DAYCLI Message

Table of Contents

1 Definitions

National Meteorological and Hydrological Services (NMHSs) have different methods and different capacities on climatological data control. Without a current world standard on climatological data control processes it is not possible to have a strict and objective way to qualify data for the DAYCLI message.

The approach followed to qualify in terms of quality the data that will be exchanged at the global level with the DAYCLI message is:

- the NMHS is the one that has the best knowledge on the quality of the data it holds;
- to remain as simple as possible
- to allow everyone to give the best possible information accompanying the data;
- to propose a framework on data quality coding that can eventually evolve towards homogeneous procedures.

Data checked
Someone or something with relevant authority has looked at the data and declared to be fit (or unfit) for at least immediate operational use.

Doubtful data
Is representative of a low confidence. Only suspect data is supposed to be exchanged through the DAYCLI message.

Data provider

Suspect data
Is representative of a high and medium confidence.

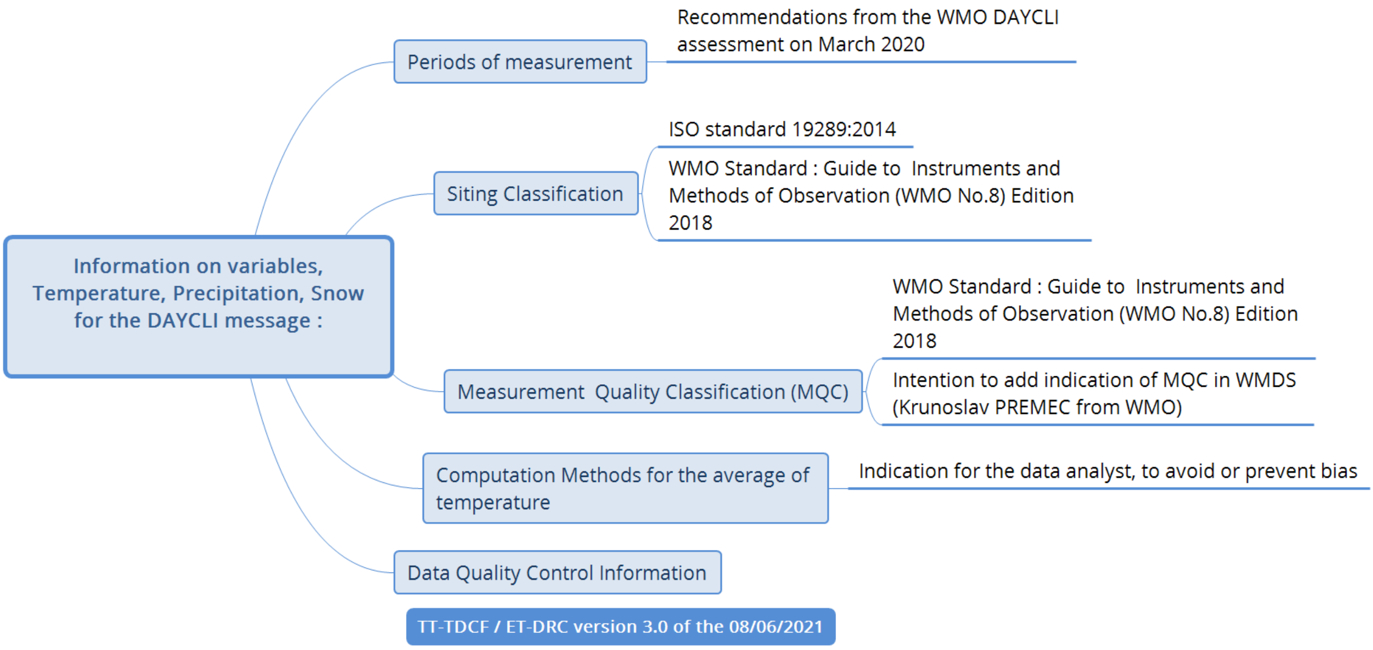
DS
RR
TM
TN
TSD
TX

Time Zone or TZ (Wikipedia, 25/06/2021)

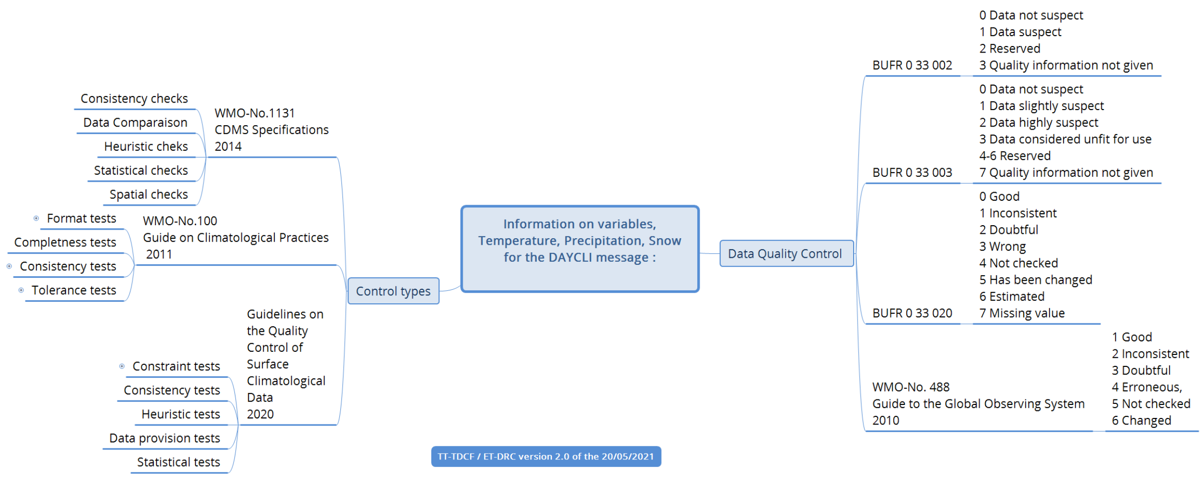
A time zone is an area that observes a uniform standard time for legal, commercial and social purposes. Time zones tend to follow the boundaries between countries and their subdivisions instead of strictly following longitude, because it is convenient for areas in frequent communication to keep the same time.

All time zones are defined as offsets from Coordinated Universal Time (UTC), ranging from UTC−12:00 to UTC+14:00. The offsets are usually a whole number of hours, but a few zones are offset by additional 30 or 45 minutes, such as in India and Nepal.

2 DAYCLI: Type of information



3 DAYCLI: References



4 Periods of measurement

The difference in hours and minutes from Coordinated Universal Time (UTC), from the westernmost (−12:00) to the easternmost (+14:00). As, Baker island UTC -12 and Kiribati +14 UTC.

