Fact Sheet 2 of 4 - Project 1075

Optimising Instrumentation for Better Process Performance

Module 2: Development of an Online Platform for Visualisation and Analysis of Online Instrument Data

Background

Project 1075 was developed as an industry collaboration to better understand, and begin to address, issues encountered by water utilities in using online instrumentation. It was separated into four modules focusing on key topics of interest. This fact sheet summarises the work in Module 2.

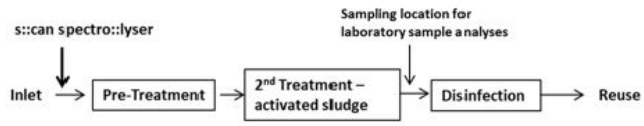
Within this project a case study was conducted by applying a data analytic approach to developing a real-time process warning system using online measurements. An online spectrophotometer was installed for an 18 month monitoring study (between 2013 and 2015) at the inlet of a wastewater treatment plant. During this time a web-based prototype portal with data integration, visualsation, prediction and anomaly detection functions for complex online data sets was developed in-house to assess the spectral data acquired by the spectrophotometer together with other databases (such as rainfall and temperature). Several data analytic options were triailled to extract useful operational information from the acquired data.

In this module the anomaly detection function, which includes pattern learning and comparison algorithms and a powerful user interface, was described in detail. By using the function, process upsets were successfully detected from the spectrophotometer data at the inlet of the treatment plant. The detected events / upsets were then compared with the treatment plant logs and were found to align well, demonstrating that the anomaly detection technique was effective and has the potential to inform decisions and assist plant operators. In addition, the proposed anomaly detection technique used a flexible algorithm.

Materials and Methods

Whyalla Water Reclamation Plant involves activated sludge treatment via sequencing batch reactors and chlorine disinfection. The biological process is prone to periodic upsets which seem to be due to some form of toxicity in the influent which is not easily detected by current monitoring techniques. Routine grab samples from the effluent were analysed for Ammonia (NHx-N), Phosphate (POx-P), Nitrite (NO2-N), Nitrate (NO3-N) and Suspended Solids (SS) using laboratory equipment and standard methods. Process engineers used these effluent analytical results to assess plant performance and defined the period when toxic shock occurred.

To investigate the source of this apparent toxicity, a 5 mm s::can spectro::lyser™, a sensitive double beam UV-Vis-spectrophotometer (200nm to 730nm), was setup in the inlet channel after the screens.



The raw spectral data files in CSV format were imported to a relational database (MySQL). Then an error correction procedure was applied to identify problematic data (negative values, null values and extremely large values) where the data values were checked and corrected. The error-corrected spectral data collected from this case study were used for further spectral analysis, one of the key aims of this project.

Data visualisation tool development

For the integration interfacing modules, an agile development approach to gradually build the following components was applied:

1. data visualisation and comparison tool;
2. connectors needed to access remote data for virtual integration;
3. data extraction tool for physical integration;
4. data summarisation tool.

This web-based prototype system is developed with a client/server system architecture, which is shown in the figure below - where a client is a web browser.

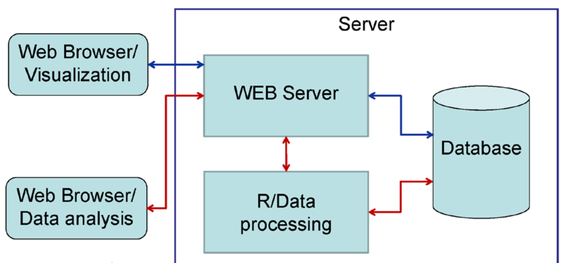


Figure 1: Client/Server system architecture

The prototype system is divided into two parts - the browser and the server. Commonly used web browsers, like Internet Explorer and Google Chrome, which are provided as part of modern computing operating systems (no need to develop specific software for graphical display), collect user requests (visualisation, or analysis) and parameters (time duration, or parameter names), and send this information to the server. After the requests are received, the server intercommunicates with the database and data processing server (R - a free software environment for statistical computing and graphics) to prepare data, generates the required visualisation and then sends the resulting graphs to the clients.

Applying pattern analysis to detect abnormal inlet water quality

This analysis aims to identify major patterns, normal (regular) and abnormal, in the spectral data. The normal patterns, when applied to new data, can be used to detect the states of the data quality in a different time / measurements.

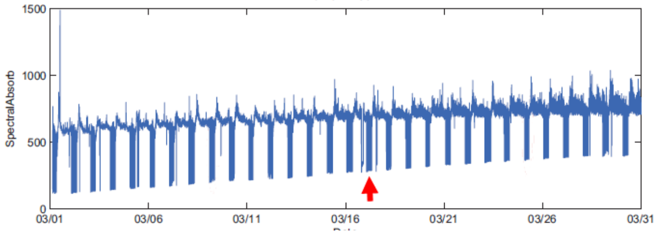


Figure 2: Raw data at wave length w=230nm during 2014 Mar

Abnormal patterns can be used to give warnings and also to help investigate/explain why an abnormal situation is occurring. Observations acquired from visualising the data showed clear daily patterns and a continuous cycle.

