Catchment Data

Data Validation

Historical Infrared Turbidity records.

Step 1: Minor Corrections to Turbidity Data

All Collected Turbidity pre-installed SO7027 sensors.

1. McVan
2. Greenspan
3. Analite

Version 2010-1

horizons
regional council

Introduction.

The specification for Turbidity data is contained within the Water Quality portfolio documents.

These notes have been slightly modified following a review of data correction procedure for historical turbidity by Murray Hicks of NIWA. The notes also reflect comments and suggestions made by Marianne Watson following initial review of the Historical mathematical bolting of historical turbidity data flow chart

\\councildata\hydrology\hydrology web\...

Validation Methodology.

The methodology employed in the validation of historical turbidity data has been adapted in house from the following references:

- USGS guidelines for Continuous Water Quality
- EPA180.1
- ISO 7027

This Data Validation document identifies the procedures to follow when undertaking the initial processing of historical Turbidity data.

Currently (circa 2010) no international standards or well-defined processes are available for the collection and validation of continuous turbidity data; also instrument methods can vary from defined discrete standards (ISO, EPA). Pre 2010, all the inspection/lab results have been analysed in terms of the EPA180.1 ALPH 180.1 (white light spectrum 400-680nm) method; all field measurements have been undertaken typically to ISO7027 (monochrome "red light"- infra-red spectrum 780-900nm).

Please review the following documents to provide details on the differences and effects that these point have on the historical data.

USGS, office of Water Quality:

- Technical Memorandum 2004.03.

- USGS Implements new Turbidity Data reporting procedures. 2004
- USGS OWQ Field Manual Chapter 6, Turbidity. 2005

These corrections notes are for reference during the data correction process and assume that you are relatively proficient at the use of Hilltop software and procedures used to correct data. 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.

Creating Your Working Directory

The working directory for historical turbidity is located in:

\\Hera\Council Data\Hydrology\Hydrology Data Correction\Turbidity\Historical Sensors\Manawatu at Teachers College\100\

This directory structure is subject to modification for system requirements

- Copy the working files from this directory in your working file:

\\Hera\Council Data\Hydrology\Hydrology Data Correction\Turbidity\Historical Sensors\

  - *File Details Template.docx (FDT)*
  - *Audit.mdb*
  - *Comment Sheet.xlsx*
- Rename each file so that the batch number is in front i.e. 100 File Details Template
- It is critical the audit.mdb be renamed the same as the Hilltop file being used for data correction i.e. 100.mdb
- Open Hilltop Manager and create a new file within your working directory. The Name of this file should be the Batch Number i.e. 100.hts and the same name as the 100.mdb
  - If done correctly hilltop will indicate an audit trail is working in the background. Failure for Manager to recognise an audit trail will fail the corrected batch of data, as unlike other data sources, Historical Turbidity Sensor corrections contain no registers, and are solely reliant on the information contained in the audit trail!

Populating the Hilltop File

- Populate the Manager File with the raw data for the correction period from:

\\zeus\orig\MHP.hts

- Or

\\zeus\orig\old hilltop files\mhp.cmb

  - If you also require older data pre 10/10/2003
- Data Requirements are such:
  - Both high and low range Historical Turbidity Sensors Data (if both are available)
  - Water Level/flow data (Use SubArchive, or Hydrometric Archive Data if available)
  - Sediment Sensor Data (if available)
  - Qualarc Data. This can be either exported form the Qualarc.mb (as a .csv file) or using the QualarcDataMerge.dsn (Post 7/10/2010)
  - The Qualarc Data required is:
    - Turbidity and Sediment
      - TSS (HRC)
      - Turbidity EPA (HRC)

- During the copy process, rename the orig file name to underscore Raw ie Manawatu at Hopelands_Raw
- This informs the reviewer that this is raw data

Populate the File Details Template.docx (FDT)

