CATCHMENT DATA

DATA CORRECTION NOTES

HRC SOIL TEMPERATURE AND SOIL MOISTURE

NON NEMS STANDARD.

Version: 2012_V2 [Under development]

Date: September 2014

horizons regional council

Validation Methodology

This Data Validation document identifies the process to follow when undertaking the correction and review of Soil Temperature and Soil Moisture data.

It is suggested that if there is a change to the Sensor or collection in this period then the period should be ended at the time of change.

The Catchment Data Team does not undertake any field verification of the Soil Temperature & Moisture sensors nor do they undertake any pre-deployment calibration checks on the sensor (no acceptance testing). The methodology employed in the validation of Soil Temperature & Moisture data is one of spike removal, commenting and Quality Coding:

- Soil Temperature & Moisture Data will be compared to at site or nearby rainfall records. Unusual records will be quality coded and commented. Use the highest quality data available.
- Spikes (less than 24 hours, but greater than 1 hour, showing no rainfall for Soil Moisture) will be removed (gap closed) and Quality Coded as Synthetic.
- Spikes less than an hour do not quantify a Quality Code change. Fill with visually interpolated data, or close the gap and allow Hilltop to interpolate over the gap. Comment in the Comment sheet for review.
- Spikes (greater than 24 hours or periods showing Rainfall) will be gaped and commented.
    - Missing record **will not be filled or have gaps closed** if there is any indication of rainfall.
- Any sensor change (or movement), significant logger configuration or modification that changes the stationarity of record, recording interval or the type/format/calculation of data collected will result in an instrumentation comment.
- For sites with both datasets Soil Temperature (set at 30cm below ground level) will be processed first, followed by Soil Moisture. A temperature compensation virtual measurement can then be applied to correct the final soil moisture data; this is a global Virtual Measurement which automatically populates when the Soil Temperature and Moisture data sources are copied into the site.
- The temperature compensation Virtual Measurement has been produced by the Water Quality Portfolio holder. The data team will check this has been correctly applied but will make no changes to the Virtual Measurement itself.
- Quality coding, lowest quality prevails except for synthetic data (QC 300). All Soil Temperature and Soil Moisture data is cautionary and non-verified and is quality coded as such.

Clean/clear file workflow enables the reviewer to easily understand the steps undertaken and data modification. This also enables the reviewer to adjust any incorrect data corrections without full re-processing.

If you are not fully familiar with the validation of this data source or are unsure of the procedures to follow please read this document in full and/or consult with a Data Analyst or Coordinator, to obtain further assistance.

- Soil Temperature is corrected first, followed by Soil Moisture for the same period: both are part of Soil Processing.

- Soil Moisture uses a Virtual Measurement which utilises the corrected Soil Temperature to apply the Virtual Measurement Correction.

The current process for Soil Moisture is subject to adjustment – awaiting soils analysis for the grain size, composition and conductivity of the soils which is also site-specific to the processing. These parameters will then be added to the Virtual Measurement and the batch re-processed.