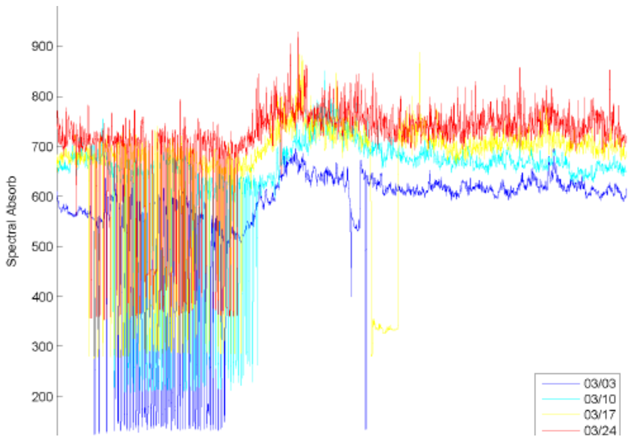


Figure 3: Raw daily data for each Monday

The developed data processing technique could allow an operator to understand whether the observed pattern is just a normal daily cycle or something very different at that time compared with the typical pattern. The challenges of this part are from the irregularity of impacts of the factors like device drifting and noise. It was interesting to see that there are persistent spikes with consistent readings everyday between 01:00 and 08:00 in the morning. This prompts us to believe that this daily pattern corresponds to some operational events, such as low flows (daily behaviour). We then excluded the pattern for further data analysis. The k-means clustering algorithm was applied and detected two patterns. The classification results of the observations in 02-Aug-2014 are shown below.

The normal data are labelled in green, the low flow data are blue, and the abnormal data are marked in red.

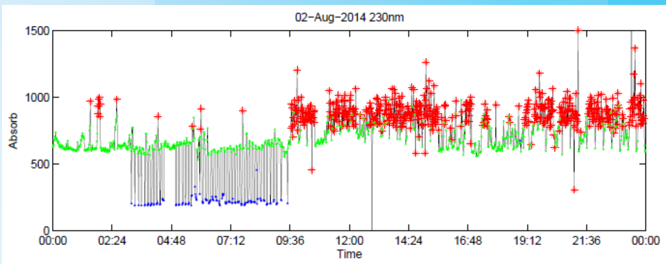


Figure 4: Classification result of 02-Aug-2014 data

Water Reclamation Plant process assessment

During all these periods high numbers of anomalies were detected by the anomaly model.

For example, the figure below shows the increased anomalies around a deterioration in performance in January and August 2014. The anomalies, associated with plant upsets, corresponded to high wavelengths, which are attributed to suspended solids. Periods of low anomalies generally corresponded with periods of stable plant operation, e.g. March 2014.

This has identified that a probable cause of the plant instability is periodic high levels of suspended solids. As this is predominantly a domestic catchment, an industrial source has been ruled out. Either cleaning of the network or scouring during high flows are the suspected source of the suspended solids. As a result of this knowledge, network cleaning activities are more carefully planned to minimise the risk of adverse impact on plant operation. In the last year or so the frequency of upsets has been significantly reduced.

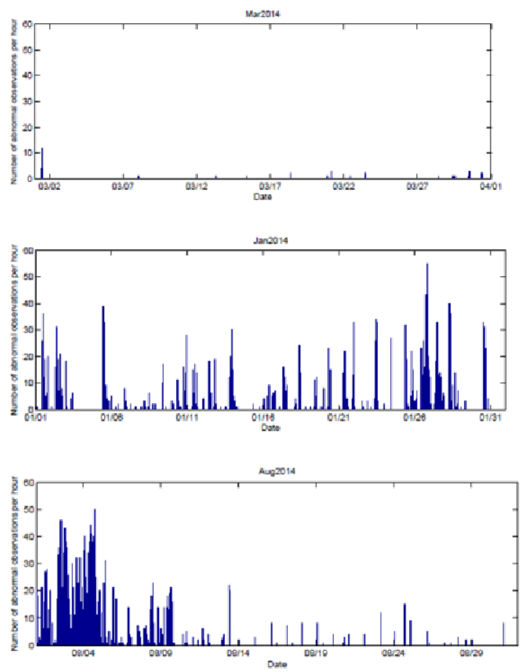


Figure 5: Plant performance variation at Whyalla in 2014

Module 2 Project Leaders

UniSA - Jixue Liu, Jiuyong Li, Guangsong Wang, Chen Zhan, Chris Saint, Xiaotian Guo, Chris Chow

SA Water - Nick Swain, Katherine Reid