water e-journal
Online Journal of the Australian Water Association
ISSN 2206-1991
Volume 3 No 1 2018
https://doi.org/10.21139/wej.2018.002

SUCCESS IN DATA ANALYTICS

SYDNEY WATER AND DATA61 COLLABORATION

D Vitanage, C Doolan L Maunsell, B Cameron, F Chen, Y Wang, Z Li

ABSTRACT

Sydney Water and Data61 are collaboratively researching advanced analytics approaches to solving water industry challenges, including water pipe failure prediction, customer segmentation, demand analysis, sewer corrosion prediction, optimising water quality, predicting sewer chokes and prioritising active leakage detection areas. The aim is to achieve better services for customers and to deliver world class network performance.

Data61 is Australia's leading data innovation group which was officially formed in 2016 from the integration of CSIRO's Digital Productivity flagship and the National ICT Australia Ltd (NICTA). The collaboration is a partnership arrangement, with Data61 providing data analytics research and development expertise and Sydney Water providing data and industry knowledge.

Both organisations have partnered to understand complex data sets that can be translated into knowledge, where the knowledge adds business insights. They help to see further, understand deeper and see it sooner about the data driven benefits from the collaboration, such as reducing cost, optimizing resources, minimizing uncertainty and mitigating risks, and improving services to customers. Within this partnership, both organisations have developed skills about data driven analytic on how to think about it, how to use it, and how to value it.

This paper outlines how Sydney Water has progressed on predictive analytics to develop capabilities of using machine learning and valuable tools for operators, shareholders and customers, with the collaborative effort from Data61.

Knowledge transfer has been applied as a key part of the collaborative partnership to facilitate the implementation of project outcomes.

INTRODUCTION

Sydney Water is taking an enterprise wide approach to building an analytics capability and culture within the organisation. The motivation behind this is two-fold: Firstly, it enables predicting the likelihood of certain scenarios or events, such as a pipe failure. This supports a pro-active response to these situations as quickly and efficiently as possible; Secondly, implementing analytics can help to identify and interpret contributing factors in these scenarios or events. Once these factors are investigated and comprehended, it is possible to mitigate them if they lead to detrimental or disruptive events or enhance them if they result in positive situations.

Machine learning, as a subfield of computer science, constructs systems that can learn from data, rather than follow explicitly programmed instructions.

Machine Learning is being used for six collaborative research initiatives between Sydney Water and Data61:

1. Improving the prediction of the likelihood of failure for critical water pipes and small reticulation pipes;
2. Customer segmentation and demand analysis;
3. Predicting critical factors related to preventing corrosion in concrete sewers;
4. Optimising water quality in delivery systems - a case study;
5. Predicting sewer chokes;
6. Prioritising active leakage detection areas.

In these projects, the use of data analytics techniques has demonstrated a higher level of confidence for the asset performance prediction. For example, the improved prediction tools for demand analysis and optimisation have a potential to improve customer services and regulator confidence

METHOD

The collaborative effort on data analytics in these projects has used machine learning to predict a number of core requirements on pipes or processes for Sydney Water. The focus of the research is to learn from the current operation data and identify previously unknown or unconfirmed relationships. The aim of doing this is to improve the prediction of the required needs. In these projects, integrating current knowledge and expertise with data analytics has demonstrated promising values in predicting asset performance. The six projects are detailed as below.

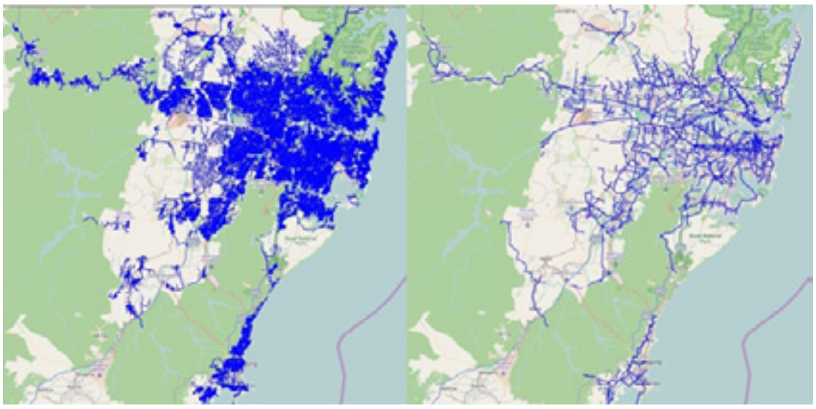

Figure 1. a: Reticulation mains. b: Critical mains

1. WATER PIPE FAILURE PREDICTION

Opportunity

Pipe failure prediction is being investigated for two classes of assets: critical water pipes (Figure 1b) and the reticulation network (Figure 1a).

