Quality control tests applied to the digitisations of surface meteorological data collected by the Chilean Air Force during the period 1950-1958

The handwritten Chilean meteorological data, collected by the Chilean Air Force during the period 1950-1958, was digitised in Excel spreadsheets and saved in text files to be quality controlled.

Input data:

1) 1 metadata file per station-year in .txt format: "**StationName_Obs_YYYY.txt**". The metadata file has the following constants per station-year (and also comments about the record):
   - **st_name** – Meteorological station name
   - **year** - Observations year
   - **uid** - Unique identifier
   - **lon** – Longitude
   - **lat** - Latitude
   - **alt** - Altitude
2) 12 digitisation tables per station-year, i.e. 1 table for each month, in .txt format: "**StationName_Mth_YYYY.txt**". The digitisation tables have 13 columns which correspond to the following variables:

- **day** - Day of the observation
- **hour** - Observation time in the format HHMM
- **dewpt** - Dew point temperature in Celsius degrees (ºC)
- **cloud** - Cloud cover in oktas: {0,1,2,3,4,5,6,7,8,9}
- **windir** - Wind direction: {N,NNE,NE,ENE,E,ESE,SE,SSE,S,SSW,SW,WSW,W,WNW,NW,NNW,C}
- **windsp** - Wind speed in knots (≈0,514 m/s)
- **press_msl / press_st** - Atmospheric pressure (at mean sea level or station level) in hectopascals (hPa)
- **temp** - Air temperature in Celsius degrees (ºC)
- **prec1** - Accumulated precipitation in millimetres (mm) usually measured at 12:00 UTC
- **prec2** - Accumulated precipitation in millimetres (mm) usually measured at 23:00 UTC
- **min_temp** - Minimum temperature in Celsius degrees (ºC) usually measured at 12:00 UTC
- **max_temp** - Maximum temperature in Celsius degrees (ºC) usually measured at 23:00 UTC
- **rel_hum** - Relative humidity in percent (%)

Quality control standard tests of World Meteorological Organization (WMO) were applied to the digitisations.

Additionally, empirical tests were applied to detect errors that occur specifically in those data: human errors that occurred during the original record or during the digitisation process; errors resulting from poor instrument calibration and/or internal consistency errors, visually perceived.

Each meteorological observation was codified with a quality control flag according to the following table:

Flag	Meaning
0	Correct
1	Erroneous
9	Suspect
-999	Default value / Missing check

Output:

- Several error files per station-year: "ErrVar_StationName_YYYY.txt", "ErrCategory_ StationName_YYYY.txt"
- 1 file per station-year with the missing observations: "MissObs_StationName_YYYY.txt"
- 1 final output per station-year according to ECMWF format. The output "**ECMWF_Form_StationName_YYYY.txt**" contains the following 12 columns:

uid | lon | lat | alt | year | month | day | hour | time_code | var_code | observation | flag

Time codes:

Observation time	time_code
4 times a day	0
2 times a day	12
1 time a day	13

Variable codes:

Variable	var_code
dewpt	18
cloud	20
windir	2
windsp	1
press_msl	6
press_st	5
temp	8
prec1	22
prec2	22
min_temp	10
max_temp	9
rel_hum	15

The quality control tests applied to each meteorological variable or to a set of variables can be divided by error category: I – Gross errors; II – Internal consistency errors; III – Time consistency errors.

