Advances in
Science & Research
Open Access Proceedings

Quality control of 10-min soil temperatures data at RMI

C. Bertrand, L. González Sotelino, and M. Journée

Royal Meteorological Institute of Belgium, Brussels, Belgium

Correspondence to: C. Bertrand (cedric.bertrand@meteo.be)

Abstract. Soil temperatures at various depths are unique parameters useful to describe both the surface energy processes and regional environmental and climate conditions. To provide soil temperature observation in different regions across Belgium for agricultural management as well as for climate research, soil temperatures are recorded in 13 of the 20 automated weather stations operated by the Royal Meteorological Institute (RMI) of Belgium. At each station, soil temperature can be measured at up to 5 different depths (from 5 to 100 cm) in addition to the bare soil and grass temperature records. Although many methods have been developed to identify erroneous air temperatures, little attention has been paid to quality control of soil temperature data. This contribution describes the newly developed semi-automatic quality control of 10-min soil temperatures data at RMI.

1 Introduction
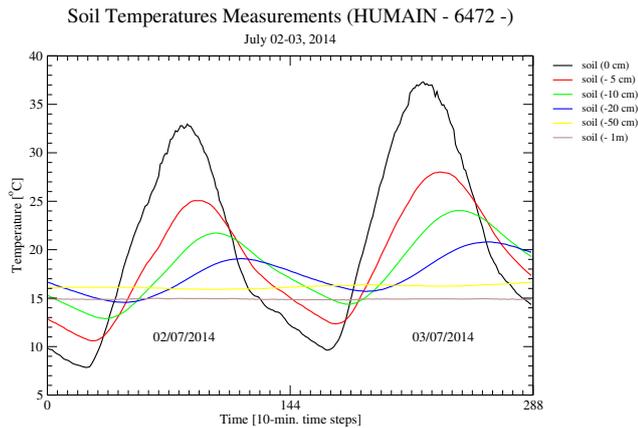
Of great importance in agriculture, soil temperature affects plant growth directly, e.g. in seed germination, root growth and nutrient uptake as well as indirectly in soil water and gas flow, soil structure and nutrient availability (Hillel, 1998). Soil temperature is an important parameter in energy balance applications such as land surface modeling, numerical weather forecasting, and climate prediction (e.g. Best et al., 2005). It is also important in radiative transfer applications, such as in the retrieval of land surface properties with satellite sensors, and especially in the retrieval of surface moisture with microwave sensors (e.g. de Jeu et al., 2008). Soil temperature varies in response to exchange processes that take place primarily through the soil surface. These effects are propagated into the soil profile by a complex series of transport processes, the rates of which are affected by time-variable and space-variable soil properties. At each succeeding depth, the peak temperature is dampened and shifted progressively in time (see Fig. 1). The degree of damping increases with depth and is related to the thermal properties of the soil and the frequency of the temperature fluctuation. Due to the much higher heat capacity of soil relative to air and the thermal insulation provided by vegetation and surface soil layers, soil heat anomalies of daily or weekly timescales in shallow layers near the surface do not propagate to the deeper layers. Only persistent long-term anomalies (e.g. at the in-

ter annual and decadal scale) affect temperature variations in those layers (e.g. Lachenbruch and Marshall, 1986).

While many methods have been proposed to identify erroneous air temperatures, only little attention has been paid to quality control of soil temperature data. This situation imposes a significant effort toward providing quality control and assurance (QC/QA) especially devoted to soil temperature measurements. Here we propose a semi-automatic quality control method to check 10-min grass and soil temperature records. In developing quality control tests for the automatic weather stations (AWSs) operated by the Royal Meteorological Institute of Belgium (RMI), we adapted some of the tools developed in Hu et al. (2002) and Hu and Feng (2003) to examine daily and hourly soil temperature data and expanded their approach to include additional tests introduced for some of them in quality control of air temperatures at RMI (Bertrand et al., 2013). The framework is similar to Gandin's concept of complex QA (Gandin, 1988), in that it approaches the question of the validity of a given datum from several different angles and considers errors of different types. However, in the present approach, the automated QA functions are included in a larger QA protocol involving manual inspections similarly to the method implemented at RMI for the quality control of 10-min air temperature data (Bertrand et al., 2013).

Table 1. List of the 13 RMI's Automatic Weather Stations performing at least one soil temperature record and the associated measurements.