Critical water pipe failure prediction (2013-2015): Sydney Water manages 5,000 kms of critical water pipes, investing about $32 million a year in renewals. Better targeting condition assessment will improve the efficiency of the investment in pipe renewals. Data61 and Sydney Water have developed an advanced machine-learning technique based conceptual model for Sydney Water which improves probabilistic prediction of high-risk failures on critical water pipes.

Reticulation water pipe failure prediction (2015-16): Sydney Water manages nearly 20,000 kms of reticulation water pipes less than 300 mm in diameter. Data61 and Sydney Water have been collaborating on innovative data-driven analysis and have developed a stochastic point process-based model for water pipe failure prediction in the reticulation network. Correct identification of pipes at risk of failure can assist in better allocation of the maintenance budget by avoiding inspections or renewals of pipes in working condition.

To target these two asset groups, two available categories of datasets were used:

- Pipe network: the attributes of all water pipes in the region being investigated (Figure 1a and b), including laid date, length, material, diameter size, location, protective coating, surrounding soil type, etc.; The oldest pipes were laid before 1900 and the average pipe age across the regions is about 45 years;
- Failure records: failure records from 1999 to 2015, including report date, type of failure, and failure location.

Methodology

Models determining which pipes are at risk of failure were created using non-parametric machine learning techniques for critical water pipes (Li, 2014, Vicky, 2013). A stochastic point process-based model is used to assess reticulation water pipe failure (Lin, 2016) in both short-term and long-term future modes. Based on the derived list of pipes at risk of failure, Sydney Water could prioritise pipes which require further ‘condition assessment’ inspections or renewals.

The models also use a range of pipe data, including material, laid year, coatings, failure history, etc. By investigating how these factors impact the previous pipe failures, the relative likelihood of failure between the set of pipes is included in the model to aid the prediction.

Outcomes and Benefits

The predictions were validated by separating the data into “in-sample” data and “out-of-sample” data. The in-sample data are used to build and train the models. Based on the prediction model, a list of high risk factors is derived. Then the out-of-sample data are matched with the derived list for model validation. The validation of models are carried out on both critical water pipes and reticulation water pipes through Sydney Water.

Critical water pipe: The model developed from the in-sample data was tested and benchmarked against the model which was being used by Sydney Water for estimating probability of water pipe failure at the time of the project. The validation demonstrated that the new conceptual model would have identified significantly more potential failures with the same inspection effort by using the out-of-sample data. This is highlighted in Figure 2.

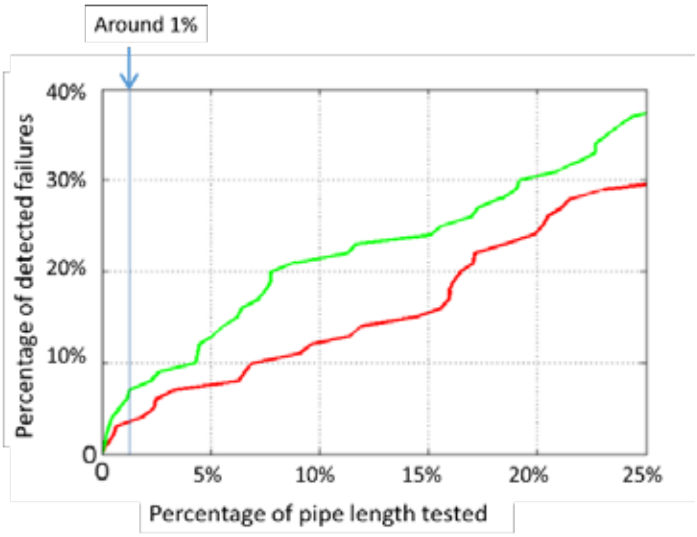

Figure 2. Predicted pipe length tested and corresponding number of failures detected, Data61 concept (green) vs water industry practice (red).

If 1% of the network was inspected, then 100% more actual failures would have been identified with the same effort.

With better targeting of high-risk pipes for critical water pipe renewals, Sydney Water has the potential to reduce maintenance and renewal costs by several million dollars over a four-year price determination period and minimise inconvenience to customers from pipe breaks.

Reticulation water pipe: The Water Model was also validated with reticulation water pipes for both short-term and long-term prediction. It also significantly outperforms other methods for the long-term prediction on Sydney Water reticulation water pipes (Figure 3).

With better targeting of high-risk pipes for reticulation water pipe renewals, Sydney Water can reduce maintenance costs each year and minimise inconvenience to customers from pipe breaks. Long-term failure prediction also better informs maintenance scheduling.

Future Collaboration

Further validation of the prediction outcomes is being implemented over the next couple of years by incorporating additional data, including soil, pressure transients and topography for selected areas of operation. Implementing the technique into Sydney Water's practices is the next step in the implementation of this project.

