Available online at www.sciencedirect.com

ScienceDirect

Procedia Engineering 119 (2015) 299 – 308

Procedia Engineering

www.elsevier.com/locate/procedia

13th International Conference on Computing and Control for the Water Industry, CCWI2015

Interpreting and estimating the risk of iron failures.

Kate Ellis[a]*, Stephen R. Mounce[a], Jonathan Edwards[a], Vanessa Speight[a], Natalie Jakomis[b] and Joby Boxall[a]

[a]*Pennine Water Group, Department of Civil and Structural Engineering, University of Sheffield, S1 3JD, UK*
[b]*Dwr Cymru Welsh Water, Pentwyn Road, Nelson, Treharris, Mid Glamorgan, CF46 6LY, UK*

Abstract

Metals and particulates accumulate in the distribution system and are mobilised by hydraulic events which can result in discolouration and exceedance of regulatory standards. Traditional decision tools for targeting preventive work are single parameter, based for example on proportion of unlined iron pipe or the number of customer contacts per district metering area (DMA). We show that this approach is too simplistic and propose a multivariate Decision Tree process, using the Random Under-Sampling ensemble method. The outputs gave a classification of High or Low risk per DMA. Initial findings demonstrate an 80 % success rate in identifying high risk DMAs across the supply area for a UK water company.

Peer-review under responsibility of the Scientific Committee of CCWI 2015

Keywords: Decision Trees; District metering areas; Geographical Information Systems; Iron; Self-organising maps; Water quality.

1. Introduction

Concentrations of iron, manganese and turbidity are regulated in the UK. The regulatory limits at customers' taps are: iron, 0.2 mg l^{-1}; manganese, 0.05 mg l^{-1}; and turbidity, 4 nephelometric turbidity units (NTU) (Drinking Water Inspectorate, 2014). The primary sources of metals and turbidity in water distribution systems are carryover from water treatment (Vreeburg and Boxall, 2007; Vreeburg et al., 2008) and corrosion by-products within pipes (Peng and Korshin, 2011; Prasad and Danso-Amoako, 2014). Once in the pipe network, metals and particulates

* Corresponding author. Tel.: +44 114 222 5761.
E-mail address: kate.ellis@stream-idc.net

Peer-review under responsibility of the Scientific Committee of CCWI 2015
doi:10.1016/j.proeng.2015.08.889

accumulate and are then mobilised by hydraulic events, which can result in discolouration and exceedance of regulatory standards (Lytle et al., 2004; Burlingame et al., 2006; Drinking Water Inspectorate, 2014).

Distribution systems are subject to a variety of site-specific hydraulic, chemical and microbiological influences that make it difficult to isolate individual reactions and fully understand the mechanisms at work (Husband and Boxall, 2011; Aisopou et al., 2012). This complexity coupled with the uncertainties surrounding the UK's ageing, buried pipe infrastructure mean that a deterministic modelling approach for estimating iron failures is rarely possible. Hence it is of interest to explore whether data-driven techniques can be applied, such as analytics, modelling and visualisation, to generate new insight and value from large amounts of complex multidimensional data.

In this paper we present our work with Dŵr Cymru Welsh Water Ltd. (DCWW) exploring the likely causes of iron failures across their region and developing a predictive model to enable proactive preventive measures. We sought to determine the value of single parameter and multivariate data analysis for the prediction of risk of iron failure across the whole DCWW water distribution system. The results of this work were then incorporated into a predictive model for iron failures at the district metering area (DMA)-level.

Nomenclature

cc	customer contacts
cc_clusters	customer contacts in water quality zone-level clusters
DCWW	Dŵr Cymru Welsh Water Ltd.
DMA	district metering area
Fe_lined	lined iron pipe
Fe_unlined	unlined iron pipe
GIS	geographical information system
Iron_avg	Median iron concentration
Mn_avg	Median manganese concentration
NTU	nephelometric turbidity units
SOM	self-organising map
Turb_avg	Median turbidity at the customers' tap
WTW_turb_avg	Median turbidity at the WTW
WQZ	water quality zone
WTW	water treatment works