The period of measurement of each parameters is specify in the DAYCLI message by 2 fields:

- The beginning of the period of measurement (TX_HOUR, TN_HOUR, RR_HOUR, DS_HOUR, and TSD_HOUR) that represents the beginning time of the measurement in **hour UTC**. To get the ending time you just have to add 24 hours.
- The time displacement that gives from which Day **considering in the time zone of the station**, the beginning of the period of measurement belongs to:
  - -1 means that it belongs to the Day before, D-1
  - 0 means that it belongs to the current Day, D
  - +1 means that it belongs to the next Day, D+1

In this document we consider that TZ Day D is always referring to a time zone day from 00:00 time zone to 23:00 time zone.

The definition of a climatological day for a parameter is not the same for all NMHSs.

The way to store hourly data in a climatological database differs from NMHSs. Some are storing all in UTC while other are using for each station its time zone. See the definition of the time zone in the Definition paragraph. That is the reason why UTC and Time Zone are given in this document.

The most common to represent a climatological day for a location is to refer to its time zone. We usually find two "big" families when a NMHS wants to compute a daily data from its hourly data:

- from 00:00 to 23:00 time zone and
- from 01:00 to 00:00 time zone

Nevertheless, some parameters do not follow those computation rules. It is frequently for the parameters as daily minimum temperature, daily maximum temperature and daily precipitation. This is due to some historical reasons: instrument and/or observer not available at some hours.

Some examples:

Minimum Temperature

COUNTRY	Time Slot
GERMANY	00 TZ / Day D to 23 TZ / Day D
GREECE	01 TZ / Day D to 00 TZ / Day +1
SWITZERLAND	06 TZ / Day D-1 to 06 TZ / Day D
AUSTRIA	18 TZ / Day D-1 to 18 TZ / Day D
AUSTRALIA	09:00 TZ / Day D-1 to 08:59 / Day D

Maximum Temperature

COUNTRY	Time Slot
GERMANY	00 TZ / Day D to 23 TZ / Day D
FRANCE	06 TZ / Day D to 06 TZ / Day D+1
HUNGARY	01 TZ Day D to 00 TZ / Day D+1
AUSTRALIA	09:00 TZ / Day D to 08:59 / Day D+1

Precipitation

COUNTRY	TIME SLOT

NETHERLANDS	01 TZ / Day D to 00 TZ / Day D+1
FINLAND	06 TZ / Day D to 06 TZ / Day D+1
SWAZILAND	08 TZ / Day D to 08 TZ / Day D+1

4.1 Example: TX Max Temperature

In Tahiti, the FAAA station (98714002), UTC -10, is measuring the maximum temperature (TX) from 18:01 UTC (08:01 TZ) the **TZ DAY D** to 18:00 UTC (08:00 TZ) the **TZ DAY D+1**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
TZ	**9**	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	07	08

TX_HOUR=**08:01** ! TX - Hour of Measurement in TZ
TX_DT=**0** ! TX - Time displacement (typical value -1, 0 or 1)

In France, the Blagnac station (31069001), UTC +0, is measuring the maximum temperature (TX) from 06:01 UTC the **TZ Day D** to 06:00 UTC the **TZ Day D+1**. That means taking all those 24 following hours in UTC:

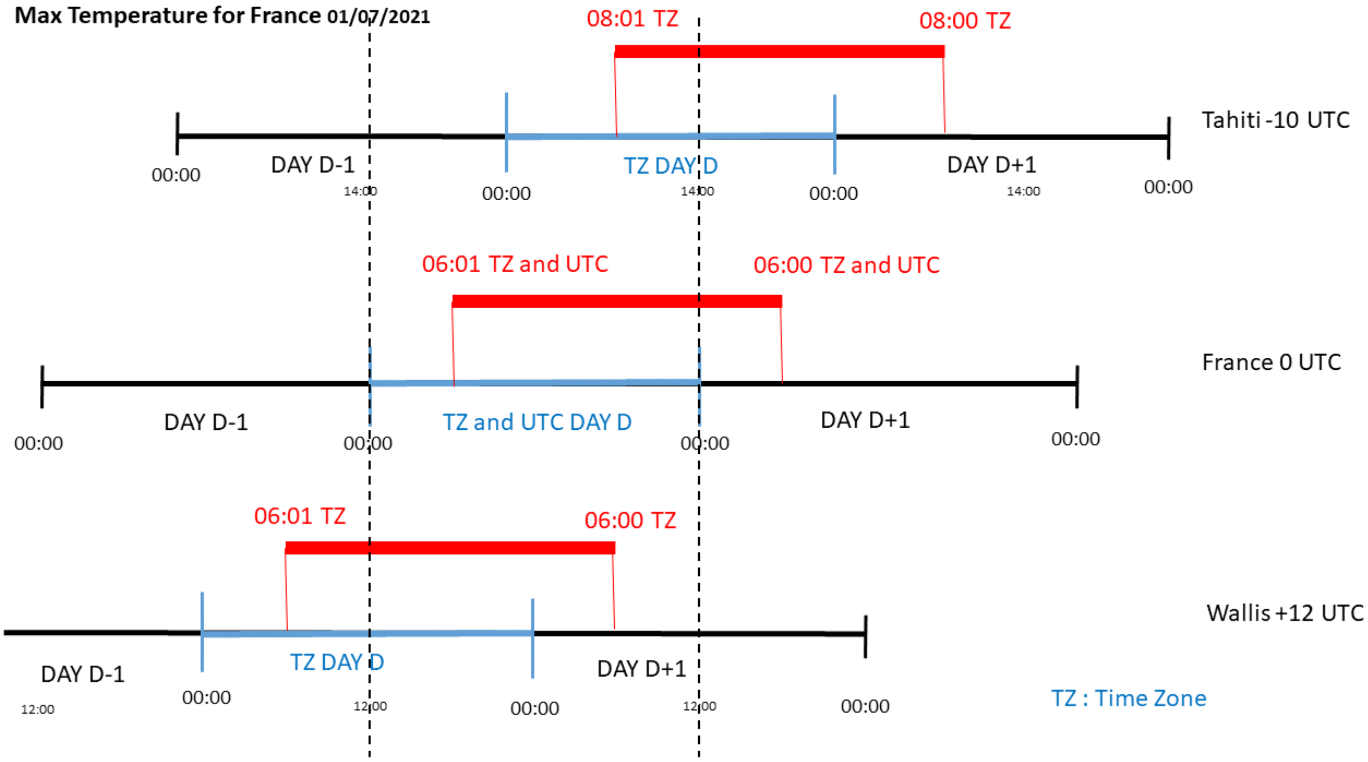
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	**07**	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06

TX_HOUR=**06:01** ! TX - Hour of Measurement in TZ
TX_DT=**0** ! TX - Time displacement (typical value -1, 0 or 1)

In Wallis and Futuna, the HIHIFO station (98613001), UTC + 12, is measuring the maximum temperature (TX) from 18:01 UTC (06:01 Time Zone) the **TZ Day D** to 18:00 UTC (06:00 Time Zone) the **TZ Day D+1**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
TZ	**07**	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06