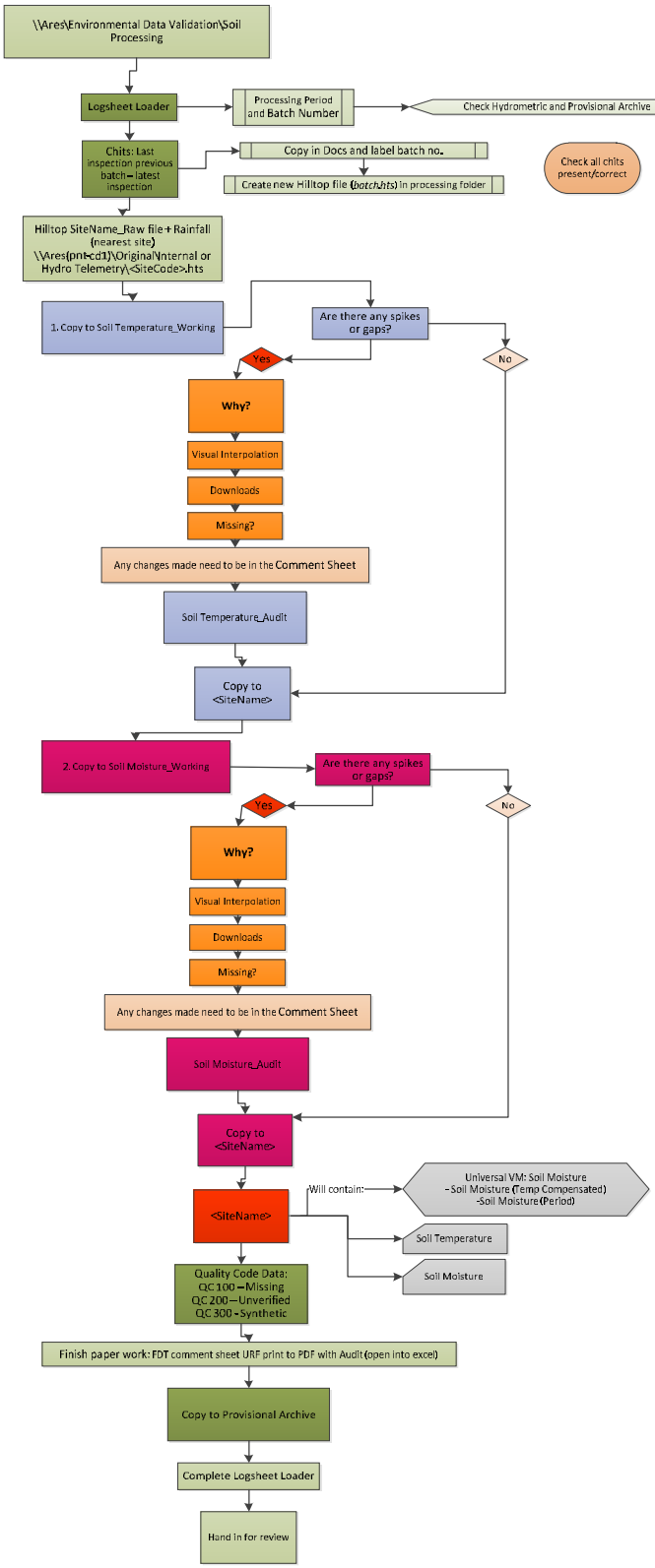


Figure 1: Flow chart of Soil Processing steps (see attached full sheet if this is not legible)

1.0 Creating Your Working Directory

1.1.1 Open the Logsheet Loader and fill in the processing register for the correction period and the data sources: Soil Temperature and Soil Moisture.

Due to inherited processing periods the batch numbers for Soil Temperature and Soil Moisture may not be identical, although the batch processing for both files together will have a combined batch number. You will need to do this first to obtain a batch number for the new period of processing and where to begin for Soil Temperature and Soil Moisture.

Begin the processing from the end DATE of the last batch (unless the data has been processed incorrectly and a full reprocessing has been requested or completed within another batch file).

Open Hilltop Manager and the Provisional Archive to check the final point of processed record for that site, as a second confirmation of the period start date for your batch and to check that all data has been processed to this point.

1.1.2 Collect all the chits (Logsheets) for the current period of data correction. The data is processed from inspection to inspection.

1.1.2.1 Check that all chits are filled in correctly & accounted for and photocopy them. Note the final inspection information in the Logsheet Loader - this will be the end of the batch for the period of processing. Both Soil Temperature and Soil Moisture will finish on the same inspection.

The processing period for Soil Moisture will be the same as that defined by Soil Temperature. The two data sources are processed in conjunction with each other for sites which contain both data sources.

1.1.2 Create a working directory in \\Ares\Environmental Data Validation\Soil Processing...

The file path should be:

\\Ares\Environmental Data Validation\ Soil Processing\<SiteName>\<Batch #>\....

1.1.3 The above processing folder needs to contain the following (obtained in the document directory of the data source):

FILE DETAILS TEMPLATE_ST.DOC *(ST = Soil Temperature)*

FILE DETAILS TEMPLATE_SM.DOC *(SM = Soil Moisture)*

AUDIT.MDB

URF_SOIL PROCESSING.XLS – CONTAINS A TAB FOR THE COMMENT SHEET:

1.2 Configure the Hilltop processing files

1.2.1 Rename the Audit file (Audit.mdb) to the batch number i.e. 102.mdb

1.2.2 In Hilltop Manager, create a new .hts (hilltop file) file and call it the same batch number as your mdb audit file, i.e. 102, save this to your working directory: \\Ares\Environmental Data Validation\Soil Processing\<SiteName>\<Batch #>..... If you have done this correctly, Hilltop will indicate that there is an audit trail in the background. If it does not, then you need to make the audit.mdb (access database) and label the access database the same name as the Hilltop file. If an audit file fails to generate a connection then contact an Analyst.

If you complete your Validation/Corrections and there is no audit trail, there will be problems during the review process and the URF may be declined.

1.3 Add the batch number to the beginning of the other documents within the working files i.e. 102 File Details Template_ST and the URF, i.e. URF_Soil Temperature_102_BAL with the 3 letter site code.

1.4 Obtain all the chits for the current period (batch) of data validation. Check all chits are present and in order. Photocopy these if you have not done so already.

Check the chits which have not been entered into the Logsheet Loader (they will not have an "Entered" stamp at the top of the sheet). If you have time, enter Logsheets to the Loader which have not already been loaded.

The Logsheet Loader is used to extract the inspection information into Hilltop.

It is suggested that you have access to the Calibration information relating to the batch period for any sensors or reference sensors [if applicable] this includes any survey information and non-conformance reports.

The reason for indexing the batch number and the folder/file structure is that the system databases (Hilltop, Logsheet Loader, Audit trails) are now linked. This enables connections and reporting between the different systems and tracing/accounting for the periods of processed data.