AWS			Air temperature measurement (s)							QC group	Coordinates	
Code	Name	Climate zone	Grass	Bare soil 0 cm	Soil −5 cm	Soil −10 cm	Soil −20 cm	Soil −50 cm	Soil −1 m		Lat. (° N)	Lon. (° E)
6414	Beitem	1		X	X	X	X	X		3	50.91	3.12
6431	Zelzate	1		X						5	51.18	3.81
6434	Melle	1	X	X	X	X	X	X		2	50.98	3.83
6438	Stabroek	1	X	X	X	X	X	X		2	51.33	4.37
6439	Sint Katelijn Waver	2	X	X		X	X	X		4	51.08	4.53
6447	Uccle	1	X	X	X	X	X	X		2	50.80	4.36
6455	Dourbes	3	X	X	X	X	X	X		2	50.10	4.60
6459	Ernage	3	X	X	X	X	X	X		2	50.58	4.69
6464	Retie	2	X	X	X	X	X	X		2	51.22	5.03
6472	Humain	4	X	X	X	X	X	X	X	1	50.19	5.26
6477	Diepenbeek	2	X	X	X	X	X	X		2	50.92	5.45
6484	Buzenol	4		X	X	X	X	X		3	49.62	5.59
6494	Mont Rigi	5		X	X	X	X	X		3	50.51	6.07

Figure 1. 10-min soil temperature at various depths vs time of the day for 2–3 July 2014, at the Humain station (RMI's AWS 6472, see Table 1).
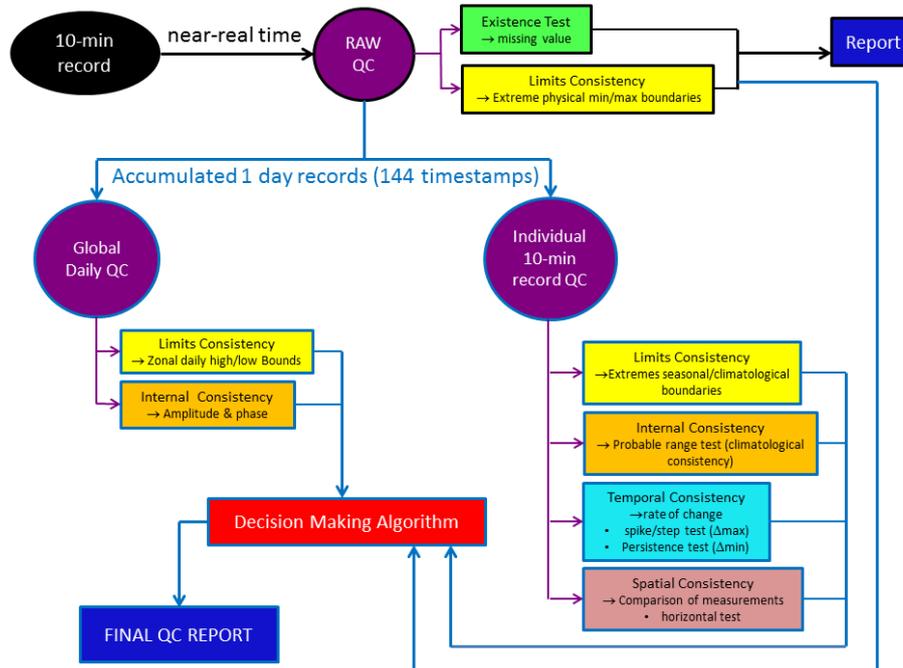
2 Soil temperature records at RMI

Automatic weather stations operated by RMI range from basic climatological stations to fully equipped synoptic stations performing a complete set of meteorological observations. Soil temperature measurements are performed in 13 of them to provide soil temperature profile in different regions across Belgium for agricultural management as well as for climate research. At each station, soil temperature can be measured at up to 5 different depths (from 5 to 100 cm) in addition to the bare soil and grass temperatures records (see Table 1). Due to the large heterogeneity within the RMI's AWSs, five groups based on the recorded temperature parameters are distinguished for the automated data QC and the stations are gathered in five climatic zones.