TX_HOUR=**06:01** ! TX - Hour of Measurement in TZ
TX_DT=**0** ! TX - Time displacement (typical value -1, 0 or 1)


Max Temperature for France 01/07/2021
08:01 TZ
08:00 TZ
Tahiti -10 UTC
00:00
DAY D-1
14:00
00:00
TZ DAY D
14:00
00:00
DAY D+1
14:00
00:00
06:01 TZ and UTC
06:00 TZ and UTC
France 0 UTC
00:00
DAY D-1
00:00
TZ and UTC DAY D
00:00
DAY D+1
00:00
06:01 TZ
06:00 TZ
Wallis +12 UTC
DAY D-1
12:00
00:00
TZ DAY D
12:00
00:00
DAY D+1
12:00
00:00
TZ : Time Zone

4.2 Example: TN Min Temperature

In Tahiti, the FAAA station (98714002), UTC -10, is measuring the minimum temperature (TN) from 06:01 UTC (20:01 TZ) the **TZ DAY D -1** to 06:00 UTC (20:00 TZ) the **TZ DAY D**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10
TZ	**21**	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20

TN_HOUR=**20:01** ! TN - Hour of Measurement in TZ
TN_DT=**-1** ! TN - Time displacement (typical value -1, 0 or 1)

In France, the Blagnac station (31069001), UTC +0, is measuring the minimum temperature (TN) from 18:01 UTC the **TZ Day D-1** to 18:00 UTC the **TZ Day D**. That means taking all those 24 following hours in UTC:

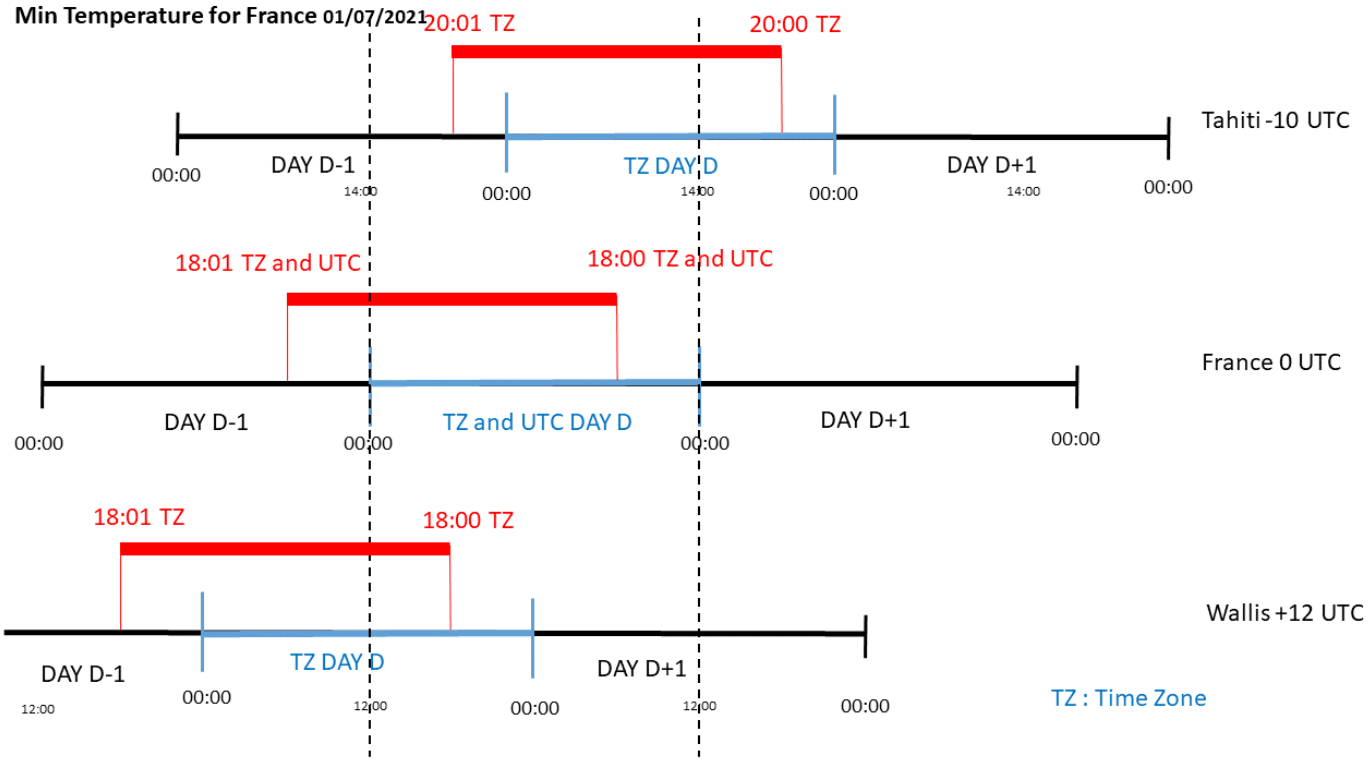
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	**19**	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18

TN_HOUR=**18:01** ! TN - Hour of Measurement in TZ
TN_DT=**-1** ! TN - Time displacement (typical value -1, 0 or 1)

In Wallis and Futuna, the HIHIFO station (98613001), UTC + 12, is measuring the minimum temperature (TN) from 06:01 UTC (18:01 Time Zone) the **TZ Day D-1** to 06:00 UTC (18:00 Time Zone) the **TZ Day D**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06
TZ	**19**	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18

TN_HOUR=**18:01** ! TN - Hour of Measurement in TZ
TN_DT=**-1** ! TN - Time displacement (typical value -1, 0 or 1)



4.3 Example: TM Mean Temperature

In Tahiti, the FAAA station (98714002), UTC -10, is measuring the mean temperature (TM) from 10:01 UTC (00:01 TZ) the **TZ DAY D** to 10:00 UTC (00:00 TZ) the **TZ DAY+1**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10
TZ	**01**	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00

TM_HOUR=**00:01** ! TM - Hour of Measurement in TZ
TM_DT=**0** ! TM - Time displacement (typical value -1, 0 or 1)

In France, the Blagnac station (31069001), UTC +0, is measuring the mean temperature (TM) from 00:01 UTC the **TZ Day D** to 00:00 UTC the **TZ Day D +1**. That means taking all those 24 following hours in UTC:

