly observations) < 0.5% and both values being less than 95%. The alert is given after the error has been identified at least 20 times. Values are flagged as SUSPICIOUS. This check is not active for the observational networks only reporting integer values.

- Atmospheric pressure
- Values outside 500 – 1100 hPa are flagged as WRONG.
- Difference between two consecutive records (either hourly or 3-hourly) > 15 hPa. The value is flagged as SUSPICIOUS.

- MSL atmospheric pressure
- Values outside 950 – 1060 hPa are flagged as WRONG.

- Wind direction
- Values outside 0-360 degrees are flagged as WRONG.

- Difference between two consecutive records (hourly observations) < 5 degree conditional on having both wind speed values greater than 4.6 m/s. The alert message is given after the 20-th identified value and only if the altitude of the station is less than/equal to 1000m. Values are flagged as SUSPICIOUS.
- Difference between two consecutive records (hourly observations) > 40 degree, conditional on having both wind speed values greater than 4.6 m/s. The alert message is given after the 5-th identified value and only if the altitude of the station is less than/equal to 1000m. Values are flagged as SUSPICIOUS.

- Wind speed
- Values outside 0 – 75 m/s are flagged as WRONG.
- Difference between two consecutive records (hourly observations) < 0.5 m/s, conditional on having both wind speed values greater than 4.6 m/s. The alert message is given after the 15-th identified message. Values are flagged as SUSPICIOUS.
- Difference between two consecutive records (hourly observations) > 15 m/sec. The alert message is given after the 10-th identified value. Values are flagged as SUSPICIOUS.
- Values > 3.5 m/s and direction NA. Values are flagged as SUSPICIOUS.
- Absolute difference between the observed wind speed and the associated MOS/External data source wind speed greater than 40 knots. The value is flagged as WRONG.
- Absolute difference between the observed wind speed and the associated MOS/External data source wind speed greater than 25 knots and less than 40 knots. The value is flagged as SUSPICIOUS.
- Observed wind speed equal to 0 and the associated MOS/External data source wind speed greater than 18 knots. The value is flagged as SUSPICIOUS.
- Ratio between the observed wind speed and the associated MOS/External data source wind speed between 1.75 and 2.25. The values is flagged as SUSPICIOUS.

- Total Cloud cover
- Total cloud cover is replaced with low cloud cover value when total cloud is missing and low cloud value is valid.
- Values outside 0-8. The value is flagged as WRONG.
- Value equal to 0 and precipitation > 0. The value is flagged as SUSPICIOUS.
- Value equal to 8 and sunshine duration > 0. The value is flagged as SUSPICIOUS.

- Low Cloud cover
- Values outside 0-8 are flagged as WRONG.

- Solar radiation
- Values outside 0 -1600 W/m2 are flagged as WRONG.
- Value equal to 0 and sunshine duration > 0. The value is flagged as SUSPICIOUS.

- Sunshine duration
- Values > 60 minutes are flagged as WRONG.
- Sunshine daily duration > 24 hours are flagged as WRONG.

- Precipitation
- Values outside 0-400 mm are flagged as WRONG.
- Hourly values > 200 mm. The value are flagged as SUSPICIOUS.
- For each precipitation property (1H, 6H, 12H, 24H) the following two parameters are derived when the associated MOS/External data source data are available:
- Standard deviation (SD): min[(5*max[0.2 , RRRmos]), max[1, abs(RRRX1/6/12/24 - RRRmos)]

- Scaled difference (F): (RRRmos – RRR)/RRRmos
- Once these two parameters are calculated, the following conditions are applied:
- When abs(RRRobs – RRRmos) > (6 * SD) AND RRRobs > 20 AND F > 5, the observed precipitation property is flagged WRONG
- When abs(RRRobs – RRRmos) > (6 * SD) AND RRRobs > RRRmos AND ((RRRobs <= 10 AND RRRmos >= 0.5) OR (RRRobs <= 15 AND RRRmos >= 2.5)). The value is flagged as SUSPICIOUS.
- Precipitation cross checks:
- if at 6am of the next day the 24H precipitation is valid, it uses the interval 06 current day – 06 next day for the precipitation cross checks
- If at 6am of the next day the 24H precipitation is not available, it checks if at 00 of the next day the 24H precipitation is valid and uses the interval 00 current day – 00 next day for the precipitation cross checks
- Assuming that a 24H precipitation value can be derived by using the corresponding time interval, the following cross checks are performed:
- If the PR06 values in the specified interval (12, 18, 00 of the next day, 06 of the next day or 06, 12, 18, 00 of the next day) are all valid, it compares the sum of these PR06 values with PR24. If the difference is greater than 10 mm, a SUSPICIOUS flag is assigned to the PR24 value and to all PR06 values
- If not all the PR06 values are valid, the same check is performed by using RR/TR combination. The system tries to build a 24-accumulated value by looking at the availability of TR=6 and TR=12. Then, this value is compared with PR24 and if the difference is greater than 10 mm, a SUSPICIOUS alert is assigned to the PR24 value and to all RR values.
- If all 1H precipitation values are available and valid, their sum is compared with PR24. If the difference between them is greater than 10 mm a SUSPICIOUS flag is assigned to the PR24 value and to all the 24 hourly precipitation values.

The following variables are derived during the WeakChecks step:

- D_E Saturation vapour pressure. It is based on TT (2-m air temperature) and TD (dew point temperature) and on the following equation (ESAT (TT, TD)):

  6.1078 * exp (LVAP(TT) / RV * (1./Celsius2Kelvin(0.) - 1./Celsius2Kelvin(TD)),

  Where, Celsius2Kelvin transforms temperature from Celsius degree to Kelvin; RV is equal to 461.51; LVAP (Latent Heat of Vaporization) is derived by:

  (597 - 0.566* TT) * 4186
- D_VPD Vapour pressure deficit. To calculate this property the system uses TT and TD. The vapour pressure deficit is calculated as the difference between the saturation vapour pressure based on TT and the one on both TT and TD:

  ESAT(TT, TT) – ESAT(TT, TD)

- D_SLOPE Slope of the saturation vapour pressure. The slope is calculated by using the following equation:

  238.102 * 17.32491 * (D_E + D_VPD) / (TT + 238.102)**2

- D_RH Relative humidity. It is calculated by using the following equation:

  100 * ESAT(TT, TD) / ESAT(TT, TT)

The WeakChecks module can also use in input the error file (XML format). The "-r" option must include at least one KO file name and maximum two, separated by ";". If the KO files are empty (contains only the tag "<Observations />") or are missing, the module does not execute the checks.

Another input possibility for the WeakChecks module is the "-m <hour>" option. This option allows to append the input CSV file and the H.<day in elaboration>.hist file (History Path). The input CSV file must contain the hourly observation from <hour> + 1 of the day in elaboration; while the history file contains all hourly observation recorded at or before than <hour>. For instance, a command line of WeakChecks with –d "20191101" –m "06" has the following meaning:

- day in elaboration 01.11.2019
- from the H.20191101.hist file all the observations with DayTime <= 2019110106 are taken
- the CSV file must contain observations with DayTime >= 2019110107

The output file of the WeakChecks module contains the parameters listed in Table 4.

