Data-Driven Approaches for Near Real-Time Forecasting of Discolouration Events in Pipe Networks

Submitted by Gregory Michael Meyers to the University of Exeter

as a thesis for the degree of

Doctor of Engineering in Water Engineering

in July 2018

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: ……...

Abstract

Water discolouration in potable drinking water networks is an increasingly important and expensive issue due to rising customer expectations, tighter regulatory demands and ageing Water Distribution Systems (WDSs) in the UK and abroad. The increase in real-time data acquisition systems installed in WDSs has opened the door to using data-driven methods to achieve what physically based models have to date been unable to.

The research presented in this thesis describes the development of three novel data-driven methodologies to aid in the reduction of discolouration risk and further understand discolouration issues in a WDSs. These methodologies are: a) a continuous turbidity forecasting methodology capable of forecasting if and when a downstream turbidity event will occur by only taking current and past flow and turbidity measurements at a number of upstream locations in the network; b) a methodology for estimating the percentage of downstream turbidity observations that can be linked to an upstream pipe in a network and thus identify network areas (i.e. pipes) where discolouration material accumulates in a WDS; c) an on-line turbidity event forecasting methodology that predicts if a hydraulic event occurring upstream will cause a turbidity event downstream by analysing the current and historic hydraulic forces in WDS pipes.

The results of applying these methodologies to data from three real trunk main networks in the United Kingdom (UK) over a period of two years and 11 months are also reported in this thesis. The results obtained illustrate that it is possible to reliably forecast the occurrence of discoloration events in real WDSs by using a data-driven (i.e. non-physically based) methodology only.

The results obtained additionally show the potential of the methodologies presented here to be used as an early warning system for discolouration events. This would enable a multitude of cost beneficial proactive discolouration risk management strategies to be implemented as an alternative to expensive trunk mains cleaning programs and thus enabling water companies to save money while improving their customer service and reputation.

Acknowledgements

I would first like to express my deepest thanks to my two supervisors, Professor Zoran Kapelan and Professor Edward Keedwell, for their continued support and guidance throughout this research. Their tireless assistance and advice has been invaluable to the completion of this thesis.

I am grateful to the UK Engineering and Physical Sciences Research Council (EPSRC) for providing the financial support for this research through the STREAM Industrial Doctorate Centre.

I would like to thank my family for their encouragement and especially my wife Olivia for her love and support.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AC	Asbestos Cement
AI	Artificial Intelligence
ANN	Artificial Neural Network
AUC	Area Under the Curve
BBO	Black, Brown, Orange
CEV	Current Exceeded Velocity
CTM	Cohesive Transport Model
DI	Ductile Iron
DMA	District Metered Area
DPM	Discolouration Propensity Model
DRM	Discolouration Risk Management
DWI	Drinking Water Inspectorate (UK)
FDR	False Detection Rate
FM	Flow Meters
FN	False Negative
FNN	False Nearest Neighbours
FP	False Positive
FPR	False Positive Rate
FTU	Formazin Turbidity Units
HMTP	Hydraulically Mobilised Turbidity Percentage
MCC	Matthews Correlation Coefficient
NOM	Natural Organic Matter
NRSE	Naïve Relative Squared Error
NSE	Nash-Sutcliffe Model Efficiency Coefficient
NTU	Nephelometric Turbidity Units
OFWAT	Office of Water Services (UK)
PODDS	Prediction of Discolouration in Distribution Systems
PSM	Particle Sediment Model
PVC	Polyvinyl Chloride

RBF	Radial Basis Function
RF	Random Forest
ROC	Receiver Operating Characteristic
ROC AUC	Receiver Operating Characteristic Area Under the Curve
RPM	Resuspension Potential Method
SCADA	Supervisory Control And Data Acquisition
SIM	Service Incentive Mechanism
SVM	Support Vector Machine
SWW	South West Water
TM	Turbidity Meter
TN	True Negative
TP	True Positive
TPR	True Positive Rate
UK	United Kingdom
VCDM	Variable Condition Discolouration Model
WDS	Water Distribution System
WHO	World Health Organisation
WRZ	Water Resource Zone
WTW	Water Treatment Works

Chapter 1 Introduction

1.1 Background and Motivation of Thesis

With the advancements in science and technology, customer expectations of the service they receive from their water provider has also risen. This has been most recently reflected through regulatory bodies placing heavier incentives, penalties and fines for water quality related issues (OFWAT, 2009). However even though the economic regulator of the water sector in England and Wales, OFWAT, introduced financial penalties for water companies that exceed an acceptable number of customer contacts for discoloured water in 2009, in 2016 alone over 1.7 million UK customers were still estimated to have been affected by water discolouration issues (DWI, 2017; OFWAT, 2009).

