**Water Use Data Assessment Part 2: Time-series visualisation and analysis**

**Aims**

Part 2 of the project aims to create analytical tools that can be used to visualise and assess the entire water use time series for selected WAPs. The tools are intended to support work such as catchment modelling, which requires detailed assessment of historic water use data. Consequently the tools are designed:

* to be run one WAP at a time
* to be run intermittently (eg, during catchment modelling projects)
* to focus on the entire time series of water use data within the Hilltop web server.

**Requirements**

In order to run the associated Python programs the user will need to have installed:

* Python Anaconda
* The hilltop-py package developed by Mike Exner-Kitteridge.

**Instructions for running Python programs**

***Create Time-series Plots.py***

* On running the program you will be prompted to enter the WAP that you are interested in.
* Once a valid WAP is entered the program will execute and will generate two files:
  + An Excel spreadsheet containing monthly summary statistics
  + A PDF file containing a collection of time series plots
* NOTE: some WAPs have extensive time series in Hilltop and, for these WAPs, the program may take several minutes to complete.

**Summary of steps within the program**

Once a valid WAP has been entered the program will work through the following steps:

***Initial extraction and processing***

1. Extracts a list of measurement types that are available for the WAP in the Hilltop WaterUse.hts file.
2. Iterates through the measurement type list, extracting water use data. NOTE: for some WAPs, multiple measurement types are available for a point in time. In order to avoid double-counting water use, the program only extracts a single measurement type for each time period.
3. Converts all measurement types to a common unit (volume extracted in cubic metres).
4. Filters out negative values in the dataset as it is assumed that these are errors.

***Create monthly summary statistics and export to Excel***

1. Detects spikes (extreme values) in the time series by:
   1. focusing on datapoints where extraction occurred (datapoints with zero extraction are filtered out)
   2. deriving measures of central tendency for the resulting distribution
   3. adding flags to datapoints that exceed certain thresholds (> 5 standard deviations from the mean, > 10 standard deviations from the mean, and > 20 standard deviations from the mean)
2. Derives various monthly statistics (including the measurement type at the start of the month, the number of days in the month with data, the number of meter reports received, and the number of spikes detected).
3. Generates an Excel spreadsheet that contains:
   1. a summary of the water use data that is available in the Hilltop WaterUse.hts file
   2. the monthly statistics derived in step 6.

***Create time series plots and export to PDF***

1. As its final step the program generates four time-series plots and exports them to a single PDF file:
   1. Meter readings received (daily)
   2. Daily volume extracted (m3)
   3. Volume extracted (m3) – 30 day moving average
   4. Average daily extraction rate (L/s) – this is essentially a duplicate of plot B but with a different unit.

**Troubleshooting**

1. ***ElementTree.ParseError***

Occasionally, on running the program, you may get the error message below. This seems to be a temporary issue with the Hilltop web server and is normally resolved when you re-run the program.

xml.etree.ElementTree.ParseError: not well-formed (invalid token)

1. ***Oddities in historic Hilltop water use data***

During testing of this program, various oddities were observed in the historic Hilltop data. Examples of these oddities included:

1. WAPs with large quantities of negative datapoints. For example, BY20/0088-M1 has over 22K negative values in its time-series. (In order to deal with this oddity the program filters out all negative datapoints).
2. Some WAPs have incomplete measurement lists in the Hilltop WaterUse.hts file. For example, M34/5722-M1 has no start date for its ‘Compliance Volume’ data. (In order to deal with this oddity the program inserts a default start date if the start date is missing).
3. WAPs with water use stored using the wrong measurement type. For example, N34/0342-M1 has a year of ‘Compliance Volume’ data – from 2015-07-01 until 2016-06-30 - which looks like it should be classified as ‘Water Meter’ data. (NOTE: the program doesn’t attempt to deal with this problem and it will definitely lead to erroneous results).

Please let me know if the program ever crashes, or returns warning messages, as this may indicate there are additional oddities that need to be considered.

**Contact details**

Alan Ambury

Email: [alan@whiterockconsulting.co.nz](mailto:alan@whiterockconsulting.co.nz) Cell: 0274 942 263