CATCHMENT DATA

DATA CORRECTION NOTES

NEMS DISSOLVED OXYGEN

"Version Number"

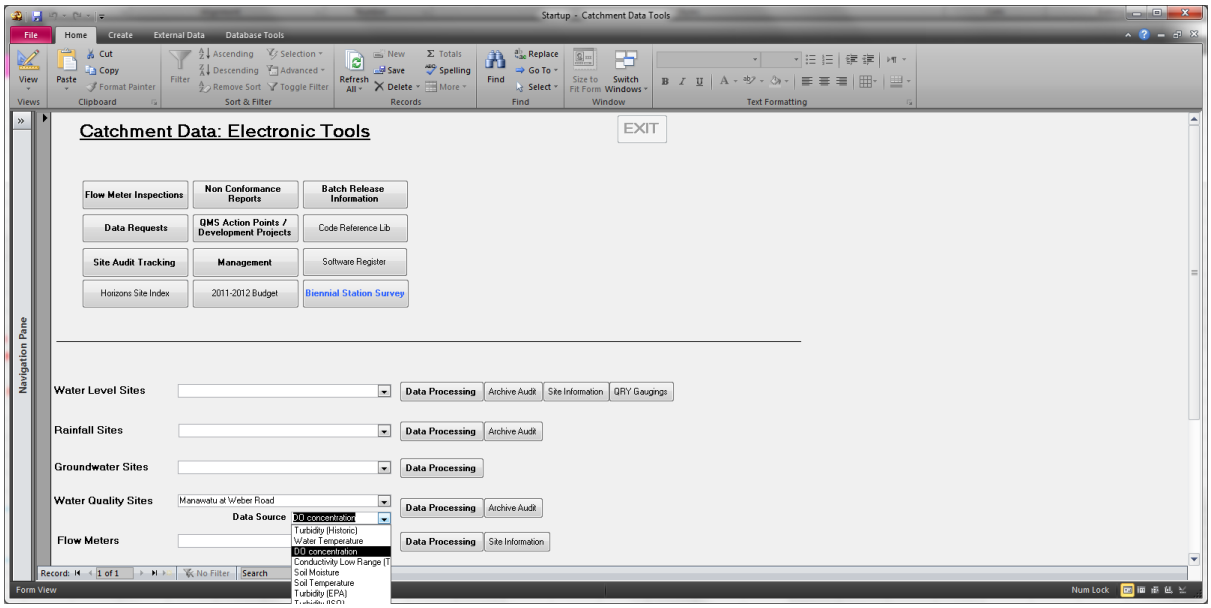
"Year"

These corrections notes are for reference during the data correction process. If you have not corrected the following data source previously or are still unsure of how it all works, it would be a wise idea to get the analyst to spend some time going through the process with you.

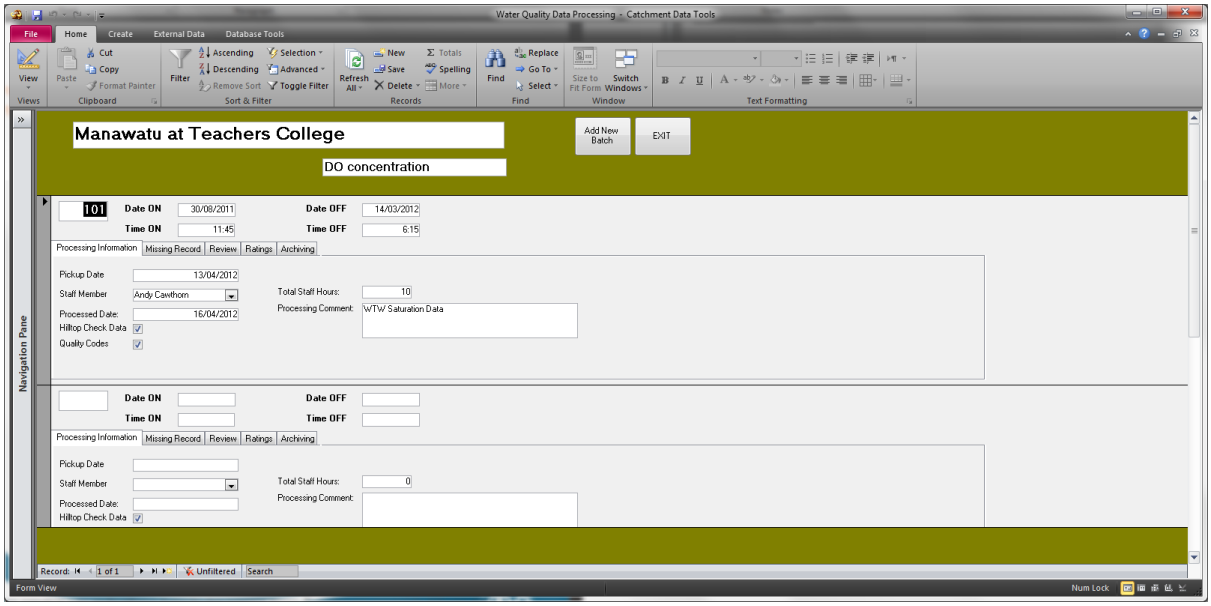
horizons regional council

Complete the CDTools Register

Open the CDTools register. This can be found on \\ares\HydrologySoftware\Catchment Data Tools\Catchment Data Tools.mdb enter the groundwater site that you are working on. Hit the Data Processing tab



File the field in the register that pops up once you have selected the site



Fill in the following fields:

1. The **batch number**. This should be pretty obvious, in increasing order
2. **File the DateON Time ON and DateOFF TIMEOFF** fields. The Date/TimeON fields should be the dates/times of the last batch Date/TimeOFF. The current batch END Date/Times are the last inspection completed for the period.
3. **Pickup date** is the date processing began on the batch
4. **Staff Member:** Obviously this is you
5. *The other fields do not require completion till the processing is finished*

Create the working directory

Groundwater processing, like all other data sources, contains its own parent directory for processing/correcting the data. This directory is located:

\\ares\Environmental Data Validation\Dissolved Oxygen\

This directory, and its sub-folders, contains all the currently corrected groundwater water sites, and their batch numbers.

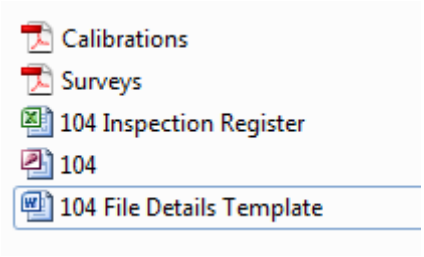
Within the parent directory is a folder, aptly named, Docs. Within this folder is the following processing documents:

1. Audit.mdb (this is the audit trail for the hilltop software)
2. File Details Template.docx, and (this is a rough summary template)
3. Inspection Register.xlsx (field inspection register, logger data and analysis sheet)

Create a new folder within the site folder renaming the folder to the correct batch number i.e.:

\\ares\Environmental Data Validation\Dissolved Oxygen\Manawatu at Teachers Collgege\104\

Copy the documents from the Docs folder into the above working directory. Rename the copied doc files so that they contain the batch numbers at the beginning i.e.



The Audit.mdb needs to be same as the Hilltop Manager File name; otherwise they cannot be linked together

Create the Hilltop Manager Working file

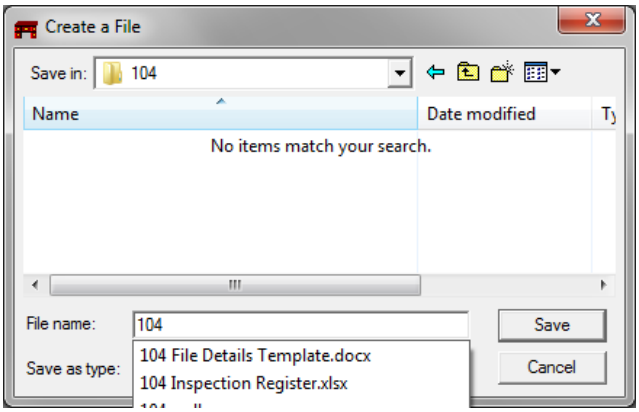
Open Manager

Click File > New

Navigate to your working directory i.e.