The appearance of discoloured water at a customer's tap has been shown to be the result of a sufficient hydraulic change in the water distribution network that mobilises accumulated discolouration material from the pipe walls and transports it through the network (Boxall et al., 2003; Vreeburg et al., 2005). Thus, while water utilities have traditionally placed heavier emphasis on improving water quality through catchment management and water treatment processing, water distribution networks are now being seen as a crucial component in improving water quality. This is in line with the Drinking Water Inspectorate (DWI) and the World Health Organisation (WHO) recommendations of having risk-based water quality management strategies that account for every stage in delivering water to a consumer's tap (DWI, 2005; WHO, 2004).

Unfortunately, reducing discolouration risk in water distribution networks is difficult not only due to the complex nature of discolouration material accumulation and mobilisation not being fully understood, but also due to the limited ability to monitor these networks as they are almost entirely buried.

Recent research has indicated that a significant number of discolouration events in downstream distribution networks can actually be attributed to further upstream trunk mains (Cook et al., 2015; Cook and Boxall, 2011). Trunk mains can be categorised as having especially high discolouration risks due to their significant size, enabling them to act as a form of reservoir for discolouration material build up. When considering that turbidity is akin to a concentration and is significantly diluted by the high flow rates

typical of trunk mains, then even relatively low turbidity observations in trunk mains should be of somewhat concern. This is because the discoloration material that does not directly reach a customer's tap can still resettle in a downstream network. Then, that same discolouration material, when remobilised in a smaller distribution pipe with a fraction of the flow rate, could result in a significantly higher turbidity reading.

The significant consequences and logistical complexities associated with trunk mains mean that research on active trunk mains has been significantly limited. The lack of understanding in this area has meant that while improvements have been made to reduce discolouration issues, this is still mostly dealt with in a reactive way by water companies (Blokker, 2010a; Cook et al., 2015).

Water companies primarily deal with discolouration by cleaning, i.e. flushing WDS mains, once a sufficient number of discolouration complaints have been reported in an area. The company may decide to reline (or replace) old mains believed to be the cause of significant discolouration, particularly if this is going to help address additional issues (e.g. leakage) (Husband et al., 2010; Vreeburg et al., 2008). However, rehabilitating WDS mains is expensive and can still potentially have a limited effect if discolouration material was primarily accumulated from the bulk flow and not corrosion of the pipe wall. Even cleaning whole network sections of the affected area can have a severely limited effect if the discolouration material was actually mobilised from a different upstream section of the network (Cook et al., 2015).

With ever increasing regulatory pressures, tighter standards and ageing pipe networks, it is evident that reducing discolouration risk is a key challenge facing the water industry today and new proactive management strategies need to be considered.

Physically driven models of discolouration material mobilisation have been developed to aid operational staff in reducing discoloration risk (Furnass et al., 2014; Husband and Boxall, 2016). However, these models were developed for the cleaning of single pipe stretches and are unsuitable in the context of continuous turbidity prediction. Furthermore, due to unknown pipe conditions and discolouration material build up rates, these models require hydraulic models with specialised onsite model calibration.

The requirement of a well calibrated and accurate hydraulic model has been noted as a major limiting factor in many existing water quality models (Machell et al., 2014, 2009; Skipworth et al., 2002; Vreeburg, 2007). This is similarly a common theme found when developing a burst detection model, water demand model or anomaly detection model

that must build upon a well calibrated and accurate hydraulic model (Arad et al., 2013; Blokker et al., 2009; Machell et al., 2010). Whether this has prevented the use of the model in certain areas (Tao et al., 2014) or resulted in a clear decrease in model accuracy away from the calibrated area (Leeder et al., 2012), the dependency of a satisfactorily calibrated hydraulic model has been shown to limit the application of these models.

These issues typically stem from the expense of developing and regularly updating hydraulic models and the fact that they are usually calibrated from "24 hour" data reflecting an average day in the water distribution system (WDS). This creates an additional problem for accurately forecasting discolouration as discolouration events are thought to be primarily a function of irregular hydraulic disturbances that mobilises the accumulated material and are not therefore part of the "average day".

