Institution: University of Exeter		
Unit of Assessment: UoA 11 Computer Science and Informatics		
Title of case study: Predictive Maintenance and Sewer Blockage Prevention through Machine Learning		
Period when the underpinning research was undertaken: 2014 – 2017		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s):	**Role(s) (e.g. job title):**	**Period(s) employed by submitting HEI:**
Edward Keedwell	Professor	2006 - present
Period when the claimed impact occurred: 2016 – 2020		
Is this case study continued from a case study submitted in 2014? N		

1. Summary of the impact

Flooding and pollution as a result of sewer blockages cause disruption and distress for the public, can be environmentally damaging and are costly to rectify. Through collaborative research between the University of Exeter and Dwr Cymru Welsh Water (DCWW), a machine learning system using data collected on prior sewer blockages has been developed to predict the location of likely blockage formation. This is enabling the blockages to be rectified before flooding occurs. The system has been embedded into the day-to-day operations at DCWW where it is now used to proactively inform sewer maintenance.

Benefits for DCWW, their 3 million customers and the environment are as follows:

- A 31% reduction in blockages since 2015
- Reduced disruption for customers (e.g. 56% reduction in repeat blockages)
- Cost savings of almost £6m
- Environmental benefits through a 14% reduction in pollution incidents across Wales and the English border counties.

2. Underpinning research

The development of the machine learning tool was underpinned by the optimisation and machine learning research of Keedwell and a longstanding collaboration with researchers (Kapelan and Djordjevic) at the Centre for Water Systems at the University of Exeter (UoE). This work includes the application of machine learning to problems in the water sector, for example **[3.1]** where collaborative work used artificial neural networks (ANNs) for urban flood prediction. The core underpinning research was developed as part of a Knowledge Transfer Partnership (KTP) between UoE and Dwr Cymru Welsh Water (DCWW), between Jan 2014 and June 2016 (9411 – rated 'Very Good' **[5.1]**). This research resulted in the award of an MPhil degree to the KTP Associate **[3.2]** (James Bailey, supervisors: **Keedwell**, Kapelan & Djordjevic) and three peer-reviewed outputs; one publication at the International Conference for Computing and Control for the Water Industry 2015, which attracted a "2nd best student paper" award **[3.3]** and two further publications at the 12th International Conference on Hydroinformatics 2016 **[3.4, 3.5]** published in *Procedia Engineering*. A follow-on 4-year EngD studentship (Draude, supervisors **Keedwell** & Kapelan) began in Sept 2016, focussing on the automated scheduling of maintenance activities, was secured as a result of this work.

The initial stages of the KTP work centred on data discovery and the exploration, cleaning and feature selection of DCWW-held asset data to determine the likely subsets of parameters involved in sewer blockages. This process also assessed the likely sewer blockage mechanisms described in the literature and the extent to which these were

reflected in the data. Siltation, ingress by tree roots, sewer collapse and the build-up of fats, oils and grease (FOG), for example, are all known contributing factors to the creation of a blockage and the different blockage modes within the data were assessed. Finalised datasets were then created for 700,000 sewers and 130,000 blockages over a period of 8 years, consisting of parameters relating to the age, construction, condition and previous reported incidents of each asset to predict the likelihood of blockage for a given sewer.

Through investigation on the above real-world historical datasets, the best performing machine learning techniques were identified through an evaluation of their ability to correctly predict the propensity of sewers to block. A robust cross-validation and testing process was adopted and through this the models were found to be a significant improvement over the existing system that utilised only historical data to predict blockages **[3.2, 3.3]**. This work identified decision trees as the best-performing methods in terms of accuracy, computational complexity and explainability. The computational complexity of models was found to be important when dealing with such large volumes of data and high levels of explainability allowed the company to scrutinise the developed models and gain insight into blockage formation mechanisms that underpin a number of the impacts described later. Additional research discovered that it was possible to increase the skill of the model through the use of an ensemble machine learning approach incorporating multiple trees (random forests), although this naturally resulted in a reduction in model explainability **[3.4]**.

Further work was carried out to embed the sewer blockage propensity system into DCWW's maintenance planning process in 2014-15 and into the Geographical Information System (GIS) software in 2015-16. The latter provides a colour-coded layer of machine-learned blockage propensity overlayed on maps within the GIS system used by the operational team to prioritise the selection of sewers to undergo proactive and planned maintenance. The visualisation and communication of complex machine-learning-based model predictions in the GIS system is an important part of ensuring that operational decisions can be supported by the evidence-based machine learning approach. The ongoing STREAM EngD project is building on this work by developing decision support tools to enable optimal scheduling of proactive and reactive maintenance through the use of an AI-based evolutionary optimisation tool. This updated system, which has undergone successful field trials in November 2020, will maximise the proactive maintenance coverage of high-risk sewers within the resources available to the company.

3. References to the research

3.1 Duncan AP, Chen AS, **Keedwell** EC, Djordjević S, Savić DA. (2011) Urban flood prediction in real-time from weather radar and rainfall data using artificial neural networks, Proceedings of WRaH 2011, Weather Radar and Hydrology Conference / IAHS / Exeter, pp 58, vol. 351 *Available on request.*

3.2 Bailey J, **Keedwell** EC, Kapelan Z, Djordjevic S, Burton C, Harris E. (2015) Predictive risk modelling of real-world wastewater network incidents, 13th Computer Control for Water Industry Conference, CCWI 2015, Procedia Engineering, volume 119, pages 1288-1298 doi.org/10.1016/j.proeng.2015.08.949