\\ares\Environmental Data Validation\Dissolved Oxygen\Manawatu at Teachers Collgege\104\

The File Name should the Batch number i.e. 100. Double check that the batch number.mdb is also in there. No audit trail = instant fail



Collecting the Evidence

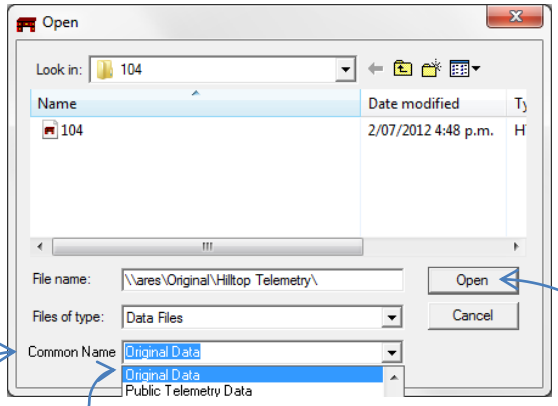
Photocopy all the field inspection chits for the period. The first inspection is the last inspection previously processed. Double check this with CDTools. The last inspection filed in the draws is the end inspection for processing i.e. the last inspection. NOTE that end inspections need to end on an actual external reference reading/check.

Calibrations need to be extracted from Hydsis

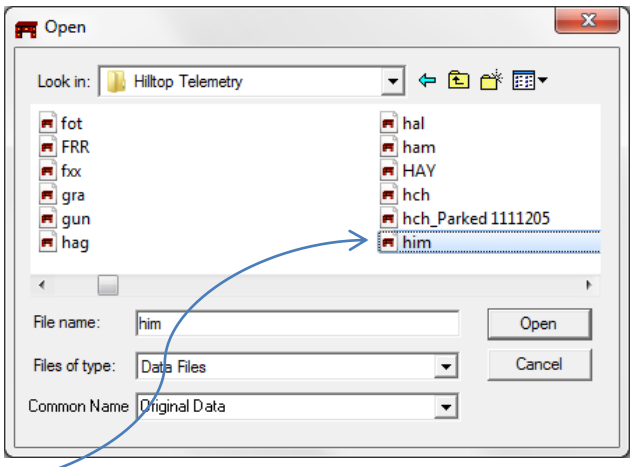
Organising the data

In Hilltop > Open

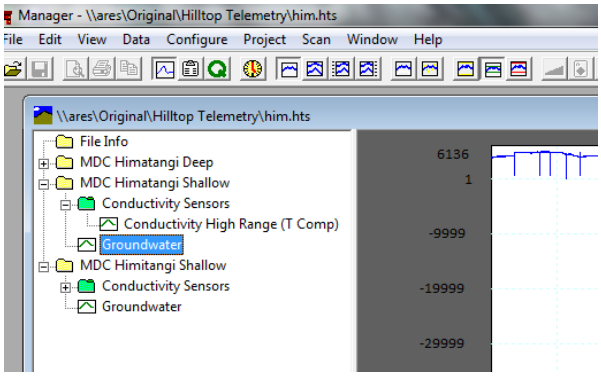
This brings up the Open dialog box



Select from the Common Name > Original Data. This opens the original data files from Telemetry. Hit Open. This bring up the following dialog box



Select the three letter code from the drop down list that you are working on. In this particular case, MDC Himatangi Shallow is HIM > Open



Copy the site data to the working Hilltop directory i.e.

Right click > Copy.

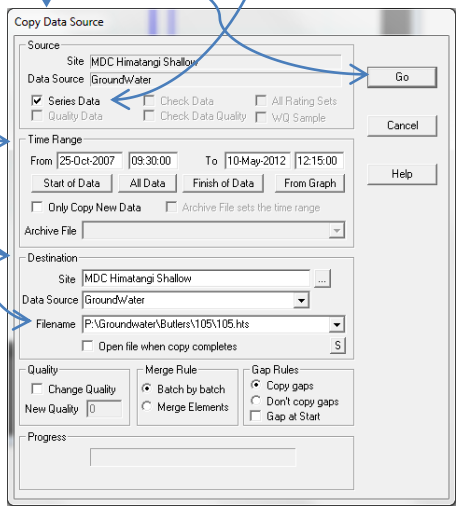
This brings up the *Copy Data Source* dialog box. Make sure that the *Series Data* check box is ticked. Set the Time Range to the period that you are processing. You should already know what this is by now!

In the *Destination*, rename the site to an underscore Raw name extension i.e. MDC Himatangi Shallow_Raw. This allows to the reviewer to know that this data is the raw data

Under the *Filename*, make sure that this is pointing at the working path. For this example, the working path is:

\\ares\Environmental Data Validation\Groundwater\MDC Himatangi\Shallow\104\104.hts

Once this is completed, hit the Go button. The data should now be in your working file

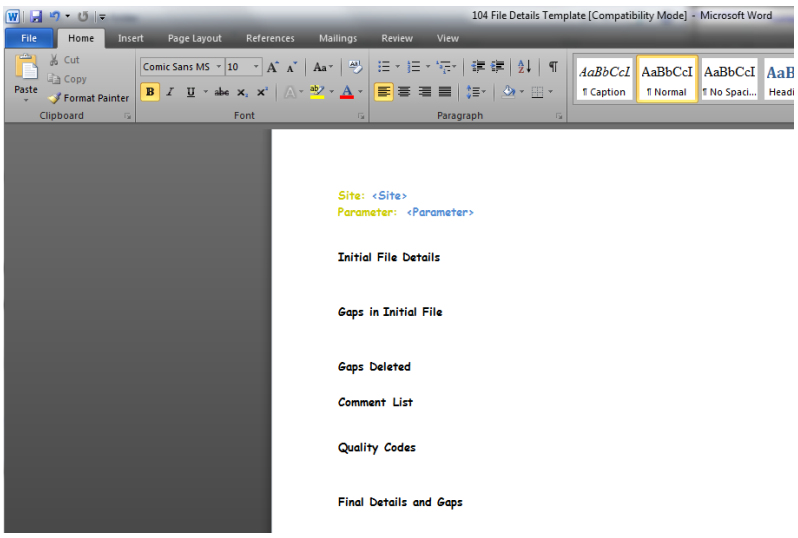


Populate the File Details Template with the basic File Information

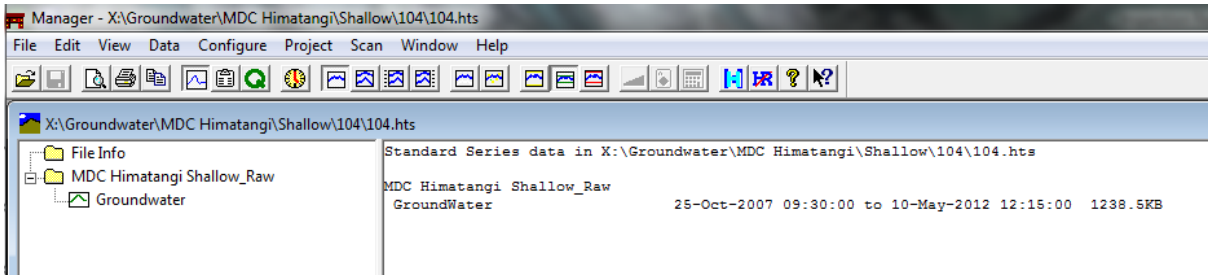
Open the # File Details Template. Most of this is pretty self-explanatory

The Site: <Site> should be the Hilltop Site Name i.e. MDC Himatangi Shallow

The Parameter: <Parameter> is Groundwater, or Water Level and/or Aquifer for some sites