I. Gross Error Limit Checks

- **N – Cloud Cover**

$$if\ (N \geq 0 \wedge N \leq 9)\ then\ flag_N = 0\ else\ flag_N = 1$$

$$N \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

(WMO, 1993: VI.6)

- **dd – Wind Direction**

$$if\ (dd \neq x)\ then\ flag_{dd} = 1\ else\ flag_{dd} = 0$$

$$x \in \{N, NNE, NE, ENE, E, ESE, SE, SSE, S, SSW, SW, WSW, W, WNW, NW, NNW, C\}$$

Were N = 360 and C = 0

(Empirical)

- **RRR_{11h} / RRR_{13h} – Precipitation**

$$if\ \left(RRR_{11h/13h} < 0\right)\ then\ flag_{RRR} = 1\ else\ flag = 0$$

(Empirical)

- **HR – Relative Humidity (HR)**

$$if\ (HR > 0 \wedge HR \leq 100)\ then\ flag_{HR} = 0\ else\ flag_{HR} = 1$$

(Empirical)

Gross Error Limit Checks were applied to observations of Wind Speed, Surface Temperature, Minimum Temperature, Maximum Temperature, Dew Point Temperature, Station Level Pressure and Mean Sea Level Pressure, considering the meteorological seasons of the Southern Hemisphere (MetOffice, 2018; Wikipedia, 2018).

Meteorological Seasons of the Southern Hemisphere	
Austral winter:	1st April – 31th September
Austral summer:	1st October – 31st March

The limit values considered for each variable are given by the following tables (WMO, 1993: VI.6-VI.7), and were defined as function of the latitude and the meteorological season.

- **ff – Wind Speed**

Table 6.4
Limit values for surface wind speed
(The value is considered suspect when MAX 1 < *ff* < MAX 2;
the value is considered erroneous when *ff* > MAX 2)

Area	*Winter*		*Summer*	
	MAX 1	*MAX 2*	*MAX 1*	*MAX 2*
45°S – 45°N	60 m s^{-1}	125 m s^{-1}	90 m s^{-1}	150 m s^{-1}
45°N – 90°N and 45°S–90°S	50 m s^{-1}	100 m s^{-1}	40 m s^{-1}	75 m s^{-1}

- **TT – Surface Temperature**

Table 6.5
Limit values for surface temperature
(The value is considered suspect when MIN 2 ≤ *T* < MIN 1 or MAX 1 < *T* ≤ MAX 2;
the value is considered erroneous when *T* < MIN 2 or *T* > MAX 2)

Area	*Winter*				*Summer*			
	MIN 2	*MIN 1*	*MAX 1*	*MAX 2*	*MIN 2*	*MIN 1*	*MAX 1*	*MAX 2*
45°S – 45°N	–40°C	–30°C	+50°C	+55°C	–30°C	–20°C	+50°C	+60°C
45°N – 90°N and 45°S – 90°S	–90°C	–80°C	+35°C	+40°C	–40°C	–30°C	+40°C	+50°C

- **Td – Dew Point Temperature**

Table 6.6
Limit values for surface dew-point temperature
(The value is considered suspect when MIN 2 ≤ T_d < MIN 1 or MAX 1 < T_d ≤ MAX 2; the value is considered erroneous when T_d < MIN 2 or T_d > MAX 2)

Area	*Winter*				*Summer*			
	MIN 2	*MIN 1*	*MAX 1*	*MAX 2*	*MIN 2*	*MIN 1*	*MAX 1*	*MAX 2*
45°S – 45°N	–45°C	–35°C	+35°C	+40°C	–35°C	–25°C	+35°C	+40°C
45°N – 90°N and 45°S – 90°S	–99°C	–85°C	+30°C	+35°C	–45°C	–35°C	+35°C	+40°C

- **$ppp_{station}$ – Station Level Pressure**

Table 6.7
Limit values for station pressure
(The value is considered suspect when MIN 2 ≤ p (station) < MIN 1 or MAX 1 < p (station) ≤ MAX 2; the value is considered erroneous when p (station) < MIN 2 or p (station) > MAX 2)

Area	*All year*			
	MIN 2	*MIN 1*	*MAX 1*	*MAX 2*
45°S – 45°N	300 hPa	400 hPa	1 080 hPa	1 100 hPa
45°N – 90°N and 45°S – 90°S	300 hPa	400 hPa	1 080 hPa	1 100 hPa