- Open the FDT
- Under *Initial File Details*
  - In Manager, Right click on Manawatu_Raw > *Details* and paste this information into the FDT
- Under Gaps in FDT
  - In Manager, Right click on Manawatu_Raw > Gaps and paste this information into the FDT
- In the Gaps menu within Manager, select all gaps < 3 hours in duration
  - Delete these gaps from the record
  - Transform the data back to 15 minutes
  - List out the gaps as above and paste this information into the FDT under gaps Deleted
    - Gaps < 3 hours in duration are removed from the record as they are not overly critical, in the greater scheme of things. Repacking the data to 15 minutes after the removal of gaps < 3 hours in duration removes a lot of un-necessary time wasting filling the gaps manually

Adding Check Data

- Copy the Manawatu at Hopelands_Raw to another file called Working. Use yearly Working Names when dealing with large volumes of data
  - This Working file is where the check data is added, and corrections are made to the data
- Photocopy all log sheets for the period. Older log sheets may be difficult to obtain
- Using either Qualarc.mdb or the QualarcDataMerge.dsn, add the lab turbidity values to the inspection chits
- Highlight comments regarding turbidity in the field i.e. sensor cleans, replacement, relocation etc.
- In Manager, click the Check Data Icon (the clipboard icon) to start entering check data to the Working file

- Enter the check data into their respective fields. If you are unsure what to do, ask an analyst
- For inspections that do not contain lab samples, enter -1 into the check data field
- Save the file once completed
- Right click on the working file, and list out the comments into the FDT

All Qualarc data needs to be treated with caution and needs to be fully verified in this process (as the system contains numerous entry errors) like the following:

- *Samples taken on weekends (though this can be checked with log sheets).*
- *Samples taken at midnight (system has lost its time format).*
- *Other water quality parameter in the turbidity data source*

For most sites, more than one council team has taken samples, so there might be occurrences where there are multiple samples taken on the same day or samples without hydro log sheets. Some instances you will be required to resolve these issues by locating sample sheets or reviewing the sample project codes. If this cannot be achieved, use the average value (the lab sample has no real value to Historical Turbidity Sensor Corrections anyway)

Editing Data – Minor Edits

- When dealing with both high and low range sensors, combine both data sets together. Use the over-range low range data as an indication of the high range start points
    - Note that not all sensors share common over-range indicators
    - Use the Virtual Measurements (VM) in the Turbidity Data Sources for combining both high and low range sensors (if you are unsure how to do this, ask and analyst)
    - Some degree of caution should be used when combining high and low range sensors as they may introduce artificial "steps" in the data. However, in some cases, this cannot be helped
    - The low range sensor, as a general rule, is the primary
    - It is not always the case that the low range sensors provide better quality data than the high range sensor. This is a judgement call
    - Ideally, this should be done in the Manawatu at Hopelands_Raw file. This aids when copying large volumes of data around
- Create a series of smoothing VM's in the Working file under the Turbidity Data Source. These are the following:
    - 1 hour minima (QC 500 NEMS) or QC 18000
    - 3 hour minima (QC 400 NEMS) or 15000
    - 6 hour minima (QC 350 NEMS) or 13000
    - 12 hour minima (QC 300 NEMS) or 12500
    - 3 day minima (QC 250 NEMS) or 7000 or 9000 depending on adjustment
    - 7 day minima (QC 200 NEMS) or 4500
    - 7 day minima (as last resort)
        - During periods of noise in the data, use the above sequentially from highest to lowest quality
        - Transform the VM data to the noise in the Working file. DO NOT use the COPY function
        - During the transform process, *add comments to the audit trail* in the Transform Data dialog box
        - Failure of the VM's to solve observed issues, delete the noise in the data and leave as missing
- Remove individual/sporadic spikes when required (QC 500). Larger spike removals greater than 1 hour are synthetic data (QC 100)
- Missing record in the data requires a corresponding comment in the Comment Sheet.xlsx

- Ramp Correct all data during significant offset changes to data i.e. during sensor cleans; logger programme offset adjustments, installation of new sensors, addition of the VM data etc. Use adjacent data and current ramp correction guidelines to do this