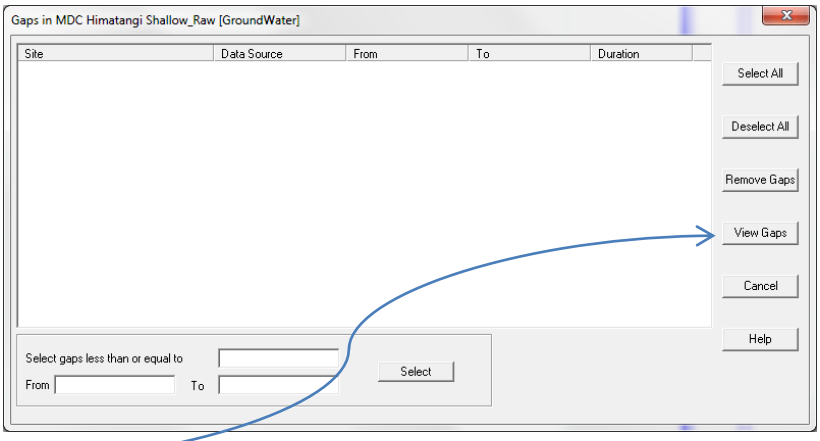


Initial File Details information comes from Manager. In Manager right click > File Info > Details. Paste this information into the **Initial File Details**



Gaps in the Initial File can be listed from Manager also. Right click on the data source ie Groundwater > Gap.

This brings up the Gaps dialog box.



List all the gaps identified in the data set by hitting the View Gaps button and paste these into the **Gaps in Initial File**. In most cases there should be no gaps in the raw file, but there can be…

Select all gaps < 15 minutes in duration using the Select Gaps Less than or Equal to box

Hit the Select all box

Hit the remove gaps. Paste this information into the **Gaps Deleted** information in the **File Details Template**

Site: <MDC Himatangi Shallow>
Parameter: <Groundwater>

Initial File Details

Standard Series data in X:\Groundwater\MDC Himatangi\Shallow\104\104.hts

MDC Himatangi Shallow_Raw
GroundWater 25-Oct-2007 09:30:00 to 10-May-2012 12:15:00, 1238.5KB

Gaps in Initial File

Gaps in X:\Groundwater\MDC Himatangi\Shallow\104\104.hts

0 gaps

Gaps Deleted

0 gaps

Loading check data

Before you even start loading check data, you will need to construct the Datum Continuity Tables in the Inspection Register.xlsx. Open this file i.e. \\ares\Environmental Data Validation\Groundwater\MDC Himatangi\Shallow\104\104 Inspection Rewgister.xlsx

Under the Datum Table Tab use the surveys to construct the Datum Continuity Table. Once this is finished,

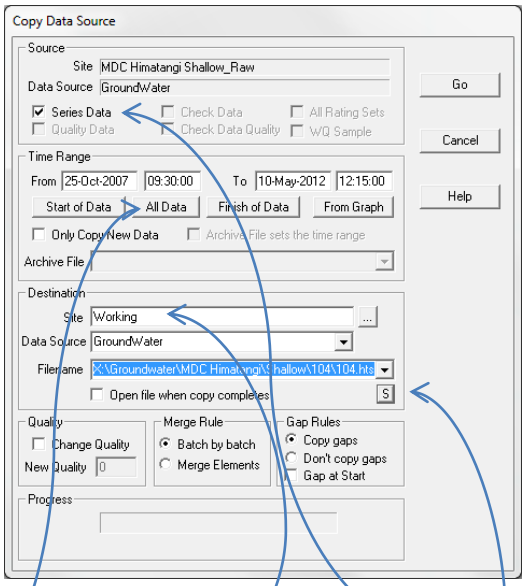
Check data is without doubt the most time consuming and frustrating part of the data processing! You have been warned!

The purpose if check data is technical; to track site performance, assist with maintenance issues, and to guide the processing of recording problems out of the data

Open Manager

Copy the Site_Raw file to another file within the object tree called Working. This is where all the check data and minor editing to the data takes place

Right click MDC Himatangi Shallow_Raw > Copy. This brings up the Copy Data Source dialog box



Under Data Source, make sure that the Series Data check box is checked

Under Time Range > All Data

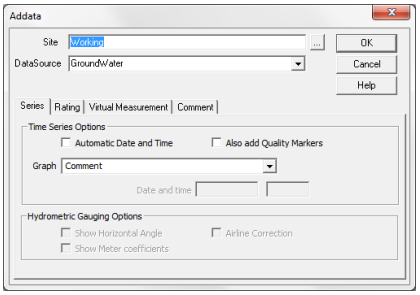
Under Destination > Site > "Working"

Data Source should remain the same

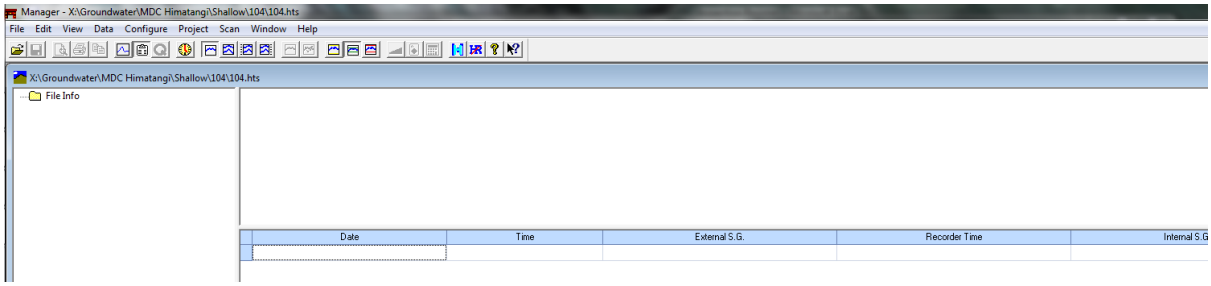
File Name > hit the S box below the drop down menu. This sets the copy destination the same as the source. Useful tool – use it often!

Hit > Go

In the working file, hit the clipboard (View Check Data)  . From the top menu > Data > Add. This opens the Add Data Box.



Not much to be done here. Hit > OK. This opens the check data entry form



Use the log information on the log sheets to populate the respective columns. But before you do this, you need to double check everything on the log sheets:

- Check the inspection date with the Julian Date. Do they agree? If not, find out why!?
- Do the inspection times represent the actual time. Check the NZDT and NZST
- Cross reference the EPB, ExSG information with the surveys. Are they correct?
- Check for logger downloads, purges/pumps, and any relevant inspection comments.

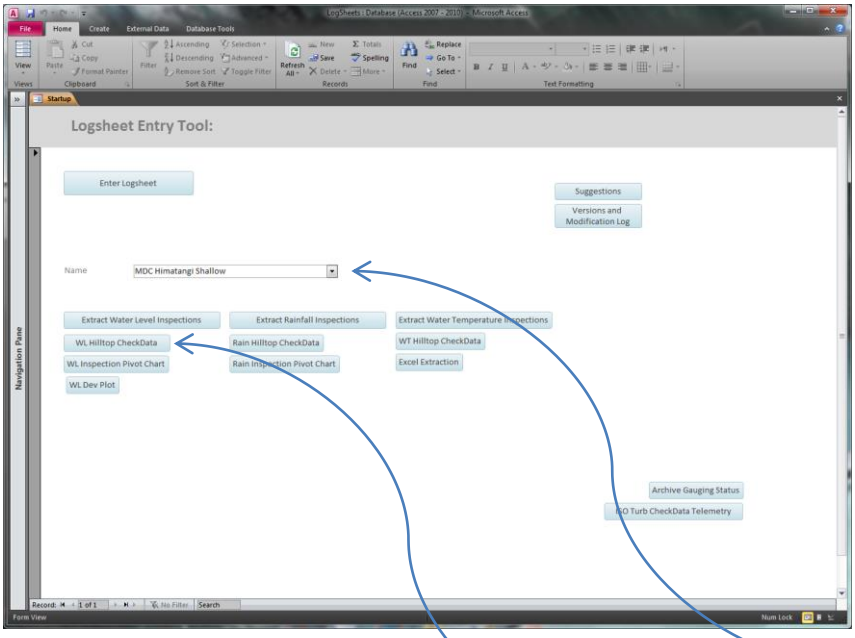
Once the log sheets have been formally checked for errors, enter the information into the check data. All primary references are entered into the ExSG field (yes, even if it is an EPB). The EPB is for the backup check. If no backup check is in place, populate the column with -1 (old Tidedar legacy). Add any relevant comments from the chit into comments column. Sometimes you may have to decipher comments so that they make sense, not only to you, but someone a 100 years' from now (it's highly unlikely that they will want to go hunting for paper chits!). Try to summarise long comments, the comments field is limited to 255 characters (old 16 bit technology)