- **ppp_{MSL} – Mean Sea Level Pressure**

Table 6.8
Limit values for mean sea-level pressure
(The value is considered suspect when MIN 2 ≤ p < MIN 1 or MAX 1 < p ≤ MAX 2; the value is considered erroneous when p < MIN 2 or p > MAX 2)

Area	*Winter*				*Summer*			
	MIN 2	*MIN 1*	*MAX 1*	*MAX 2*	*MIN 2*	*MIN 1*	*MAX 1*	*MAX 2*
45°S – 45°N	870 hPa	910 hPa	1 080 hPa	1 100 hPa	850 hPa	900 hPa	1 080 hPa	1 100 hPa
45°N – 90°N and 45°S – 90°S	910 hPa	940 hPa	1 080 hPa	1 100 hPa	920 hPa	950 hPa	1 080 hPa	1 100 hPa

II. Internal Consistency Checks

- **Wind Direction and Speed**

$$if\ (dd = 0 \wedge ff \neq 0)\ then\ \left(flag_{dd} = 1\ \wedge\ flag_{ff} = 1\right)$$

WMO (1993: VI.10)

$$if\ (dd \neq 0 \wedge ff = 0)\ then\ \left(flag_{dd} = 1\ \wedge\ flag_{ff} = 1\right)$$

WMO (1993: VI.10)

$$if\ (dd = -999 \wedge ff = 0\ or\ ff \geq 5\ m/s)\ then\ \left(flag_{dd} = 1\ \wedge\ flag_{ff} = 1\right)$$

WMO (1993: VI.10)

- **Surface Temperature, Maximum and Minimum Temperature**

$$if\ (Tmin > T)\ then\ (flag_{Tmin} = 1\ \wedge\ flag_{T} = 1)$$

WMO (1993: VI.12)

$$if\ (Tmax < T)\ then\ (flag_{Tmax} = 1\ \wedge\ flag_{T} = 1)$$

WMO (1993: VI.12)

- **Surface Temperature and Dew Point**

Empirical complementary tests were programmed to evaluate the internal consistency between T and Td:

$$if\ (Td > T)\ then\ (flag_{Td} = 9\ \wedge\ flag_{T} = 9)$$

(Empirical)

The following tests were defined based on the Dew Point table for station altitude bellow 340 m. They allow to detect errors which result from digitisation, usually can be corrected and also are confirmed by other consistency tests between T and Td:

$$if\ ((altitude < 340) \wedge (T - Td > 61.5))\ then\ flag_{Td} = 9$$

$$if\ ((altitude < 340) \wedge (T - Td > 80.6))\ then\ flag_{Td} = 9$$

(Empirical)

- **Surface Temperature, Dew Point and Relative Humidity**

The HR value often seems suspect by visual analysis. We admit that in several cases the HR value could have been read directly from the hygrometer chart, instead of being calculated from T and Td or the hygrometer was not properly calibrated. For that reason the following tests were programmed:

$$if\ (T = Td \wedge HR \neq 100)\ then\ flag_{HR} = 9$$

$$if\ (T \neq Td \wedge HR = 100)\ then\ flag_{HR} = 9$$

$$if\ ((T - Td < 0.6) \wedge (HR < 90))\ then\ flag_{HR} = 9$$

(Empirical)

The values of T seem to be ok in general but Td seems suspect in some cases. In several stations the column that should contain Td, contains Tw (Wet Bulb Temperature) instead. That change isn't indicated and it's sometimes difficult distinguish whether it is Td or Tw. In order to assess the consistency of the relationship between T, Td and HR, estimates for these variables and deviations from the observed values were determined. The estimates were calculated from a Simple Approximation and also from formulas based on the August-Roche-Magnus Approximation. A flag of suspect will be applied to inconsistent T, Td and HR after we have decided what is "inconsistent".