2. Methods

2.1. Data collection and manipulation

The following raw data were collected from DCWW: water supply system connectivity – which WTWs serve which water quality zones (WQZs), which service reservoirs serve which DMAs and which DMAs customer tap samples originate from; population of each DMA; customer contact data for discolouration; pipe materials in each DMA – length of unlined iron, lined iron and other pipe materials and total length (km); water chemistry data for customers' taps, service reservoirs and WTWs – iron (mg l^{-1}), manganese (mg l^{-1}), turbidity (NTU), free and total chlorine (mg l^{-1}), water temperature (°C) and pH (pH units); and the dates and locations of bursts, main repairs and routine and reactive pipeline flushing. A complex data processing and formatting process involving multiple MS Excel (Microsoft Corporation, New Mexico) and MATLAB® 2014b (The Mathworks Inc., Massachusetts) operations led to the development of year-by-year data matrices for the period January 2008 to September 2014. Each year's matrix consisted of 56 parameters populated for DCWW's 1,312 DMAs. Overall, 25 % of data were missing. Missing data occurred because a) the regulations do not require all DMAs to be sampled for all parameters every year and b) some datasets were only brought on-line later in the study period such as burst events (2010 onwards), customer contacts (from 2011) and pipeline flushing (2012 onwards).

2.2. Single parameter data analysis

2.2.1. Geographical Information Systems (GIS)

ArcMap 10.1 (ESRI, California) was used to plot the number of iron failures against a backdrop of the proportion of unlined iron pipe per DMA. The percentage of unlined iron pipe was calculated within ArcMap. The outputs from the predictive model were also plotted to enable DCWW to focus their preventive strategies on the DMAs with greatest risk of failure.

2.2.2. Customer contacts in WQZ-level clusters per DMA

Individual customer contacts are rarely indicative of a wide-scale water quality problem, but multiple contacts within a short time-frame could be (Husband et al., 2010). The Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm was used in MATLAB to identify temporal clusters of contacts (Ester et al., 1996; Mounce et al., 2012). A cluster was defined as having three or more contacts at the DMA-level ('minimum number of points to form a cluster') within a 48 h rolling window ('neighbourhood distance'). The results were exported to Excel and ordered by WQZ and then limited to WQZs with at least two DMAs containing clusters. This field was then added to the matrices for each year.

2.2.3. R^2

A scatter plot of iron samples exceeding DCWW's internal iron standard (0.1 mg l^{-1}) against the number of customer contacts per WQZ was generated in Excel. A linear trend-line was added and the R^2 value extracted.

2.3. Multivariate data analysis

2.3.1. Self-Organising Maps (SOMs)

SOMs were used to identify trends and potential relationships among the different water quality and system management parameters within the matrix. In this paper, a SOM was generated for the following parameters: number of iron fails per DMA, customer contacts and customer contacts in WQZ-level clusters per DMA, and proportions of unlined iron, lined iron and pipe of other materials per DMA. The SOM was generated using MATLAB® 2012a and the SOM Toolbox version 2.0 (Laboratory of Computer and Information Science, Finland).

SOMs are a type of artificial neural network that is trained using unsupervised learning; this means that when the input data are presented to the network it forms its own clustering of the training data thus allowing the potential to derive information from data without any previous knowledge. SOMs were developed by Teuvo Kohonen (Kanga and Kohonen, 1996; Kohonen, 1998) and provide an intuitive representation of complex datasets which can be used to inform distribution system management practices. SOMs are commonly used for visual data mining/exploration, pattern classification and have potential for use in process control. SOMs have been used for the analysis and modelling of drinking water microbiological and physico-chemical quality at laboratory-scale (Mounce et al., 2012) and full-scale (Ellis, 2013; Ellis et al., 2014; Mounce et al. 2014; Ellis et al., 2015).

2.3.2. Random Under-Sampling (RUSBoost) Decision Tree analysis

A decision tree approach to identifying DMAs at risk of iron non-compliance was selected because it can provide utility operators with an output explaining how a particular decision was made, unlike a 'black box' model. Due to the complexity of the problem, and since there are comparatively few failures upon which to base a predictive model (approximately 4 % failure rate across the period of study), a single decision tree is too simplistic. For this reason an ensemble method was used. Ensemble methods use multiple learning algorithms to improve performance (Rokach, 2010). Results from multiple 'weak' decision trees were melded into one high-quality ensemble predictor using the RUSBoost algorithm (in MATLAB), which is designed for situations where one class (results for iron <0.2 mg l^{-1}) has many more observations than another (results for iron >0.2 mg l^{-1}) (Seiffert et al., 2010). A protocol for equalising classes and randomly removing data points for particular model subsets was used; this resulted in weighted average predictions from multiple trees. The target for the ensemble was the classification of risk for each DMA as follows: High (H) which corresponds to one or more iron failures in the DMA or Low (L) which corresponds to zero iron failures in the DMA (for the yearly period under consideration). Only an outline of RUSBoost and the models developed has been provided here. An in-depth description of the ensemble methodology for this application (including a full description of the models) and its validation can be found in Mounce et al., (2015). In this study, the Decision Tree analyses were trained using data from 2011 to 2013 and tested on data from 2014 to predict iron compliance in 2015 ('Futurecast 2015'). The matrix parameters were:

- Median iron concentration from all samples per DMA per year, 'Iron_avg'
- Median manganese concentration from all samples per DMA per year, 'Mn_avg'
- Median turbidity from all samples per DMA per year, 'Turb_avg'
- Percentage of unlined iron pipe per DMA, 'Fe_unlined'
- Percentage of lined iron pipe per DMA, 'Fe_lined'
- Median turbidity from supplying WTWs per year and, where applicable, calculated as a weighted average for DMAs with multiple supplying WTWs, 'WTW_turb_avg'
- Number of customer contacts per DMA per year, 'cc'
- Number of customer contacts in WQZ-level clusters per DMA per year, 'cc_clusters'

The results were obtained as predictions of likelihood of risk of iron failure per DMA: High or Low, and as a ranking of risk across all DMAs. The model also provides the weighting of High vs Low based on the relative weighting across the set of Decision Trees.

3. Results

3.1. Overview

DCWW's supply area is divided into 1,312 DMAs and distributed among 86 WQZs. During the study period, 4.08 % of iron samples exceeded the regulatory limit of 0.2 mg l^{-1}; 1.85 % of manganese samples exceeded the limit of 0.05 mg l^{-1}; and 1.35 % of turbidity samples failed the standard of 4 NTU at customers' taps. From 2011 until the end of the study period, 22,487 customer contacts were received regarding discoloured drinking water. DCWW is investigating the causes of these failures and customer contacts and is investing in remedial work to improve compliance and customer satisfaction. Traditional analysis of variables commonly associated with iron compliance (e.g. extent of unlined iron pipe or customer contacts) revealed that no single parameter could reliably predict iron failures, thus necessitating a multivariate approach.

3.2. Single parameter data analysis results

3.2.1. Iron compliance and pipe material

The proportion of unlined iron pipe varies by DMA at DCWW, as illustrated by the excerpt map in Fig. 1. The number of iron failures did not correlate strongly with the proportions of unlined iron pipe. In the excerpt map it can be seen that whilst the higher numbers of iron failures per DMA (4 – 6 for the excerpt view) were found in DMAs with at least 61 % unlined iron pipe, there were instances where DMAs with less than 20 % unlined iron pipe had failures too (1 – 3). Across the whole of the DCWW region, it was observed that whilst some failures occurred in DMAs with high proportions of unlined iron pipe, there were many DMAs with greater than 61 % unlined iron pipe which had compliant iron results. The proportion of unlined iron pipe was not a clear indicator of iron failure risk.

3.2.2. Iron compliance and customer contacts

Fig. 2 shows the relationship between the number of iron samples exceeding DCWW's internal (non-regulatory) iron standard (0.1 mg l^{-1}) and the number of customer contacts for discolouration received per WQZ. The number of customer contacts is higher than the number of iron samples exceeding the internal standard for the majority of WQZs. A plot of iron samples exceeding the internal standard against customer contacts per WQZ produced a linear trend-line with an R^2 value of 0.21 (not shown). Thus, whilst there is some relationship between the two parameters, it only accounts for 21 % of failures. The number of customer contacts cannot be considered a strong predictor of iron non-compliance for DCWW.