3 Complex automatic QA

RMI's AWSs are built around a programmable data logger that measures the sensors, then processes, stores and transmits the data to the central database (DB) in Uccle, Brussels. Once converted to digital values a first processing is performed on the raw data at the data logger level allowing calculation of 10-min temperatures values from the 5-s measurements. To ensure that gross errors are trapped before being further transmitted in the central DB a first basic QC is performed on all temperature values once acquired centrally. Automated procedures monitor the data to make sure they are collected and that the system performance is acceptable. After an existence test, a module checks for physical limits and flags the data violating these limits (erroneous when data lie outside physical limits and suspect when lying outside basic long-term climatological extremes that do not take into account the time of year and location). A list of missing and flagged data is automatically produced after each control cycle and transmitted to the AWS network maintenance team for further intervention. Note that values flagged as erroneous fail immediately and do not require further testing.

Second, each night automated QA procedures check the previous day 10-min temperatures values for more subtle errors. Implementation of the complex automated QA is diagrammed in Fig. 2. Daily temperature values are first checked using annual variation expectancies envelopes and the shape of the diurnal variation of the soil temperature at the different depths are compared to ensure consistency between them (see Sect. 3.1 for details). Individual 10-min records are further scrutinized for plausibility (using adjusted limits to reflect climatic conditions more precisely than in the first near real time range test), internal consistency, temporal consistency and spatial consistency (see Sect. 3.2 for details).

Figure 2. Complex QA of 10-min grass/soil temperatures records.

At the end of the checks, when all 10-min daily temperature time series recorded in a given station have been analyzed, a decision algorithm (that is applicable to all variables and all sites) interprets the scores obtained at each of the individual tests and attributes a final flag (i.e. erroneous, suspect or valid) to each particular data given the weight of evidence. At the end of the process, a report is automatically generated for each AWS and sent to the QC staff.

3.1 Global daily QC

In nature, soil temperature varies continuously in response to the ever changing meteorological regime acting upon the soil-atmosphere interface. That regime is characterized by a regular periodic succession of days and nights, and of summers and winters. The regular diurnal and annual cycles are perturbed by irregular episodic phenomena as cloudiness, cold wave, warm waves, rain storms or snow storms, and periods of drought. In addition to these external influences, there are the soil's own changing properties (i.e. temporal changes in reflectivity, heat capacity, and thermal conductivity as the soil alternately wets and dries, and the variation of all these properties with depth), as well as the influences of geographic location and vegetation. While the thermal regime of soil profiles is very complex, a simple mathematical representation of the fluctuating thermal regime in a soil profile is obtained by assuming that at all depths in the soil the temperature oscillates as a pure harmonic (sinusoidal) function of time around an average value (van Wijk and de Vries, 1963).
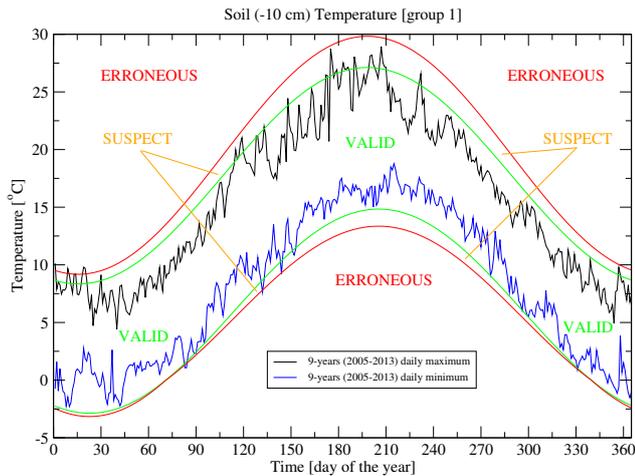
To rapidly identify data outside the variation range of soil temperatures at each depth, daily values (i.e. computed from the 10-min measurements) are compared to lower and upper bounds for soil temperature at given depths. Similarly to the LIM test in Hu and Feng (2003), lower and upper bounds for a given temperature data series (i.e. grass and soil temperatures) are constructed for each of the five climate zones by retrieving the highest and lowest daily values on each calendar day of the year from 9 years (2005–2013) of manually quality controlled historical data. Assuming that annual soil/grass temperature variations follow a sinusoidal curve, envelopes of annual variation of these extreme temperatures were then defined using wave functions of the form:

$$T_{L/U}(z, d) = T_{Lo/Uo}(z) + A_{Lo/Uo}(z) \cdot \sin(\omega_o d) \qquad (1)$$