Simple Approximation (Lawrence, 2005):

- Valid for $0°C < T < 30°C$ and $50\% < HR < 100\%$
- Accurate better than 1°C for Td and better than 5% for HR $Td = T - \frac{100-H}{5}$

$$HR = 100 - 5 \times (T - Td)$$

August-Roche-Magnus Approximation (McNoldy, 2001; McNoldy, 2017):

- Valid for $0°C < Td < 50°C,\ 0°C < T < 60°C$ and $1\% < HR < 100\%$

Td_e = (237.7*(log(HR/100)+(17.271*T/(237.7+T)))/(17.271-log(HR/100)-(17.271*T/(237.7+T))))

T_e = (237.7*(((17.271*Td)/(237.7+Td))-log(HR/100)))/(17.271+log(HR/100)-((17.271*Td)/(237.7+Td)))

HR_e = 100*(exp(17.271*Td/(237.7+Td))/exp(17.271*T/(237.7+T)))

III. Time Consistency Checks

Time Consistency Checks were applied to observations of **Surface Temperature (*T*)**, **Dew Point Temperature (*Td*)**, **Station Pressure (*Pstation*)** and **Mean Sea Level Pressure (*P*)**, considering the limit values given in the following table (WMO, 1993: VI.21):

Table 6.16
Suggested tolerances for the temperatures and the tendency as a function of time period between consecutive reports

Parameter	*dt = 1 hour*	*dt = 2 hours*	*dt = 3 hours*	*dt = 6 hours*	*dt = 12 hours*
T TOL	4°C	7°C	9°C	15°C	25°C
T_dTOL	4°C	6°C	8°C	12°C	20°C
*pp*TOL	3 hPa	6 hPa	9 hPa	18 hPa	36 hPa

For the time intervals present in the Chilean records, the limits given in the table below were considered. Tolerances for T_{tol} and Td_{tol} were interpolated using the values in the Table 6.16. For pp_{tol} a linear variation of 3 hPa per hour was considered, also based on Table 6.16.

dt [hours]	T_{tol} *[°C]*	Td_{tol} *[°C]*	pp_{tol} *[hPa]*
4	11	9.5	12
5	13	11	15
11	25	20	33
13	25	20	39
16	34	28	48

- **Surface Temperature Time Consistency**

$$if\ (|T(t) - T(t-dt)| > T_{tol})\ then\ (flag_{T(t)} = 9\ \wedge\ flag_{T(t-dt)} = 9)$$

(WMO, 1993: VI.21)

- **Dew Point Time Consistency**

$$if\ (|Td(t) - Td(t-dt)| > Td_{tol})\ then\ \left(flag_{Td(t)} = 9\ \wedge\ flag_{Td(t-dt)} = 9\right)$$

(WMO, 1993: VI.21)

- **Pressure Time Consistency**

$$if\ (|P(t) - P(t-dt)| > pp_{tol})\ then\ (flag_{P(t)} = 9\ \wedge\ flag_{P(t-dt)} = 9)$$

(WMO, 1993: VI.21)

References

Lawrence, M.G., 2005: The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications. Bull. Amer. Meteor. Soc., 86, 225–233, https://doi.org/10.1175/BAMS-86-2-225

McNoldy, B., 2001: Humidity conversions. http://andrew.rsmas.miami.edu/bmcnoldy/humidity_conversions.pdf (Accessed on 30th June 2017)

McNoldy, B., 2017: Calculate Temperature, Dewpoint, or Relative Humidity. http://andrew.rsmas.miami.edu/bmcnoldy/Humidity.html, Rosenstiel School of Marine and Atmospheric Science - University of Miami (Accessed on 30th June 2017)

MetOffice, 2018: When does spring start. https://www.metoffice.gov.uk/learning/seasons/spring/when-does-spring-start (Accessed on 26th June 2018)

Wikipedia, 2018: Season. https://en.wikipedia.org/wiki/Season (Accessed on 26th June 2018)

WMO, 1993: Chapter 6 - Quality Control Procedures. Guide on the Global Data-processing System, World Meteorological Organization, Geneva, No. 305, VI.1-VI.27, ISBN 92-63-13305-0