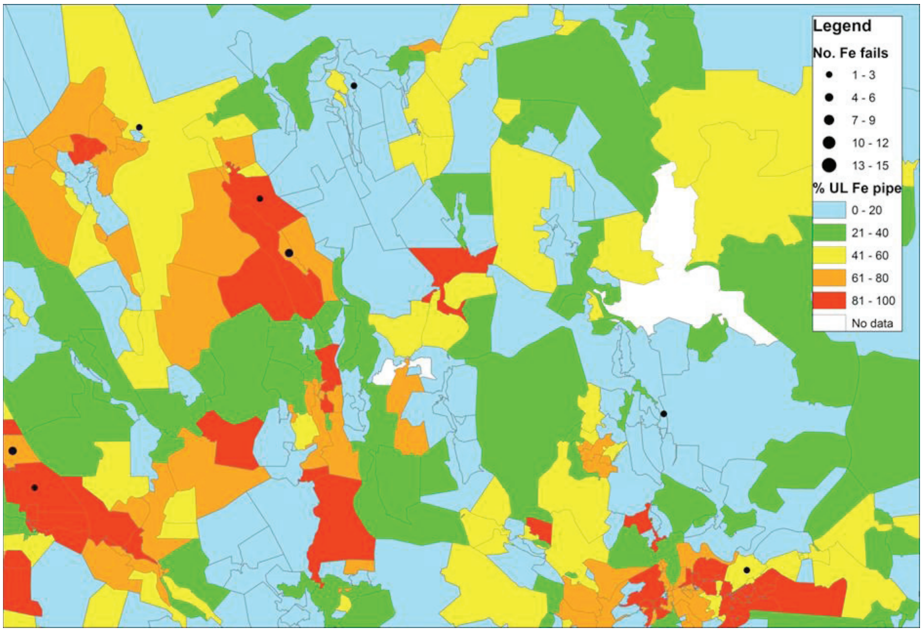

Fig. 2: Number of iron fails overlain on proportions of unlined iron pipe per DMA for 2014, excerpt of DCWW region.

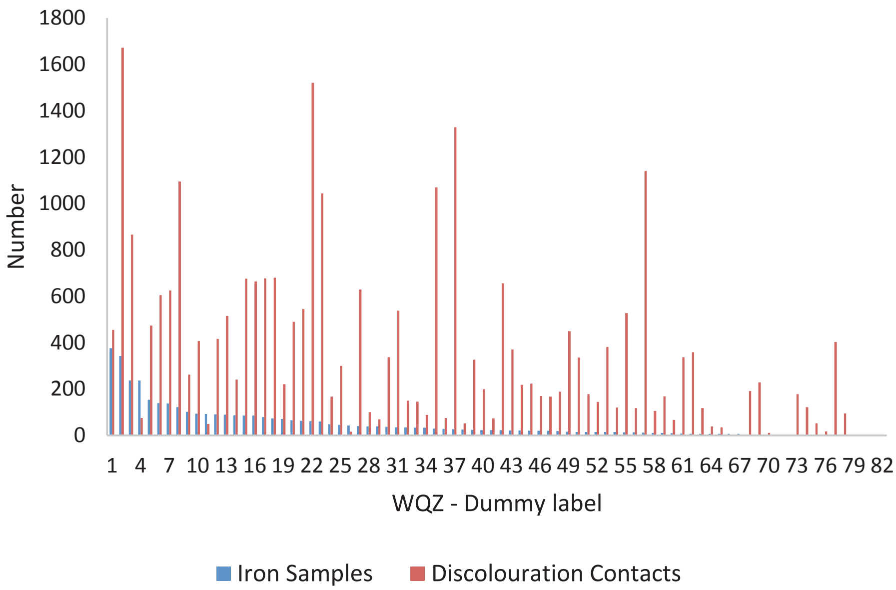

Fig. 1: Number of iron samples with results >0.1 mg l^{-1} and number of customer contacts for discoluration per WQZ, 2008 – 2014.

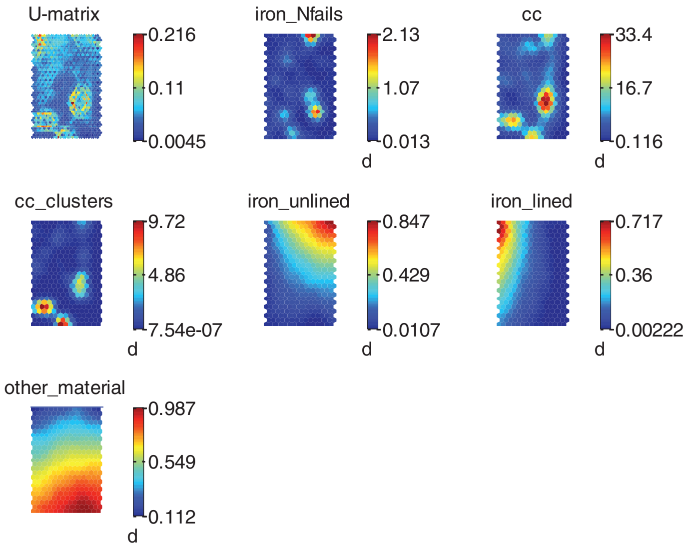

Fig. 3: SOM of compliance, customer contacts and proportions of different pipe materials, 2008 – 2014.