2. CUSTOMER SEGMENTATION AND DEMAND ANALYSIS

Opportunity

Sydney Water manages water supply and demand of over 1.8 million properties in the Sydney area, including both residential and non-residential properties. Better prediction of water demand in the next few years will provide significant value to the price making decisions and supply security planning at Sydney Water. In this project, the proposed model will use three datasets available from Sydney Water:

- Attributes of customer's properties, including a pool indicator, recycled water indicator, tenancy indicator, area, delivery system, build date, demand management indicator, water saving indicator and property type;
- External factors, including water price in the previous years, weather records and season indicator;
- Historical water consumption data.

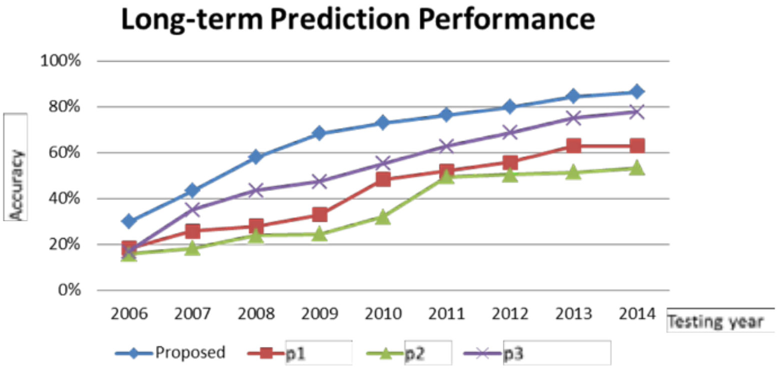

Figure 3. Performance comparison. Here p1 – p3 are three projections based on the current practice. For all the methods, 2000-2005 data were used for training and 2006-2015 data for testing. For each testing year, the top 200 pipes with highest cumulative risk of failure are shown, and the accuracy is measured by the percentage of the selected pipes that actually failed between 2006 and testing year.

Methodology

The modelling input includes the historical customer consumption records (e.g. monthly apportioned meter readings) and a number of factors/attributes of the properties. The flowchart of the project includes three steps:

1. Segment customers based on their historical consumption patterns.
2. Discover factors influencing consumption and behaviours for each segment.
3. Forecast future consumption using customer historical consumption records and other factors influencing customer behaviours.

Outcomes and Benefits

Customer Segmentation: Data61 has developed an approach to customer segmentation based on historical consumption patterns. Customers are grouped based on similarities in their historical consumption patterns.

Demand Forecasting: This work has resulted in a new machine learning model built on the data driven customer segments for demand forecasting (Li, 2016). The model adopts latent variables to predict unknown real monthly consumptions. The model can achieve less than a 0.5% prediction error and improved ability in tracking total consumption trend (Figure 4).

3. SEWER CORROSION PREDICTION

Opportunity

Predicting sewer corrosion is a critical task for water utilities worldwide in order to improve efficiency and reduce the cost of chemical dosing, sewer pipe rehabilitation and sensor deployment. The presence of sulphuric acid generated by gaseous hydrogen sulphide (H_2S) derived from the sewage attacks concrete structures in the sewerage system. A new and reliable toolkit, which enables spatiotemporal estimation of H_2S gas concentrations within the sewer network, is being developed.

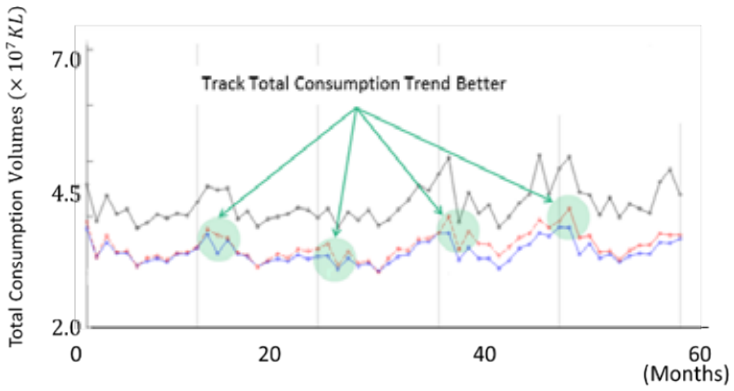

Figure 4. Consumption Training and Forecasting Curves (Black solid line represents the actual total consumption volume; blue solid line represents the accumulated consumption volume; red dashed line represents training consumption volume and red solid line represents predicted consumption volume).

Analytical modelling of spatiotemporal H2S distribution over the entire sewer network is challenging. Increasing the number of H2S monitoring stations is often infeasible due to cost and accessibility. Therefore, in this work an attempt was made to use data analytics to estimate the spatiotemporal distribution of H2S with a limited number of observations. The model not only estimates the H2S concentration, but also the uncertainty associated with the prediction, which is an important measure in decision making. These H2S concentrations are planned to be used in the overall data-driven corrosion model. The final outcome of the prediction model can be used to prioritise high risk areas, recommend chemical dosing locations, and suggest deployment of sensors.