2.0 Populate Data to your File

2.1 Copy the Soil Temperature and Soil Moisture for the Site you are processing:

2.1.1 In your Hilltop file, copy the original (raw) data for the period of processing from \\Ares\Original\Hilltop Telemetry\<SiteCode>.hts or \\Ares\Original\Internal\<SiteCode>.hts to your Hilltop processing file and call it *<[SiteName]_Raw>*

2.2 Copy Rainfall data from the Site you are processing, or the closest Rainfall Site to your processing Site.

2.2.1 Copy the rainfall data from the site you are correcting (or nearest neighbour). This can either be from the Provisional Archive or raw (Hydro Telemetry). It is not critical, *but the more corrected the data, the better.*

2.2.2 Note when copying the data into your Hilltop file *where* the Rainfall data has come from, i.e. *<[SiteName]_Archive>*, *<[SiteName]__Provisional>*, or *<[SiteName]__Telemetry>.*

2.2.3 Copy in Water Level and Flow data for the site or the nearest neighbouring site.

2.2.4 Copy in the voltage and logger status.

2.2.5 If you have not processed the current site before, or are unfamiliar with the site, copy the last batch of processing to the file. This way you should have some indication of what was done in the past or issues observed.

3.0 Populate the File Details Template (FDT) *(for each data source..._ST and ..._SM)*

3.1 Document processing

3.1.1 Fill the FDT with the Site, Correction parameter (Soil Temperature or Soil Moisture) and Batch Number from the Logsheet Loader, as well as the combined Batch number (next sequential number system in the Soil Processing folder from the previous batch). The FDT should reflect what is held in the .hts file. *If the FDT, Logsheet Loader and .hts do not match then your URF will fail and require re-processing.*

3.2 Documenting Site/Data Source/Period

3.2.1 Under File Info within Hilltop Manager > Right click the *<[SiteName]_Raw>* and select *Details* and copy that information into your FDT under Initial File Details

3.3 Data Gaps

3.3.1 Right click on the data source and select *gaps* and list, if any, the Gaps in the raw file under Gaps in Initial File and list any gaps deleted in the *Gaps Deleted field for each data source. If there are no gaps, then enter 0 gaps and delete the Gaps Deleted Field within the FDT.* If there are any single point gaps (15 minute) these can be selected and deleted at this stage with no change to the Quality Code, listing all gaps deleted under the Gaps Deleted field.

3.4 Comments

3.4.1 Currently there is no check data/verification for either Soil Temperature or Soil Moisture. However, within the data it needs to be noted when site visits are made and the corresponding data response.

- For sites which do provide an Ebro check for the Soil Temperature this can be entered as Check Data with comments for each of the inspections
- Otherwise Soil Temperature checks (-1) listed in the Temperature check field, with comments for site inspection – or blank. This 'check' data provides an indication of onsite visits and potential response within the data.
- Soil Moisture – comments are added rather than checks and used in a similar manner to above for tracking changes in the data and response at the site.
- Also pertinent to comments are if the site was sprayed or weed-whacked and if there was a corresponding response with the removal of vegetation shown in the data (and commented into the comment sheet.

4.0 Editing the Data

4.1 Soil Temperature

4.1.1 Copy *<[SiteName]_Raw>* into *<Working_ST>* site in Hilltop.

The <Working_ST> site is where you are going to make minor edits to the data.

4.1.2 Here are some more popular methods:

1. Use the Soil Temperature Spike removal Virtual Measurement to remove any spikes in the record. See Appendix for an example of the code used (the spike value is specific to the site)

2. Run the Gap marker VM to identify (if any) hidden gap markers in the record greater than the recording interval i.e. 15 minutes. See Appendix for an example of the code used

Any corrections made to the data need to be listed in the Comments sheet including the SiteName, Data Source, Date and Time for the beginning and end of the adjusted period [this is defined as the last good data point and the next good data point following the period being considered], Comment Type (Synthetic, Backup, Data Correction, Missing Record, Instrument Change...) and Comment – what you have done and why to the data. This is located in the URF form.

If during your processing there is a gap in the Orig file which is filled by a download, convention is to label the download_sitecode_date (from inspection/chit) within the control tree. The Orig file also needs to be updated. Missing record from downloads does not require a quality code change.

4.1.3 Create another site called *< Audit_ST>* as a Virtual Measurement. This shows the changes made to the data during the correction process and helps in the review of the data. See Appendix for an example of the code used to write the Virtual Measurement.