The log sheet loader tool can also be employed here also. This can be found on:

\\pnt-cd1\HydrologySoftware\Catchment Data Tools\LogSheets.accdb

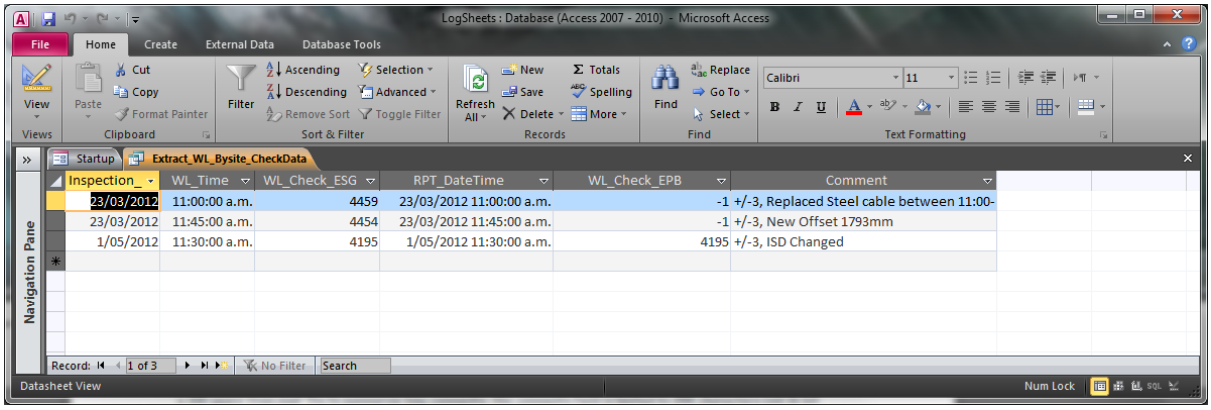
Open this file up. This brings up the following screen/form

Take the time to check ExSG and EPB zeros and confirm RL's of other reference points used for external reference check readings. Use all the readings provided, calculating them out if necessary. Field technicians usually do the utmost to provide something in the way as a reference if none exist or are damaged; all such readings should be in the check data. Maintain a datum consistency table for each site and update it after each survey run so that shifts to references are noted, confirm adopted levels are reliable and correct, and quickly calculate consistent reference data. All this should be referenced back to the ExSG/EPB readings



Under name > use the drop down menu and select the site you're working on i.e. MDC Himatangi Shallow.

Hit the WL Hilltop Check Data button, this brings up the following screen/table



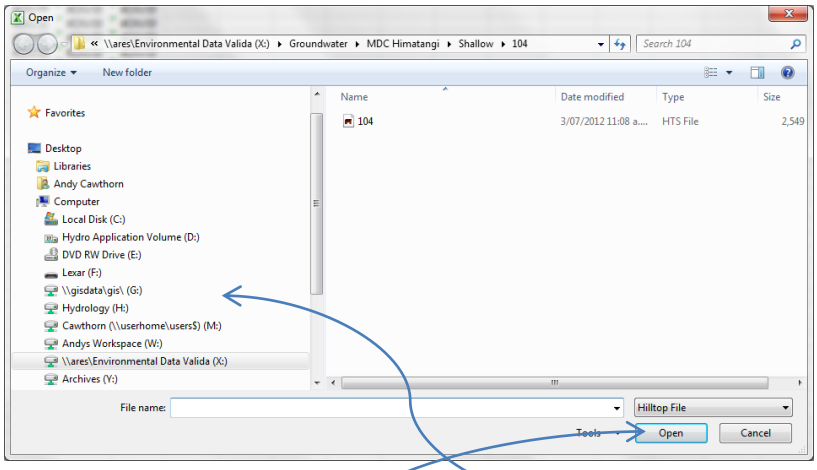
From here, you can simply *Select All > Copy >, and Paste* these data into Manager Check Data table. One caveat to this is that all inspections in the Log Sheet Loader still need double checking for errors. Most inspections, given the share volume made since 2004, and the smart idea of installing pump rigs for water quality, are entered on-mass.

Populating the Register

Open the Inspection Register on \\ares\Environmental Data Validation\Groundwater\MDC Himatangi\Shallow\104\104 Inspection Register.xlsx

In the Dump Tab, use the Get Hilltop Check Data Script. If you do not have this, ask a data processor to add this to your PERSONAL workbook.

Hit the Get Hilltop Check Data Icon. This will open up the following dialog box:

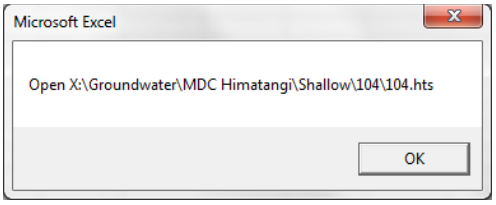


Locate your working Manager file; this should be in the working directory i.e.

\\ares\Environmental Data Validation\Groundwater\MDC Himatangi\Shallow\104\104.hts

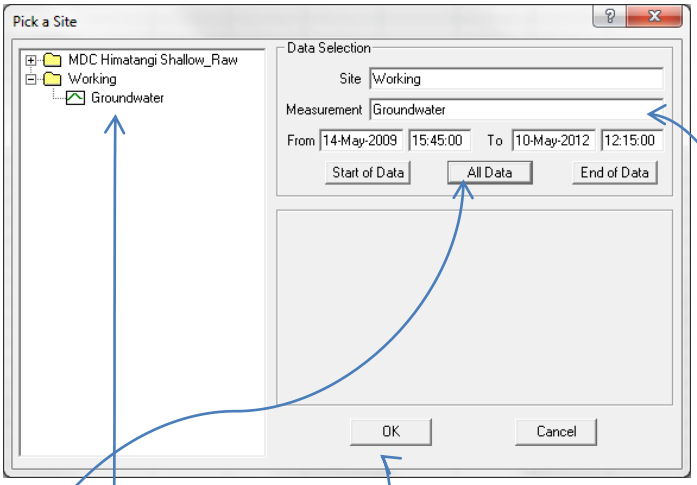
Hit the open button

This opens the following dialog box:



Obviously the only option you have here is OK

This then opens the following dialog box:



The site should be the Working file

The Measurement should be Groundwater (or the other relevant data source)

Hit the All Data button

Hit OK

This will populate the Dump Tab with the check data and the corresponding logger value.

Copy and Paste the Date/Times from the Dump Tab into the Date/Time cells in the Adjustment Register Tab

Copy and Paste the Check Groundwater values from the Dump Tab into either the ExSG/EPB cells in the Adjustment Register Tab

Copy and Paste the Logger Groundwater values from the Dump Tab into the Raw Logger reading cells in the Adjustment Register Tab

If there is a +/- error associated with the check reading, add this to the +/- cell

Editing Data

Editing data is pretty subjective. One person's interpretation of data will always, in some way or another (mostly the other), differ from some else's. Just the nature of the job

This section makes the huge assumption that the user processing data is reasonably confident and familiar with the editing features in Manager. If not, get training!