Methodology

The expected toolkit is a desktop application with a user-friendly interface for querying and producing output on the geographic information system (GIS). The inputs of the toolkit include the sewer network system (i.e. GIS plan), monitored/sampled factors, and hydraulic information. The toolkit can also incorporate existing corrosion model's results (Wells & Melchers, 2016) as prior knowledge. The toolkit will use data analytics techniques to enable: (1) Spatiotemporal corrosion prediction over the entire sewer network; (2) H_2S concentration (and other parameters) estimation; (3) Chemical dosing optimisation and (4) Optimal sensor deployment.

Current Success and Vision

Data61 has successfully developed and evaluated models for spatiotemporal H_2S estimation and corrosion prediction. Chemical dosing optimisation and optimal sensor deployment for monitoring are ongoing as scheduled.

- **Spatiotemporal H_2S concentration estimation:** the proposed analytics model was tested in a Sydney sewer subsystem. There were 17 H_2S observation sites at a monitoring frequency of 15 minutes from Jan 2011 to Dec 2015. The aim is to estimate the spatiotemporal dynamics of H_2S on the entire network over time and visualise it on a map in video format. Figure 5 illustrates a frame of the video which plots the H_2S distribution on the network at 01:15:00, 15-Sep-2015.

- **Corrosion Prediction:** Based on the estimated H_2S concentration and other monitored or estimated factors, a second model, which integrates physical model (or experts' domain knowledge), has been developed to predict corrosion levels with a stated level of uncertainty over the entire sewer network.
- **Dosing:** The estimation of H_2S concentration and predicted corrosion levels can assist in the development of the dosing strategy. The locations and amounts of chemical dosing can be optimised according to the H_2S concentration, sewer corrosion level and hydraulic information on the sewer network.
- **Monitoring:** A good deployment of sensors can maximise the monitoring capability on the sewer network. New sensors can be installed at locations with high uncertainty of H2S estimation obtained from spatiotemporal factor prediction and corrosion prediction phases.

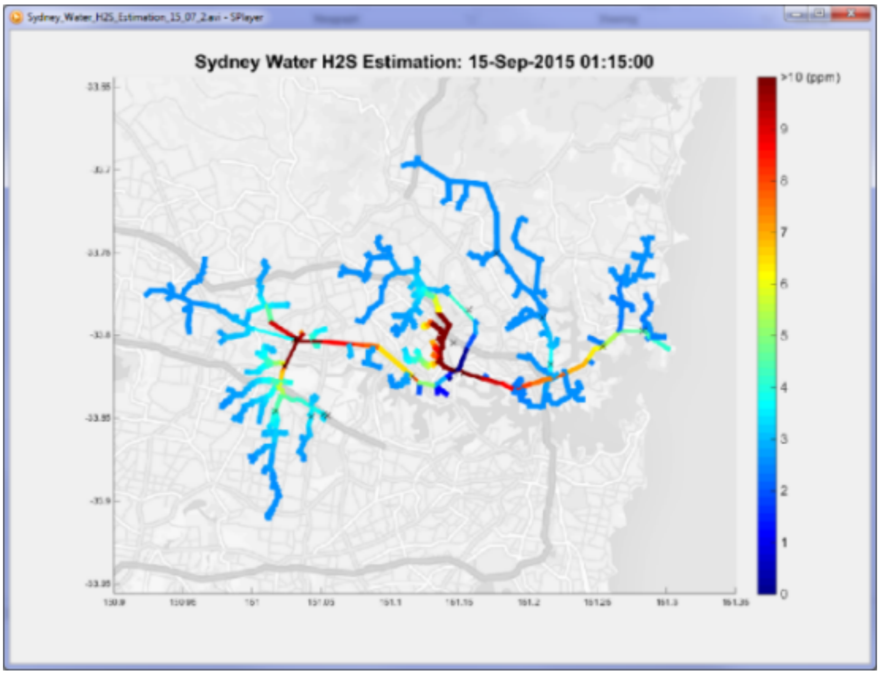
Sydney_Water_H2S_Estimation_15_07_2.avi - SPlayer
Sydney Water H2S Estimation: 15-Sep-2015 01:15:00
>10 (ppm)
9
8
7
6
5
4
3
2
1
0
-33.7
-33.75
-33.8
-33.85
-33.9
151
151.05
151.1
151.15
151.2
151.25
151.3

Figure 5. Spatiotemporal estimation of H2S concentration in the entire network over time visualised via video. The plot illustrates a frame of the video which plots the H2S distribution on the network at 01:15:00, 15-Sep-2015.

Conclusion and Future Work

The implementation stage will see the Toolkit become a desktop application with a user friendly interface. Without a need of specialised training, asset management staff will be able to input specific queries related to asset corrosion and have the choice of both GIS and non-GIS (e.g., spreadsheets, look-up tables) output formats.