Use the decision making trees below as guidance while editing and making changes to the data:

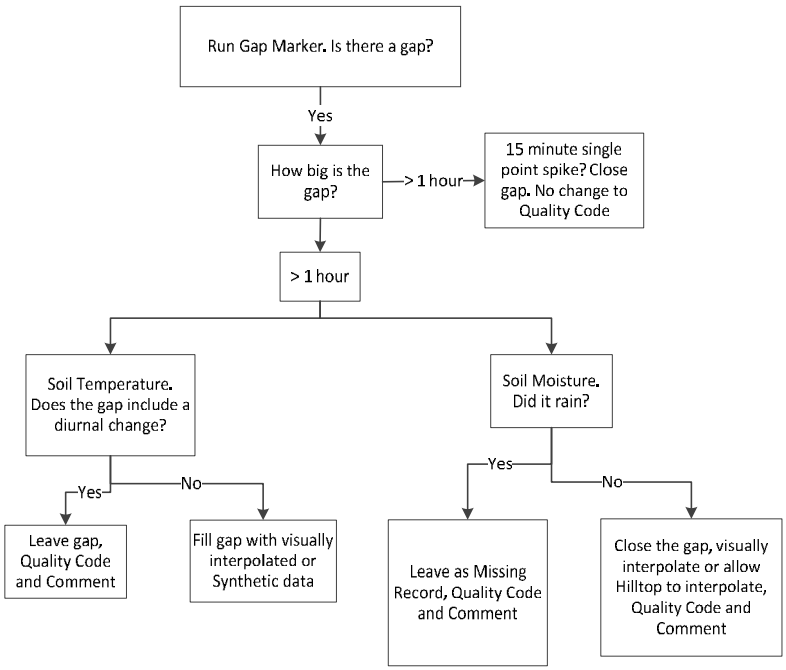


Figure 2: Dealing with gaps in the data. This figure details some of the decisions which have to be made concerning the correction of this data

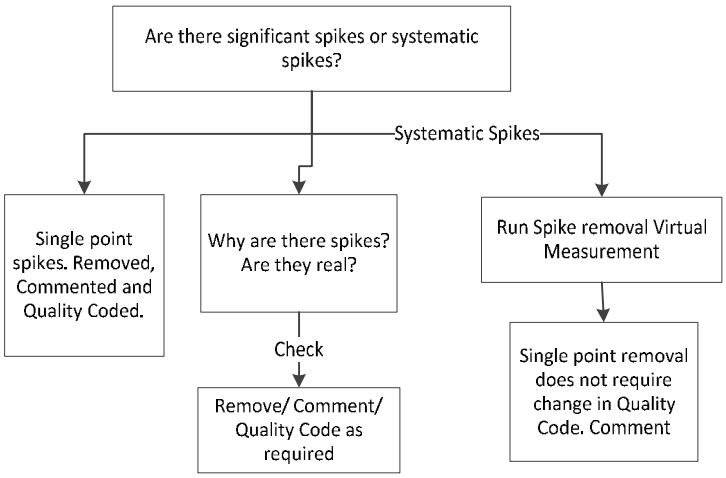


Figure 3: Dealing with spikes in the data. This figure details some of the decisions which have to be made concerning the correction of this data

4.2 Soil Moisture

4.2.1 Copy *<[SiteName]_Raw>* into *<Working_SM>* site in Hilltop.

The <Working_SM > site is where you are going to make minor edits to the data.

4.2.2 Here are some more popular methods:

1. Use the Soil Mositure Spike removal VM to remove any spikes in the record. See Appendix for an example of the code used (the spike value is specific to the site)

2. Run the Gap marker VM to identify (if any) hidden gap markers in the record greater than the recording interval i.e. 15 minutes. See Appendix for an example of the code used.

Any corrections made to the data need to be listed in the Comments sheet including the SiteName, Data Source, Date and Time for the beginning and end of the adjusted period, Comment Type (Synthetic, Backup, Data Correction, Missing Record, Instrument Change…) and Comment – what you have done and why to the data.

If during your processing there is a gap in the Orig file which is filled by a download, convention is to label the download_sitecode_date (from inspection/chit) within the control tree. The Orig file also needs to be updated. Missing record from downloads does not require a quality code change.

4.2.3 Create another site called *<Audit_SM>* as a Virtual Measurement. This shows the changes made to the data during the correction process and helps in the review of the data.

Use the decision making trees above (figures 2 and 3) as guidance while editing and making changes to the data:

5.0 Quality Coding the Data

5.1 Once you have completed the minor corrections copy *<Working_ST>* and *<Working_SM>* [the cleaned/spike removed] sites into *<[SiteName]>* site.

5.2 The*<[SiteName]>* is where the Quality Code is applied both Soil Temperature and Soil Moisture.

There are no offset corrections to be made (as there are no reference check data).

Quality coding is required for all the changes that have been made to the data. The Quality Codes are as follows:

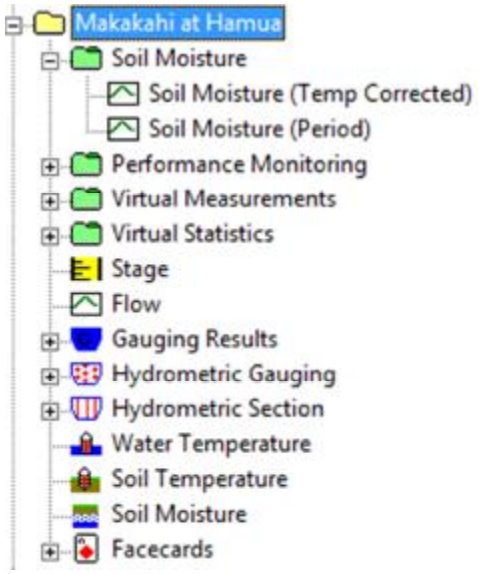
Unverified Data QC = 200

Synthetic Data QC = 300 *(corrections made greater than single 15 minute spikes or gaps)*

Missing record QC = 100 *(gaps which cannot be filled, over a diurnal or during a rainfall event)*

The lowest quality code prevails however synthetic data needs to be QC300 as the assumptions in filling this data/how it was filled are different and need to be differentiated within analyses.

