Improving operational service through Real Time Analytics.

Paper Number – 10

Presenter

Iwan Jones (MWH Stantec)

Contributors

Rob McTaggart (MWH Stantec)

1.0 Background –The Need for Data Integration and Analytics

The need to move to a position where our performance is dominated by planned interventions instead of reacting to asset failure.

In recent AMP periods there has been significant investment in improving water company data systems. The majority of UK water company operational decision-making however remains largely reliant on after-the-fact information, based on alarms, work history and customer contact. Interventions are therefore predominantly driven by reactive behaviours.

It is generally accepted that that improved visibility of real time asset performance would enable water utilities act quickly and proactively, potentially reducing the need to deal with:

- high operational costs of asset failure.
- high capital costs to build redundant capacity, and
- customer complaints and increased risk of non-compliance.

This paper provides three examples demonstrating how the combination of analytical software, open source programming, telemetry data and open source data has been used to generate continuous analytics to gain an insight to real time performance of assets.

1.2 Analytics within the Water Industry

Analytics is the discovery, interpretation, and communication of meaningful patterns in data. Especially valuable in areas rich with recorded information, analytics relies on the simultaneous application of statistics, computer programming and operations research to quantify performance (Wikipedia).

Figure 1 shows three essential elements required in order to undertake analysis of data.

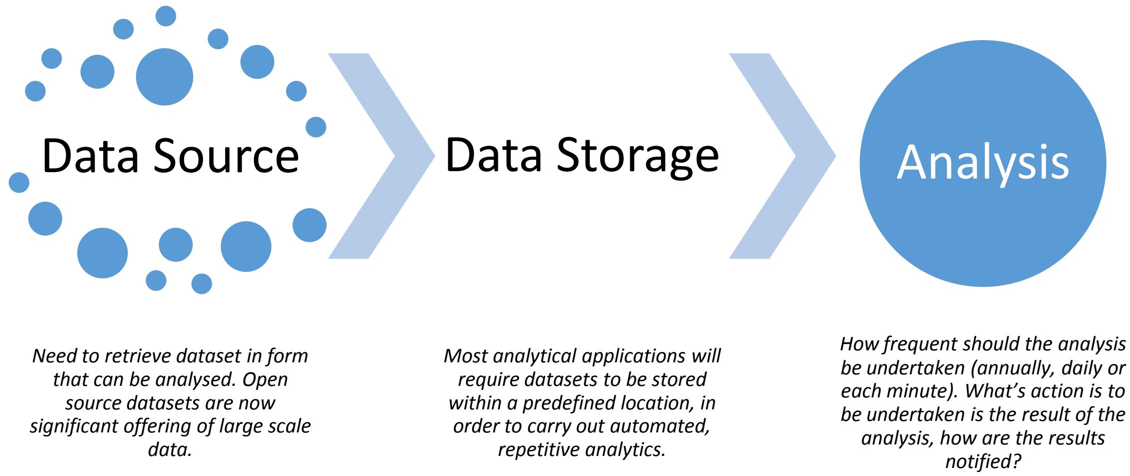

Figure 1 – Principals of Data Integration and Analytics

1.3 Data Integration

1.3.1 Data Sources

There are two main types of data sources associated with asset performance being:

- Water Companies Telemetry Datasets - already store established large datasets that provide an insight to the historical performance of networks. Datasets often represent historical performance of assets for number of years, if not decades.
- Open Source Datasets - Increase availability of open source environmental datasets published by government departments and agencies, public bodies and local authorities, a summary of which is shown in Table 1.

Data Type	Address	Description
Historical Tide Data	https://environment.data.gov.uk/flood-monitoring/doc/tidegauge	15 minute historical tide level data (chart data) at key ports.
Rainfall Data	https://environment.data.gov.uk/flood-monitoring/doc/rainfall	Mix of tipping bucket and
Water Framework Directive	https://data.gov.uk/dataset/catchment-management	Measured in river sanitary determinands for WFD sample sites.
Bathing Water Samples	https://environment.data.gov.uk/bwq/profiles/data.html	Historical E.coli and Intestinal Enterococci sample data.

Table 1 – Example open source environmental datasets

Open source datasets used within the examples provided within this paper are generally hosted in web address using the JSON format (JavaScript Object Notation). Ruby Script (and open source programming language) has been used to download open source datasets at a predetermined frequency.

1.3.2 Data Storage

There are many applications currently available to store geographical and time based data. In the examples provided below Infinity System has been used to store open source datasets and water company telemetry systems.