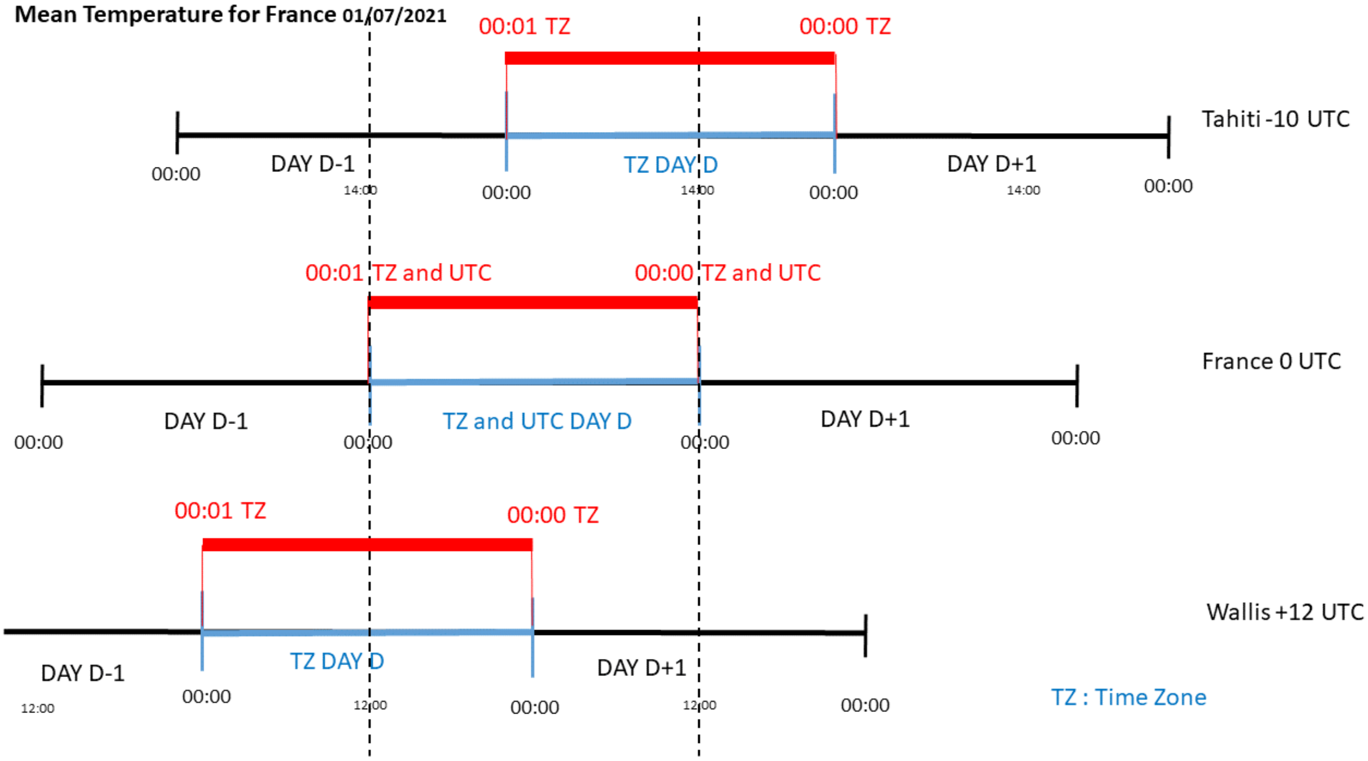
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	**01**	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00

TM_HOUR=**00:01** ! TM - Hour of Measurement in TZ
TM_DT=**0** ! TM - Time displacement (typical value -1, 0 or 1)

In Wallis and Futuna, the HIHIFO station (98613001), UTC + 12, is measuring the mean temperature (TM) from 12:01 UTC (00:01 Time Zone) the **TZ Day D** to 12:00 UTC (00:00 Time Zone) the **TZ Day D+1**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12
TZ	**01**	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00

TM_HOUR=**00:01** ! TM - Hour of Measurement in TZ
TM_DT=**0** ! TM - Time displacement (typical value -1, 0 or 1)


Mean Temperature for France 01/07/2021
00:01 TZ
00:00 TZ
Tahiti -10 UTC
00:00
DAY D-1
14:00
00:00
TZ DAY D
14:00
00:00
DAY D+1
14:00
00:00
00:01 TZ and UTC
00:00 TZ and UTC
France 0 UTC
00:00
DAY D-1
00:00
TZ and UTC DAY D
00:00
DAY D+1
00:00
00:01 TZ
00:00 TZ
Wallis +12 UTC
DAY D-1
12:00
00:00
TZ DAY D
12:00
00:00
DAY D+1
12:00
00:00
TZ : Time Zone

4.4 Example: RR Precipitation

In Tahiti, the FAAA station (98714002), UTC -10, is measuring the precipitation (RR) from 18:01 UTC (08:01 TZ) the **TZ DAY D** to 18:00 UTC (08:00 TZ) the **TZ DAY D+1**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
TZ	**9**	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	07	08

RR_HOUR=**08:01** ! RR - Hour of Measurement in TZ
RR_DT=**0** ! RR - Time displacement (typical value -1, 0 or 1)

In France, the Blagnac station (31069001), UTC +0, is measuring the precipitation (RR) from 06:01 UTC the **TZ Day D** to 06:00 UTC the **TZ Day D+1**. That means taking all those 24 following hours in UTC:

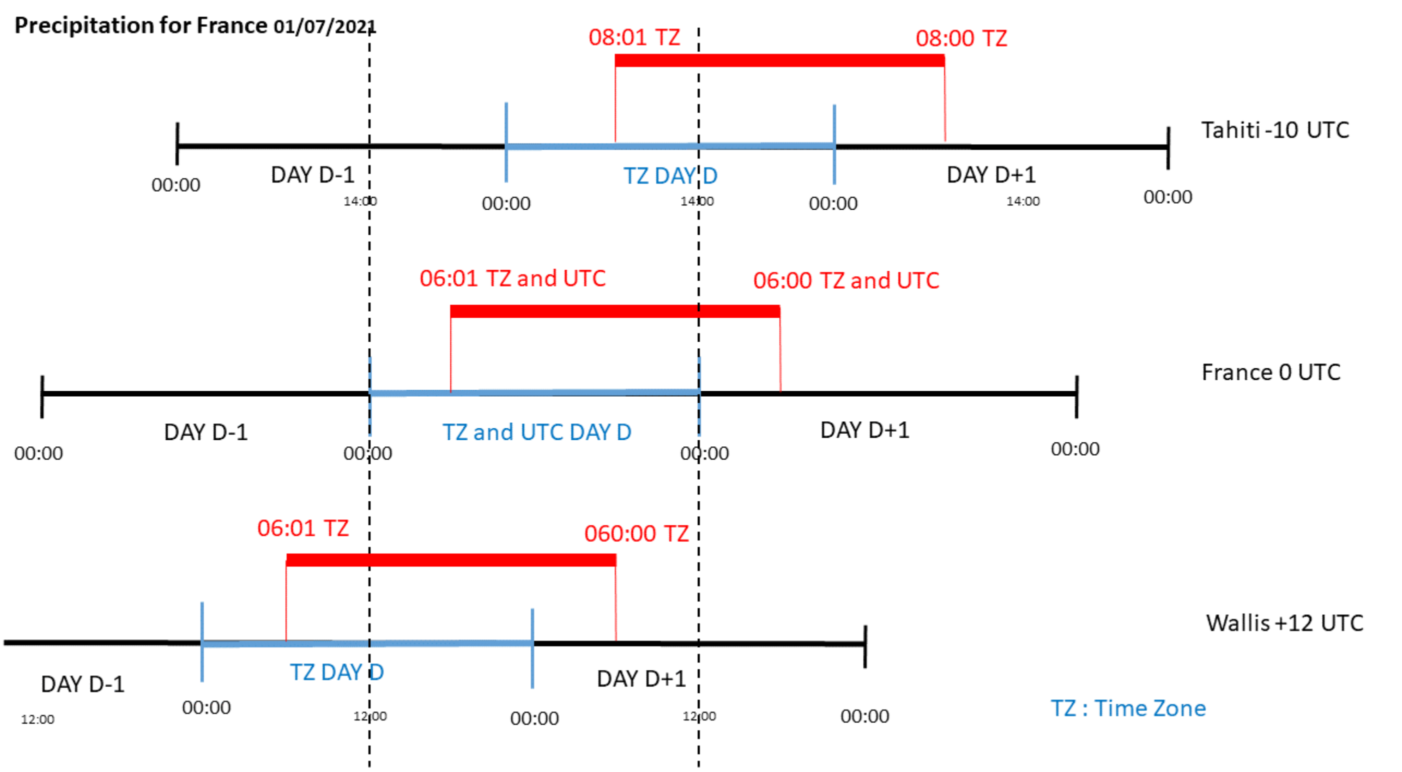
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	**07**	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06