Table 4 Parameters in the output file of the WeakChecks module

Property description	**Acronym**	**Note**	**Unit of Measure**
Station number	Station	Numerical value	
Day time	DayTime	Numerical value (YYYYMMDDHH24)	
2-m air temperature	TT	Numerical value	Celsius Degree
Dew point temperature	TD	Numerical value	Celsius Degree
Maximum temperature 1 Hour	TX1	Numerical value	Celsius Degree
Minimum temperature 1 Hour	TN1	Numerical value	Celsius Degree
Maximum temperature 6 Hour	TX6	Numerical value	Celsius Degree
Minimum temperature 6 Hour	TN6	Numerical value	Celsius Degree
Maximum temperature 12 Hour	TX12	Numerical value	Celsius Degree
Minimum temperature 12 Hour	TN12	Numerical value	Celsius Degree
Hourly Precipitation	PREC	Numerical value	mm
24 Hour Precipitation	PR24	Numerical value	mm
Precipitation	RR	Numerical value	mm
Precipitation interval	TR	Numerical value (Hour number)	
Snow depth	Snow	Numerical value	cm
Wind direction	DIR	Numerical value	Degree
Wind speed at 10 m	FF	Numerical value	m/s
Cloud cover	N	Numerical value	Octas
Low cloud cover	L	Numerical value	Octas
Radiation 1 Hour	RD	Numerical value	J/cm^2
24 Hour Radiation	RD24	Numerical value	J/cm^2
Atmospheric pressure at station level	AP	Numerical value	
Atmospheric pressure at station level	QFF	Numerical value	
Sunshine 1 Hour	SH	Numerical value	minutes
Sunshine 24 Hour	SH24	Numerical value	minutes

Property description	Acronym	Note	Unit of Measure
Visibility	VIS	Numerical value	km
Relative Humidity	RH	Numerical value	%
Saturation vapor pressure	D_E	Numerical value, Calculated	hPa
Relative Humidity	D_RH	Numerical value, Calculated	%
Vapor pressure deficit	D_VPD	Numerical value, Calculated	hPa
Slope of saturation vapor pressure	D_SLOPE	Numerical value, Calculated	hPa/Celsius Degree
Soil status	SOIL	Numerical value	

3.2.1 Interpolation

The last step of the WeakChecks module is the interpolation of hourly missing data when feasible. The interpolation is driven by a configuration file ("Interpolations.config") having the following structure:

Property	StartDay	StartTime	EndDay	EndTime	AllowNegativeValues	Decimals
N	00	00	01	06	NO	1
L	00	00	01	06	NO	1
TN1	00	00	01	06	YES	1
TX1	00	00	01	06	YES	1
TD	00	00	01	06	YES	1
TT	00	00	01	06	YES	1
FF	00	00	01	06	NO	1
RH	00	00	01	06	NO	2
D_RH	00	00	01	06	NO	2

Where:

- *Property* is the name of the property to be interpolated
- *StartDay* is a code for the start day (00 – current day, 01 – next day)
- *StartTime* is the time relative to the start day
- *EndDay* is a code for the end day
- *EndTime* contains the time relative to end day
- *AllowNegativeValues* (YES/NO) allows to generate negative values during interpolation
- *Decimals* specifies the number of decimals for rounding

The interpolation process uses the hourly values presents in the interval <StartDay:StartTime> - <EndDay:EndTime>.

3.3 Aggregation

The main scope of this step is the aggregation of hourly observations to daily scale.

As the WeakChecks, the Aggregation level uses a specific configuration file (Aggregation.XML). It contains the definition of the daily properties to be generated by using the hourly observations of the current day and, when needed, of the previous and next days.

3.3.1 Aggregation.XML structure

The Aggregation.XML file contains the configuration of the aggregation process from hourly to daily values. It is an XML file containing definitions for each property to be processed. A property is identified by the tag PROPERTY, a simple one without attributes. In the PROPERTY tag, it is possible to define the following additional tags:

- *Name* (mandatory) It defines the name of the property in the output file.
- *FormulaValue* – (mandatory) This tag can get the following attributes:
    - *Type* (mandatory) with two possible values: *Fixed, Formula*. The latter one is used when it is necessary to call a method to calculate the value. *Fixed* is used when the property value is retrieved at a specific datetime, to be specified by adding other attributes:
        - *Time* hour of the observation (e.g., 00, 06 or 23)
        - *Day* day of the observation. It contains the following possible values: 0 day under analysis; 1 next day; -1 previous day
        - *ReferenceProperty* The name of the property to be read
        - *RoundedDigits* It specifies the number of digits to round the daily value.
    - When *Type* is equal to *Formula*, the following attributes and tags must be added
        - *Parameter* it can get only two values: V – vector; M – matrix
        - *Formula* – (mandatory) name of the formula, that must correspond to a method implemented into the "Aggregation.Formulas.R" file
        - *Parameter* tag – (optional) allows to specify an acronym that passes to the method some additional parameters. The following ones are defined at the moment:
            - ⇒ Current.Date – day in elaboration
            - ⇒ Inter.Data – interpolated data
            - ⇒ Day.Agg – all the properties aggregated for the day in elaboration
            - ⇒ Expected.Values – number of expected values for the MEAN method
        - For each formula, sometimes it is possible to specify the time intervals. That can be realized in two ways:
            - ⇒ Using one or more "Specific" tags like: <Specific Time="00, 03, 06, 09, 12, 15, 18, 21" Day="00" />, where "Time" contains the time of hourly observations, "Day" contains the day of the observation (relative value). If values from more days are needed, then add a "Specific" tag for each day.
            - ⇒ Using a combination of tags "Start" and "End" like : <Start Time="03" Day="00"/> <End Time="06" Day="01" />, where the "Start" tag defines the start of the interval and the "End" tag defines the end of the interval. With these tags the application retrieves all the observations between 03 of the day in elaboration until 06 of the next day.
        - When the "Parameter" tag is "V" there is a need to specify the "ReferenceProperty" tag with the name of the property to extract and use as an input to the method. When the value is "M" the matrix with all the properties is given to the method
        - Independently of the "Parameter" tag (either "V" or "M"), the temporal tags (Specific, Start, End) are applied to extract only the values for the specific times of observation.

- Another tag that can be set within the “Property” tag is “NAFormula” (optional). It defines an alternative formula to be applied if the formula specified by “FormulaValue” returns NA.

Examples:

1.
```
<Property>
    <Name>TT_09</Name>
    <FormulaValue Type="Fixed" Time="09" Day="00" ReferenceProperty="TT" RoundedDigits=”1”/>
</Property>,
```

The “Aggregation” module creates a property with name “TT_09”. The value is retrieved from the property TT for the day in elaboration at time 09:00 (from the input file 0.WeakChecks.<day>.dat)

2.
```
<Property>
    <Name>N</Name>
    <FormulaValue Type="Formula" Parameter="V" ReferenceProperty="N" RoundedDigits=”2”>
        <Specific Time="00, 03, 06, 09, 12, 15, 18, 21" Day="00" />
        <Formula>Aggregation.Mean.80</Formula>
        <Parameter>Expected.Values</Parameter>
    </FormulaValue>
</Property>
```

The “Aggregation” module creates a property with name “N”. The value is calculated using the formula “Aggregation.Mean.80” that receives in input an array with values of the property “N” in correspondence of the hourly observations 00, 03, 06, 09, 12, 15, 18, 21 of the day in elaboration. At the same time, the method receives in input the number of expected values for the mean calculation (in this case Expected.Values = 9)

3.
```
<Property>
    <Name>RH06</Name>
    <FormulaValue Type="Fixed" Time="06" Day="0" ReferenceProperty="RH" />
    <NAFormula>
        <FormulaValue Type="Fixed" Time="06" Day="0" ReferenceProperty="D_RH" />
    </NAFormula>
</Property>
```

the “Aggregation” module creates a property with name “RH06”. Its value is retrieved from the property RH at time 06 of the day in elaboration. If the returned value is NA, the application tries to retrieve the value of the property from the property D_RH at time 06 of the day in elaboration (NAFormula).

The Aggregation.XML changes according to the ROI. At the moment, there are two versions, one for Europe and one for China.