1.3.3 Data Analysis

Infinity Systems' 'Ask' functionality has been used to analyse data within the three examples provided below. The 'Ask' functionality generates a secondary derived stream from the predetermined analysed streams based on conditions set by the user.

Additional bespoke analytics have also been generated through the use of Ruby Scripts accessing the data stored within Infinity System via ODBC data exchange.

2.0 Analytics Examples

2.1 Example 1 – Assessing DWF Compliance at STW

UK water companies often assess measured treated dry weather flow (DWF) at Sewage Treatment Works (STW) against permitted DWF. Areas with significant planned growth are often analysed to review whether future DWF is likely to exceed permit limits as set by 'Water Discharge and Groundwater Activity Schedule' EPR7.01[1].

The current DWF is therefore assessed from the measured flow record as the largest 20th percentile of the previous years certified flow data to the STW analysed in three consecutive one-year sets. The example attempts to generate an analytic that compares measured DWF against permitted limits.

2.1.1 Analysis

An analytic was derived through the using Infinity System to:

- retrieve STW flows data from water company telemetry system and to calculate annual Q80 flows.
- The validity of Q80 flows are initially evaluated againast available corroborative streams if available (Flow to Works/ Flow to Treatment or Flow from Works).

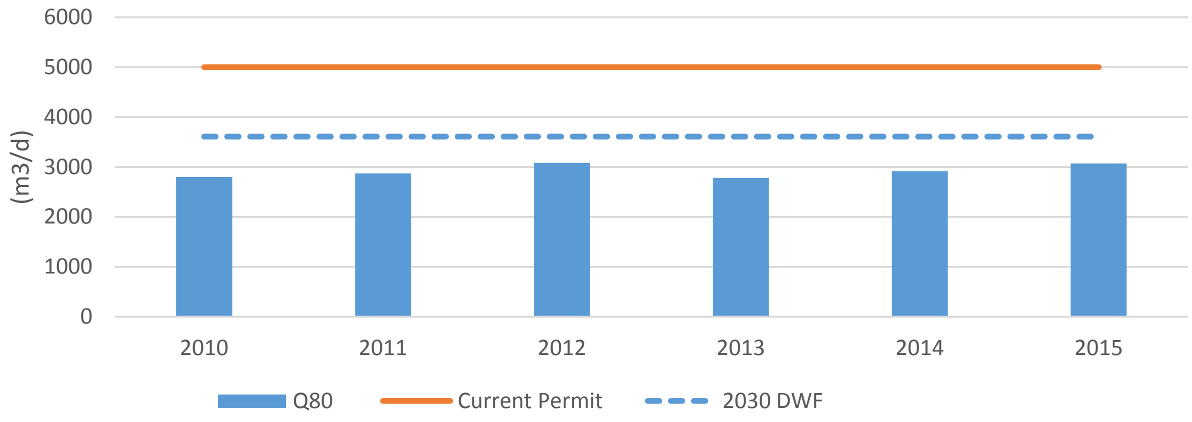

Figure 2 – An example output of a simplified analytic to continuously calculate Q80 flow from a STW

Direct link into telemetry system enables continuous calculation of annualised Q80 flows which is in turn assessed against current permit limits.

2.1.3 Benefits

The adoption of analytics provides an efficient process to continuously review the treated flows observed at STW. The process also eliminates the need for handing of data, and therefore reducing the potential for error.

Analytics also provides a significant reduction in timescales associated in processing the data of several assets. The adoption of similar analytics applied to in river sample data (which is available as open source dataset) used in combination with asset based analytics provide further insight to the impact of critical assets.

2.2 Example 2 – Use of Statistical Analysis with Bathing Water datasets

The revised Bathing Water Directive (rBWD) has been fully implemented in the UK since 2015. It sets more stringent water quality standards, puts a stronger emphasis on the management of bathing waters by the beach operator and greater provision of public information than in the original 1976 directive.

The rBWD uses an overall assessment of all water quality taken over a four year period to provide a clearer understanding of the long term bathing water quality. Bathing waters are classified into four classes: Poor, Sufficient, Good or Excellent. Classifications are based on levels of Escherichia Coli (E.Coli) and Intestinal Enterococci (IE), both of which are faecal indicator organisms (FIO).

Bathing water sampling data is for UK bathing waters are available through open source datasets.

This example attempts to generate an analytic that:

- derives bathing water classification from historical samples which in turn provides an insight to the likelihood of future change to bathing water status.
- review historical distribution of percentiles from historical bathing water sample results to provide an insight to whether bathing water are generally improving or deteriorating and also consider whether bathing waters are impacted by irregular exceedance events or many lower concentration pollution events.
- Evaluates whether bathing water exceedances events can be associated with predominant environmental conditions.