RR_HOUR=**06:01** ! RR - Hour of Measurement in TZ
RR_DT=**0** ! RR - Time displacement (typical value -1, 0 or 1)

In Wallis and Futuna, the HIHIFO station (98613001), UTC + 12, is measuring the precipitation (RR) from 18:01 UTC (06:01 Time Zone) the **TZ Day D** to 18:00 UTC (06:00 Time Zone) the **TZ Day D+1**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
TZ	**07**	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06

RR_HOUR=06:01 ! RR - Hour of Measurement in TZ
RR_DT=0 ! RR - Time displacement (typical value -1, 0 or 1)



Example: DS Depth of Fresh Snow

In Saint Pierre et Miquelon, the St-Pierre station (97502001), UTC -4, is measuring the depth of fresh snow (DS) from 10:01 UTC (06:01 Time Zone) the **TZ Day D** to 10:00 UTC (06:00 Time Zone) the Day **TZ D+1**. That means taking all those 24 following hours in UTC:

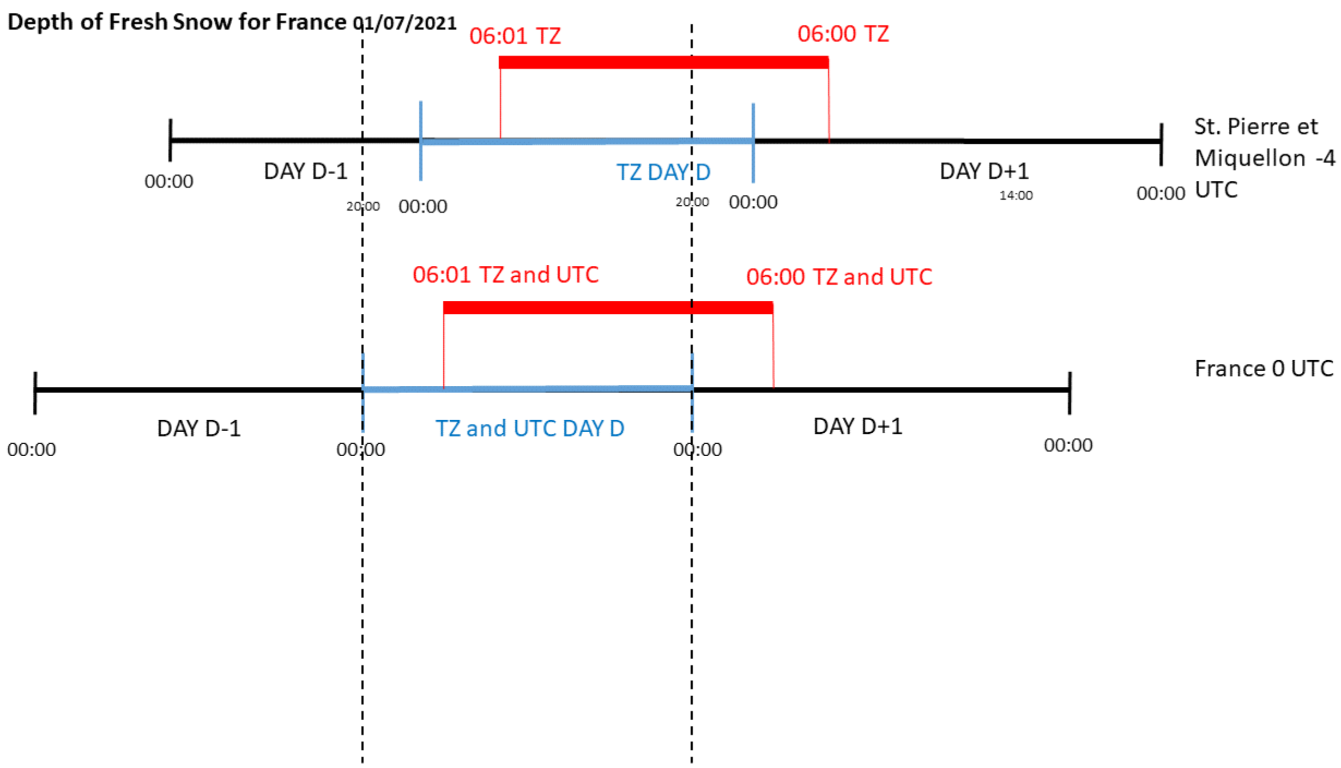
#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06	07	08	09	10
TZ	**07**	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06

DS_HOUR=**06:01** ! Depth of Fresh Snow - Hour of Measurement in TZ
DS_DT=**0** ! Depth of Fresh Snow - Time displacement (typical value -1, 0, 1)

In France, the Blagnac station (31069001), UTC +0, is measuring the depth of fresh snow (DS) from 06:01 UTC the **TZ Day D** to 06:00 UTC the Day **TZ D+1**. That means taking all those 24 following hours in UTC:

#	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
UTC	**07**	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	00	01	02	03	04	05	06

DS_HOUR=**06:01** ! Depth of Fresh Snow - Hour of Measurement in TZ
DS_DT=**0** ! Depth of Fresh Snow - Time displacement typical value -1, 0, 1)



4.5 Example: TSD Total Snow Depth

In Saint Pierre et Miquelon, the St-Pierre station (97502001), UTC -4, is measuring the total snow depth (TSD) at 10:00 UTC (06:00 Time Zone) the **TZ Day D**.

```
TSD_HOUR=06:00          ! total snow depth - Hour of Measurement in TZ
TSD_DT=0                ! total snow depth - Time displacement (typical value -1, 0, 1)
```