3.3.1.1 Aggregated properties for ROI EUROPE

Properties produced by the Aggregation step for the ROI – Europe are shown in Table 5

Table 5 Description of the output variable of the aggregation step

Property description	Acronym	Note	Unit of Measure
Station number	Station	Numerical value	
Day Time	DayTime	Numerical value (YYYYMMDD)	
Total cloud cover	N	Daily mean using observation from 00, 03, 06, 09, 12, 15, 18, 21 UTC of current day, 00 UTC next day	Octas
Low cloud cover	L	Daytime mean using observation from 06, 09, 12, 15, 18 UTC	Octas
Cloud cover daytime mean	NDT	Daytime mean using observation from 06, 09, 12, 15, 18 UTC	Octas
Measured Sunshine	MSUN	See the formula	Hours
Measured Radiation	MRAD	See the formula	MJ/m^2
Calculated Radiation	CRAD	See the formula	MJ/m^2
Minimum Temperature	TN	See the formula	Celsius Degree
Maximum Temperature	TX	See the formula	Celsius Degree
Daily Mean Vapor pressure	MVP	Daily mean using observation from 00, 03, 06, 09, 12, 15, 18, 21 UTC	hPa
Daily Mean Wind speed	FF	Daily mean using observation from 00, 03, 06, 09, 12, 15, 18, 21 UTC of current day, 00 UTC of next day	m/s
Precipitation	RRR	See the formula	mm
Air temperature	TT<hour>	Air Temperature at 06, 09, 12, 15 and 18 of current day	Celsius Degree
Relative Humidity	RH<hour>	Relative Humidity at 06, 09, 12, 15 and 18 of current day	%
Daily Mean Vapor Pressure Deficit	VPD	Daily mean using observation from 00, 03, 06, 09, 12, 15, 18, 21 UTC of current day, 00 UTC of next day	hPa

Property description	Acronym	Note	Unit of Measure
Daily Mean of Slope Saturation	SLOPE	Daily mean using observation from 00, 03, 06, 09, 12, 15, 18, 21 UTC of current day, 00 UTC of next day	hPa/Celsius Degree
ANGOT Radiation	ANGRAD	See the formula	KJ/m^2
Angstrom Prescott Radiation	APRAD	See the formula	KJ/m^2
Supit-Van Kappel Radiation	SVKRAD	See the formula	KJ/m^2
Hargreaves Radiation	HGVRAD	See the formula	KJ/m^2
Penman potential evapotranspiration from a free water surface	E0	See the formula	mm/day
Penman potential evapotranspiration from a moist bare soil surface	ES0	See the formula	mm/day
Penman potential evapotranspiration from a crop canopy	ET0	See the formula	mm/day
Snow depth	SNOW	Instantaneous value at 06 UTC	cm
Visibility	VIS	Formula using observation from 06, 09, 12, 15, 18 UTC	km
Absolut daily minimum temperature	TND	See the formula	Celsius Degree

3.3.1.2 Aggregated properties for ROI CHINA

Properties produced by the Aggregation step for the ROI – CHINA are shown in Table 6

Table 6 Description of the output variable of the aggregation step

Property description	Acronym	Note	Unit of Measure
Station number	Station	Numerical value	
Day Time	DayTime	Numerical value (YYYYMMDD)	
Total cloud cover	N	Daily mean using observation from 18, 21 UTC of previous day, 00, 03, 06, 09, 12, 15, 18 UTC current day	Octas
Low cloud cover	L	Not available	Octas

Property description	Acronym	Note	Unit of Measure
Cloud cover daytime mean	NDT	Daytime mean using observation from 00, 03, 06, 09, 12 UTC	Octas
Measured Sunshine	MSUN	Not available	Hours
Measured Radiation	MRAD	Not available	MJ/m^2
Calculated Radiation	CRAD	See the formula	MJ/m^2
Minimum Temperature	TN	See the formula	Celsius Degree
Maximum Temperature	TX	See the formula	Celsius Degree
Daily Mean Vapor pressure	MVP	Daily mean using observation from 18, 21 UTC of previous day, 00, 03, 06, 09, 12, 15, 18 UTC current day	hPa
Daily Mean Wind speed	FF	Daily mean using observation from 18, 21 UTC of previous day, 00, 03, 06, 09, 12, 15, 18 UTC current day	m/s
Precipitation	RRR	See the formula	mm
Air temperature	TT<hour>	Air Temperature at 00, 03, 06, 09, 12 of current day	Celsius Degree
Relative Humidity	RH<hour>	Relative Humidity at 00, 03, 06, 09, 12 of current day	%
Daily Mean Vapor Pressure Deficit	VPD	Daily mean using from 18, 21 UTC of previous day, 00, 03, 06, 09, 12, 15, 18 UTC current day	hPa
Daily Mean of Slope Saturation	SLOPE	Daily mean using observation from 18, 21 UTC of previous day, 00, 03, 06, 09, 12, 15, 18 UTC current day	hPa/Celsius Degree
ANGOT Radiation	ANGRAD	See the formula	KJ/m^2
Angstrom Prescott Radiation	APRAD	See the formula	KJ/m^2
Supit-Van Kappel Radiation	SVKRAD	See the formula	KJ/m^2
Hargreaves Radiation	HGVRAD	See the formula	KJ/m^2
Penman potential evapotranspiration from a free water surface	E0	See the formula	mm/day

Property description	Acronym	Note	Unit of Measure
Penman potential evapotranspiration from a moist bare soil surface	ES0	See the formula	mm/day
Penman potential evapotranspiration from a crop canopy	ET0	See the formula	mm/day
Snow depth	SNOW	Instantaneous value at 00 UTC	cm
Visibility	VIS	Formula using observation from 00, 03, 06, 09, 12	km
Absolut daily minimum temperature	TND	See the formula	Celsius Degree

3.3.2 Aggregation formulas

For each property produced by the Aggregation step, the module is calling a formula.

- Air temperature (TT)
- Dew point temperature (TD)
- Maximum 1h temperature (TX1)
- Minimum 1h temperature (TN1)
- Relative humidity (RH)
- Calculated relative humidity (D_RH)
- Wind speed (FF)
- Slope (D_SLOPE)
- Cloud cover (N)
- Low cloud cover (L)
- Vapor pressure deficit (D_VPD)
- Vapor pressure (D_E)
- Visibility (VIS)

Before applying those aggregation formulas, some hourly properties may undergo a temporal interpolation procedure. To apply the interpolation is necessary to have at least 80% of valid values in the interval configured in the Aggregation.XML file.

3.3.2.1 Aggregation rules for ROI – EUROPE

We describe below the rules to derive the aggregated values for the ROI - Europe:

- Total cloud cover = mean of hourly cloud cover observations, (property N) : 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC, 21 UTC – current day, 00 UTC – next day
- Cloud Cover Daytime mean = mean of cloud cover values for specific times, using the observations (property N): 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC – current day
- Low Cloud Cover Daytime mean = mean of low cloud cover values for specific times, using the observations (property L): 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC – current day

- Measured sunshine. The system first looks for a 24-hour sunshine observation between the observation at 06 of the next day and sunset time of the current day. If it does not return a valid value, the system tries to derive the measured sunshine as follows: it retrieves the sunrise and the sunset time for the current day. If not available, it considers the interval 06 – 21 of the current day.

If all the hourly observations in the interval determined at the previous point are available, the measured sunshine is calculated as the sum of hourly sunshine observations.

- Measured radiation. The system first looks for a 24-hour radiation observation between the 00 of the next day and the sunset time of the current day. If it does not return a valid value, the following algorithm is applied: it retrieves the sunrise and the sunset time of the current day. If not available, it considers the interval 06 – 21 of the current day. If all the hourly observations are available in the interval determined at the previous step, the measured radiation is calculated as the sum of hourly radiation observations.

- Calculated radiation. If the measured radiation is correctly calculated it is also used to return the calculated radiation. Second alternative. If the measured sunshine is valued, then the Angstrom Radiation is calculated and used. Third alternative. If cloud cover, maximum and minimum temperatures are all valid then the Supit radiation is derived and used. Fourth alternative. If cloud cover is not available but minimum and maximum temperatures are, then the Hargreaves radiation is computed and used.