2.2.1 Analysis

Analytics have been generated to continuously obtain and review historical bathing water samples. This was undertaken through:

- Inclusion of Ruby Script activated on a daily basis to import Bathing Water Sample data.
- Timeseries analytics - to determine four year cumulative 90%ile/95%ile E.coli and Intestinal Enterococci values, and hence bathing water classification.
- Statistical Analysis - self generating histograms to assess whether bathing water exceedance events have occurred during certain environmental conditions, examples shown in Figure 3.

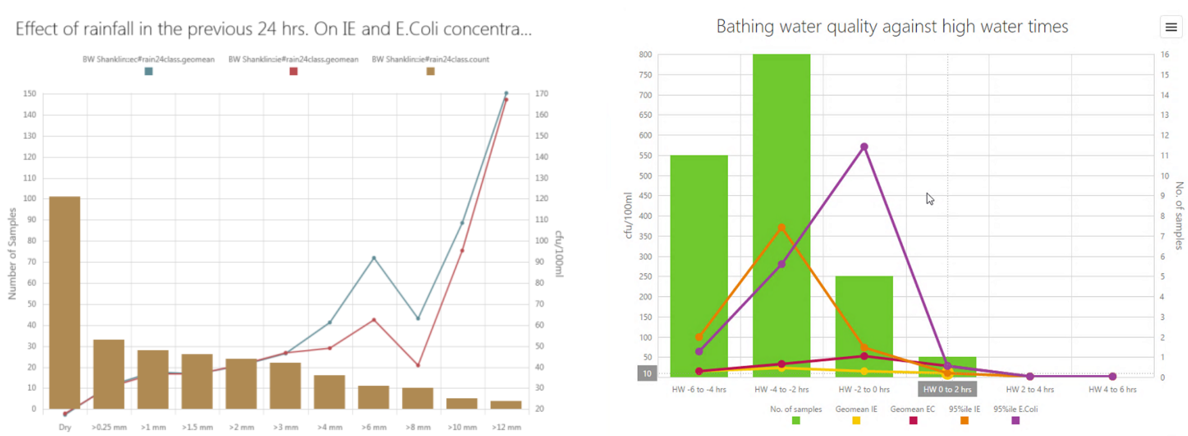

Figure 3 – Example statistical analysis of bathing water sample against environmental datasets

2.2.2 Benefits

Use of continuous analytics to determine assess historical bathing water samples provided an insight to whether bathing waters were generally improving or deteriorating. This also enabled an efficient process to quickly determine the impact of potentially discounting up to 15 % samples over a four year period[1].

A review of environmental conditions combined with historical bathing water samples may provide an indication to the likely source of bathing water pollution (i.e. whether the pollution source is diffuse, or likely to time based/ seasonal or during a particular tide condition). The findings statistical analysis is best used to prioritise the risks associated with known bathing water pollution sources as identified based on desk top land use assessments.

2.3 Example 3 – Identifying High Spill Frequency CSOs

Discharges from storm overflows are a reputational issue for the water industry. Population growth, urban creep, infiltration and changing rainfall patterns will further increase the pressure on storm overflows.

The Urban Waste Water Treatment Regulations[2] require sewer networks for agglomerations with a population equivalent of 2,000 or more to be designed, constructed and maintained according to best technical knowledge not entailing excessive costs. This includes the volume and characteristics of the wastewater and the limitation of pollution of receiving waters due to storm water overflows. The regulations supplement the duty imposed on sewerage undertakers by the Water Industry Act 1991 to provide, improve, and extend a system of public sewers.

The need to monitor the performance of storm overflows was set out by Government in 2013, with the expectation that the majority of storm overflows be monitored by 2020 (Benyon, 2013[3]). There have also been concerns expressed regarding the frequency of discharge as well as the environmental impact and population growth, urban creep, infiltration and changing rainfall patterns will further increase the likelihood of discharges from storm overflows.

This example attempts to generate an analytic to continuously assess the validity of recorded CSO spill events to determine annual and bathing season spill frequency based on the EA's spill block method.

2.3.2 Analysis

Create an analytic to continuously:

- retrieve CSO Spill event data as well as corroborative asset stream data (i.e. nearby measured level and flows).
- Create a series a rules /conditions to validate spill events.
- Determine spill event frequency using the Environment Agency's 12/24 hour spill block method.
- Create a Ruby Script that imports open source rainfall.