where, T_L and T_U are the lower and upper bounds of soil temperature variations, respectively, d is the day of year, ω_o is the angular frequency, which is 2π times the actual frequency (i.e. $2\pi/365$ in case of an annual forcing), z is depth, and $T_{Lo/Uo}(z)$ and $A_{Lo/Uo}(z)$ are the annual mean of these extreme soil temperatures and their amplitude of variations, respectively. To include the extreme values within the derived expectancy envelopes from Eq. (1), the boundaries are adjusted as follows (see Fig. 3 for an illustration): data satisfying

$$T_L(z, d) - 2 \leq T(z, d) \leq T_U(z, d) + 2$$

succeed the limits consistency test, and data failing it are soft flagged if they are less than 10 % outside the range delimited by the adjusted boundaries and hard flagged otherwise.

Figure 3. Example of annual variation expectancy envelopes used in the Global Daily QC Limits Consistency test. Illustrated limits apply to the daily mean soil temperature recorded at -10 cm for stations of climate zone 1 (see Table 1)

To check the reduction in amplitude of the diurnal temperature cycle and the phase shift of the temperature maximum and minimum with depth (see Fig. 1), the recorded diurnal temperature cycle at any depth z, and time t, is modeled as follows:
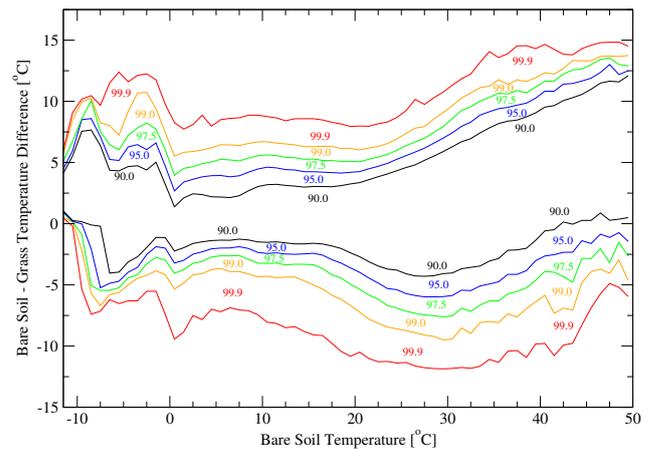
$$T(z, t) = \overline{T}(z) + A(z) \cdot \sin(\omega t - \phi(z)) \qquad (2)$$

where, $\overline{T}(z)$ is the daily mean temperature at depth z, ω is the angular frequency (diurnal forcing), $A(z)$ is the amplitude of the temperature wave at depth z, and $\phi(z)$ is the phase constant at depth z (aligns soil temperature variation with the forcing). The different $A(z)$ and $\phi(z)$ which are functions of z but not of t are then compared and temperature data series which do not respect the damping and retarding of the temperature waves with depth fail the internal consistency test.

Note that these two tests are denoted as global because they cannot discern which observations within the daily time series are responsible for the offense.

3.2 Individual 10-min record QC

The question of the validity of a given datum is approached using a number of additional tests applied to each particular 10-min temperature record. To verify whether the values are within acceptable range limits depending on the climatic conditions of the measurement site, individual data values are compared with upper and lower seasonal bounds derived for each temperature parameters and zones from several years of previous manually controlled data. The check provides information as to whether the values are erroneous or suspect. To minimize the possibility of a false positive identification, the algorithm does not report an anomaly in case where the

Figure 4. Probable bare soil–grass temperature differences as a function of the bare soil temperature.

majority of the recorded parameters in a given site at time t are flagged as being either to cool or to warm. Grass temperature and bare soil temperature are further examined for climatological consistency.

Bare soil temperature is measured by a PT100 sensor in contact with the ground on a horizontal surface fully exposed to the open sky. Similarly, grass temperature is measured by a PT100 sensor, fully exposed to the open sky, suspended horizontally over an area covered with short cropped turf and in contact with the tips of grass blades. With the advent of automation and the lack of daily attention by an observer or caretaker, this set up has proved limitations (e.g. lack/lost of contact between the temperature probe and the ground surface or the grass blades, probe fully covered by grass, . . .). The probable range test provides a more stringent constraint than simple valid maximum/minimum limit test by requiring consistency among temperature parameters as well as consistency with historical data. Basically, the differences between bare soil and grass temperatures as a function of the bare soil temperature are compared to probable difference determined from several years of previous data. Contours in Fig. 4 indicate which combinations of grass and bare soil temperatures fall within a given percentile of joint probability density. Following a review of the values that fall outside the 99.9 % boundary, the 99.9th percentile was selected as the boundary of acceptability. Combinations are hard flagged when falling outside the 99.9 % boundary and soft flagged when falling between the 99.0 and 99.9 % boundaries. Because two comparisons (involving three parameters) are necessary to unambiguously identify which parameter is problematic, similar joint probability densities were established involving the soil temperature at -10 cm. This last parameter has been chosen for the probable range test as it is systematically recorded in stations where both grass and bare soil temperatures are measured (see Table 1).