The decreasing costs of sensors and data acquisition systems and increasing recognition of the benefits in real-time monitoring of WDSs has led to an unprecedented amount of real-time data now generated by water companies. To take advantage of this substantial increase in real-time data, data-driven models could be developed to forecast discolouration events and further understand the underlying processes of discolouration. This in turn would enable new proactive management strategies to reduce discolouration contacts. Furthermore, these data-driven models would be free from the aforementioned issues associated with using hydraulic models that are expensive to build and maintain, and potentially allow rapid application to almost any WDS with the required data acquisition systems.

1.2 Research Scope and Objectives

The overall aim of this thesis is to develop and test data-driven methodologies for the analysis and real-time forecasting of discolouration events in WDSs that enable proactive interventions in response to these discolouration events. In so doing water companies can improve their customer service through reducing discoloured water contacts which undermine consumer confidence and negatively impact a water company's reputation.

The aforementioned research scope is achieved in this thesis through the following specific objectives:

1. To undertake a critical review of discolouration in potable water distribution systems, managing discolouration risk and the current state of the art methods for

modelling discolouration in potable water distribution systems as described in the academic literature.

2. To develop a data-driven methodology for the short-term forecasting of downstream turbidity using only historical and real-time data collected from flow, pressure and turbidity monitors in a real life WDS in the UK.

3. To explore the performance of different data-driven methods for a regression based and a classification based approach to forecasting downstream turbidity in trunk water mains systems.

4. To develop a data-driven methodology to determine the approximate mobilisation origin of discolouration material that had been mobilised by preceding hydraulic forces.

5. To conduct a long-term continuous study of discolouration mobilisation to determine the approximate percentage of turbidity in trunk mains that can be linked to a preceding upstream hydraulic mobilisation process.

6. To develop an on-line turbidity event forecasting methodology that can predict if a hydraulic event will cause a downstream turbidity event.

7. To investigate the feasibility of preventing or limiting the magnitude of the resulting discolouration event through real-time control of the flow rates in WDSs.

8. To derive conclusions, summarise key findings and recommend future work from results obtained in this thesis.

1.3 Thesis Structure

Including this introduction chapter, the thesis is divided into seven chapters.

Chapter 2: Literature Review

This chapter provides a review of relevant literature and is divided into two parts. The first part of this chapter focusses on what is known about discolouration, the impacts of

discolouration and the management strategies typically used to reduce discolouration risk. The second part of this chapter assesses all known discolouration management related models and compares them and their specific area of application. This chapter addresses objective 1.

Chapter 3: Case Study Data

This chapter presents detailed information on the real trunk mains and their associated hydraulic and turbidity data used in this thesis. The data is presented here for the sake of conciseness and ease of understanding as the novel methodologies presented in later chapters train and test using different aspects of this data.

Chapter 4: Short-Term Turbidity Forecasting

This chapter presents and verifies a data-driven continuous turbidity forecasting methodology capable of forecasting if and when turbidity will exceed a predefined threshold. Three data-driven turbidity forecasting models based on Artificial Neural Networks, Random Forests and Support Vector Machines are assessed. Each of these models takes current and past flow and turbidity measurements at a number of upstream locations in the WDS to either directly predict turbidity (regression based model) or classify the turbidity at the downstream location as being above (or not) a pre-specified threshold (classification based model). This chapter addresses objectives 2 and 3.

Chapter 5: Data-Driven Study of Discolouration Mobilisation

This chapter investigates the extent to which typical hydraulic conditions can mobilise discolouration material and presents a methodology to determine the approximate origin of hydraulically mobilised discolouration material. This would identify areas or pipes in a network that discolouration material accumulates in and thus could aid in targeted proactive cleaning or replacement of problem pipes. This chapter addresses objectives 4 and 5.

Chapter 6: On-Line Turbidity Event Forecasting

This chapter presents an on-line turbidity event forecasting methodology capable of predicting if a hydraulic event will cause a downstream turbidity event within the subsequent 24 hours. If flowrates in the network can be automatically controlled, then proactive interventions may be possible to prevent or limit the magnitude of the resulting discolouration event through the use of this methodology. This chapter addresses objectives 6 and 7.