5.3 In the *<[SiteName]>* site a new folder should automatically appear: Soil Moisture as shown below. There should be two global Virtual Measurements: *Soil Moisture (Temp Corrected) and Soil Moisture (Period),* which are used to derive the period and the temperature compensated Soil Moisture. They are global Virtual Measurements and adjusted only by the portfolio holder.



<[SiteName]> should contain the Soil Moisture Folder [containing Soil Moisture (Temp Corrected) and (Period) Virtual Measurements] as well as your corrected Soil Temperature and Soil Moisture Time series.

6.0 Final Steps

6.1 Fill in the rest of the File Details Template for ST and SM: Final Details and Gaps. Any significant comments pertinent to the data processing period should be listed in the **Additional Comments** section as well, at the bottom of each FDT. These are comments which are pertinent to the entire data set, site and period.

6.2 Open and complete the URF form located in your working directory: \\Ares\Environmental Data Validation\Soil Processing\ <SiteName>\<Batch#>...

6.3 Copy the Corrected data (including ***Check*** and ***Quality Data***) to the Provisional Archive. Make sure you have the '*Gap at Start*' check box ticked!

6.4 List the Audit trail in the audit tab of the URF excel spreadsheet

7.0 Print to PDF for Review

7.1 Print the URF

7.2 Print the File Details Template_ST and File Details Template_SM

7.3 Print the Comments spreadsheet (if comments have been made)

7.4 Graph for each of the data sources showing the after, i.e. final Soil Temperature/Soil Moisture with Quality data

7.5 Graph showing the before and after i.e. raw and final with the Audit Virtual Measurements, one for Soil Temperature and one for Soil Moisture.

7.6 Graph showing the before and after i.e. raw and final with the Quality code, one for Soil Temperature and one for Soil Moisture.

7.7 Graph with the before and after/ Raw and Final with the stage/flow data and rainfall data for the site, one for Soil Temperature and one for Soil Moisture.

7.8 Save the PDF with the batch number and site three letter code, at the end of the URF into your working directory (\\Ares\Environmental Data Validation\Soil Processing\ <SiteName>\<Batch#>...) and print the PDF.

8.0 Complete the Logsheet Loader for the period of processing for both Soil Temperature and Soil Moisture.

9.0 Hand in the corrected data file to the Analyst for review

10.0 Reviewing the Data

Reviewing the data is a relatively simple process if the above has been followed. Basically, the review process is to double check that the above expectations have been followed. The below gives an indication of what to look for when reviewing the data:

1. Check the URF form on the front. Is all the information needed found there? Cross check CDT
2. Check the initial file details template. Is it all correct? Have the check comments been made, spelt correctly or missing? Note the mistakes/errors on the review template
3. Are all the changes made to the data complete and accurately reflect the correction processes used? Have they been commented and quality coded accordingly?
4. Does the audit show all the changes made. If not, then add comments to the review template
5. Are all the graphical printouts within the processing? Are there any visual inconsistencies'?
6. Fill out the non-conformance register with all errors found in correction process
7. If minor errors exist, correct them and make note of the changes made
8. *Keep in mind when filing the review sheet, that correction batches only fail if they require complete or substantial reprocessing*