To examine the temporal consistency of the data, two tests involving the rate of change of the variables from a preceding acceptable level are applied: the spike/step (ΔMax) test and the persistence (ΔMin) test. Both, the maximum and minimum probable changes for each analyzed parameters (i.e. grass and soil temperatures) are based on the 99.9th percentile change for several years of previous data. Because the rate of change for deep soil temperatures can be very small, the persistence test does not apply to soil temperature below -20 cm. Values are checked for 10-min, 1, 2, 3 and 6 h time steps. To minimize the possibility of a false positive identification, the data must fail in at least 3 of the 5 tested time steps prior to be flagged as suspect or erroneous. Moreover, because in case of extreme meteorological conditions, unusual variability in the air temperature may occur, grass and bare soil temperatures data may be flagged as suspect, although correct. To prevent from this, the algorithm does not report a spike/step anomaly for grass and bare soil temperatures if both temperatures at the same site fail the spike/step test. Similarly, the algorithm does not report a persistence anomaly for the grass and bare soil temperatures in case of snow cover. Below $0\,^\circ$C in snow free condition, a persistence anomaly is reported based on the assumption that freezing conditions can affect multiple sensors. Note that for stations where only one of the grass or bare soil temperatures is recorded, the -5 cm soil temperature value is used (when applicable) in place of the lacking parameter to adjust the ΔMax and ΔMin tests.

Finally, horizontal comparisons of the same measurement at different stations are performed for all recorded temperature parameters. As for the quality control of 10-min air temperature data implemented at RMI (Bertrand et al., 2013), the horizontal check works in two steps. First, an outlier detection is performed on both the station data being quality controlled and the data of the surrounding stations using the daily 10-min temperature time series of each stations.

Let $T_{i,t}$ be a 10-min temperature record at station i ($i = 1, M$ with $M \leq 13$) at time t ($t = 1, N$ with $N \leq 144$) in a given day. $Y_{i,t} = T_{i,t} - \overline{T_i}$ with $\overline{T_i} = \frac{1}{N} \sum_{t=1}^{N} T_{i,t}$ the daily mean temperature at station i. $Z_{i,t} = Y_{i,t} - \overline{Y_t}$ with $\overline{Y_t} = \frac{1}{M} \sum_{i=1}^{M} Y_{i,t}$ the stations' mean at time t. We test whether or not the $Z_{i,t}$ values fall within the confidence interval defined by:

$$\overline{Z_t} - C\sigma_{Z_t} \leq Z_{i,t} \leq \overline{Z_t} + C\sigma_{Z_t} \tag{3}$$

where $\overline{Z_t} = \frac{1}{M} \sum_{i=1}^{M} Z_{i,t} = 0$, $\sigma_{Z_t} = \sqrt{\frac{1}{M-1} \sum_{i=1}^{M} (Z_{i,t} - \overline{Z_t})^2}$ is the estimated standard deviation at time t, and C is an adjustment parameter function of the considered soil/grass temperature parameter. Values $T_{i,t}$ that do not satisfy the relation in Eq. (3) are considered as outliers.

If an outlier is detected for the station being quality checked, then the data fails the horizontal consistency test.

Table 2. Overall performance of the automatic QC. The evaluation is performed over the full month of November 2014. A total of 298 512 (100 %) 10-min records including all soil/grass temperature parameters recorded within the RMI AWS network have been analyzed.