4. OPTIMISING WATER QUALITY

Opportunity

Data analytics research will systematically study the impact of network operations on water quality (characterised by indicators, such as chlorine residuals) and energy costs (mainly generated by pumping stations). Historical network operation records, including reservoir levels, pumping status, pressures, water demand, energy consumption and tariffs, and water quality, will be used as learning data. Data analytic techniques will help determine optimal network operation strategies, including chemical dosing strategies and operational reservoir protocols, which can in turn help reduce energy use and improve water quality throughout the network. The Woronora delivery system in Sydney's south is used as the case study area in this project.

Challenge: Quantify uncertainty in water quality and demand predictions for supply networks

Intelligent networks include heterogeneous data ranging from reservoir levels, pumping and controlling valve status, flow rates, pressures, and chlorine residuals. This data comes in a variety of formats, time resolutions and volumes, making it complex to aggregate and analyse by traditional methods. Moreover, predicting chlorine residuals and water demand is important for network optimisation, but the dynamic characteristics of the data make it difficult to predict future values accurately, even in a next 24-hour temporal window. A major challenge is to consider the assets jointly when optimising water quality and energy cost.

Methodology

The research efforts will investigate the water distribution systems using the network and smart sensing data to develop a decision making model based on data analytics to optimise and manage different operational parameters.

The case study will use machine learning techniques to:

- understand the important features that may cause or explain the variations in water quality, energy operations, and dosing responses.
- optimise normal system operations in terms of cost using the energy consumption and pumping costs, reservoir operating windows, system demand, network flow and pressure.
- optimise the Chlorine dosing program using the water quality data, historical dosage records and system demand data.

Current Success and Vision

Data61 has successfully constructed the water quality prediction model in the reticulation network, which is capable of predicting downstream water quality in a real-time manner.

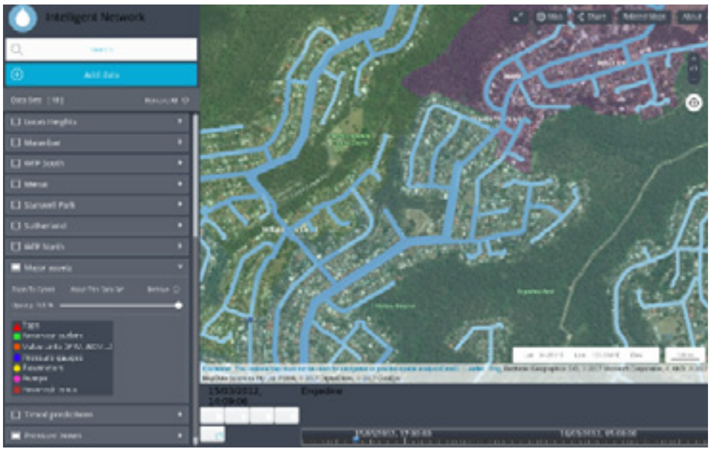

Figure 7. Water quality situational awareness tool

The expected future outcomes from the intelligent water quality predictive tool for the Woronora delivery system would demonstrate whether machine learning based data analytics could provide improved decision making for water system management. The expected outcomes are shown in Figure 6.

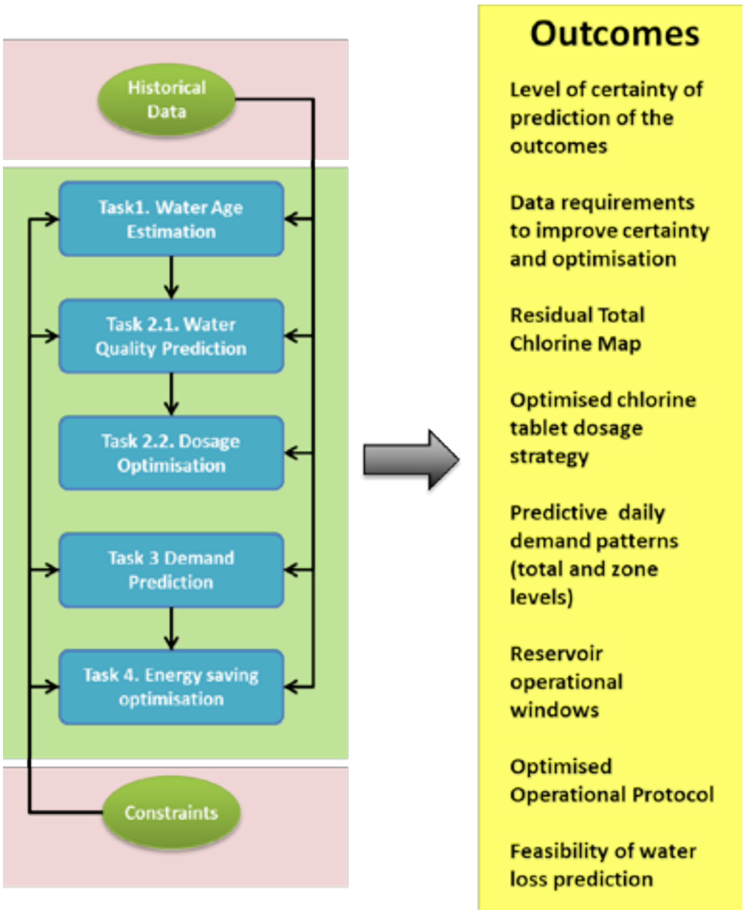
Historical Data
Task1. Water Age Estimation
Task 2.1. Water Quality Prediction
Task 2.2. Dosage Optimisation
Task 3 Demand Prediction
Task 4. Energy saving optimisation
Constraints
Outcomes
Level of certainty of prediction of the outcomes
Data requirements to improve certainty and optimisation
Residual Total Chlorine Map
Optimised chlorine tablet dosage strategy
Predictive daily demand patterns (total and zone levels)
Reservoir operational windows
Optimised Operational Protocol
Feasibility of water loss prediction