- Minimum temperature. If the TN12 property is available at 06 of the next day, its value is used. If the previous step does not return a valid value, and if the TN6 property is available at 06, 12, 18 ,24 of the current day and 06 of the next day, then their minimum value is taken. If the previous step does not return a valid value and if the hourly observations are available, it gets the minimum from the TN1 values in the interval 18 of the previous day and 06 of the current day. If the previous step does not return a valid value, the application calculates the minimum of all TT values in the interval 18 of the previous day – 06 of the current day. If there is an air temperature value in the interval 18 of the previous day – 06 of the current day that is less than the calculated daily minimum temperature and higher than the minimum temperature from the MOS file at 06 UTC, then the daily minimum temperature is replaced by the minimum temperature from the MOS file at 06 UTC.

- Maximum temperature. The system checks whether the TX12 property is available at 18 of the current day. If not available, it checks for the TX12 property at 15 of the current day. If the previous step does not return a valid value, it checks whether the TX6 property is available at 12, 18 and 24 of the current day and 06 of the next day and gets the maximum value of these four observations. If the previous step does not return a valid value and if the hourly observations are available, it gets the maximum from these TX1 values in the interval 00-23 of the current day. If the previous step does not return a valid value, the application calculates the maximum between all TT values in the interval 06 – 18 of the current day. If there is an air temperature value in the interval 06 of the current day – 18 of the current day that is higher than the calculated daily maximum temperature and lower than the maximum temperature from the MOS file at 18 UTC, then the daily maximum temperature is replaced by the maximum temperature from the MOS file at 18 UTC.

- Daily mean vapor pressure = mean of hourly valid vapor pressure values (D_E) from 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC, 21 UTC –current day, 00 UTC next day.

- Daily mean wind speed = mean of hourly valid wind speed values(FF) from 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC, 21 UTC – current day,00 UTC next day.

- Precipitation. First, the system looks for the 24-hour precipitation value (PR24) at 06 of the next day. Second alternative: sum of the 6-hourly precipitation values (PR06) in the interval 06 UTC – 06 UTC of the next day, if all these values are valid. Third option: sum of the hourly observation (PREC) if all the 24 observations (in the interval 07 of the current day– 06 of the next day) are valid. Fourth option: it uses PR24 at 00 UTC of the next day (if it exists and it is valid). Fifth option: sum of all the

6-hourly observations (PR06) in the interval 00 UTC – 00 UTC of the next day (if all values are valid). Sixth option: sum of all the hourly observations (PREC) in the interval 01 UTC – 00 UTC of the next day (if all 24 observations are valid). If none of the previous step returns a valid value, the algorithm looks for the first hourly observation having both TR & RR (in the interval 06 of the next day and 06 of the current day). According to the time of the identified observation it adapts the length of the interval as follows:

  - If TR & RR is found at either 06 or 03 of the next day, the interval time is 24 hours
  - If TR & RR is found at either 24:00 or 21:00 of the current day the interval time is 18 hours
  - If TR & RR is found at either 18 or 12 of the current day the interval time is 12 hours
  - If all the previous steps fail, the algorithm cannot be applied

- Air Temperature: values at the following time of observation: 06, 09, 12, 15, 18 of the current day.

- Relative Humidity: values at the following time of observation: 06, 09, 12, 15, 18 of the current day. In case those values are either missing or not valid the system provides an estimated relative humidity (D_RH).

- Daily mean vapor pressure deficit = mean of hourly valid vapor pressure deficit values (D_VPD) at 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC, 21 UTC – current day, 00 UTC – next day.

- Daily mean of slope saturation = mean of hourly valid slope saturation values (D_SLOPE) at 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC, 15 UTC, 18 UTC, 21 UTC – current day, 00 UTC – next day.

- Potential evapotranspiration:
  - Penman potential evapotranspiration from a free water surface (E0)
  - Penman potential evapotranspiration from a moist bare soil surface (ES0)
  - Penman potential evapotranspiration from a crop canopy (ET0)

- Snow depth = Instantaneous value of SNOW at 06 of the current day

- Visibility = daytime visibility. Hourly observations in the interval 06 – 18 of the current day.

- Absolute Daily Minimum Temperature. First option: it gets the minimum of TN12 at 12 UTC of the current day and 00 UTC of the next day. Second option: it gets the minimum of TN6 at 06, 12, 18 of the current day and 00 UTC of the next day. Third option: it gets the minimum value of TN1 between 01 of the current day and 00 of the next day. Last option: it gets the minimum value of TT between 01 of the current day and 00 of the next day.

Concerning the mean values, they are calculated only if at least 80% of valid values (in relation to the number of expected values) are available.

3.3.2.2 Aggregation rules for ROI – CHINA

- Total cloud cover = mean of hourly cloud cover observations (property N): 18 UTC, 21 UTC previous day, 00,03, 06, 09, 12, 15, 18 UTC – next day.

- Cloud Cover Daytime mean = mean of cloud cover values from the observations (property N): 00 UTC, 03 UTC, 06 UTC, 09 UTC, 12 UTC – current day.

- Calculated radiation. If the measured radiation is correctly calculated, it is also used as calculated radiation. Second option: if the measured sunshine is valued, the Angstrom Radiation is calculated. Third option: if cloud cover, maximum and minimum temperatures are valid then the Supit radiation is computed. Fourth option: if cloud cover is not available but minimum and maximum temperatures are available, then the Hargreaves radiation is computed.

- Minimum temperature. The system checks if the TN12 property is available at 06 of the current day. If the previous step does not return a valid value, it checks if the TN6 property is available at 06, 12, 18 of the previous day and at 06, 12 of the current day, and gets the minimum value of all these observations. If the previous step does not return a valid value, in case hourly observations are available, it gets the minimum from all the TN1 values in the interval 06 of the previous day and 06 of the current day. If the previous step does not return a valid value, the application calculates the minimum between all TT values in the interval 06 previous day – 06 current day.

- Maximum temperature. The system checks if the TX12 property is available at 18 of the current day. If the previous step does not return a valid value, it checks if the TX6 property is available at 18 of the previous day and at 00,06,12, 18 of the current day and gets the maximum value between the four observations. If the previous step does not return a valid value and in case hourly observations are available, it gets the maximum from all the TX1 values in the interval 18 of the previous day and 18 of the current day. If the previous step does not return a valid value, the application calculates the maximum between all the TT values in the interval 18 previous day - 18 current day.

- Daily mean vapor pressure = mean of hourly valid vapor pressure values (D_E) from 18 UTC, 21 UTC – previous day, 00,03, 06, 09, 12, 15, 18 UTC – next day.

- Daily mean wind speed = mean of hourly valid wind speed values(FF) from 18 UTC, 21 UTC – previous day, 00,03, 06, 09, 12, 15, 18 UTC – next day.

- Precipitation. If available, the system takes the 24-hour precipitation (PR24) at 00:00 of the next day. If not, it uses the PR24 at 21 of the current day.

- Air Temperature: it takes the values at 00, 03, 06, 09, 12 of the current day.

- Relative Humidity. It takes the values at 00, 03, 06, 09, 12 of the current day. In case the reported values of relative humidity are missing or not valid (RH), the system retrieves the calculated relative humidity (D_RH).

- Daily mean vapor pressure deficit = mean of hourly valid vapor pressure deficit values (D_VPD) at 18 UTC, 21 UTC – previous day, 00,03, 06, 09, 12, 15, 18 UTC – next day.

- Daily mean of slope saturation = mean of hourly valid slope saturation values (D_SLOPE) at 18 UTC, 21 UTC – previous day, 00,03, 06, 09, 12, 15, 18 UTC – next day.

- Potential evapotranspiration:

- Penman potential evapotranspiration from a free water surface (E0)
- Penman potential evapotranspiration from a moist bare soil surface (ES0)
- Penman potential evapotranspiration from a crop canopy (ET0)

- Snow depth = Instantaneous value of SNOW at 00 of the current day.

- Visibility = daytime mean of visibility hourly observations using the interval 00 – 12 of current date.