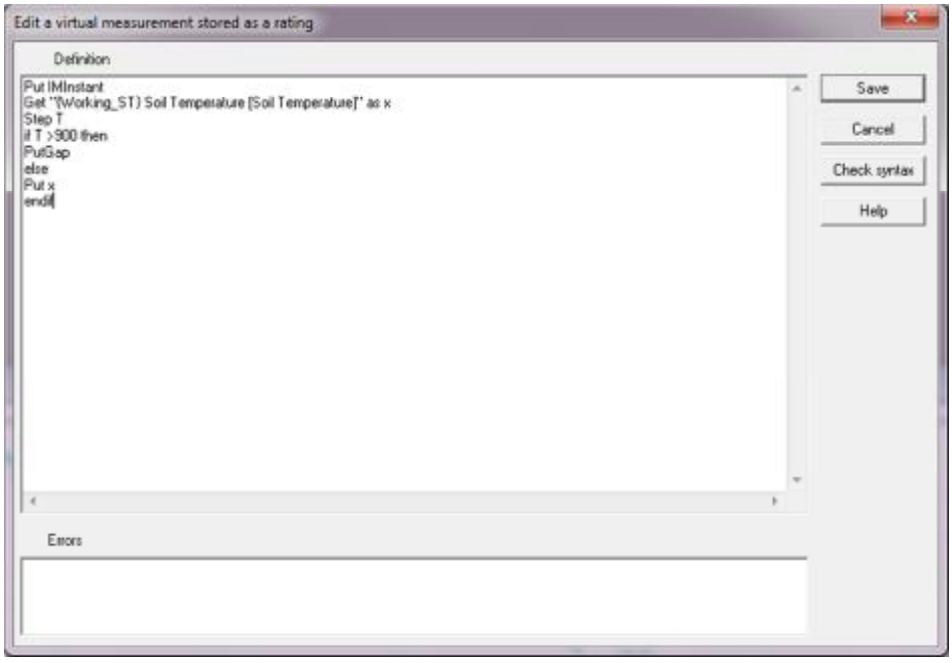
Appendix

Virtual Measurements

Gap Removal Virtual Measurement. Applicable to both Soil Temperature and Soil Moisture:

Put IMInstant
Get "{Working_ST/SM} Soil Temperature/Moisture [Soil Temperature/Moisture]" as x
Step T
If T > 900 then
PutGap
Else
Put x
Endif

Example:



Spike Removal Virtual Measurement. Applicable to both Soil Temperature and Soil Moisture:

Put IMInstant
Get "Soil Temperature/Moisture" as x
Get "Soil Temperature/Moisture movav (3 hour,centre) as y
Put y
Q = x-y

Z = min(x,y)

If Q > ? then

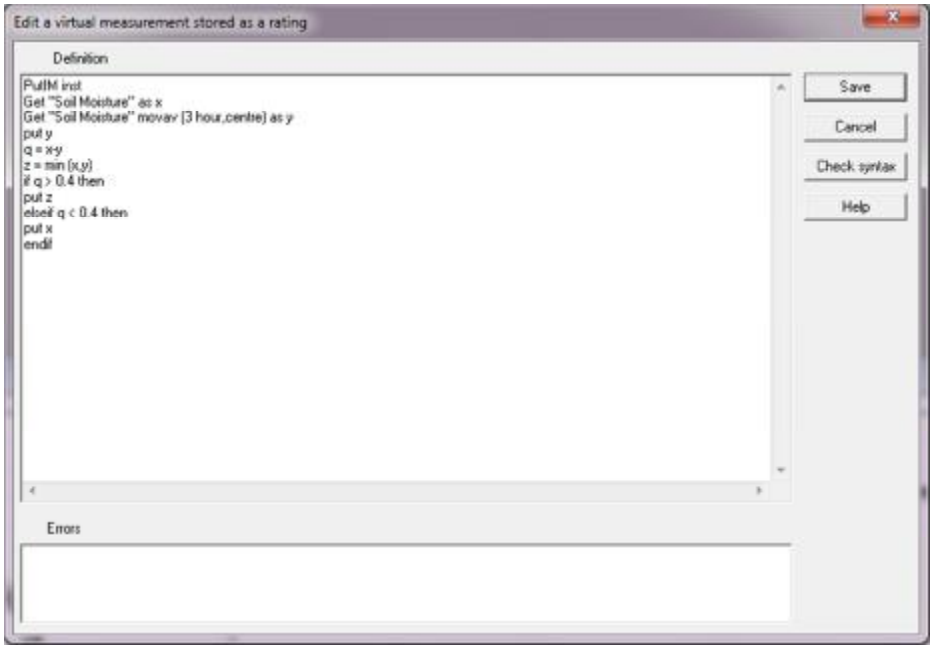
Put Z

Elseif Q < ? then

Put x

Endif

Example:



Audit

Audit_ST:

Select New Site, label as Audit_ST, Datasource: Audit and Virtual Measurement: Audit

Write in as shown in the box below:

Get "{SiteName_Raw} Soil Temperature [Soil Temperature]" as x

Get "{SiteName} Soil Temperature [Soil Temperature]" as y

Z = y-x

Put z

Audit_SM:

Select New Site, label as Audit_ST, Datasource: Audit and Virtual Measurement: Audit