3.3. Multivariate data analysis results

3.3.1. SOM

The SOM for numbers of iron failures, customer contacts and customer contacts in WQZ-level cluster for all years, plus the different pipe materials is shown in Fig. 3. This map contains colour-coded hexagons that summarise all of the component planes that represent individual variables. There are two separate parts of the SOM display. These are the summary U-matrix and the component planes for each individual variable. The U-matrix shows the distances between the reference vectors of adjacent cells, therefore ridges in the U-matrix delineate clusters in the trained SOM. In the component planes for individual variables, the colouring corresponds to actual numerical values for the input variables that are referenced in the scale bars adjacent to each plot. Blue shades show low values and red corresponds to high values. Clusters of a given colour that occur at the same location in different component planes provide evidence of a relationship between those variables.

The SOM shows that a high number of iron failures (shown in red in the SOM) correlated with medium-high (green-red) numbers of customer contacts; low-medium (blue-green) numbers of customer contacts in WQZ-level clusters; low (blue) and high (red) proportions of unlined iron pipe; low (blue) proportions of lined iron pipe; low (blue) and high (red) proportions of pipe of other materials.

In addition, it can be seen that high numbers of customer contacts and contacts in WQZ-level clusters were observed regardless of the presence of iron failures, and were more common with low-medium (blue-green) proportions of unlined or lined iron pipe and medium-high (yellow-red) proportions of pipe of other materials. The SOM corroborates the observations made of Fig. 1 and Fig. 2; neither a high proportion of unlined iron pipe nor a high number of customer contacts, or even clusters of contacts, is a conclusive predictor of iron failures.

3.3.2. RUSBoost Decision Tree

A Decision Tree can be read in a similar manner to a flow chart by reference to variable values at each decision node. An example Decision Tree is shown in Fig. 4. In this example, DMAs were first separated based on their median iron concentrations (this parameter was identified as dominant during further SOM analyses too, but not shown in this paper). The concentration cut-off was 0.02 mg l^{-1}. DMAs with median iron concentrations of less than the cut-off were then classified by percentage of unlined iron pipe and those greater than the cut-off were divided by median turbidity. These divisions continued until the DMAs were all classed as either High or Low risk of iron failure (class label at the leaf node). This Decision Tree was one of, potentially, a thousand used in the RUSBoost ensemble to produce the model outputs. Each Decision Tree has a weighting in producing the final result, and in this example, the three strongest contributed ~7 % to the final result. Therefore, exploring these larger weighted Decision Trees can enable an understanding of what is most significant in determining a class label for a particular DMA.

The RUSBoost Decision Tree process determined that 22.9 % of DMAs were at high risk of iron failure, and the remaining 77.1 % were at low risk for Futurecast 2015 (Table 1). At the end of 2015, it will be possible to fully compare the predictions with actual iron compliance. This could be achieved by compiling a table similar to Table 2, which shows the performance of the model in predicting 2014's DMA risk ('Futurecast 2014'). The model correctly identified 74.8 % of DMAs as having either High or Low risk of iron failure in 2014 (overall accuracy). Of DMAs that had experienced iron failures, 60.5 % were correctly predicted by the RUSBoost Decision Tree process. These predictions for 2014 were trained on data from 2011 to 2012 and tested on 2013's data; it is expected that with a larger training dataset, the predictions for Futurecast 2015 could have a higher proportion of correct classification.