Figure 6. Intelligent water system data analytics expected outcomes.

In addition, a water quality situational awareness tool is developed. This platform is intended to map the forecasting results on a geographical map, which is capable of providing the decision support tool to aid re-chlorination and pumping scheduling, as shown in Figure 7.

5. PREDICTING SEWER CHOKES

Opportunity

Every year, Sydney Water responds to approximately 12,000 to 21,000 sewer blockages, known as chokes. A project has been developed to gain a better understanding of the causes of sewer chokes and develop an analytical tool to predict the likelihood of future chokes. The aim of the project is to provide data driven decision making to assist Sydney Water to shift from the current reactive choke management approach towards a predictive management approach. The data used in this project include:

- **Sewer network:** the attributes of all sewer pipes, including asset identification number, construction date, length, material, diameter size, location and etc.;
- **Choke records:** choke records from 2000 to 2015, including asset identification number, choke date, choke type, choke location and etc.;

Environment data: tree canopy[1], climate[2] and soil data.

Methodology

A stochastic point process-based statistical model for short-term (one year) choke prediction is proposed. Specifically, it is a new variant of Hawkes process[3](Hawkes, 1971, Vere-Jones 1988 and Diggle, 1994). It is an interaction point process that considers not only the choke frequency of a pipe itself (determined by pipe characteristics), but also the triggered chokes caused by other chokes (spatially and temporally, e.g. previous chokes and nearby chokes). Intuitively, a sewer water pipe that has been subject to chokes in the past will block more frequently in the future. The model parameters of the proposed new variant of Hawkes process are learned from historical chokes from 2000 - 2015 and pipe characteristics (including material, length, laid year, etc.). The model is trained and validated on separated groups of pipes for different choke causes, including root, debris, grease, soft choke, etc.

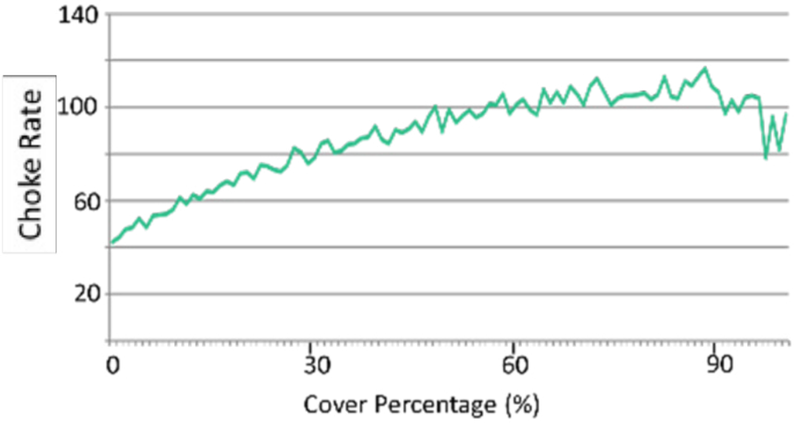

Figure 8. Tree canopy mapping example and the overall influence to choke rate[4].

Outcomes and Benefits

Detailed below is an example for verifying the effectiveness of the proposed prediction model on chokes caused by tree roots.

Tree roots: The majority of Sydney Water's chokes are caused by tree roots blocking sewer pipes. The exact percentage of chokes due to tree roots is quite variable and can be anywhere between 50% and 80% in any given year. From Figure 8, it can be observed that choke rate is higher in the highly tree covered areas. To improve the effectiveness of the sewer choke prediction, Data61 has taken the tree canopy mapping data into consideration (as one weight parameter) within the stochastic point process-based statistical model to more accurately predict chokes.

The data validation demonstrated that the Data61's conceptual model for the total Sydney Water system can identify about 10% of tree root caused chokes with only 1% of the total length of the network inspected (Figure 9).