Chapter 7: Summary, Conclusions and Future Work Recommendations

This chapter summarises the work done and conclusions of each chapter, and discusses their significance and limitations. The novel contributions of this thesis are highlighted and future work recommendations are identified. This chapter addresses objective 8.

1.4 Publications

Some elements of this thesis have been published in the following journals and conference proceedings:

Meyers, G., Kapelan, Z., Keedwell, E., Randall-Smith, M. (2016). Short-term Forecasting of Turbidity in a UK Water Distribution System. *Procedia Engineering, 154*, 1140-1147.
https://doi.org/10.1016/j.proeng.2016.07.534
(This paper in part addresses objective 2)

Meyers, G., Kapelan, Z., Keedwell, E. (2017). Data-driven Approach to Short-Term Forecasting of Turbidity in a Trunk Main Network. Computing and Control for the Water Industry, Figshare, Sheffield.
https://doi.org/10.15131/shef.data.5363824.v1
(This paper in part addresses objective 2)

Meyers, G., Kapelan, Z., Keedwell, E. (2017). Short-term forecasting of turbidity in trunk main networks. *Water Research*, *124*, 67–76.
https://doi.org/10.1016/j.watres.2017.07.035

(This paper in part addresses objectives 2 and 3)

Meyers, G., Kapelan, Z., Keedwell, E. (2017). Data-Driven Study of Discolouration Material Mobilisation in Trunk Mains. *Water, 9*, 811. https://doi.org/10.3390/w9100811

(This paper in part addresses objectives 4 and 5)

Chapter 2 Literature Review

2.1 Introduction

This chapter provides an overview of the core themes and most pertinent research related to understanding discolouration in potable water distribution systems in order to assess the impact on water utilities and their customers, compare the current state of the art methods for modelling discolouration in potable water distribution systems, as described in the academic literature, and identify areas that have either been overlooked or not addressed with newer methods.

The literature review is split into three distinct sections. Section 2.2 provides a background of the processes of discolouration and the various factors that may impact how and where discolouration events occur. Section 2.3 shows the scale and impact of discolouration issues as it concerns water utilities and their customers. Section 2.4 briefly introduces what a data-driven modelling approach is and assesses known discolouration modelling methods. Section 2.5 introduces real-time modelling of water distribution systems. Finally, Section 2.6 provides a summary of this chapter and identifies the current research gaps and thus identify matters of further research interest.

2.2 Discoloured Water in Potable Water Distribution Systems

2.2.1 Overview

Through detailed and high frequency monitoring of treatment works, confidence in the consistent quality of water entering distribution systems is increasing. However, determining the water quality of water within the water distribution systems is difficult. This is because the physical, chemical and biological processes that occur within water distribution systems are difficult to study due to their complex topologies and subterranean nature.

Discolouration formation is complex and not completely understood, with bulk water quality, temperature, network layout, pipe material and age all believed to be factors (Abe et al., 2012; Husband and Boxall, 2011; Van Thienen et al., 2011; Vreeburg et al., 2008). Discolouration has been seen to vary even between different parts of the same water distribution network and yet is still similarly experienced throughout the world regardless

of the wildly differing factors between their Water Distribution Systems (WDSs) (Armand et al., 2015; E. J. M. Blokker and Schaap, 2015; Husband et al., 2008; Vreeburg and Boxall, 2007).

From the accumulation of foreign material in pipes and the corrosion of pipe materials itself, usually old unlined ferrous pipes, 'pockets' of various materials accumulate along the insides of pipes. Then most often upon some sort of significant hydraulic change in the pipe, the stored material is mobilized. If concentrations are sufficiently high to be visible to the human eye, this results in discoloured water at an outlet, usually by a customer. Discoloured water is also known as brown water, red water or BBO (Black, Brown, Orange).

2.2.2 Formation and Mobilisation of Discolouration Material

2.2.2.1 Sources of Discolouration Material

The formation of discolouration material in water systems is still not completely understood, this is due to the large amounts of possible factors involved in the formation of discolouration deposits (Armand et al., 2017). While the exact make-up of discolouration material can be quite complex, the origins of the particulates can be put into one of two categories, external or from the potable system itself.

External particulates can enter the system as:

- Minute concentrations of organic or inorganic material from the treatment works either by acceptable levels of residual compounds or unintentional incomplete removal of compounds.
- Accidental addition of particulates from the treatment works itself such as residual coagulation chemicals.
- Intentional addition of particulates from treatment works such as residual disinfection products and their by-products.
- Bursts and leaks in the system resulting in material directly entering the system.
- Residual material from inspections or repair/replacement work.