	10-min records		
QC$_{AUTO}$	Manual QC TRUE	Manual QC FALSE	Total
Valid	97.71 %	1.60 %	99.31 %
Suspect	0.17 %	0.09 %	0.26 %
Erroneous	< 0.01 %	0.43 %	0.43 %
Total	97.88 %	2.12 %	100 %

Otherwise, the algorithm tests on a 10-min basis whether the analyzed station value, $T_{i,t}$, falls inside a confidence interval formed from surrounding stations data that were not classified as outliers. Measurements that fail the test are soft or hard flagged depending upon the departure of the data from the confidence interval. Note that the outlier check in Eq. (3) can lead to false positives if one or several of the comparison measurements are spatial outliers able to influence the stations mean, $\overline{Y_t}$, in such a way that the measurement under the test is erroneously flagged as an outlier while being valid. In such cases either the decision algorithm at the end of the checking process identifies the false positives as valid based on the scores obtained at the other tests involved in the automated QA system or they will be reviewed during the manual follow up (see Sect. 4) and set to valid if justified.

In developing quality control methods for the US Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) Soil Moisture-Soil Temperature (SM-ST) network, Hu et al. (2002) established a soil heat diffusion model to screen and identify erroneous soil temperature data. Because such kind of model was found to perform well only in sunny and clear days situations, in our case, modeled data are not used to examine the soil temperatures records. Instead a soil model is used to assist the QC staff in their corrections and estimations (see Sect. 4).
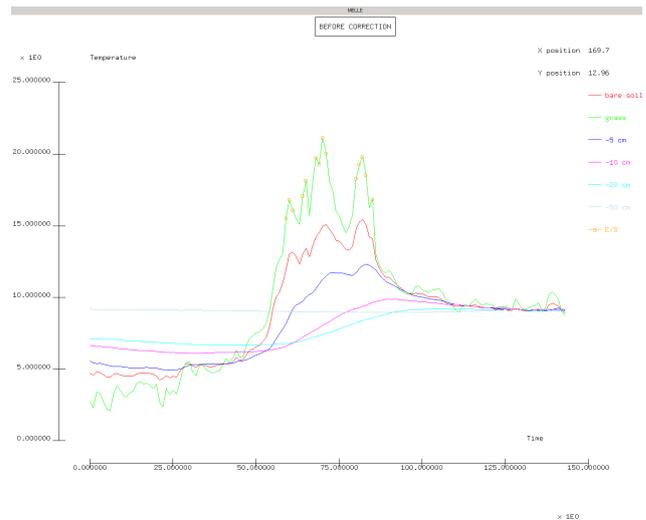
3.3 Automated QC performance

Quality assurance consists of procedures or rules against which data are tested. Each procedure will either detect the data as being valid, suspect or erroneous. False positives (i.e. type I error) increase the burden on the manual QC, and false negatives (i.e. type II error) reduce the quality of the data. One month of data (e.g. November 2014) has been used to determine the overall performance of the automated QA system. Independent manual QC applied on the recorded 10-min soil temperatures during the same month has been considered as reference for the evaluation. Table 2 presents a general overview of the performance of the newly developed

Table 3. Quantitative evaluation of the different tests involved in the automated QC of the 10-min soil/grass temperatures records (QC1 = physical limits test, QC2 = Min–Max range test, QC3 = spike/step − persistence test, QC4 = spatial horizontal test and QC5 = probable range test). The evaluation is performed over the full month of November 2014. A total of 298 512 (100 %) 10-min records including all soil/grass temperature parameters recorded within the RMI AWS network have been analyzed (v = valid, s = suspect, e = erroneous, and nt=no check).