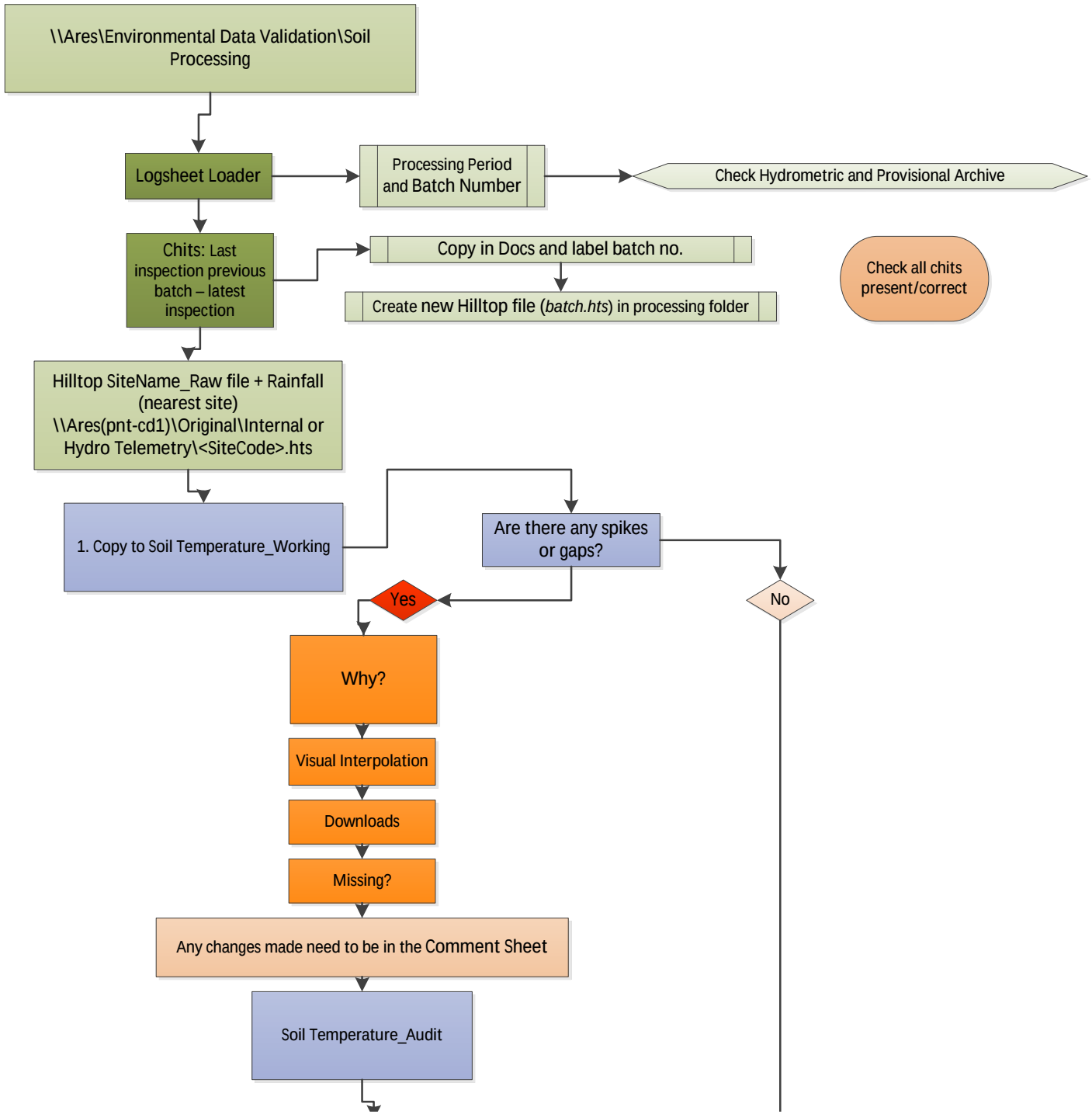
Write in as shown in the box below:

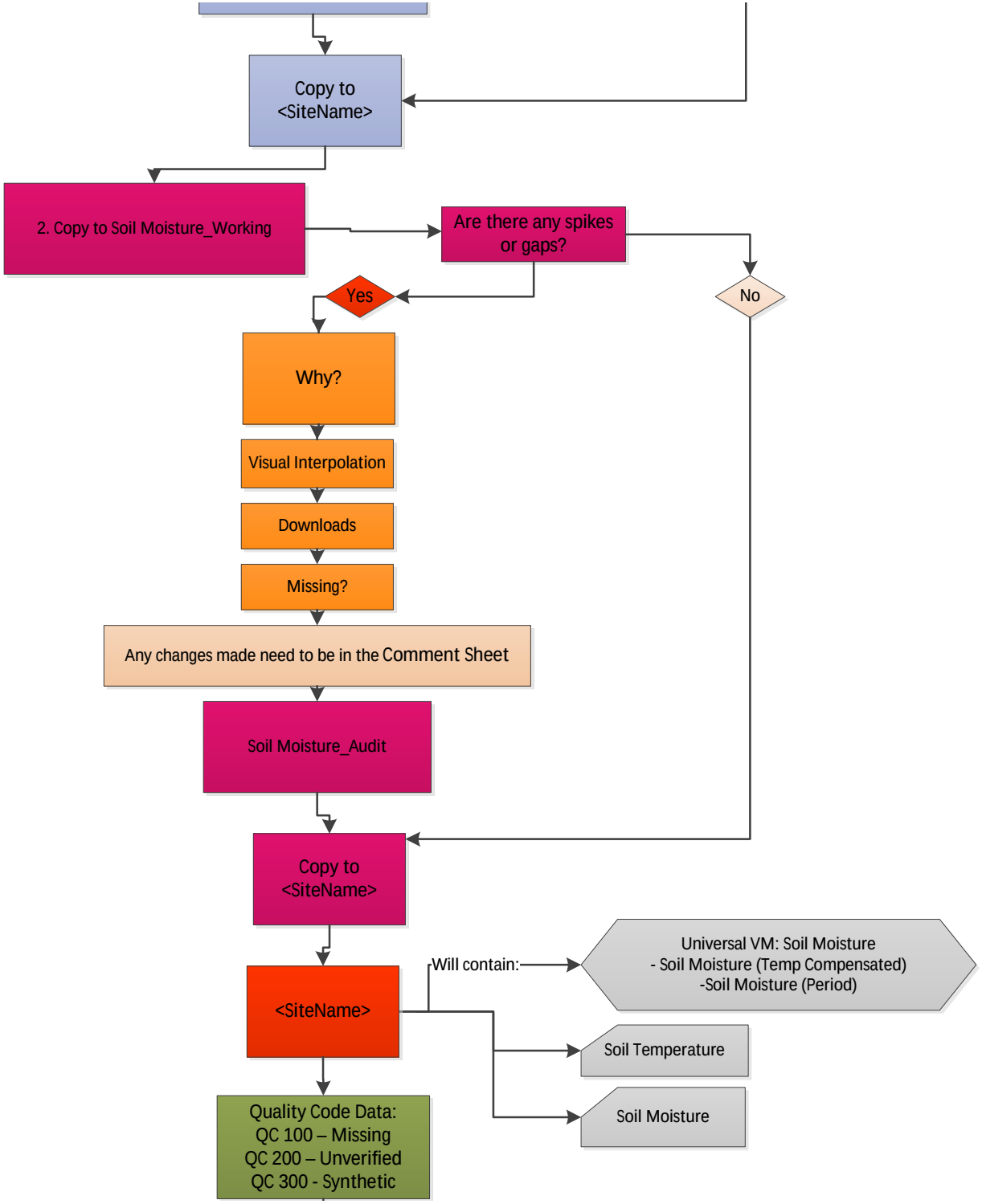
Get "{SiteName_Raw} Soil Moisture [Soil Moisture]" as x