The potable system itself can also produce particles such as:

- Corrosion of pipes and their fittings
- Corrosion of pipe lining
- Chemical by-products

- Biological growth

The mix of particulates further change and vary as they travel through the system, with the age, size and water flow rate of pipes greatly effecting the make-up of the slow accumulation of particulates.

2.2.2.2 Gravity Settling and Cohesive Wall layers

Traditionally discolouration material has been viewed as sediment that settles along the bottom of pipes. However, studies have shown that the typical hydraulic conditions of WDSs prohibits discolouration material from settling and thus other processes must facilitate the transport of material from the bulk flow to the pipe wall (Boxall et al., 2003, p. 200; Van Thienen et al., 2011; Vreeburg, 2007).

Van Thienen et al. (2011) showed that turbulent diffusion is the process that dominates radial transport as particle diameters are typically too small and flow velocities are typically too low for other mechanisms such as turbophoresis (i.e. the transport of particles across a turbulence gradient due to particle inertia) to be a significant factor. However, Vreeburg (2007) showed that discolouration material is not held at the wall by radial transport processes alone as some material remains bound to the pipe wall after draining, thus indicating that discolouration material can have cohesive properties and accumulates on the pipe wall in cohesive layers. This concurred with earlier work by Boxall and Saul (2005) that theorised that discolouration material binds to the pipe wall in cohesive layers and are conditioned by the hydraulic shear stress. Subsequent work by Pothof and Blokker (2012) also similarly concluded that non-cohesive discolouration material can only accumulate within pipes if the daily peak hydraulic forces are constantly very low.

2.2.2.3 Hydraulic Forces

Discolouration mobilisation is widely believed to be primarily caused by sufficiently large hydraulic changes in the WDS resulting in the detaching and transportation of the accumulated discolouration material through the network and producing discoloured water at the consumer's tap (Boxall et al., 2003; Prince et al., 2001; Vreeburg et al., 2005).

The exact contributions of individual hydraulic processes in driving material detachment and transport from the pipe wall are currently unknown. However, the shear stress along the pipe wall (i.e. at the material/fluid boundary), is typically considered to be the primary hydraulic force required for the erosion and mobilisation of cohesive layers (Husband and

Boxall, 2009; Pothof and Blokker, 2012; Vreeburg and Boxall, 2007). However, as shear stress along the pipe wall is a function of not just velocity, but also hydraulic radius and boundary roughness, a calibrated hydraulic model is typically required to derive the shear stresses in a water distribution network.

2.2.3 Factors that can Impact Discolouration Risk

2.2.3.1 Network Composition

While potable water systems are complex with near infinite possible layouts and setups, there is actually very little difference in pipe diameters and the materials used. Aside from network size and configuration, the main difference is the percentage of material types used. The make-up of the country's potable water network is usually determined by what era the network saw significant growth in and the amount of money put towards rehabilitating older parts of the network.

The UK is a country with an ageing network, with most of its network installed before the 1940s when cast iron was the main material used for water pipes. This has left the UK with a large iron corrosion problem that is recognized as a key contributing factor of discolouration in the UK (Husband et al., 2008; Prasad and Danso-Amoako, 2014). This is in stark contrast to the Dutch's potable water networks, where large growth and heavy investment in the last 50 years has resulted in less than 10% of their networks being made out of iron pipes.

One might believe that with a much newer network composed primarily of Polyvinyl Chloride (PVC) and Asbestos Cement (AC) pipes, where little to no pipe corrosion exists, investing in preventing discolouration events and mitigating their effects with discolouration risk modelling is unnecessary. However, while these networks experience reduced discolouration, it has been shown that background concentrations of organic and inorganic materiel in the bulk flow from water treatment plants is more than sufficient to cause regular discolouration deposits in the network (Blokker et al., 2011; Vreeburg et al., 2008). Discolouration has even been seen to be significantly affected by factors such as: if iron pipes are cast, ductile or galvanised; the type and condition of pipe lining; if protective scaling of iron pipes has been removed through rough cleaning activities such as pigging; and the age and roughness of pipes (Sarin et al., 2004; Slaats et al., 2003; Verberk et al., 2006; Vreeburg et al., 2009).