QC TESTS		Manual TRUE 292 187 = 100 %	Manual FALSE 6325 = 100 %
		10-min records	
QC1	v	100 %	100%
	s		
	e		
	nt		
QC2	v	98.15 %	71.40 %
	s	1.81 %	26.12 %
	e	0.04 %	2.34 %
	nt		0.14%
QC3	v	99.36 %	89.38 %
	s	0.51 %	7.67 %
	e	0.13 %	2.81 %
	nt		0.14 %
QC4	v	91.45 %	96.13 %
	s	7.07 %	1.44 %
	e	< 0.01 %	2.43 %
	nt	1.48 %	
QC5	v	99.96 %	78.95 %
	s	0.04 %	1.56 %
	e	< 0.01 %	19.34 %
	nt		0.15 %
QC_{AUTO}	v	99.83 %	75.38 %
	s	0.17 %	4.49 %
	e	< 0.01 %	20.13 %
	nt		
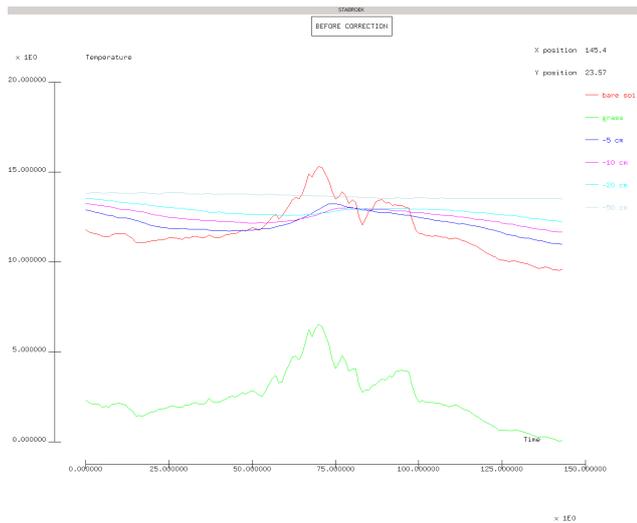
complex QA system while Table 3 provides a quantitative evaluation of the various tests involved in the data checking. It is worth pointing out that both tables refer to the 10-min data tests (daily tests cannot identify which observations are responsible for the offense) and that the results could differ for a given temperature parameter or a station type (i.e. QC group in Table 1). Table 2 indicates that type I errors generated by the automated QA system are very low (less than 0.01 % of the true 10-min records were detected as erroneous by the algorithm). By contrast the percentage of type II errors is very large (more than 75 % of the false 10-min records were found as valid by the automated QC as indicated in Table 3). However, this apparent very bad performance of the automated QA system in term of type II errors has to be handle with caution. First, it often occurred that while

Figure 5. Visualization of the automated QC applied on the 10-min soil/grass temperatures records performed on 22 November 2014 at the Melle station (AWS 6434, see Table 1). Erroneous grass temperature data are indicated by a orange circle on the green curve.

the algorithm effectively detected 10-min erroneous/suspect measurements in a daily parameter time series, the operators corrected more records than the ones found problematic by the system. As an example, Fig. 5 indicates that the 10-min grass temperature records were found erroneous by the automated QA system 14 times (orange circles on the green curve) on 22 November 2014 at the Melle station. After visualization of the station grass temperatures time series on 22 November 2014, 33 corrections were performed by the operator on the 10-min records (i.e. the full time segment where problematic measurements were detected by the algorithm was manually corrected). Second, Tables 2 and 3 only deal with the 10-min tests and do not account for the daily tests. It is worth pointing out that during the month of November 2014, the grass temperature measurements performed in the Stabroek station were found systematically wrong by the QC staff during 28 days. Over this time period, the grass temperature parameter in this station was detected as erroneous (suspect) 23 (4) times on a daily basis while the 10-min tests did not necessarily reported any erroneous/suspect measurements for this parameter as illustrated in Fig. 6.

When accounting for both daily and 10-min automated QC results, the algorithm succeeded to identify the stations, days and parameters on which corrections were made by the QC staff. A detailed analysis of the type II errors revealed that they mainly concern grass temperature measurements. We strongly suspect that the grass temperature database (and to a lesser extend the bare soil temperature database) used to derive the tests was not validated as it should have been. Typically, probable range test aims at detecting problematic situations as the one illustrated in Fig. 6 for the grass tem-

Figure 6. Visualization of the automated QC applied on the 10-min soil/grass temperatures records performed on 3 November 2014 at the Stabroek station (AWS 6438, see Table 1). No erroneous or suspect data were found in the daily 10-min temperatures time series.

perature 10-min daily time series. Because erroneous data have been involved when defining the boundaries used in the automated tests, probable differences between bare soil and grass temperatures given in Fig. 4 are certainly too permissive. This drawback has a direct impact on the detection performance of problematic grass, bare soil and −10 cm soil temperatures records by the probable range test. As an example, Table 3 indicates that this test (i.e. QC5) produced type II errors in about 79 % of the cases while this category of test has proven to be one of the most efficient in the detection of erroneous 10-min air temperature records performed by the RMI's AWSs.