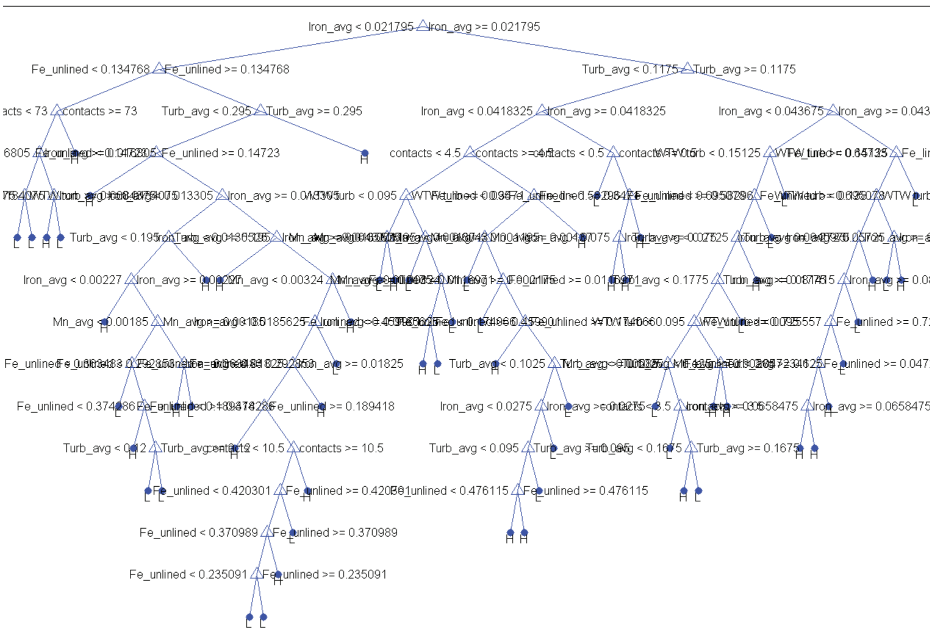

Fig. 4: Example decision tree (one of hundreds) for Futurecast 2015.

Table 1: Risk predictions for Futurecast 2015.

Classification	Number of DMAs	Percentage of DMAs
Low	1012	77.1
High	300	22.9
Un-sampled	-	-

Table 2: Confusion matrix comparing predicted risk of iron failure against actual performance for 2014.

	Percentage of DMAs predicted as:		
Actual	Low risk	High risk	Un-sampled
Low risk	**75.9 (True L)**	24.1 (False H)	0.0
High risk	39.5 (False L)	**60.5 (True H)**	0.0
Un-sampled	74.3	25.7	0.0

Overall accuracy: 74.8%

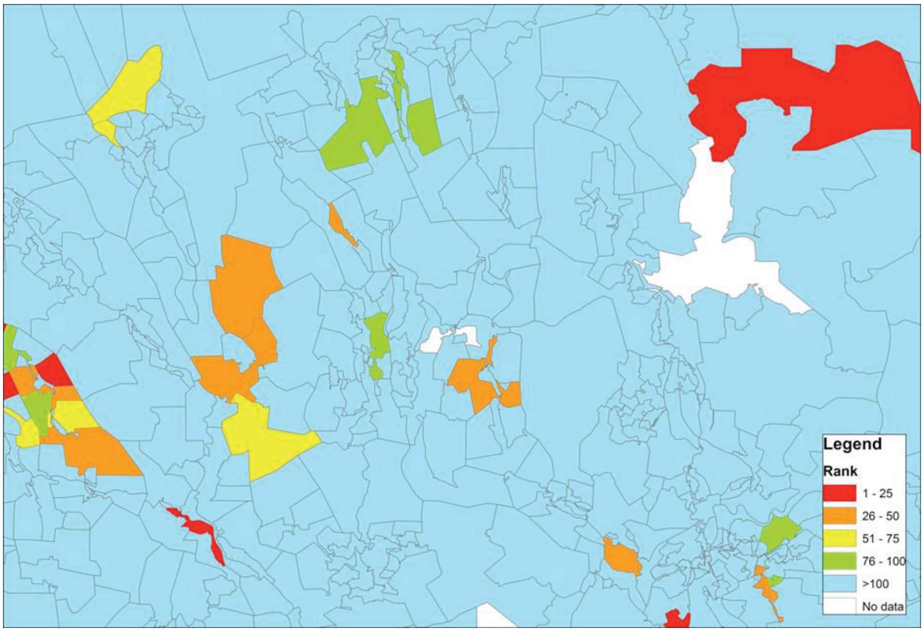

Fig. 5: Futurecast 2015 ranking results plotted on excerpt of DCWW region.

Classifying DMAs as either High or Low risk presents DCWW with a list of DMAs requiring attention; there are 300 High risk DMAs for 2015 (Table 1). It is unrealistic to expect DCWW to take action to prevent failures in all High risk DMAs. For this reason, the model output also ranks the DMAs so that the highest risk sites can be targeted first (using the generated weightings). Fig. 5 shows the Futurecast 2015 rank predictions for the excerpt featured in Fig. 1; the ranks have been colour-coded, with the 25 DMAs having the highest risk in red.

Early findings from 2015's compliance for the live network showed that of the five iron non-compliances that had been detected (as of May 2015), four were in DMAs predicted as High risk by the RUSBoost Decision Tree process. The fifth failing DMA was one where no past sampling data were available for training or testing the model (Natalie Jakomis, pers. comm., 12th May 2015).