Mean values are calculated only if at least 80% of valid values (w.r.t. the number of expected values) are available.

3.4 Heavy Checks

More complex checks are executed after the aggregation step. At the same time, it is also possible to use a MOS file to fill the gaps. All the errors to be managed are configured in the Messages.HeavyChecks.xml. Each message has a warning level (flag): The following flags can be configured in the XML file: W wrong, S suspicious. The "HeavyChecks" step produces two files. The XML file with a list of stations/observations for which at least an error message exists (either W or S). The DAT file with all stations/observations, transformed as follows:

- If a property has at least one wrong flag, then the property has the value NA and status = "W" (Wrong)
- If a property has only suspicious flags, then the property maintains the value and get the status = "S" (Suspicious)
- If a property has NA and does not have any alerts, then a value from the model file can be used (if it exists). Then, in the XML file this property gets the status = "R" (replaced).

The DAT file is then copied to the history path for future checks (name of the file changed to D.<YYYYMMDD>.hist).

The following heavy checks are currently available:

- Cloud check
  - Values outside the interval [0:8] Octas are considered WRONG.
  - Cloud cover is 0 and precipitation > 0. The value is considered SUSPICIOUS.
  - Cloud cover = 8 and sunshine duration > 0.5 * day sunlight duration. The value is considered SUSPICIOUS.

- Minimum temperature check
  - Minimum temperature outside the interval [-35:35] Celsius Degrees. The value is considered WRONG.
- Maximum temperature checks
  - Maximum temperature outside the interval [-20:50] Celsius Degrees. The value is considered WRONG.
  - Difference between maximum and minimum temperatures outside the interval [0:30] Celsius Degrees. Maximum temperature is considered wrong

- Snow check. If mean temperature is greater than 5 Celsius Degree, the SNOW value is considered WRONG

- Measured Sunshine check
    - If the Measured Sunshine is outside the interval [0:24] hours, its value is considered WRONG
    - Measured Sunshine > 1 and Measured Radiation is 0. The value is considered SUSPICIOUS.

- Measured Radiation. If the value is outside the interval (0:36] MJ/m2, then it is considered WRONG.

- Daily Mean Vapor Pressure. If the value is outside the interval [0:38] hPa, it is considered WRONG.

- Wind speed check. If the value is outside the interval [0:25] m/s, then it is considered WRONG.

- Precipitation check. If the value is outside the interval [0:1400] mm, then it is considered WRONG.

- Air Temperature Checks. All values need to be in the interval [-50:50] Celsius Degrees, otherwise they are considered WRONG.

- Relative Humidity Checks. All values need to be in the interval [0:100] %, otherwise they are considered WRONG.

- Vapor pressure deficit Check. If the value is outside the interval [0:60] hPa, then it is considered WRONG.

- SLOPE check. If the value is outside the interval [0:30] hPa/Celsius Degree, it is considered WRONG.

- Penman Evaporation check. If the value is outside the interval [0:25] mm/day, then it is considered WRONG.

When the HeavyChecks module generates a KO file it is possible to fix the reported errors and run again the procedure by feeding the HeavyChecks module with the new KO file instead of the output from the Aggregation module. If the KO file is empty (i.e., contains only the tag "<Observations />") the module ends the run without executing any checks.

3.4.1 Exceptions management

The HeavyChecks module can manage exceptions for all its checks. All these exceptions are configured in the file HeavyChecks.Exceptions.xml, stored in the configuration path (specified into the command line of the module through the "-c" option). The structure of the XML exceptions is the following one:

```
<Exceptions>
    <Exception>
        <Description>Mountain Station - Maximum Temperature</Description>
```

```
            <Error Property="TX" Code="001" />
            <Property Name="TX">
                <Condition Type="Include" Target="Value" MinEqual="-40" MaxEqual="50" />
            </Property>
            <Condition Type="Include" Target="Altitude" MinEqual="1000" />
        </Exception>
        <Exception>
            <Description>Rusia &amp;amp; Kazakhstan - Maximum Temperature</Description>
            <Error Property="TX" Code="001" />
            <Property Name="TX">
                <Condition Type="Include" Target="Value" MinEqual="-40" MaxEqual="50" />
            </Property>
            <Condition Type="Include" Target="Latitude" MinEqual="40" />
            <Condition Type="Include" Target="Longitude" MinEqual="42" />
        </Exception>
</Exceptions>
```

Tags explanation:

- Error => it configures the error for which the exception is valid. Two attributes (mandatory) can be used for this tag:
  - Property: name of the property for which the error is raised
  - Code: error code
- Property => it defines the property for which the exception is configured. More than one Property tag can be configured. In case the exception contains more tags Property their evaluation is performed by using an AND logical relation. The Property tag must contain the attribute Name with the name of the property.
- Condition => it defines a condition to check, related to a property or to a specific target.
- Target: it defines the attributes related to the property or to the station to be verified.
- Value = value of the property specified in the Property@Name
- Latitude = latitude of the station
- Longitude = longitude of the station
- Altitude = altitude of the station
- Type: it defines how to interpret the condition: “Include” or “Exclude”. The “Exclude” type represents the opposite of “Include”.
- Min: it defines a minimum value to compare with (the value to check must be > Condition@Min)
- MinEqual: minimum or equal comparison (the vale to check must be >= Condition@MinEqual)
- Max: it defines a maximum value to compare with (the value to check must be < Condition@Max)
- MaxEqual: maximum or equal comparison (the value to check must be <= @ConditionMax)
- Equal: fixed value

In case both Min and Max (or MinEqual and MaxEqual) are present the control checks that the value lies in the interval (Min, Max) (or [MinEqual, MaxEqual]). All Condition tags in the exceptions are managed by using a logical AND. When at least one exception is verified, the error is not raised and the value is consider correct.

3.5 Threshold checks

On the output of the “HeavyChecks” step, more complex checks are executed by using the “threshold” files. Those files are created once per year considering historical data from past years and the module TGenerator.

The errors to manage are configured in two files:

- Messages.Daily.ThresholdChecks.xml (file with daily threshold values).
- Messages.Seasons.ThresholdChecks.xml(file with seasonal threshold values)

All files have the same structure w.r.t. the messages. Each message has a warning level that determines how the property value is transferred to the output. Two possible flags can be configured in the XML file: W wrong, S suspicious.

In input, the "ThresholdChecks" module uses two data files that contain daily and seasonal values:

- Threshold.Daily.dat: daily threshold data for each station.
- Threshold.Seasons.dat: seasonal threshold data for each station.

Another input of this module consists of history files, produced by the "HeavyChecks" or Aggregation modules (name D.<YYYYMMDD>.hist). Only the file relative to the previous day of elaboration is used.

The "ThresholdChecks" step produces two files:

XML file with the station observations for which at least one error message exists

DAT file with all stations' observations, transformed as follows. If one property has at least one wrong alert then the property gets value NA. If the property has only suspicious alerts than the property keeps its value.