4 Manual QA

Each day, the QC staff analyses the preceding day 10-min temperature records in the light of the assigned quality flags from the automated QA system. Results of the automated QA system can be graphically plotted on the operator terminal screen as illustrated in Figs. 5 and 6. In that case, all the analyzed station 10-min soil/grass temperatures records of the inspected day are illustrated in a graphic window and erroneous or suspect data are indicated in the corresponding parameter daily time series (e.g. orange circles on the green line in Fig. 5). Visual inspection of all records flagged by the automated decision making algorithm is done to distinguish instrumental problems from plausible behaviors. It is the human decision whether or not a value is accepted. When errors are verified or visually detected, faulty records are eliminated and "trouble tickets" are issued where needed to the maintenance team so that sensors can be replaced or repaired. More than simply deleting erroneous measurements, human

operators supply corrections and estimations (i.e. when values are missing) where possible. They are supported in this task by automated procedures. As an example, assuming that the thermal properties are constant with depth, soil temperature at any depth below the ground surface (i.e. $0 < z < \infty$) can be estimated using a soil heat diffusion model (van Wijk and de Vries, 1963).

The correction/estimation process is fully interactive, operators directly visualize on screen the corrections they applied on the parameters time series (the graphic window displaying the station temperatures time series being automatically updated after each modification). They have the opportunity to visualize different corrections on the problematic time series in order to determine the most appropriate in their specific case. When the correction/estimation process is completed, all modifications introduced by the operator are automatically implemented in the central RMI database. Note that the original parameters values are kept in the database and still accessible by the QC staff if required.

5 Conclusions

Automation of the RMI's AWSs data quality control is ongoing. After the automated quality control of 10-min air temperature data (Bertrand et al., 2013), automated quality assurance procedures devoted to 10-min grass/soils temperature records have been operationally implemented to support the QC staff in their work. The purpose of this automated data screening is to objectively identify abnormal data values for subsequent review by an experienced data analyst. The review is necessary to determine whether an anomaly results from a problem with instrumentation or whether it accurately reflects unusual meteorological conditions. Validation exercises have revealed that the complex automatic QA system is able to correctly identify problematic parameters in a particular station on a given day. However, automated tests applied to 10-min temperature records produce a very high percentage of type II error. In depth analysis of type II errors indicates that because the database of grass temperature records (and to a lesser extent bare soil temperature records) used to derive the boundaries involved in the automated tests was not validated as it should be, the probable range test fails to perform correctly. To overcome such a limitation an extensive validation of our historical records of 10-min grass and bare soil temperatures will be undergone as soon as possible. Once available, the new validated database will be used to refine the automated tests in general and in particular the probable range test involving the grass, bare soil and -10 cm soil temperatures. This forthcoming version of the algorithm will be evaluated using test data days from a whole year as the use of one single month of data could have masked sensitivities of the automated QA system to seasonal variations.

Edited by: I. Auer
Reviewed by: C. Sigg and another anonymous referee

References

Bertrand, C., Gonzalez Sotelino, L., and Journée, M.: Quality control of 10-min air temperature data at RMI, Adv. Sci. Res., 10, 1–5, doi:10.5194/asr-10-1-2013, 2013.

Best, M. J., Cox, P. M., and Warrilow, D.: determining the optimal soil temperature scheme for atmospheric modeling applications, Bound.-Lay. Meteorol., 114, 111–142, 2005.

de Jeu, R. A. M., Wagner, W., Holmes, T. R. H., Dolman, A. J., van de Giesen, N. C., and Friesen, J.: Global soil moisture patterns observed by space borne microwave radiometers and scatterometers, Surv. Geophys., 29, 399–420, 2008.

Gandin, L. S.: Complex quality control of meteorological observations, Mon. Weather Rev., 116, 1137–1156, 1988.

Hillel, D.: Environmental soil physics. Academic press, London, 771 pp., 1998.

Hu, Q. S. and Feng, S.: A daily soil temperature dataset and soil temperature climatology of the contiguous United states, J. Appl. Meteorol., 42, 1139–1156, 2003.

Hu, Q. S., Feng, S., and Schaefer, G.: Quality control for USDA NRCS SM-ST Network soil temperatures: a method and a dataset, J. Appl. Meteorol., 41, 607–619, 2002.

Lachenbruch, A. H. and Marshall, B. V.: Changing climate: geothermal evidence from permafrost in the Alaskan Arctic, Science, 234, 689–696, 1986.

van Wijk, W. R. and de Vries, D. A.: Periodic temperature variations in homogeneous soil, in: Physics of plant environment, edited by: van Wijk, W. R., North-Holland Publ. Co., Amsterdam, 102–143, 1963.