Conclusion

Factor analysis and sewer choke prediction have been conducted in a collaborative project between Sydney Water and Data61 based on Sydney Water's asset and choke data. Various insights about factor influences have been derived from factor analysis. These insights help understand the importance of different factors and help build an accurate choke prediction model based on stochastic point process. The model considers intrinsic choke patterns caused by physical attributes and previous chokes together for predicting future chokes. Experiments have been conducted on Sydney Water data for justifying the effectiveness of the proposed model. Data61 is developing software tools for sewer choke prediction which can analyse multiple data sources, visualise factors, cohort analysis results, and support prediction result online generated and downloadable. This project is a finalist in the 2017 AWA NSW Water Awards Research and Innovation.

[1] The plan view of a tree canopy is representative of the majority of the root zone. It is estimated that the root zone will extend to at least the tree canopy coverage.

[2] There are a number of climate indicators that are considered to influence the likelihood of a tree root blockage. The conditions include rainfall, temperature, evaporation and soil moisture.

[3] Hawkes process is a classic spatial-temporal statistical model which only considers the spatial-temporal relationship between points while our developed method considers both intrinsic pipe characteristics (e.g., laid year, length, and previous chokes) and external factors for predicting future chokes.

[4] The choke rate from tree roots decrease between 90% and 100% tree coverage as a large proportion of these sewers are in undeveloped areas and often downstream in the system. There are two potential causes of this phenomenon: (1). there are less junctions on sewers with high tree coverage, and (2). downstream sewers are often deeper.

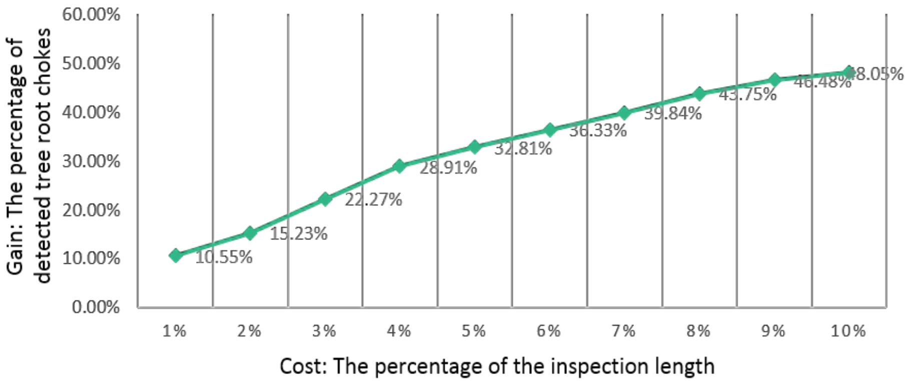

Figure 9. Short-term prediction for Ryde subcatchment for 2014. The model is trained by using all the choke data from 2000 to 2013 and tested by using the choke data in 2014. The selected pipe (Ryde) was correctly predicted if it choked in the following one year.

6. PRIORITISING ACTIVE LEAKAGE DETECTION AREAS

Opportunity

Each year Sydney Water inspects a subset of pipes through an active leak detection (ALD) program in order to find leaks. In 2013-14 Sydney Water spent around $1.5M on 15,000 km of pipes for ALD. The ALD program involves scanning of reticulation pipes and fittings using acoustic leak detection equipment to locate leaks. Then pressure zones are selected for inspection based on the predicted leakage volume, cost of water and cost of intervention program.

The efficiency of the ALD program can be improved by segmenting large pressure zones into smaller segments with different leakage behaviours and better prioritisation of zones/segments. To achieve this, Data61 uses two sets of data:

- Pipe network and reported leakage/break records, as described in water pipe failure prediction project.
- ALD records: more than 3,000 ALD inspection records from 2001 to 2015, including inspection date, zone, and different types of leaks found.

Methodology

The technique uses data mining and machine learning methods, including Generalized Linear Model (Madsen & Thyregod, 2011) and Model-Based Recursive Partitioning (Zeileis et al, 2008), to first segment large pressure zones into segments, and then prioritize zones/segments to carry out active leak detection.

For zone segmentation, the method takes pipe location and leakage/break history into consideration, aiming to find segments of pipes that expose different leakage behaviour.

For zone/segment prioritisation, the model uses a range of pipe properties, including age, material, failure history and the previous ALD records. By looking at how these factors impact the outcomes of ALD inspections, the volume of leakage for each zone/segment can be predicted, and the inspection candidates can be selected based on the prediction.

Current Progress and Outcomes

For zone segmentation, the algorithm has been validated on several large pressure zones. The merits of the algorithm include:

- Finer-grain inspection targeting: by segmenting large zones into segments of homogeneous leakage behaviour, different inspection strategies can be applied to each segment (e.g., lower inspection frequency for less leaky segments). One example of the segmentation is shown in Figure 10.
- Compatible to current practice: Sydney Water engages contractors for ALD based on catchment management authorities (CMA). The algorithm produces zone segments that aligns with CMA (or predefined areas), so that the change to the current ALD practice will be minimized.