Remove large spikes when you find them. Any edits to data require a comment to be added to the comment tab in the Inspection Register.xlsx. The comment fields are pretty straight forward. All edits require a corresponding Quality Marker also. This is described in greater detail in the quality data section

All edits also require a corresponding quality marker added the comments sheet (See the Quality Schema for classification)

Check around inspections for hidden gap markers, spikes from the technician's on site, and other funny stuff. You can also run a hidden gap marker Virtual Measurement. Refer to the Hilltop VM documentation

Noise in the data can be interpolated using a range Virtual Measurements to 'Smooth' the data. Refer the Hilltop VM Documentation

Delete all data that is obviously wrong i.e. sensor faults. No synthetic data is can be applied to any dataset over events. Leave it as missing

Equally, if missing record is present in the dataset, locate the down-loads (if any). Check the data on Public Telemetry and even Hydro Telemetry as a first port of call. Don't just start making data up without checking in every location where data may be hiding. It is not always where it should be!

All ramp corrections to the data require evidence of the ramp correction applied in the Adjustments Register. If you are not familiar with this, ask a data processor

Other areas to look at are:

- Sensor changes, which relates to;
- Offset/multiplier changes. Find out what these are. Offset and multiplier changes are a real headache, often resulting in significant modification to the data and over significant time periods. This may lead to re-correcting older data files
- 'Weird' data
- Sudden drops and rises in water level, without inspections

Use the Inspection Register also when looking through the data. Look for large deviations of check data from logger data. If this occurs, find out why. Take some note of the +/- error (if any) and factor this into the equation

You may also want to run some really basic statistics also during your processing such as

- Reference deviation plots
- Sensor drift plots

All edits to data require comments that should contain no jargon or spelling mistakes (some people are quite anal about spelling). Nor should they contain overly specific technical information. You need to numb it down, people/scientists will not know what an ISD is, nor a RL0, which also means- NO abbreviations i.e. ESG, EPB, CG etc. this just adds to the confusion.

Ramp Corrections: Why the stigma or misconception on this? Ramp corrections are linear transformations of data as a function of time i.e. R (t). Ramp corrections often are required where a regular transformation fails to solve the problem. Since hydraulic data varies over bath space and time (rivers are not in a controlled environment, no matter how much operations tries), often the best practice is to make linear transformations to known references largely to correct

recording faults. Ramps minimise the effect of the fault. Ramp corrections do not distort data, nor do they introduce false record; they merely adjust data linearly that already exists.

Synthetic data should be classified as data obtained from indirect measurement i.e. using visually interpolations over extended periods of time, modelled data, instantaneous record created from discrete data points etc. The true test for synthetic data is: "Have I used or modified pre-existing data, to achieve the desired result (whether this is VM's, linear offsets etc.)?" if the answer is yes, then it cannot be called "synthetic", if the answer is no, then it most perceptibly is synthetic data

All data that there is no confidence in, or there is reason to believe that the data is not recording what it should be, should be deleted from the record. *Any unreliable data that cannot be repaired should be removed from the record*

Evidently, there is room for further editing tools. Will add to this latter

Correcting the Data

Once all the editing to the data has been finalised, or that you're happy with the corrections made, copy the working file to another file within the object tree called "Corrected". This is where you are going to apply the offset transforms and ramp corrections

Groundwater is measured to +/- 50 mm tolerances. Keep this in mind when applying transforms and/or adjusting offsets. Use some brain power to decide if there are shifts in the offset based on the information provided by the check Inspection Register and the logger values

Ramp corrections are applied when changes are made to offsets. Ramp corrections need to be applied from one offset to another to prevent steps in the data sets

There are a plethora of basic statistical tools available in both Hilltop Hydro and Microsoft Excel. Use these for justification for offset changes also

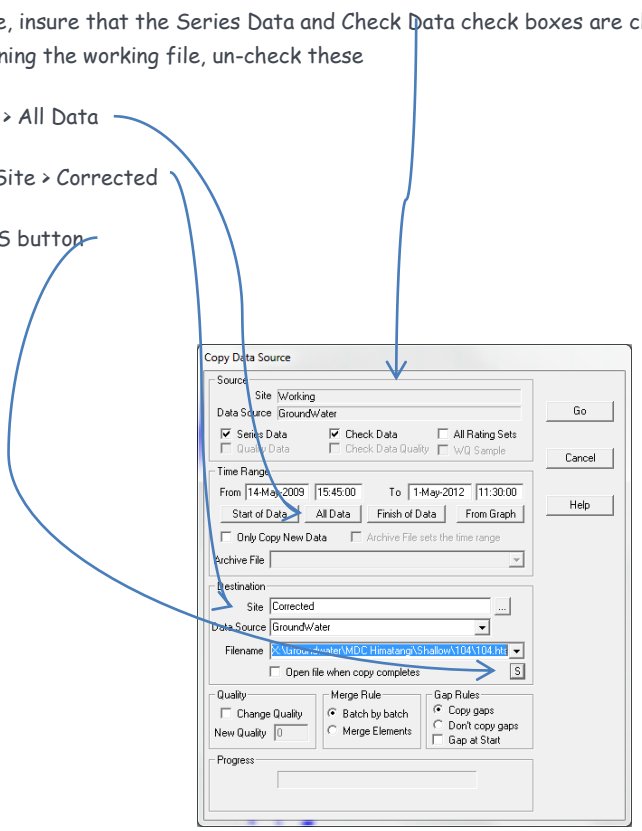
Right click > Copy

In the Data Source, insure that the Series Data and Check Data check boxes are clicked. If you have Virtual Measurements running the working file, un-check these

In the Time Range > All Data

The Destination > Site > Corrected

Filename > hit the S button



Hit the Go button once you're done

In the Inspection Register, hit the Dump Tab and re-run the Get Hilltop Check Data script, this time for the Corrected series data. Follow the same steps as described above

Copy and Paste the Logger Groundwater values in the Dump Tab, into the Absolute Difference cells in the Adjustments Register

Complete the basic maths in the Absolute difference row (Adjusted Logger – External ExSG)

Determine any offsets in the Offset/multiplier cells and/or Ramp adjustment cells

Once the offsets and ramp corrections (if any) have been established, open manager and correct the data in the Corrected Series Data (using the method that has been described throughout this document). If you are unfamiliar with transforming and/or ramping data, ask a data processor to help you through it

Once the transforms and/or ramp corrections have been applied to the data, copy the Corrected file to another file within the object tree and rename the file to the Hilltop Site Name i.e. MDC Himatangi Shallow (using the copy method that has been described throughout this document)

Quality Coding Data

Use the Inspection register to define the base quality codes. Populate the Corrected logger and Final Difference cells in the Inspection Register.

The NEMS (not even sure that this is correct yet) schema is as follows:

- All Transformations and Ramp Corrections ≤ 50 mm QC 600.
- All final differences between the ExSG/EPB and logger ≤ 50 mm QC 600
- All Transformations and Ramp corrections ≥ 50 – 100 mm QC 500
- All final differences between ExSG/EPB and logger ≥ 50 – 100 mm QC 500
- All Transformations and Ramp Corrections ≥ 100 mm QC 400
- All final differences between ExSG and logger ≥ 100 mm QC 400

Add the quality codes to the inspection register in the Baseline QC cells.