Currently, the following checks are implemented within the "Threshold Checks" step:

Checks based on daily thresholds

- Daily mean temperature greater than its 99th percentile associated with the day under analysis (MT99 column from the daily threshold file).
- Daily maximum temperature greater than its 99th percentile associated with the day under analysis (MTX99 column from the daily threshold file)
- Daily minimum temperature greater than its 99th percentile associated with the day under analysis (MTN99 column from the daily threshold file).
- Daily mean temperature lower than its 1st percentile associated with the day under analysis (MT1 column from the daily threshold file)
- Daily maximum temperature lower than its 1st percentile associated with the day under analysis (MTX1 column from the daily threshold file).
- Daily minimum temperature lower than its 1st percentile associated with the day under analysis (MTN1 column from daily the threshold file).
- Daily precipitation greater than its 10-year return level of the season under analysis (RRR5Y column from the season threshold file).
- Daily wind intensity greater than its 10-year return level of the season under analysis (FF5Y column from season threshold file).
- Comparison of daily mean temperature predicted by an autoregressive of order 1, AR(1), model having mean daily temperatures from the 5 closest stations as additional covariates. [NOT AVAILABLE YET]

Checks on seasonal thresholds

- Difference in the mean temperatures of two consecutive days greater than the 99th percentile value associated with the season under analysis (MT95 column from season threshold file).
- Difference in the maximum temperatures of two consecutive days greater than the 99th percentile value associated with the season under analysis (MTX95 column from season threshold file).
- Difference in the minimum temperatures of two consecutive days greater than the 99 th percentile value associated with the season under analysis (MTN95 column from season threshold file).
- Difference in the mean temperatures of two consecutive days lower than the 1st percentile value associated with the season under analysis (MT5 column from season threshold file).
- Difference in the maximum temperatures of two consecutive days lower than the 1 st percentile value associated with the season under analysis (MTX5 column from season threshold file).
- Difference in the minimum temperatures of two consecutive days lower than the 1 st percentile value associated with the season under analysis (MTN5 column from season threshold file).

- Daily wind speed value greater than the 10-year return level of the season (RRR5Y column name). In this case a SUSPICIOUS flag is raised.
- Precipitation daily value greater than the 10-year return level of the season (FF5Y column name). In this case a SUSPICIOUS flag is raised.

As the ThresholdChecks module generates a KO file, it is possible to fix the reported errors and run the module again. Then the input is the KO file with fixed errors. If the KO file sent as input is empty, (contains only the tag "<Observations />") the module ends without executing any checks.

3.5.1 Threshold file generator

The generator of the threshold files (daily and seasonal) is a module that does not run in NRT. It runs once per year and creates two files.

Daily threshold file: it contains for each station and for each day, the properties shown in Table 7.

Table 7 Description of the daily threshold file

Property description	Column	Note	Unit of Measure
Station number	Station	Numerical value	
Day	Day	Julien day value for referenced year	
Mean temperature 99 percentile	MT99	Numerical value	Celsius Degree
Mean temperature 1 percentile	MT1	Numerical value	Celsius Degree
Maximum temperature 99 percentile	MTX99	Numerical value	Celsius Degree
Maximum Temperature 1 percentile	MTX1	Numerical value	Celsius Degree
Minimum Temperature 99 percentile	MTN99	Numerical value	Celsius Degree
Minimum Temperature 1 percentile	MTN1	Numerical value	Celsius Degree
Mean temperature prediction indicators using the mean temperature of 5 closest station [NOT AVAILABLE YET]			
1st indicator	MTC5_1	Numerical value	Celsius Degree
2nd indicator	MTC5_2	Numerical value	Celsius Degree
3rd indicator	MTC5_3	Numerical value	Celsius Degree
4th indicator	MTC5_4	Numerical value	Celsius Degree
5th indicator	MTC5_5	Numerical value	Celsius Degree
6th indicator	MTC5_6	Numerical value	Celsius Degree

The values for the temperature columns are calculated by using a complex approach described in the companion paper with data in the interval 1985 – 2018. The daily threshold file contains, for each station, the mean temperatures values for each day of the analyzed year.

The Seasonal threshold file contains for each station and for each season (winter, spring, summer, autumn), the properties shown in Table 8.

Table 8 Description of the seasonal threshold file

Description	Column	Note	Unit of Measure
Station number	Station	Numerical value	
Season	Season	Text value with possible values: winter / spring / summer / autumn	
Mean temperature 99 percentile	MT95	Numerical value	Celsius Degree
Mean temperature 1 percentile	MT5	Numerical value	Celsius Degree
Maximum temperature 99 percentile	MTX95	Numerical value	Celsius Degree
Maximum Temperature 1 percentile	MTX5	Numerical value	Celsius Degree
Minimum Temperature 99 percentile	MTN95	Numerical value	Celsius Degree
Minimum Temperature 1 percentile	MTN5	Numerical value	Celsius Degree
Precipitation 10-year return levels of the season	RRR5Y	Numerical value	mm
Wind speed 10-year return levels of the season	FF5Y	Numerical value	m/s

The values of all columns are calculated by using the approach described in the companion paper with data in the interval 1985 –2018 for each station and for each season. For each ROI (Europe, China) a set of threshold files (daily, seasonal) exist.

3.5.2 Convex Hull Exceptions management

The threshold checks module does not generate warnings when the same warning message is identified at many stations located within a convex hull with certain attributes. This step requires the presence of the file ConvexHull.Threhsolds.xml into the CONFIG PATH of the module. The file contains a configuration for each error for which the convex hull approach must be applied. This is an example of the configuration file:

```
<Exception>
   <Area>Daily</Area>
   <Error Property="TX" Code="001" />
   <Stations_Counter>10</Stations_Counter>
   <Hull_Radius>150000;200000</Hull_Radius>
   <Hull_Area>71000</Hull_Area>
</Exception>, where
```

- <Area> Either Daily or Season
- <Error> contains the attributes to identify the error.
- Property: the name of the property for which the error is raised.
- Code : error code.
- <Stations_Counter> the minimum number of stations.
- <Hull_Radius> a list of radius (in meters), separated by ';', to be used for the cluster calculation.
- <Hull_Area> the minimum limit of the convex-hull area.

The convex-hull approach is applied at the end of the Threshold Checks module. Checks are run only if the number of stations having the error is greater/equal the value configured in the tag "Stations_Counter". Once the stations' number is reached the module derives clusters (based on stations' locations) and for each cluster estimates the convex hull. Then, it checks whether the area of the convex hulls is greater than a value configured into the tag "Hull_Area" (measured in Km2). Once a valid convex hull area is identified, the SUSPICIOUS flag assigned to all stations within the convex hull is removed. To keep track of those stations, the list of stations for which the flag was removed is saved into the module log file.

3.6 Flags management

Quality checks flags provide useful additional information. Quackme produces the following quality flags:

- H (human change): the value of a property has been modified by an expert (e.g., meteorologist).
- W (wrong value)
- S (suspicious value)
- I (interpolated value)
- R (replaced by MOS value). Currently, this flag can be assigned only to minimum and maximum temperature daily values when the MOS value replaces the calculated station value (Aggregation module).

Flags are produced first by the WeakChecks module at hourly level. The flags associated with the sub-daily values used in the aggregation determine the flag for the daily value. For example, if a daily value is derived by using nine hourly values having flags, the daily value will contain a combination of all these hourly flags separated by the character '|'.

The HeavyChecks manages the flags, and it is based on the flags generated by the Aggregation module. At this point the flag 'H' (if entered following changes done by a meteorologist) resets all the other flags associated with the value.

In the same way, the ThresholdChecks manages the flag starting with the flags generated by the HeavyChecks module. The flags file produced by the ThresholdChecks module (Flags.ThresholdChecks.<YYYYMMDD>.dat) represents the final set of flags for the elaboration date.

3.7 SConverter module

The main scope of this module is to convert the aggregation results (after the heavy checks and possibly the threshold checks) into the S-File format as shown in Table 9.