In France, the Blagnac station (31069001), UTC +0, is measuring the total snow depth (TSD) at 06:00 UTC the **TZ Day D**.

```
TSD_HOUR=06:00          ! total snow depth - Hour of Measurement in TZ
TSD_DT=0                ! total snow depth - Time displacement typical value -1, 0, 1)
```

5 Siting Classification & Measurement Quality Classification

Code	
0	Reserved
1	1A (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
2	1B (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
3	1C (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
4	1D (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
5	1E (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
6	2A (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
7	2B (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
8	2C (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
9	2D (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
10	2E (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
11	3A (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
12	3B (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
13	3C (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
14	3D (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
15	3E (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
16	4A (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement

Code	
	Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
17	4B (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
18	4C (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
19	4D (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
20	4E (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
21	5A (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
22	5B (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
23	5C (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
24	5D (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
25	5E (Siting Classification according to ISO/WMO standard 119289:2014(E) and Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition)
26	1 (Siting Classification according to ISO/WMO standard 119289:2014(E), Measurement Quality Classification is missing
27	2 (Siting Classification according to ISO/WMO standard 119289:2014(E), Measurement Quality Classification is missing
28	3 (Siting Classification according to ISO/WMO standard 119289:2014(E), Measurement Quality Classification is missing
29	4 (Siting Classification according to ISO/WMO standard 119289:2014(E), Measurement Quality Classification is missing
30	5 (Siting Classification according to ISO/WMO standard 119289:2014(E), Measurement Quality Classification is missing
31	A (Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition), Siting Classification is missing
32	B (Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition), Siting Classification is missing
33	C (Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition), Siting Classification is missing
34	D (Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition), Siting Classification is missing
35	E (Measurement Quality Classification according to the Guide to Instruments and Methods of Observation (WMO-No. 8), 2020 Edition), Siting Classification is missing
36 - 254	Reserved

Code	
255	Missing

6 Computation Method for the average of temperature

CODE TABLE 0-08-094
0-08-094 Method used to calculate the average daily temperature

Code Figure	--
0	Average of maximum and minimum values: Tm = (Tx + Tn) / 2 (see Note 1)
1	Average of the 8 tri-hourly observation
2	Average of 24 hourly observation
3	Weighted average of 3 observations: Tm = (aT1 +bT2 + cT3) (see Note 1)
4	Weighted average of 3 observation and also maximum and minimum values: Tm= (aT1 +bT2 + cT3 +dTx + eTn) (see Note 1)
5	AWS complete integration from minute data
6 - **254**	Reserved
255	Missing value

Add “other computation method”

Note (1) : The letters "a", "b", "c", "d" and "e" generically represent the weight associated with the respective temperature T. The sub-index of T: "1", "2", "3", "x" and "n" represent the values measured at different times or maximum (x) or minimum (n) values.

7 ET-DRC Proposition for Data Quality Control (QC) information

7.1 Code Table for Data Quality Control information

0	Data checked and declared good
1	Data checked and declared suspect
2	Data checked and declared aggregated
3	Data checked and declared out of instrument range
4	Data checked and declared aggregated and out of instrument range
5	Parameter is not measured at the station
6	Daily value not provided
7	Data unchecked
8	-
9	-
10	-
11	-
12	-
13	-
14	
…	…
255	Missing (QC info not available)

7.2 Design of the Data Quality Control information

The intention is to have a maximum of “Data checked and declared good”, code **0**.

The assumption is that the data provider, generally a National Meteorological and Hydrological Service (NMHS), is the entity having the best knowledge on its data status.

The strategy with the DAYCLI message is to share mainly checked data and avoid sharing data that is highly suspect. Nevertheless, suspect data could be shared if the data provider considers that this could have an interest.

The DAYCLI message is intended to be simple to manage and has been designed to do not include all the complexity of the data quality management in place in most Climate Data Management Systems.

The Data Quality Control information is designed for both machine-to-machine exchanges and systems using manual processes.

In the case of a data obviously false, without any interest for DAYCLI users, it is recommended that the data provider does not share this value and declares it missing e.g.

TMin	Associated QC information	QC code
Missing	Daily value not provided	6

In the case of a suspect data that the data provider wants to share e.g.

TMin	Associated QC information	QC code
32.5	Data checked and declared **suspect**	1

In the case that the data provider did not have enough time to control the data, then the data could be coded unchecked, e.g.

TMin	Associated QC information	QC code
11.2	Data unchecked	7

In the case that the data provider did the data control, then made a modification (estimation/reconstruction/interpolation/etc.) on the data, then he could share this data with a code “Data checked and declared good”.

TMin	Associated QC information	QC code
14.5	Data checked and declared good	0

7.3 Codes explanation

7.3.1 <u>Parameter is not measured at the station</u>

The station does not have the instrument to be able to measure this parameter. E.g. no instrument for measuring snow.
it is also possible that the instrument is non-operating at that date.

7.3.2 Daily value not provided

The data is missing for several reasons: transmission problem, observer absent, a data highly suspect that has been removed by the data provider, etc.

7.3.3 Data unchecked

No data control processes have been performed on the data.

Most data should have been undergone with format test (see WMO-No.100) or constraint test (see WMO Guidelines on the Quality Control of Surface Climatological Data, 2021). When assigning “Data unchecked” to a data means no control has been performed on the data except format or constraint tests. E.g. consistency, statistical, spatial controls, etc.

7.3.4 Data checked and declared suspect

Data has been controlled (tests on consistency, statistical, spatial controls, etc.) and has been declared suspect by one of them.

7.3.5 Data checked and declared aggregated

Data has been controlled and has been declared aggregated.

That is typically the case when the amount of precipitation assigned for a day represents the amount of precipitation of several days. . This can also occur for snowfall amount.

In that case, the number of days from which the aggregated value corresponds to should be known.

7.3.6 Data checked and declared out of instrument range

This covers the special case where the total amount of precipitation cannot be determined, but at least exceeds the capacity of the gauge.

7.3.7 Data checked and declared aggregated and out of instrument range

This covers the situation when the amount of precipitation assigned for a day represents the amount of precipitation of several days, accompanied by an overflowed/underestimated case.

7.3.8 Data checked and declared good

Data checked by one or several processes (tests on consistency, statistical, spatial controls, etc.) and declared good by the data provider.

8 Use cases

Note: it is important to distinguish between **Parameter** (e.g. snowfall) not measured, **daily value** not provided, and **QC** not provided (i.e. QC missing = flag 15 in schema).

8.1 Use case: Parameter not measured at the station

Parameter not measured, e.g. snow in a tropical lowland country, or snow is only measured seasonally. In that case, QC code is 1. QC cannot be carried out.

Depth snow value	Associated QC information	QC code
Missing	Parameter is not measured at the station	5

8.2 Use case: Parameter measured but the value missing

Parameter measured, but daily **value** is missing. In this case, QC code is 2. QC cannot be carried out.