4. Discussion

4.1. Analytical approach

The development of the data matrices provided a valuable overview of annual DMA performance for a variety of parameters, including iron, manganese and turbidity. These data were then analysed using traditional methods and SOMs, which demonstrated some relationship among pipe material, customer contacts and iron compliance. These outcomes concurred with other research (Vreeburg and Boxall, 2007; Husband and Boxall, 2011). One main drawback of the traditional approaches is that they make generalisations about the DMAs which could not realistically be applied across heterogeneous distribution systems (Francisque et al., 2009). A further complication is the paucity of iron non-compliances, statistically speaking, upon which to draw conclusions. This is a problem that is not restricted to iron compliance (Piriou et al., 1997; van Lieverloo et al., 2007; Ellis et al., 2013).

The RUSBoost Decision Tree ensemble methodology has great potential for analysing multiple water distribution system parameters and producing a DMA-specific prediction of failure risk. The model presented here was tested on an incomplete dataset (January to September 2014). In basing the predictions for Futurecast 2015 on three years of training data it could be more accurate than those for Futurecast 2014. It would nevertheless have been beneficial to be able to test the model on a full year's data and it is possible that accuracy may have been compromised by having a truncated dataset. In future years, larger training datasets will be available to improve the accuracy of the model and give DCWW greater confidence in utilising the model outputs in operational decisions. The results from this work are already enabling DCWW to better target specific areas of their network to prevent iron non-compliance, with a view to extending the impact with future iterations of the model.

4.2. Operational value of results

The traditional methods and perceived wisdom would suggest that DCWW replace all their unlined iron pipes with lined iron pipes or plastic pipes and to focus first of all on DMAs with high numbers of discolouration complaints. However, these data show that the number of customer complaints can a) be misleading with regard to actual water quality and b) misrepresent the extent of buried infrastructure replacement that is required annually. It has also been shown that such an approach, based on individual parameters like customer contacts, as well as being expensive may not actually improve compliance. The Decision Tree analysis has shown the ability to incorporate a wide range of data, assisted by trend identification using SOMs, and is able to take these data into account when predicting the risk of failure.

This project has demonstrated the value of both traditional and machine-learning techniques for understanding water quality. The widespread application of this type of Decision Tree analysis within the water industry could revolutionise the way that budgets are allocated for proactive pipeline management, making them more efficient and more cost-effective.

5. Conclusions

The compilation of data matrices from a variety of DCWW sources enabled an exploration of parameters associated with iron failures. A variety of traditional and machine-learning techniques were employed. The presence of unlined cast iron pipe in a DMA is not necessarily correlated with a higher number of iron failures. Customer contacts in general had a weaker than anticipated correlation with iron failures. In general, single parameter analyses are inefficient for exploring and explaining complex, interacting behaviours. The application of SOMs provided a means of visualising data clusters within different groupings of parameters and yielded several interesting results. The SOM corroborated the observations of the graphical data exploration and also showed that the presence of a customer contacts in WQZ-level clusters appeared to be associated with some iron failures. This suggests that trunk main operations/bursts or WTW iron sources are playing a role in those failures.

The data exploration work supported the development of a predictive model for iron failures through the identification of relevant input parameters. While several options for the modelling tool were investigated, the RUSBoost (ensemble) Decision Tree model was selected for its ability to address the skewed distribution of failure data while producing accurate predictions. Furthermore, Decision Trees show operators and managers how the results were obtained, making them an accessible tool for decision-making. The Futurecast 2015 model was developed to predict 2015's performance based on historical data. Approximately 75 % of DMAs for the Futurecast 2014 run were correctly classified and it is expected that as more yearly matrices are added to the training set, higher accuracy will be achieved starting with the Futurecast 2015 predictions.

The success of this methodology has been demonstrated by 80 % correct prediction of failures (with the only missed fail being in a previously un-sampled DMA) from January to May 2015.

Acknowledgements

The authors would like to thank Dŵr Cymru Welsh Water for their involvement in this project and for their permission to publish the results. This work was supported by the Engineering and Physical Sciences Research Council's Pennine Water Group Platform Grant (EP/I029346/1).

References

Aisopou, A., Stoianov, I., & Graham, N. (2012). In-pipe water quality monitoring in water supply systems under steady and unsteady state flow conditions: A quantitative assessment. *Water Research, 46*, 235-246.

Burlingame, G., Lytle, D., & Snoeyink, V. (2006). Why red water? Understanding iron release in distribution systems. *Opflow, 32*(12), 12-16.