2.2.3.2 Network Topology

Certain features of distribution network's topology, such as loops and dead ends, can increase discolouration risk because they can experience low shear stresses for sustained periods of time (Armand et al., 2015; Schaap and Blokker, 2013). Likewise, discolouration material has also been shown to accumulate around fixtures, fittings and sharp angled bends (Husband, 2010; Neilands et al., 2012). The greater build-up of discolouration material before being rapidly mobilised by a hydraulic event in the WDS is what significantly increases discolouration risk.

2.2.3.3 Water Treatment Works (WTWs)

The composition of water leaving WTWs can greatly impact downstream discolouration risk. Vreeburg et al. (2008) showed that installing 0.1µm ultra-filtration immediately downstream of the treatment works in a DWDS containing only non-corrodible pipes (AC and PVC) significantly reduced material regeneration times by an order of magnitude. Note however that this still did not eliminate all discolouration risk in the WDS.

Properties of WTWs and its effluent water that could influence the downstream discolouration risk are:

- Sudden changes in raw water quality can increase effluent turbidity levels.
- Treatment works failures can increase effluent turbidity levels.
- Backwashing of treatment works filters have been correlated with turbidity spikes (Vreeburg et al., 2008).
- Coagulants in some treatment work processes can form discolouration material (Husband and Boxall, 2011).
- Water sources richer in natural organic matter (NOM) such as upland sources compared to groundwater.
- Changes in pH could expedite corrosion (McNeill and Edwards, 2000; Sarin et al., 2004).
- Nitrate, dissolved CO2 and phosphate can affect biofilm growth rates (Douterelo et al., 2016; Makris et al., 2014).
- Disinfectant residual concentrations have been shown to effect downstream accumulation rates (Gauthier et al., 1999; Sly et al., 1990).

2.2.3.4 Temperature

The exact effects of temperature on discolouration risk is still unclear, however several studies have noted temperature as an influence on discolouration material accumulation rates (E.J.M. Blokker and Schaap, 2015; van Summeren et al., 2015). Schaap and Blokker (2013) showed a positive correlation between bulk water temperature and discolouration material accumulation rates over an eighteen month study of a Dutch distribution main. Sharpe (2012) showed in month long experiments at a temperature controlled facility that temperature had a significant effect on discolouration material accumulation. McNeill and Edwards (2000) showed that changes in temperature can impact the corrosion of pipes, fittings and fixtures.

2.2.3.5 Flow Reversals

Husband and Boxall (2016) conducted flow reversal experiments on a trunk main by using fixed speed pumps at a downstream reservoir. The results suggested that for the low roughness main used, flow direction had little impact on material mobilisation. Furthermore, the mobilisation patterns for both forward and reverse flow increases indicated that accumulation was consistent along the length of the main.

2.2.3.6 Transience Pressure Waves and Unsteady Hydraulics

Transient flow comes about due to rapid changes in the hydraulic conditions caused by actuations in the WDS. These are typically valve open and closing operations, pump starting and stopping and sudden leaks. Each of these can generate significant transient waves that propagate upstream and downstream of the event. The speed and shape of the resulting transient wave depends on the characteristics and structure of the WDS (ex. junctions, valves, elbows, leaks, pipe material composition, etc.).

Tzatchkov et al. (2008) showed that transients generated by the operation of household water appliance in a residential property is near imperceptible in upstream distribution mains.

Both very high and low transient pressures have been shown to cause structural damage to a WDS, however, low transient pressure waves can additionally cause water to be sucked into the WDS via cracks and leaks (Besner Marie-Claude et al., 2007; Collins et al., 2012; Tang et al., 2000). This water ingress has the potential to contaminate the WDS with any manner of particulate materials (Collins et al., 2010; LeChevallier et al., 2003).

As mentioned above, discolouration material mobilisation is primarily believed to be due to sudden increases in hydraulic forces, however often only steady-state hydraulic conditions are considered where the liquid has constant density, no inertia and is transported through rigid pipes. The total shear stress at the wall under unsteady conditions could be significantly more than the steady-state shear stress, with some indications of it being even three to ten times the magnitude (Aisopou et al., 2011; Brunone et al., 2000; Naser et al., 2006).

The effects of unsteady shear stresses and transient events on water quality is not fully known. This is in part due to the difficulty in modelling the complexity of the velocity profile under such conditions and also due to high temporal resolution of data required to capture these events properly.