TMax	Associated QC information	QC code
Missing	Daily value not provided	6

8.3 Use case: QC not provided

Parameter is measured, daily value is present, but daily QC missing. There is no precision on the QC code. Therefore, QC code is 15. Note that this is separate from the cases where there is an aggregation, as in Use case 5.

TMax	Associated QC information	QC code
32.5	QC not provided	15

8.4 Use case: Overflowed/ Out of instrument range

Daily value is measured, but there is an overflow for a raingauge, or perhaps a blocked gauge in the case of snow, and therefore the best we can say is that we can specify a minimum amount, but no more than that. In that case, the QC code is 7.

In that case, we can understand with high confidence that the measurement exceeds a known minimum, but we don't know what the actual is/was.

Date	Precipitation	Associated QC information	QC code
Day 1	200 mm	Data checked and declared out of instrument range	3

8.5 Use case: Aggregation

Aggregation: In this case, observations are not available for some individual days, but the aggregated total (or highest/lowest value in the case of temperature) are known.

Precipitation: because of an observer problem the rain gauge has accumulated a precipitation over 4 days (from DAY 2 to DAY 5 in the below tables). The precipitation amount for this 4 days are reported at DAY 5

(32.5 mm). From DAY 2 to DAY 5 the QC information is set to 5 “Data checked and declared aggregated”. That allows specifying on how many days the aggregation corresponds to. For Day 6, the problem is solved and there is no precipitation.

Date	Precipitation	Associated QC information	QC code
Day 1	200 mm	Data checked and declared out of instrument range	3
Day 2	Missing	Data checked and declared aggregated	2
Day 3	Missing	Data checked and declared aggregated	2
Day 4	Missing	Data checked and declared aggregated	2
Day 5	32.5	Data checked and declared aggregated	2
Day 6	0	Data checked and declared good	8

8.6 Use case: Aggregation with instrument in maintenance

Note also the case made by Marcus: if we know that missing data over five days, with a total on the sixth, included a period where the station was undergoing maintenance, then we could distinguish in the daily messages between "data missing" during the maintenance period, where observations did not occur (say, on Days 1-2), and aggregated (where we know that on days 3-5 measurements did occur, but we don't know the daily split). Or if the maintenance took place on Days 2-3 the string of flags would be: Day1 -aggregated; days 2-3 missing; days 4-5 aggregated (with the total aggregation on Day 5).

Date	Precipitation	Associated QC information	QC code
Day 1	Missing	Data checked and declared aggregated	2
Day 2	Missing	Parameter is not measured at the station	5
Day 3	Missing	Parameter is not measured at the station	5
Day 4	Missing	Data checked and declared aggregated	2
Day 5	32.5	Data checked and declared aggregated	2

8.7 Use case: Aggregation with overflow/out of instrument range

Here the case with aggregated precipitation and at the end an amount of precipitation that exceeds the raingauge capacity.

Aggregation from DAY 2 to DAY 5, with DAY 5 overflowed.

Date	Precipitation	Associated QC information	QC code
Day 1	3.5	Data checked and declared good	0
Day 2	Missing	Data checked and declared aggregated	2
Day 3	Missing	Data checked and declared aggregated	2
Day 4	Missing	Data checked and declared aggregated	2
Day 5	200 mm	Data checked and declared aggregated and out of instrument range	4

8.8 Use case: Missing values for proper computation

Here the case where some values are missing to be able to compute “perfectly” a daily value. It could be because:

- ✓ the AWS has had a problem and on the 1440 Minute values (60mn*24) needed to compute the daily value, some are missing;
- ✓ on the 24 values of the day, 1 or several values are missing to compute the daily value;
- ✓ on the 8 main synoptic hours of the day, 1 or several values are missing to compute the daily value.

In that case, several possibilities are offered for the data provider:

1. the data provider considers that the data is good, e.g. on the 1440 Minute values 60 are missing, he will code it as **0** : “Data checked and declared good”

TMin	Associated QC information	QC code
11.2	Data checked and declared good	0

2. the data provider considers that the data is suspect but interesting to share worldwide, e.g. on the 24 Hourly values of the day 4 are missing, he will code it as 4 “Data checked and declared suspect”

TMin	Associated QC information	QC code
11.2	Data checked and declared suspect	1

3. the data provider considers that the data is highly suspect, e.g. on the 8 main synoptic hours 6 are missing, he will assume that his data does not represent the reality and he decides to do not share this value with the world. He will assign Missing to the data with code 6 “Daily value not provided”.

TMin	Associated QC information	QC code
Missing	Daily value not provided	6

9 Results on discussions / Comments

9.1 Overflowed / Out of instrument range

William
I am part of a small WMO Task Team trying to decide on the coding format for the new DayCli messages, which as you know will provide a vehicle for submitting daily observations of temperature, snowfall and precipitation, albeit on a monthly basis.

We are trying to agree on some aspects of the quality assurance coding for the DayCli message, representing QC performed prior to submission of the message. One contentious point is the case where a rain gauge overflows, as was commonly the case during the 2010-11 La Nina years. A series of possible options exist for flagging such (admittedly rare) cases:

Our options are:

1. Flag as "missing" (since a value can't be determined);
2. Flag as "doubtful" (with no information on why it is considered doubtful);
3. Flag value representing gauge capacity (203mm) as "doubtful", but with a special flag indicating overflow situation. This indicates that the value is doubtful, but establishes a lower bound, with the true amount higher. - thereby distinguishing from the case where QC operator believes the value is erroneous (e.g. 200mm reported instead of 20mm).
4. Flag as "out of instrumental range". This is essentially the same as Case #3, but is more generic, i.e can apply to snowfall, snow depth, even temperature.

Note that there are separate flags for cases of aggregation.

Since one of the major motivations for introducing DayCli messages was to provide greater visibility to extreme events for climate purposes, I would suspect you'd lean towards Options 3 or 4. But - what do you think?

Blair Trewin answer <blair.trewin@bom.gov.au>
Yes - I think options 3 and 4 are the best ones - having the extreme value is useful for some applications even if it is known to be a lower bound.

9.2 AWS being prone to short term failures

Jose
William's point 3 coincides with what I put in the chat about AWS being prone to short term failures. Depending on the length of such failures, daily data will have different degrees of uncertainty. This connects also with William's point 4 considering the convenience of having more than one degree of "doubtful", so that daily data calculated from AWS measurements with more missing sub-daily data would be more doubtful.