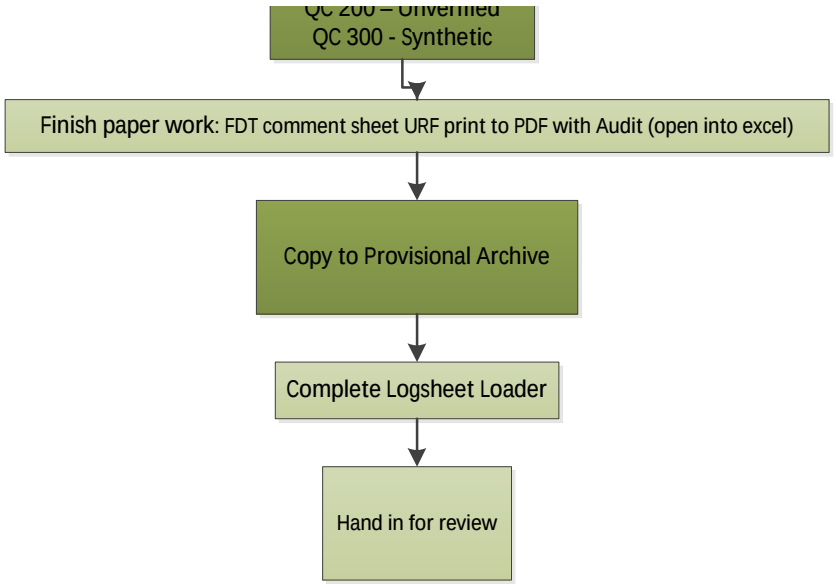
Get "{SiteName} Soil Moisture [Soil Moisture]" as y

Z = y-x

Put z


\\Ares\Environmental Data Validation\Soil Processing
Logsheet Loader
Processing Period and Batch Number
Check Hydrometric and Provisional Archive
Chits: Last inspection previous batch – latest inspection
Copy in Docs and label batch no.
Create new Hilltop file (batch.hts) in processing folder
Check all chits present/correct
Hilltop SiteName_Raw file + Rainfall (nearest site) \\Ares(pnt-cd1)\Original\Internal or Hydro Telemetry\<SiteCode>.hts
1. Copy to Soil Temperature_Working
Are there any spikes or gaps?
Yes
No
Why?
Visual Interpolation
Downloads
Missing?
Any changes made need to be in the Comment Sheet
Soil Temperature_Audit


Copy to
<SiteName>
2. Copy to Soil Moisture_Working
Are there any spikes or gaps?
Yes
No
Why?
Visual Interpolation
Downloads
Missing?
Any changes made need to be in the Comment Sheet
Soil Moisture_Audit
Copy to
<SiteName>
<SiteName>
Will contain:
Universal VM: Soil Moisture
- Soil Moisture (Temp Compensated)
-Soil Moisture (Period)
Soil Temperature
Soil Moisture
Quality Code Data:
QC 100 – Missing
QC 200 – Unverified
QC 300 - Synthetic


QC 200 – Unverified
QC 300 - Synthetic
Finish paper work: FDT comment sheet URF print to PDF with Audit (open into excel)
Copy to Provisional Archive
Complete Logsheet Loader
Hand in for review