[1] Short term pollution (STP) is bacterial pollution that occurs at a bathing water and is not expected to last for more than 72 hours. Management actions to deal with STP are designed to help protect the public from predictable pollution of short duration at a bathing water (Environment Agency[4]).

A spill frequency analytic was generated with Infinity System which in turn gained access telemetry data, validated spills recorded by event duration monitors against associated asset data.

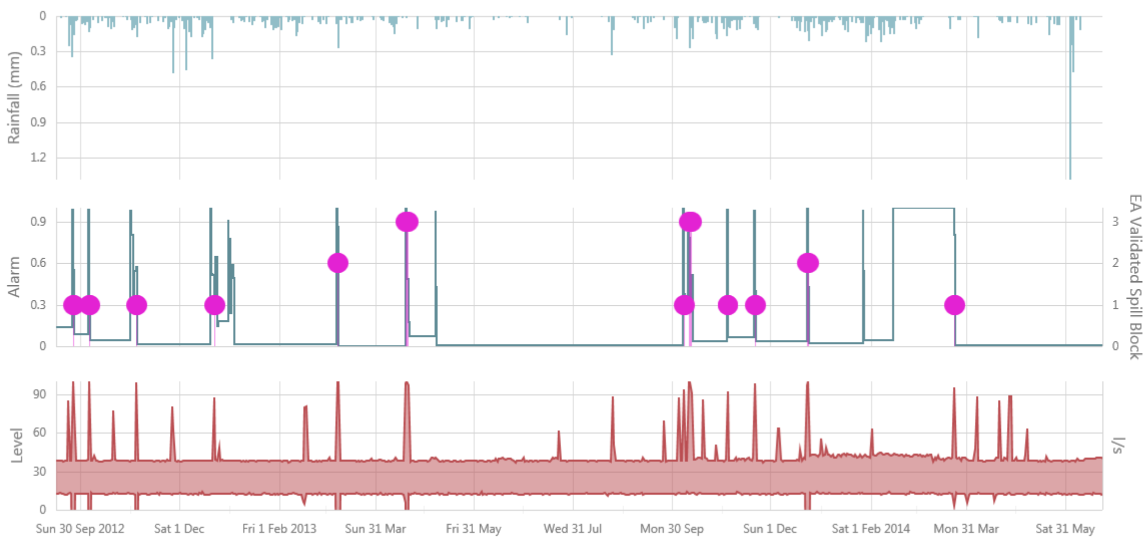

Figure x – An example validated spill data stream

Analysed Validated Spill frequency was in turn calculated using the EA's 12/24 hour spill block method.

2.3.3 Benefits

The example demonstrates that the assessment of overflow spill frequency can be automated with the inclusion of validation which in turn significantly improves the confidence associated with reported spills. The inclusion of notification (applied via Ruby Script) promotes the early warning of events enabling early interventions to address spill events discharge to sensitive waters or addressing spill events associated with operational issues.

Analytics directly linked to telemetry dataset provides modellers, engineers and asset planners a clearer view on long term performance issues, and also provides a valuable insight to asset resilience and to potentially identify irregular events (i.e. spill events associated with high groundwater).

3.0 Conclusions

Real time and predictive analytics are driving new forms of decision support, helping to generate improved levels of efficiency and enhance the services being delivered to customers of water and wastewater services. The digital revolution is providing exponentially better and quicker access to the broad spectrum of core asset data which is then being converted into information giving invaluable insights to asset performance.

Access to these real time analytics provides further opportunity to manage assets in real time and to respond to live conditions. They also enable us to predict asset performance based on forecast catchment and environment conditions and therefore implement the appropriate mitigation response.

The combination of new technologies such as open source programming, open source datasets and analytics in a seamless interface greatly assists water utilities in transforming data into useful insights, providing support for better decisions.

References

1. Water Discharge and Groundwater (from point source) Activity Permits (EPR 7.01) v3, https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/298081/LIT, Environment Agency.
2. Urban Waste Water Treatment Regulations (UWWTR), 1994, http://ec.europa.eu/environment/water/water-urbanwaste/index_en.html.
3. Letter from Richard Beynon MP sent on 18 July 2013 to CEOs of UK water and sewerage companies about spills from combined sewer overflows. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/364435/letter_2013_07_18_RB_to_CEOs_-_CSO_spills__2_.pdf.
4. https://environment.data.gov.uk/bwq/profiles/help-understanding-data.html, Definition of Short Term Pollution, Environment Agency