3.3 Bailey J, Harris E, **Keedwell** E, Djordjevic S, Kapelan Z. (2016) Developing Decision Tree Models to Create a Predictive Blockage Likelihood Model for Real-World Wastewater Networks, 12th International Conference on Hydroinformatics, Incheon, South Korea, 21st - 26th Aug 2016, Procedia Engineering, volume 154, pages 1209-1216, DOI:10.1016/j.proeng.2016.07.433

3.4 Bailey J, Harris E, **Keedwell** E, Djordjevic S, Kapelan Z. (2016) The Use of Telemetry Data for the Identification of Issues at Combined Sewer Overflows, 12th International Conference on Hydroinformatics, Incheon, South Korea, 21st - 26th Aug 2016, Procedia Engineering, volume 154, pages 1201-1208 doi.org/10.1016/j.proeng.2016.07.524

4. Details of the impact (indicative maximum 750 words)

The research and development of a machine learning (ML) approach to sewer blockage prediction described above has allowed Dwr Cymru Welsh Water (DCWW) to make a step-change in the operational management of its country-wide 36,000km sewer network. DCWW, the 4th largest company in Wales, operates throughout Wales and parts of England along the Welsh border and is a not-for-profit business, run for the benefit of its 3 million customers and the 1.4 million homes and businesses it serves **[5.7].**

The ML tool is in use by DCWW in its central operations hub in Cardiff where maintenance crews can be directed to areas of greatest blockage risk, allowing DCWW to systematically prevent blockages from forming and facilitating the operational transition from a reactive to a preventative sewer maintenance strategy. This system formed a key component of a company-wide strategy to develop analytical prediction methods as a drive towards proactivity with the attendant reductions in cost and improvement in the customer relationship this brings. As a result, the tool has delivered benefits for DCWW and their customers, as well as reducing environmental pollution across Wales and the English border counties as explained below **[5.2].**

Benefits for the water company

DCWW have confirmed that the collaboration work with Prof Keedwell's team, through the developed machine learning system, has helped drive a 31% reduction in sewer blockages since 2015 for the company, which saw the number of reported blockages fall from 28,740 in 2013-2014 to 19,670 in 2019-2020. Furthermore, the introduction of the tool has enabled DCWW to gain a better understanding of the key factors likely to lead to sewer blockage, in particular the identification of asset condition as a root cause. By using this information to appropriately focus investment it has enabled DCWW to implement a programme of ongoing preventative maintenance which has helped to reduce the number of collapsed sewers by 26%. This approach is therefore minimising the risk of more severe reoccurrences as the assets deteriorate, which is impacting positively on on-going performance.

Benefits for customers

The improved performance of DCWW enabled by the ML tool has led to the following benefits for DCWW's 3 million customers.

Reduced disruption: sewer blockages negatively impact customers in a range of ways, from minor inconvenience associated with unusable facilities, to more significant disruption associated with internal (e.g. basement or ground floor) flooding which can require homeowners to relocate to alternative accommodation whilst cleaning and reparation is carried out. Prioritisation of high-risk sewers using the machine learning approach has contributed to a 16% reduction in incidents of internal flooding caused by a blockage or asset failure from 2015 to 2020, thus significantly reducing the number of customers having to suffer these negative events. This reduction has ensured that significantly fewer properties were flooded per 10,000 DCWW sewer connections (1.02) than the average across England and Wales from 2017-2020 (1.46). **[5.3]**

Cost savings: as DCWW is a not-for-profit business, any cost savings or financial gains that the company achieves benefits customers either directly, through reduced prices, or indirectly, through investment in improved services and infrastructure. Overall, the initiative has reduced operational costs by £2m for reactive attendance and investigations **[5.2]** and contributed to cumulative rewards received by the company from the regulator of £3.95 million from 2015-2020 for pollution incident reduction **[5.4]**. In addition, although not quantified, there have been savings associated with the reduction in compensation paid to customers as a result of sewer blockage events.

Better customer service**:** the reduction in the incidence of sewage blockage or failure due to the use of the ML tool has contributed to DCWW delivering a better level of service to its customers, with customers experiencing a 54% reduction in repeat events (950 in 2015 – 435 in 2020). Customers who experience repeat events are amongst the highest priority for the company as the effects for the customer are magnified and the compensation required is significantly greater than for one-off events.

Benefits for the environment

Sewer blockages can also cause pollution incidents where a watercourse is contaminated with the sewer contents, requiring emergency rectification and the involvement of the Environment Agencies (Natural Resources Wales or the Environment Agency in England). These incidents are among the most serious that a water company can face, with the potential for significant adverse environmental impact through damage to wildlife and watercourses, in addition to severe financial penalties from the regulator OFWAT. The introduction of the system by DCWW has contributed to a 14% reduction in pollution incidents caused by a blockage between 2014 and 2018. This has led to positive benefits to the environment in Wales and the areas of England covered by DCWW as shown in the NRW 2018 performance report (company rated 'good' in 2018) **[5.5]** and in subsequent years where DCWW has risen from 2 to 3 stars (out of 4) **[5.6]**.

In summary, the collaborative development of this system by researchers in UoA11 at the University of Exeter has transformed the wastewater performance of this large utility company with significant benefits for the company, its customers, and the natural environment.

5. Sources to corroborate the impact

5.1 KTP final assessment (rated very good)

5.2 Letter of support from DCWW

5.3 DiscoverWater Sewer Flooding – Household Ratings https://www.discoverwater.co.uk/sewer-flooding

5.4 DCWW Annual Performance Report 2019-2020 (page 17)

5.5 National Resources Wales Annual performance report for Dŵr Cymru Welsh Water 2018 (page 7)

5.6 DiscoverWater Environmental Performance Ratings https://www.discoverwater.co.uk/environmental-performance

5.7 Dŵr Cymru Welsh Water website listing number of customers and households and businesses served https://corporate.dwrcymru.com/en/about-us