Drinking Water Inspectorate. (2014). *Drinking Water 2013: A Report by the Chief Inspector of the Drinking Water Inspectorate.* London: Drinking Water Inspectorate.

Ellis, K. (2013). *Improving root cause analysis of bacteriological water quality failures.* EngD thesis: University of Sheffield.

Ellis, K., Mounce, S., Ryan, B., Biggs, C., & Templeton, M. (2015). Improving root cause analysis of bacteriological water quality failures at water treatment works. *13th International Conference of Computing and Control for the Water Industry*.

Ellis, K., Mounce, S., Ryan, B., Templeton, M., & Biggs, C. (2014). Use of on-line water quality monitoring data to predict bacteriological failures. *Procedia Engineering, 70*, 612-621.

Ellis, K., Ryan, B., Templeton, M., & Biggs, C. (2013). Bacteriological Water Quality Compliance and Root Cause Analysis: An Industry Case Study. *Water Science and Technology: Water Supply, 13*(4), 1034-1045.

Ester, M., Kriegel, H., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In E. Simoudis, J. Han, & U. Fayyad, *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (pp. 226-231). California: AAAI Press.

Francisque, A., Rodriquez, M., Miranda-Moreno, L., Sadiq, R., & Proulx, F. (2009). Modeling of heterotrophic bacteria counts in a water distribution system. *Water Research, 43*, 1075-1087.

Husband, P., & Boxall, J. (2011). Asset deterioration and discolouration in water distribution systems. *Water Research, 45*, 113-124.

Husband, P., Whitehead, J., & Boxall, J. (2010). The role of trunk mains in discolouration. *Water Management, 163*(WM8), 397-406.

Kangas, J., & Kohonen, T. (1996). Developments and applications of the self-organizing map and related algorithms. *Mathematics and Computers in Simulation, 41*, 3-12.

Kohonen, T. (1998). The self-organizing map. *Neurocomputing, 21*, 1-6.

Lytle, D., Sorg, T., & Frietch, C. (2004). Accumulation of arsenic in drinking water distribution systems. *Environmental Science & Technology, 38*, 5365-5372.

Mounce, S., Douterelo, I., Sharpe, R., & Boxall, J. (2012). A bio-hydroinformatics application of self-organizing map neural networks for assessing microbial and physico-chemical water quality in distribution systems. *Proceedings of the 10th International Conference on Hydroinformatics, Hamburg, Germany*.

Mounce, S., Ellis, K., Edwards, J., Speight, V., & Boxall, J. (2015). Ensemble decision tree models for estimating and interpreting district management areas risk of iron failure. *Submitted to Water Resources Management*.

Mounce, S., Sharpe, R., Speight, V., Holden, B., & Boxall, J. (2014). Knowledge discovery from large disparate corporate databases using self-organising maps to help ensure supply of high quality potable water. *Proceedings of 11th International Conference on Hydroinformatics, New York, USA*.

Peng, C.-Y., & Korshin, G. (2011). Speciation of trace inorganic contaminants in corrosion scales and deposits formed in drinking water distribution systems. *Water Research, 45*, 5553-5563.

Piriou, P., Dukan, S., Levi, Y., & Jarridge, P. (1997). Prevention of bacterial growth in drinking water distribution systems. *Water Science and Technology, 35*(11-12), 283-287.

Prasad, T., & Danso-Amoako, E. (2014). Influence of chemical and biological parameters on iron and manganese accumulation in water distribution networks. *Procedia Engineering, 70*, 1353-1361.

Rokach, L. (2010). Ensemble-based classifiers. *Artificial Intelligence Review, 33*(1-2), 1-39.

Seiffert, C., Khoshgoftaar, T., van Hulse, J., & Napolitano, A. (2010). RUSBoost: A Hybrid Approach to Alleviating Class Imbalance. *IEEE Transaction on Systems, Man and Cybernetics-Part A: Systems and Human, 40*(1).

van Lieverloo, J., Mesman, G., Bakker, G., Baggelaar, P., Hamed, A., & Medema, G. (2007). Probability of detecting and quantifying faecal contaminations of drinking water by periodically sampling for E. coli: A simulation model study. *Water Research, 41*, 4299-4308.

Vreeburg, J., & Boxall, J. (2007). Discolouration in potable water distribution systems. *Water Research, 41*, 519-529.

Vreeburg, J., Schippers, D., Verberk, J., & van Dijk, J. (2008). Impact of particles on sediment accumulation in a drinking water distribution system. *Water Research, 42*, 4233-4242.