Table 9 Structure of the S-File

Property description	Length	Note	Unit of Measure	Digits
Station_Number	01:07	Station number		N.A.
Day	08:17	Day number of the year (Julien format)		N.A.
CLOUD_24_TOTAL	18:23	Daily mean of total cloud cover (mean over 24 hours)	Octas	1
SUNSHINE	24:29	Sunshine duration (sum of 24 hours)	Hours	1
RAD_MEA	30:35	Measure global radiation (sum	MJ/m2	1

Property description	Length	Note	Unit of Measure	Digits
		of 24 hours)		
TEMP_MIN	36:41	Minimum temperature (last 12 hours in Europe)	Celsius Degree	1
TEMP_MAX	42:47	Maximum temperature (last 12 hours in Europe)	Celsius Degree	1
VAP_PRES	48:53	Daily mean vapour pressure (mean over 24 hours)	hPa	1
WIND_10	54:59	Daily mean speed at 10 metters (mean over 24 hours)	m/s	1
RAIN	60:65	Amount of precipitation (sum over 24 hours	Mm/d	1
TEMP_06	66:71	Air temperature morning (06 UTC in Europe)	Celsius Degree	1
HUM_06	72:77	Relative humidity morning (06 UTC in Europe)	%	0
TEMP_09	78:83	Air temperature late morning (09 UTC in Europe)	Celsius Degree	1
HUM_09	84:89	Relative humidity late morning (09 UTC in Europe)	%	0
TEMP_12	90:95	Air temperature midday (12 UTC in Europe)	Celsius Degree	1
HUM_12	96:101	Relative humidity midday (12 UTC in Europe)	%	0
TEMP_15	102:107	Air temperature afternoon (15 UTC in Europe)	Celsius Degree	1
HUM_15	108:113	Relative humidity afternoon (18 UTC in Europe)	%	0
TEMP_18	114:119	Air temperature evening (18 UTC in Europe)	Celsius Degree	1
HUM_18	120:125	Relative humidity evening (18 UTC in Europe)	%	0
STATE_SOIL	126:131	State of soil		1
VAP_PRES_Def	132:137	Daily mean vapour pressure deficit (mean over 24 hours)	hPa	1
SLOPE_VP_VS_T	138:143	daily mean slope saturation vapour pressure vs. temperature curve	hPa/Celsius Degree	3
CLOUD_DAYTIME_TOTAL	144:149	daytime mean of total cloud cover (mean over daytime hours)	Octas	1
CLOUD_DAYTIME_LOW	150:155	daytime mean amount of CL clouds or, if no CL clouds are present, the daytime mean amount of CM clouds (mean	Octas	1

Property description	Length	Note	Unit of Measure	Digits
		over daytime hours)		
CLOUD_SHADOW	156:161	daytime mean amount of shadow clouds (mean over daytime hours)	Octas	1
	162:167	Calculated sunshine duration (NOT SAVED INTO DB)	%	1
	168:173	Highest possible global radiation at clear sky Crad (NOT SAVED INTO DB)	MJ/m2	1
	174:179	Potential evapotranspiration ETP (NOT SAVED INTO DB)	mm/day	1
VISIBILITY	180:185	daytime mean visibility (daytime mean)	km	1
SNOW_DEPTH	186:191	Snow depth (instantaneous value 06 UTC in Europe)	cm	1
RAIN_MOS	192:197	indicator for reference period and usage of short range forecast in daily sum of precipitation RAIN		N.A.

The SConverter module ensures that the conversion from the JRC format to the S-File format respects the following convention name: S<YYYYMMDD>.dat, where YYYYMMDD represents the day of the input file (the day of the observations contained in the input file).

3.8 RRR file generation module

This module produces the 3h/6h precipitation data files. It can run only after the WeakChecks module because it needs the hourly history files. As an input, the module takes two hourly history files:

- H.<reference date>.hist
- H.<reference date – 1 day>.hist

where both files are stored in the CONFIG path (specified through parameters into the command line).

The module also takes as an input MOS files:

- M.MG.<reference date>.csv
- M.MG.<reference date – 1 day>.csv

As output, the module produces 4 files for 6h mode and 8 files for 3h mode, starting with time 00 of the reference date (day for which the module is running) for every 3/6 hours. The file of 3/6 hour precipitation has the structure shown in Table 10.

Table 10 Structure of the RRR files

Description	Column	Note	Unit of Measure
Station	Station	Numerical value	
Day time	DayTime	Day time of observation in the format	

		YYYY-MM-HH H24:mm	
Precipation value	RRR	Numerical value	mm
Flag	Flag	Numerical value	

Flags can have the following values:

- 1=Changed by meteorologist (not applicable)
- 2=Automatically corrected (not applicable)
- 3=Observation
- 4=Linear interpolation from observations
- 5=Interpolated via MOS from observations
- 6=MOS analyses (not managed yet)

For each time interval (3 or 6 hours), precipitation values are derived as follows:

3-hourly interval:

- If for the specific time (09, 12, 15, 18, 21, 00 next day, 03 next day, 06 next day) all the observation from the previous 3 hours are not available (specific time -2 hours), the date is ignored.
- If both RR & TR are correctly valued and RR = 3 then its value is used.
- If all the 3 observations (starting for the time under elaboration) in the interval [elaboration time – 2 hours: elaboration time] are available with valid values, then it sums these values. In case one value was interpolated, it associates flag 4 else 3.
- If only one of the 3 observations is not available (with the other 2 have valid values), then it tries to retrieve the missing value from MOS. If available and valid, then it sums the two valid observations and the MOS value. In this case the associated flag is 5.
- If the RR/TR are both valid and TR is equal either to 6 or 12, the system tries to calculate 3h precipitation value as follows
    - If TR = 6, it calculates the ratio of the precipitation by using MOS 1H precipitation values:
        - Ratio for the first 3-hour interval = (sum of 1H precipitation of the first 3 hours in the interval / sum of 1H precipitation of the 6 hours interval) * RR value
        - Ratio for the second 3-hour interval = (sum of 1H precipitation of the second 3 hours in the interval / sum of 1H precipitation of the 6 hours interval) * RR value
    - If TR = 12 it applies the same logics of the previous point but the RR Value is now the difference between the current RR (the one with TR = 12) and the RR observation corresponding to 6 hours before the current time (if that one has TR = 6). In both cases the flag is 5

6-hour interval:

- If for the specific times (12, 18, 00 next day 06 next day) observations do not exist, the station is ignored and NA is reported.
- For the observation corresponding to the time in elaboration, it checks if the PR06 column contains a valid value and uses it
- If TR = 6, then it uses the value of RR if valid.
- It checks if all the 6 observations (starting from the time in elaboration) in the interval [elaboration time – 5 hours : elaboration time] (e.g. [19:00 previous day – 00:00 current day]) are available. It verifies if TR = 1. If the RR values are all valid, then the SUM(RR) is used. Else, if the PREC column contains valid values for all the 6 observations, then SUM(PREC) is used. If one of these values is interpolated, then the flag is set to 4.
- In case none of the previous steps produces a valid value, it checks the possibility of 12-hour precipitation split. Two cases are possible:
    - When TR = 12 at the current elaboration time, it considers the MOS ratio of precipitation using MOS values (RR) at the current elaboration time and at the current elaboration time - 6 hours. The reported value is calculated as follows:

        - The MOS ratio = RR MOS value at the same time of the observed data in elaboration / (sum of RR MOS values found at the current elaboration time and current elaboration time – 6 hours). The reported value = RR of observed data * MOS ratio
    - If TR = NA at the current elaboration time, it uses the MOS ratio of precipitation by using MOS values (RR) at the current elaboration time and at the current elaboration time + 6 hours, only if the observation corresponding to +6 hours has TR = 12 and a valid RR. The reported value is calculated as follows:
        - The MOS ratio = the RR MOS value at the same time of the observed data in elaboration / (sum of RR MOS values at the r current elaboration time and at current elaboration time + 6 hours). The reported value = RR of observed data * MOS ratio. When the MOS data are used, the reported flag is 5.

The names of the output files adhere the following naming convention:

rrr_<yyyyMMddHH>.txt – for 6H precipitation,

rrr3h_<yyyyMMddHH>.txt – for £H precipitation,

, where:

- yyyy = the year of reference date (4 digits)
- MM = the month of reference date (2 digits)
- dd = the day of reference date (2 digits)
- HH = the time

These files are created on the OUTPUT path specified in the command line.

3.9 GUI

The GUI has been designed to enable meteorologists to correct the identified errors, validate the flags and in case remove values. Both the input and the output are in XML format.