For zone/segment prioritization, the outcomes are validated by measuring the errors between the predicted leakage volume and the actual leakage volume found on ALD inspections. The results of the validation include:

- More accurate inspection targeting: with better leakage volume prediction, the ALD program can target high-risk zones/segments more effectively.
- Initial data validation demonstrated that the Data61 conceptual model yields more accurate leakage volume prediction. The average prediction error is 50% less than the industry practice (Lambert, 2009).

Conclusion and Future Work

This project has been successful in developing new data-driven techniques to improve the effectiveness of the ALD program. Further validation and fine-tuning of the model will be done in the next step.

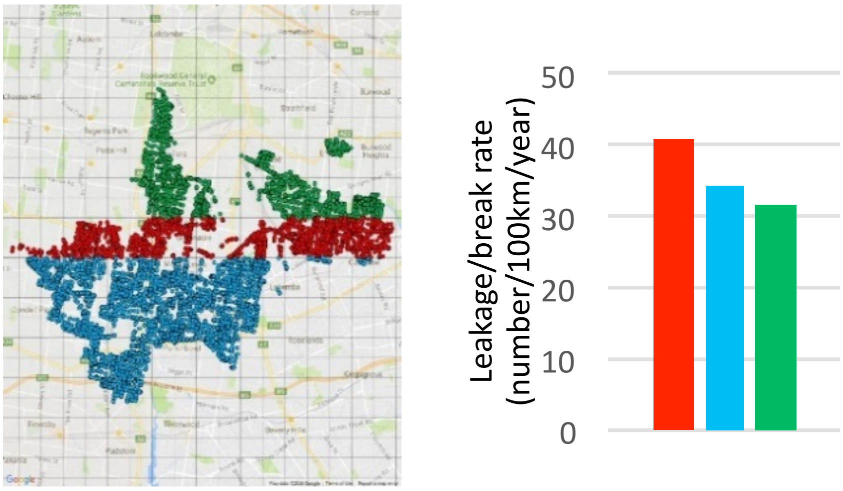

Figure 10. Segmentation of P_BANKSTOWN_REM, the grid-like pattern are CMA sheets.

OVERALL CONCLUSION

The collaborative machine learning based data analytics research will be further validated within an operational context to confirm and quantify the approach and benefits. This will be done over a 12-18 month period for all projects.

After validation, it is planned to incorporate the outcomes into normal business-as-usual practice. Work to date has demonstrated that these data analytical approaches can be used to optimise operational efficiencies and reduce costs. This will contribute to better services for customers and higher levels of confidence for stakeholders as well as regulators.

ACKNOWLEDGMENT

Sydney Water and Data61 would like to thank all staff who have contributed to these collaborative projects and provided such valuable support to ensuring outcomes are implemented.

REFERENCES

Zhidong Li, Bang Zhang, Yang Wang, Fang Chen, Ronnie Taib, Vicky Whiffin, Yi Wang. 2014. Water pipe condition assessment: a hierarchical beta process approach for sparse incident data. Machine Learning 95(1): 11-26

Vicky Whiffin, Craig Crawley, Yang Wang, Zhidong Li, and Fang Chen 2013. Evaluation of machine learning for predicting critical main failure. IWA Leading-Edge Strategic Asset Management (LESAM),*

Peng Lin, Bang Zhang, Ting Guo, Yang Wang, Fang Chen 2016. Interaction Point Processes via Infinite Branching Model. The 13th AAAI Conference on Artificial Intelligence (AAAI).

Li, B., Fan, X., Wang, Y., Wang, Y., Chen, F., Spaninks, F. 2016. Data-Driven Customer Segmentation for Water Demand Analysis. OZWater16.

Wells, T. & Melchers, R. 2016. Concrete Sewer Pipe Corrosion – Findings from an Australian Field Study. Ozwater 2016. Melbourne, Australia.

Rasmussen, C.E. 2004. Gaussian Processes in Machine Learning. Advanced Lectures on Machine Learning. Lecture Notes in Computer Science. 3176.

Diggle, P. J.; Fiksel, T.; Grabarnik, P.; Ogata, Y.; Stoyan, D.; and Tanemura, M. 1994. On parameter estimation for pair-wise interaction point processes. International Statistical Review 99–117.

Hawkes, A. G. 1971. Spectra of some self-exciting and mutually exciting point processes. Biometrika 58(1):83–90.

D. Vere-Jones 1988. An introduction to the theory of point processes. Springer Ser. Statist., Springer, New York

Madsen H; Thyregod P 2011. Introduction to General and Generalized Linear Models. Chapman & Hall/CRC.

Zeileis A, Hothorn T, Hornik K 2008. Model-Based Recursive Partitioning. Journal of Computational and Graphical Statistics, 17(2), 492–514.

Lambert AO 2009. Ten Years Experience in using the UARL Formula to calculate Infrastructure Leakage Index.