Denis answer
This is a good example for using the code: “Data checked and declared out of instrument range”

Jose answer (mail 14/05/2021)
Your answer 'This is a good example for using the code: “Data checked and declared out of instrument range”' (page 12, bottom) is not appropriate here, since a daily data calculated from incomplete sub-daily AWS measurements has nothing to do with being "out of instrument range". It would just have a degree of uncertainty which would be greater as the number of missing data increase. I suppose there exists a regulation or recommendation on how many (or what proportion) of sub-daily missing data are allowed to calculate the daily value. In the case of 10' measurements (144 values per day), 1 missing data is not a reason to declare the daily data missing, but 100 missing data would be a strong one. As I do not know which is the accepted threshold, let's assume we can calculate the daily value with a maximum of 3 hours of missing data. (That would allow providing a daily value when one of 8 possible 3 hourly records is missing). In that case, 18 missing data (3 hours of 10' measurements) would have the maximum

doubtfulness, while 1 missing data would have the maximum reliability. Should we distinguish such different levels of doubtfulness as William is suggesting? If accepted, a simple two or maximum three categories of suspect could be established based on the departure of the measurement from its expected value or the number of non-passed QC tests.

Denis answer
Yes you are right. The “Data checked and declared out of instrument range”' is not appropriate here.
What you are talking here is a data control based on data completeness or also rules on missing data. We know that today, those rules are not standard and sometimes we do not know what are the internal rules inside a AWS. For example, for the computation of a daily average it exists several WMO rules (WMO TD 341, WMO TD 1188, WMO 1203) and a lot of national practices. With AWS it’s much more complex. As you said how many minute data is allowed to compute a daily average or extreme? What is the algorithm inside the AWS?
So, the only thing we could do, in my opinion, is to trust the data provider. In case of a trouble with the AWS, if:

(1) the data provider considers that the data is correct, he will code it as 8 : “Data checked and declared correct”
(2) the data provider considers that the data is doubtful but interesting to share worldwide, he will code it as 4 “Data checked and declared doubtful”
(3) the data provider considers that the data is erroneous, he will put an empty value for the data with the code 2 “Daily value not provided”.

Jose
Should we distinguish such different levels of doubtfulness as William is suggesting? If accepted, a simple two or maximum three categories of suspect could be established based on the departure of the measurement from its expected value or the number of non-passed QC tests.

Denis
DAYCLI message should not give the complexity of the data control processes in CDMS. It only needs to give the best answer from the data provider in a very simple way. We just added Doubtful code and Doubtful with some special cases that could help for extreme values (overflowed and aggregated).

9.3 Another code for estimated

Jose
And with respect to overflown raingauge, I would be in favor of allowing "estimated" (another flag) data in the DAYCLY message. Otherwise we would have a missing data and users would not be aware of that heavy rainfall, especially significant in hydrological and extreme values studies. As different raingauges can have different capacities, the observer could provide an estimation of the rainfall in such cases. (Also the monthly total would be calculated disregarding this important event!).

Denis answer
When a data is assigned with an overflowed code the observers should put in the DAYCLI message the value of the data, which is not missing (only it is underestimated).
If the NMHS/Organisation wants to estimate the value, it should give the estimated value with a code “Checked and declared correct”. Need to write a use case on that.

9.4 Another code for “New Record”

Rachid
Suggestion to add a “New Extreme” flag for the code table to highlight the fact that the value is a new record and it is a validated one.

Denis answer
It will give more complexity to the current DAYCLI message and be redundant with the code “Checked and declared correct”.

9.5 Adding the time of the extreme

Rachid
To include the time of the extremes (TMax, TMin) into the DAYCLI message.

Denis answer
It is a very interesting information to add. It was not specified in the scope of the DAYCLI message. Taking into account the current delay to respect on the DAYCLI message this information will not be reported.

9.6 "Softer" form of uncertainty

William
In reflection of some of the existing flagging schemes: do we need a "softer" form of uncertainty than "doubtful?. "Doubtful" to me suggests we strongly suspect it is wrong. But what about the case where a value is probably correct, but we have less confidence than for data that clearly pass the QC process? Perhaps, "suspect, but high (or medium) confidence"

Denis answer
If the data is clearly in Error, is there a need to share worldwide this data?

WMO-No. 488, Guide to the Global Observing System
1 = good, 2 = inconsistent,
3 = doubtful, 4 = erroneous,
5 = not checked, 6 = changed

9.7 Aggregated data

Jose
In an example shown at our last teleconference there was a table with zeros and quality flags "missing, not measured, accumulated". Don't we have a missing data code for the data, such as -999 or similar? I do not feel comfortable seen zeros when data are missing. Moreover, as Jitsuko pointed out, quality flags are exclusive. Therefore, if we use the "accumulated" quality flag, we cannot qualify the accumulated data as "suspect", which could be the case. A possible solution would be to use a special value (such as -998) for unknown but accumulated (not measured but carried on) values, and then the value resulting from the accumulation can have its own quality flag and the number of days of the accumulation would also be clearly determined (solving William's point 2 below).

Denis answer
See use cases above.

9.8 Plain text

William
Would it be possible to add plain text in the DAYCLI message?

Jitsuko answer
Everything is possible but it is not in the BUFR "philosophy".

Sergio answer
Plain text is very difficult to manage especially if we go towards "Machine to Machine".

9.9 Confidence Flag

William

Confidence flags: I think "doubtful" by itself is too restrictive. Not just in the case of the "overflow" category above, but also to distinguish between cases where.QC processes indicate that either a value is (a) truly doubtful, in that confidence is low; (b) suspect, but there is medium to high confidence. Ideally, the same distinction should be made with Estimated value, i'e, flags represent Estimated with low confidence, estimated with high confidence. So that's probably all up another four flags.

Denis answer
I think we need to be very simple in the DAYCLI message. Remember it is the first time that we add so much metadata into a WMO message. Most value shared worldwide should be “checked and declared correct”. When a data is designated “checked and declared correct” that means that the data could have been corrected from its original value to become correct with different methods: estimation/interpolation/calculation …If not we will come back to our initial suggestions with all possible modifications, which are in my opinion too complicated for the DAYCLI and its purpose.

William
Note that Estimated flags can, and arguably should, be used where a data provider tries to break down an aggregated total into the constituent daily amounts (something we didn't discuss). For instance, if a three-day rainfall (or snowfall) total is 24mm, but I know that no precip fell on Day 2, and I believe I can reasonably estimate the distributions on Days 1 and 3, then each of the Daily totals would be flagged as "Estimated - high confidence".

Denis answer
In that case, we trust the data provider who should declare in the DAYCLI message that days 1, 2 and 3 are “checked and declared correct”. It is only in the case where the data provider has no time or not the expertise for checking and doing the correction (estimation in this case) that he will use the aggregated code.

William
My own opinion is that this would be valuable information in the case of extremes, and would distinguish between a general category of doubtful (i.e, the QC process has limited or no confidence in the value), versus doubtful (as in, we know the value isn't correct, but we do know it is above some threshhold). The alternative here, as suggested by Peng, is to have a flag for QC'd - out of instrument range, but not implausible. In either case, the advice to users would be to contact the Met Office for more info.

10 Meetings

- 8 June 2021 : ET-DRC & TT-TDCF
- 4 May 2021
- 4 April 4 2021
- 19 February 2021