The XML file has the following structure:

```
<Observations>
    <Observation>
        <Station> station code </Station>
        <DayTime>day time with the format YYYYMMDDHH </DayTime>
        <TAG Status="status of the value for current property"> value </TAG>
        .....
        <TAG Status="status of the value for current property"> value </TAG>
        <ALERT Message="alert message"
                Type="type of the alert"
               Property="property name to which the alert refers"
             Status="proposed status for the property" />
      .....
        <ALERT Message="alert message"
                Type="type of the alert"
              Property="property name to which the alert refers"
            Status="proposed status for the property" />
    </Observation>
```

</Observations>

The XML tags are:

- TAG is an XML tag with the name of the property (for example: TT – air temperature, Snow – show depth)
- Status can get the following values:
  - C – correct
  - W – wrong
  - S – suspicious
  - R - replaced
- Inside the ALERT, the LEVEL attribute provides the flag of the alert: W – wrong value, S – suspicious value, R – replaced (this value isused only for the HeavyChecks output).
- Property – contains the name of property (must exists a tag with same name in the same Observation parent node).

Example of the XML with errors:

<Observations>

<Observation>

<Station>15490</Station>

<DayTime>2017110206</DayTime>

<TTT Status="C">-4.46</TTT>

<TD Status="C">-3.46</TD>

<Prec Status="C">0</Prec>

<Snow Status="W">5</Snow>

<FF Status="C">1.7</FF>

<ALERT Level="W" Property="Snow" Message="a decrease of snow depth of more than 50cm is very unlikely under any circumstance"/>

</Observation>

<Observation>

<Station>15491</Station>

<DayTime>2017110206</DayTime>

<TTT Status="C">-4.46</TTT>

<TD Status="C">-3.46</TD>

<Prec Status="C">0</Prec>

<Snow Status="C">50</Snow>

<FF Status="S">1</FF>

<ALERT Level="S" Property="FF" Message="Suspicious value when difference relative to previous observation are great than 25 knots"/>

</Observation>

<Observation>

<Station>15491</Station>

<DayTime>2017110212</DayTime>

<TTT Status="C">0.45</TTT>

<TD Status="C">0.42</TD>

<Prec Status="C">0</Prec>

<Snow Status="C">70</Snow>

<FF Status="S">0</FF>

<ALERT Level="S" Property="FF" Message="Suspicious value in correspondence with previous and after observations"/>

<ALERT Status="A" Property="Snow" Type="W" Message="Fresh snow must originate from precipitation. Normally, 1mm precipitation results in 1cm snow. It is not likely that this ratio is highly exceeded."/>

</Observation>

<Observation>

<Station>15491</Station>

<DayTime>2017110221</DayTime>

<TTT Status="C">2.02</TTT>

<TD Status="C">1.82</TD>

<PREC Status="C">101</Prec>

<SNOW Status="S">500</Snow>

<FF Status="C">2</FF>

<ALERT Level="S" Property="Snow" Message="the values are unusually high but yet possible"/>

</Observation>

</Observations>

QUACKME engine can generate errors with two alert levels: S (suspicious), W (wrong). From the QUACKME GUI, the expert can either change the value of the reported property or decide that the value reported like “suspicious” or “wrong” is correct. After the decision, the status of the property can be changed in:

- M (modified) the expert has inserted a new value different from the reported one
- F (forced) the expert has confirmed the reported value as correct.
- W (wrong) the expert leaves unaltered the flag (W). In this case, during the subsequent run, the value is replaced with NA.
- S (suspicious) the expert leaves unaltered the original status (S).

4 Configuration

The software contains configuration files associated to each level of execution:

- Weak Checks
  - One file to configure messages in the current step (Messages.WeakChecks.xml)
  - One file to configure the routines to be executed at each step (Workflow.xml)
  - One file that contains stations' data (Stations.xml)
  - Interpolation configuration file (Interpolation.config)

- Aggregation
  - One file to configure the properties and the associated formulas to derive the aggregated properties (Aggregation.xml)

- Heavy Checks
  - One file to configure the messages in the current step (Messages.HeavyChecks.xml)
  - Use of Aggregation configuration file (Aggregation.xml)

- Threshold Checks
  - Two files to configure the messages at each step:
    - Messages.Daily.ThresholdChecks.xml
    - Messages.Seasons.ThresholdChecks.xml
  - Two files with threshold data
    - Threshold.Daily.YYYY.dat, which contains the daily threshold data for every day of the year
  - Threshold.Seasons.YYYY.dat, which contains the seasonal threshold data relative to the past year.
  - Use of Aggregation or HeavyChecks output: file of the day under analysis.
  - Use of history file (produced at the Aggregation step)

All configuration files are XML format, except for the threshold file having text format.

4.1 Message Structure

The "Messages.xml" file contains a list of MESSAGE tags and it has the following structure:

<Message>

<Property>TD</Property>

<Code>006</Code>

<Text>Dew-point temperature exceeds air temperature by 1C or less: dew-point temperature automatically changed to air temperature</Text>

<Level>A</Level>

</Message>, where:

- "Property" tag contains the name of the property to which it refers
- "Code" represents the code of the message (a sequential number related to each property)
- "Text" contains the full description of the error
- "Level" contains the warning level: A – alert, S – suspicious, W – wrong value.

- "Values" (only for Messages.HeavyChecks.xml, Messages.Daily.ThresholdChecks.xml, Messages.Seasons.ThresholdChecks.xml) contains a list of property to write (like <PropertyName>=<PropertyValue>) into the output KO XML file.

The text message can contain dynamic values such as #0#, #1# … #n# that must be transferred inside the source code. The #0# will be replaced by the first value received, the #1# by the second one, and so on.

The KO XML file, if empty, contains only the tag "<Observations />".

4.2 Message behaviour

At the start of each check step the engine merges the original input data and the events corrected by the expert though the QUACKME GUI. The following rules hold:

- If a property is reported and corrected by the expert with status F then it is considered correct.
- If a property is reported and corrected by the expert with status M then the new value (assigned by the operator) is moved to the input data and considered correct.
- If a property is reported with the status W (wrong) and left unchanged by the expert, then its value is discarded and replaced by NA (NOT DEFINED).
- If a property is reported with the status S (suspicious) and left unchanged by the expert, then its value remains the same for the next run.

For all runs after the first, it exists the possibility to generate the same message for a specific property of an observation. In this situation, two rules hold:

- If the property is corrected by the expert with status F (forced) then the message is ignored.
- If the property is corrected by the expert with status M (modified) the message is then reported again.

4.3 Workflow configuration structure

The "Workflow.xml" is organized by levels. All levels have the same structure:

```
<Check>
    <Name>Dew Point</Name>
    <Module>DewPoint_WeakChecks.R</Module>
    <Method>WeakChecks.CheckDewPoint</Method>
    <Name>ON/OFF</Name>
</Check>, where:
```

- Name: check description
- Module: name of the module that contains the checks
- Method: name of the method to call for the current check
- Status: ON – the check module needs to be considered, OFF – the check module is ignored

A parent tag specific for each level includes all the checks relative to that level:

- WeakChecks for the weak checks level
- HeavyChecks for the heavy checks level
- ThresholdChecks for the checks using threshold data

5 Conclusions

QuackMe is a flexible, dynamic, easily customisable software for the quality checks of meteorological data. It has been developed for the JRC crop monitoring and forecasting system, however it can be successfully applied in other services and for other regions of the world. As the system is operationally used by JRC, it is under continuous improvement and maintenance. The open-source code is freely available and can be retrieved in the dedicated github page of JRC.

.

List of abbreviations and definitions

AMDAC	Actual Meteorological DAtabase Construction
ECMWF	European Centre for Medium-Range Weather Forecasts
MOS	Model Output Statistics
QuackMe	QUAlity ChecKs of MEteorological Data
ROI	Region of Interest
WMO	World Meteorological Organization

List of figures

List of tables

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-01-22-047-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub

ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office of the European Union

doi:10.2760/96559

ISBN 978-92-76-46870-7