Final Steps

- Once the all the corrections have been made to the data, copy the Working file to the site file i.e. Manawatu at Hopelands
- Open the FDT
    - List out the check data comments into the Comment List in the FDT
    - List out the Quality Data into the Quality Codes in the FDT
    - In Manager > Right Click Details on the Manawatu at Hopelands site and paste this information into the FDT under Final Details and Gaps
- List out any gaps in the final file and paste this information into the Final Gaps in the FDT
- Add any (relevant) comment to the FDT under Additional Comments
- Copy the final file to the SubArchive

To Print

- Open and complete the URF (Update Request Form) located in:

\\Hera\Council Data\Hydrology\Hydrology Data Correction\URF.xlsx

- Print out the FDT
- Print out the Comments Sheet
- Full Range plot, including QC, of the final site
- Full range over plot of Raw and final data (excluding QC Data)
- Full range over plot of Raw and final data (excluding QC Data). Over plot the actual QC data on the z-axis
- Print out the Audit Trail
- Use the PDFprint Factory to collate all the paperwork into one PDF document
- Add photocopied chits and other printouts to document

Synthetic Data Creation

- Once the above editing has been completed to the data, it's time to create synthetic record to fill created and observed missing record from the previous batch
- Create a new batch for processing
- Use the above file creation standards as above
- Copy the corrected data from the previous batch from the SubArchive to the new Batch

Rules of engagement

- Creating synthetic record should follow the sequential ordering below
- Synthetic data should be copied to the period of missing record in the corrected Turbidity data
- Apply standard ramp corrections to the adjacent data that the synthetic data has been copied to, esp. if the addition of synthetic data adds large offset changes to the turbidity data

1. Using the Sediment Sensor Data (if Available)

- Over plot the Working Turbidity data with the Raw Sediment Sensor Data
- Find a common trend between the two, ideally along a recession that is characteristic of the whole period
- Prep this recession information into excel and regress the information
    - Before doing this, use the Baseline VM to only select data where the Sediment Sensor is in the water. This will vary depending on Site/sensor configuration
    - Because of this, the Sediment Sensor is only really ideal for creating synthetic turbidity peaks
- Use the Regression slope information to create a new synthetic Turbidity Data Set (as outlined above). This data should be turbidity data derived from the Sediment Sensor Data
- Copy this data to observed missing record in the turbidity data and ramp to adjacent turbidity data

2. Flow to TSS to NTU Data (Excels version of Multiple Regression)

- If the current site being worked on does not contain Sediment Sensor Data, the next option is to use Multiple Regression in Excel
- Create a Hilltop Hydro Prep File containing the flow, TSS and NTU lab data. Prep this data into excel
- Regress the relationship between Flow and TSS
- Regress the relationship between TSS and NTU
- In Manager, create a new synthetic data VM using the relationships established above
- Use the above methods established for Sediment Sensor Data to copy the synthetic turbidity data to the observed missing record in the Working file

3. Visual Interpolation

- Following the failure of the above methods to derive acceptable synthetic record, visual interpolation of the missing record during peaks and recessions can be attempted
- When dealing with peaks, use the slope information to establish a respectable peak turbidity level. Try using the above two synthetic record datasets as reference
- Recession and baseline missing record should be done following standard methodology
- Visual interpolation has no specific instantaneous time interval. The key criteria is filling the missing record
- Obviously, periods of longer missing record require greater interpolation intervals i.e. 12 hour discrete points

Cautionary Notes

- All synthetic data is QC 100, its effectively MISSING RECORD, you're basically guessing
- Do not add synthetic data events, i.e. over a flood event. Leave it as missing
- Actual missing record is QC 0

To Print

- Same as above
- In addition, print out (to the PDFfactory) the quality data graph in Hilltop Hydro > Table > Psummary
    - Unsure that the tick is in the *Quality Data* check box under *Options*

Fin