Once the base line Quality Data has been determined, add all the quality codes from the manual edits made to the data. The comments Tab on the Inspection Register (if you have been adding the QC as you go – as you should be…) will basically tell where and what the QC is for the edits made. These also require entering into Hilltop Manager. These are

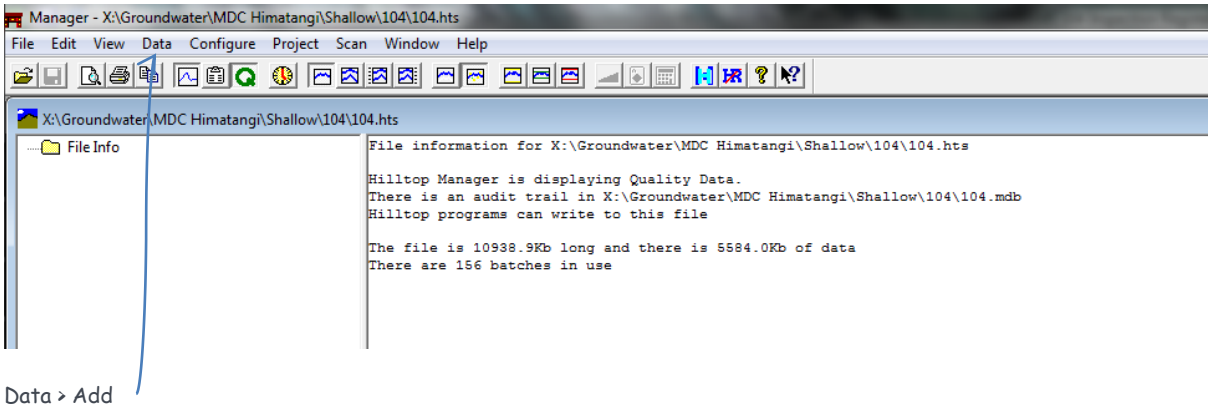
- All manual editing to the data ≤ 3 hours QC 500
- All Manual Editing to the data ≥ 3 hours QC 400
- All Synthetic editing to the data QC 300
- All unverified/cautionary data QC 200
- Missing record QC 100

Note that the Lowest Quality Code Prevails

This means exactly that: QC 100 ≥ (↔) 200 ≥ 300 ≥ 400 ≥ 500 ≥ 600

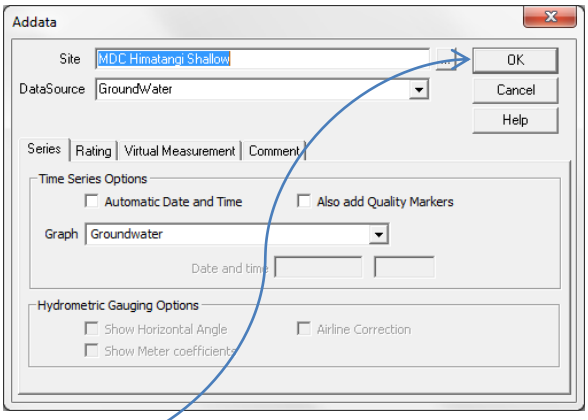
To enter the data into Manager:

Hit the Quality Data View Icon  . This brings up the following blank window:



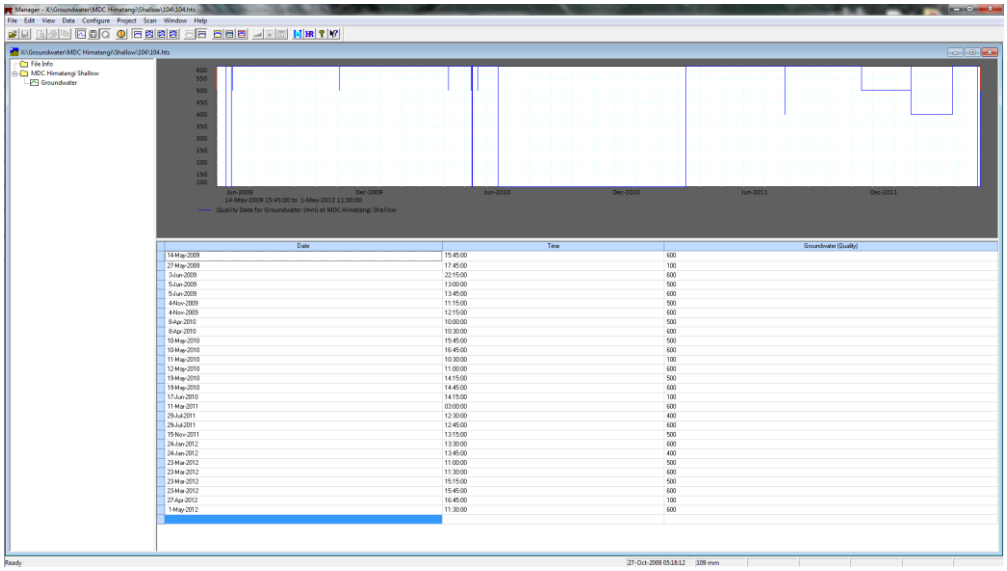
Data > Add

This brings up the following dialog box that you have seen before



All you really need to do is hit > OK

This brings up the data entry table



Add the quality data into the table. The headers should make it pretty clear what goes where.

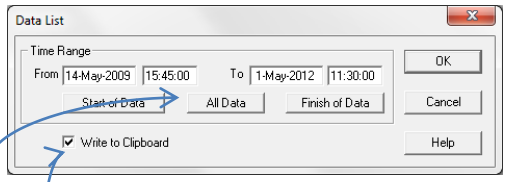
Once you have finished adding the quality data, save the file

Final Steps: Mopping Up the Paper Work

In Manager, list out the inspection comments.

Hit the  icon > Right Click > List

This brings up the dialog box



Hit > All Data

Insure that the  is checked so that the data is written to the clipboard

Paste the output into the File Details Template under Check Comments

Hit the 

Follow the instructions above for check comments (exactly the same, now just for the quality data)

Paste the output into the File Details Template under Quality Data

Right click on the Hilltop Site Name in Manager i.e.:

Right Click > Details

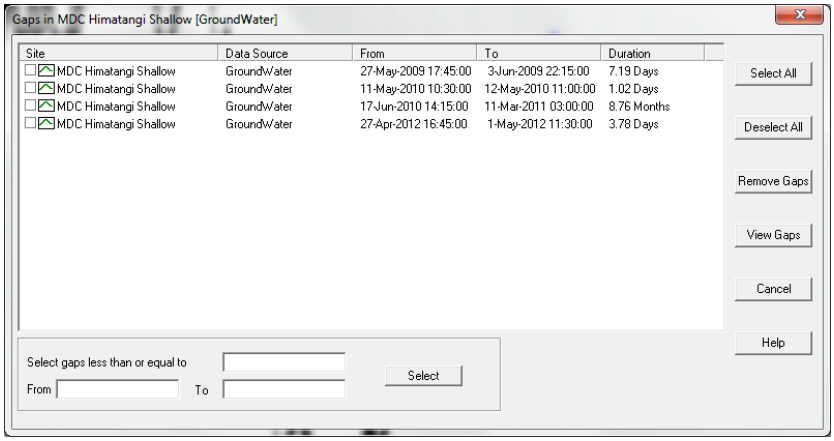
This gives a summary of the data sources in the Hilltop Site Name Dataset. You will need to physically click the text screen in Manager and Ctlr+C to copy the text output from the screen

All Data Sources, check, and quality data start and end Dates/times should be the same. If they are not, they need to be. All data should end on an actual inspection, and never on missing record

Paste this into the File Details Template under Final Details and Gaps

Click on the Data source > Right Click > Gaps

This brings up the following screen (this should really be hitting home by now…):



Select all > View Gaps

This gives a summary of the gaps in the data source(s) in the Hilltop Site Name Dataset. You will need to physically click the text screen in Manager and Ctlr+C to copy the text output on screen

Paste this into the File Details Template under Final Gaps

Under Additional Information in the File Details Template, add any relevant comments. These comments should be aimed at helping during the review

Cleaning Up and Printing Off

Populate the URF in the Inspection Register (it even has a tab, like the rest). It should be pretty obvious what goes where. Print this off. It is the Cover sheet to your processing

Print off the File Details Template

Print off the Adjustments Register

Print off any xlsx graphs you may have done

Print off the comments sheet(s)

Print out the Datum Tables

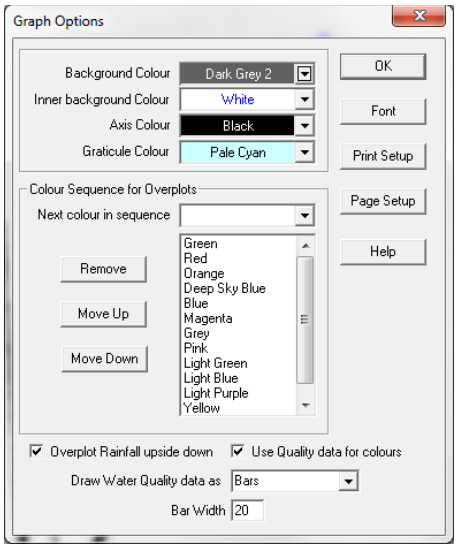
Print out the Previous Inspection Register (helps during review)

In Manager, print out the final graph at full scale at the full time range. Insure that the quality marker colours are enabled. Print this graph out

The graph quality marker colours can be toggled on/off using the Configure menu i.e.:

Configure > Graphs

This brings up the following dialog box



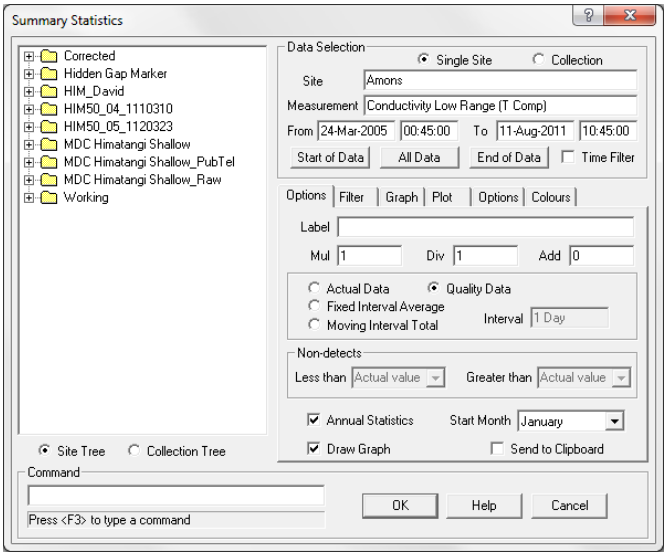
Use the  check box to toggle the quality marker colours off/on on the grapghs. It even has a comment 'Use Quality Data for Colours'.

Hit OK

Open Hilltop Hydro using hitting the  icon

Once in Hydro > Graphs > Statistics (PSummary)

This brings up the following dialog box:



Im not going into a huge amount of detail here, there is documentation on all the quirks of this programme in the Hilltop Help menu

Unsure that the Single Site  is checked

Select the Hilltop Site name and the data source you are working on i.e. Groundwater

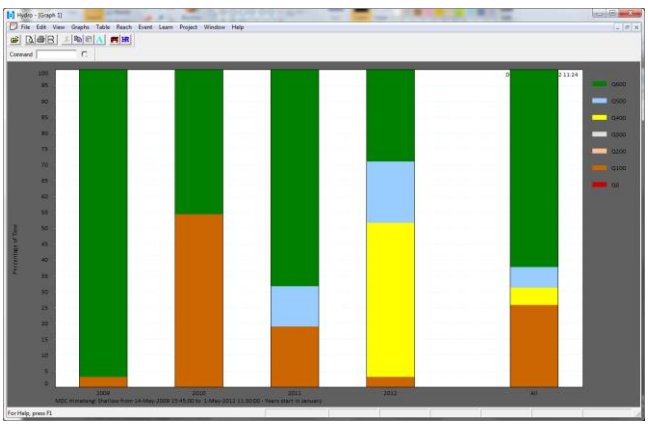
Hit > All Data

Unsure that the Quality Data  is checked

Make sure that the following are also checked:  Annual Statistics,  Draw Graph, and  Send to Clipboard

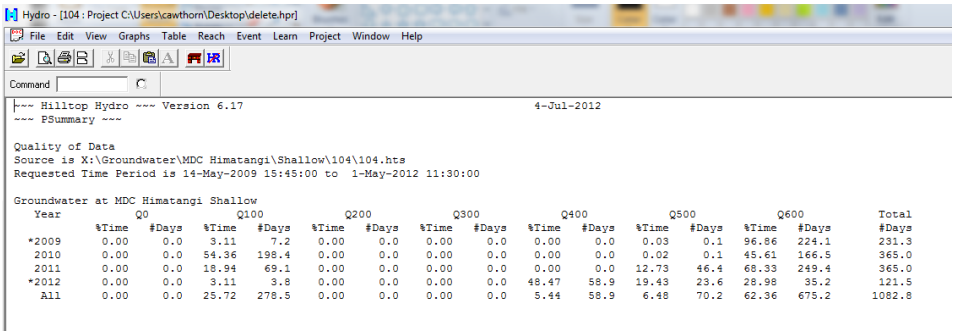
Hit OK

This brings up the quality data in bar graph form



File > Print

Close the graph in Hydro (not Hydro, the Small x in the viewing window), and you will see that there are actual generic/ numeric summary statistics in the background




```
~~~ Hilltop Hydro ~~~ Version 6.17                                  4-Jul-2012
~~~ PSummary ~~~

Quality of Data
Source is X:\Groundwater\MDC Himatangi\Shallow\104\104.hts
Requested Time Period is 14-May-2009 15:45:00 to  1-May-2012 11:30:00

Groundwater at MDC Himatangi Shallow
   Year          Q0            Q100          Q200          Q300          Q400          Q500          Q600          Total
            %Time  #Days   %Time  #Days  %Time  #Days  %Time  #Days  %Time  #Days  %Time  #Days  %Time  #Days  #Days
  *2009      0.00    0.0    3.11    7.2   0.00    0.0   0.00    0.0   0.00    0.0   0.03    0.1  96.86  224.1  231.3
   2010      0.00    0.0   54.36  198.4   0.00    0.0   0.00    0.0   0.00    0.0   0.02    0.1  45.61  166.5  365.0
   2011      0.00    0.0   18.94   69.1   0.00    0.0   0.00    0.0   0.00    0.0  12.73   46.4  68.33  249.4  365.0
  *2012      0.00    0.0    3.11    3.8   0.00    0.0   0.00    0.0  48.47   58.9  19.43   23.6  28.98   35.2  121.5
    All      0.00    0.0   25.72  278.5   0.00    0.0   0.00    0.0   5.44   58.9   6.48   70.2  62.36  675.2 1082.8
```

File > Print

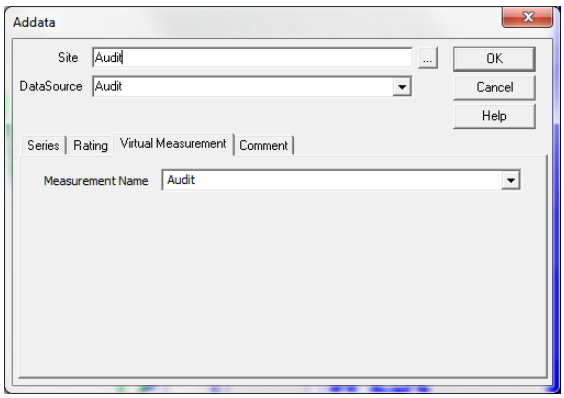
You can close Hydro now

Over plot the Raw and Final Datasets. Print this off. Enure that you have the quality colour markers OFF. Scale the y-axis to an appropriate scale so that it can be read

Create the Audit Virtual Measurement

Data > Add

This brings up the following dialog box



Site > Audit

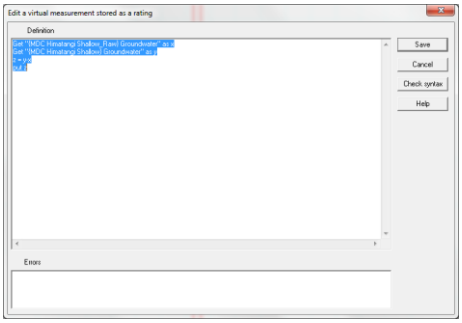
DataSource > Audit

Virtual Measurement Tab > Audit

OK

This brings up the Virtual Measurement dialog box. Essentially all that is being done in the Audit VM is Subtracting the final from the raw datasets. This shows the auditor exactly what was going on during the processing stage.

Virtual Measurements work on conditional and Boolean logic. Further information about Virtual Measurements can be found in the Virtual Measurement Documentation



Save once your code is done

This should now display the output of your Virtual Measurement Code

Over plot the Audit VM (on a new right-hand axis)

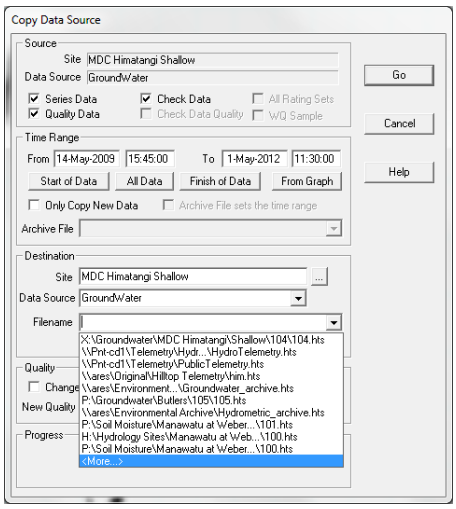
Print the graph generated in 3 month plots. You will also need to annotate the resulting outputs (currently there is no option to display check comments on graphs). Annotated plots should cover

- Adjustments made to the data
- Areas that warrant justification for edits/adjustments made
- Suspect areas that you're not confident with. The reviewer/auditor can then make decisions on best practice. It also acts as a second opinion. If you're lucky, you might only get the one…
- Inspections that you have issues with or left from offset adjustments

Copy the data to the Provisional Archive. Use the common name in the copy dialog box when doing to insure that it is ging to the right file ie

On the data source > Right Click > Copy

This opens the following dialog box:

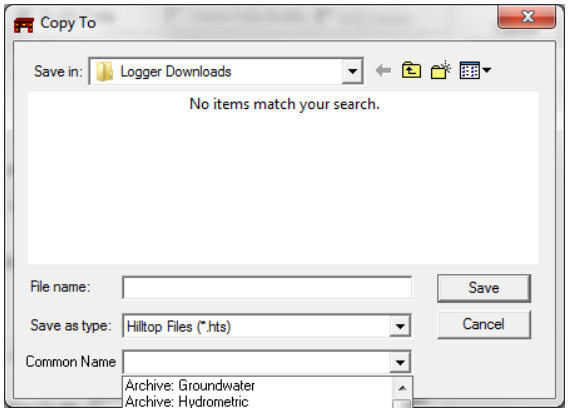


Ensure that the Series Data, Check Data, and Quality Data are checked

Time Range > All Data

File Name > Scroll to < More >

This opens the following dialog box:



Use the Common name and select > Archive: Provisional Data

Save

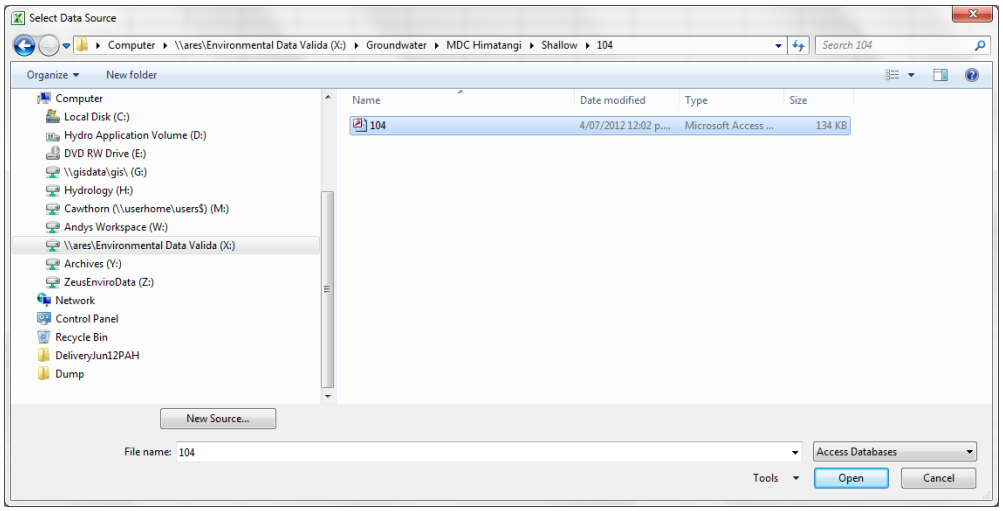
This goes back to the copy data dialog box > Go

You can now close the Manager file

In the Inspection Register under the Audit Tab

Data > From Access (Get External Data)

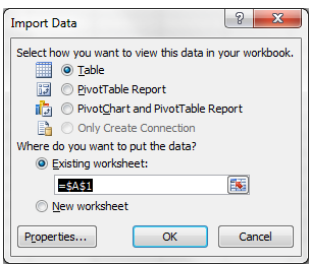
This brings up the following dialog box in Excel:



Locate and select the Audit.mdb in your working directory i.e.:

\\ares\Environmental Data Validation\Groundwater\MDC Himatangi\Shallow\104\104.mdb

This brings up the following Import Data dialog box:



Hit > OK

Format the StartDate and EndDate columns so that they contain both the Dates and Times

This will populate the Audit Sheet in the Inspection Register with data from the Audit Trail. Print this off

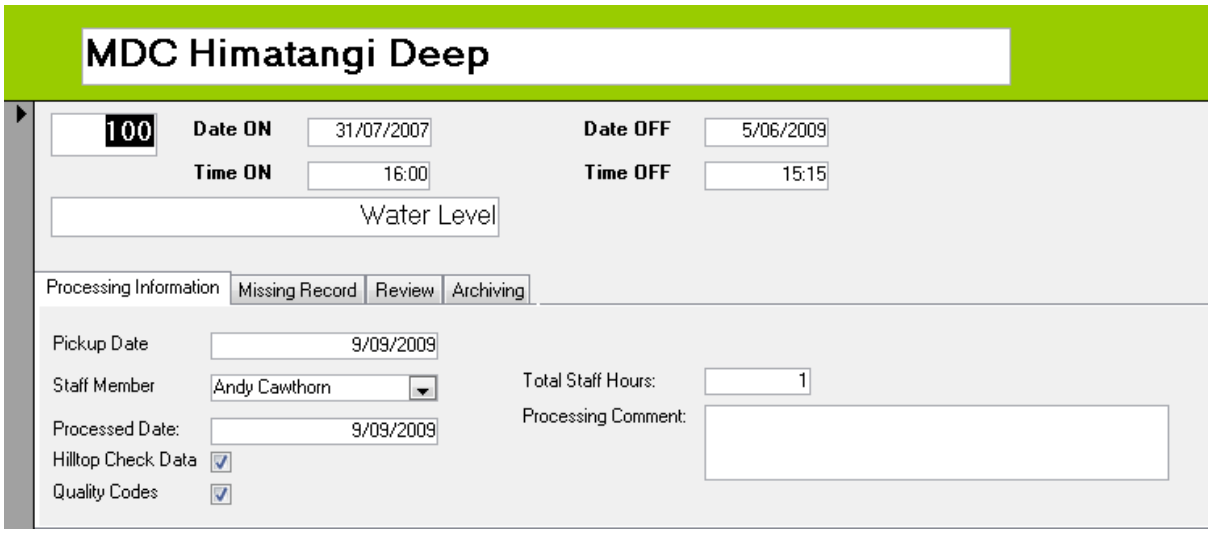
Final Touch Ups

Collate the above print offs

Add the photocopied inspection chits, surveys, calibrations etc to the above paperwork

Put in the pile with all the other data awaiting QA

Open CDTools



Add the date the processing was complete and the number of hours spent processing the data. Add any additional comment into the Procesisng Comment field. Use this aid the preceding batch ie if the data ends on missing record or the preceeding batch requires back correction of this batch

Finally, if there is mssing record, add this to the Missing recod Tab. You may also want to add all the synthetic Data you have created for this also. Check the  Missing Record and enter the number of Days