For capturing and modelling most WDS events and states, an observation time step of 5 or 15 minutes is usually deemed sufficient and thus vary rarely are permanent pressure monitors recording at the sub-second temporal resolution required to capture transient events. The required pressure sensor time step size is shown to usually be in the order of magnitude of a few milliseconds (Aisopou et al., 2011; Husband and Boxall, 2016). Additionally a non-insignificant amount of work is required to keep such pressure monitors running long term and handle the large amounts of resulting data generated.

The only significant study found to have experimentally assessed the effect of transience events on turbidity levels in a controlled setting is Mustonen et al. (2008). A pilot distribution system consisting of two 100 meter long looped pipelines (one 10 mm copper pipeline and one 12 mm plastic pipeline) was subjected to pressure shocks after first allowing material to develop on the pipe walls. The pressure shocks were artificially created by rapidly halting the flow via compressed air valves for 5 seconds before then rapidly reopening the valves. Turbidity was seen to have increased in both the copper and plastic pipelines, however this was only an average increase of 0.2 NTU over the normal bulk flow levels. Furthermore the increased turbidity levels did not last long as the turbidity levels reverted back to the pre-shocked bulk flow level within 15 minutes. Thus while discolouration material did mobilised as a result of the pressure shocks, the amount can be considered significantly limited considering the substantial nature of the pressure shocks given to the system.

While some consider transience to be overall relatively insignificant when modelling water quality at time steps greater than 1 minute (Blokker, 2010b), it is widely agreed that

the underlying physical processes and impacts of unsteady hydraulics including transient pressure waves on discolouration potential require further investigation (Aisopou et al., 2011; Furnass, 2015; Romano, 2012).

2.3 Scale and Impact of Discolouration

2.3.1 Perception and Quantification of Discolouration Water

A water company has to inform the Drinking Water Inspectorate of the number of customers that contact the water company about a water quality issue. These water quality contacts can however help identify the state of a water network. The Drinking Water Inspectorate states that a discoloured water contact is "a contact where the consumer perceives something different about the appearance of the water from the norm" and "the colour of the water is the focus of the contact" (DWI, 2006).

Because a customer contact for discolouration is based on perception, these contacts are inherently very subjective. This is especially so because typical water discolouration experienced by consumers is not due to colour but due to suspended solids in the water that can slowly settle under certain conditions (Vreeburg and Boxall, 2007). For instance, discoloured running water is generally harder to perceive compared to water gathered in a white sink or bath. This can also greatly vary from one customer to the next as their current circumstance, past experience and if they deem it is worth contacting their water supplier, are all considered factors.

Discolouration levels are thus typically quantified as turbidity and described as a measure of how incident light scatters due to the particulate matter in water. The two most common specifications for measuring turbidity is the European standard ISO 7027 and the American standard USEPA 180.1. In both standards it is specified that the light emitter and detector are positioned perpendicularly to each other. It is recommended that turbidity be measured as soon as the sample is taken because turbidity is sensitive to particle size and the particulate matter may flocculate given time (WHO, 2011). Turbidity is measured in either Nephelometric Turbidity Units (NTU) or Formazin Turbidity Units (FTU). While FTU is the ISO standard for measuring turbidity, NTU and FTU are typically considered interchangeable.

UK water providers have recently begun installing permanent turbidity monitors at WTWs, service reservoirs and key trunk mains to allow for continuous, prompt and

accurate observation through a Supervisory Control And Data Acquisition (SCADA) system.

With improvements in technology over the years, instruments that simultaneously monitor multiple water quality factors are being developed. These instruments could enable insights into how factors such as turbidity, temperature, conductivity, Cl concentration and so on correlate and interact (Aisopou et al., 2010; Gaffney and Boult, 2012; Leeder et al., 2012).

Unfortunately in the case of Aisopou et al. (2012), the sensors' performance was quickly affected by bio-fouling and deposit build-up on the electrochemical and optical sensors such that sensitivity and repeatability deteriorated in less than three weeks. The low accuracy and uncertainties in the sensor data meant that frequent and complex recalibrations would be required. Thus Aisopou et al. (2012), recommended that further and fundamental developments were necessary before water utilities can reliably use these sensors for large scale water quality monitoring.

While turbidity under 4 NTU can be detected by the naked eye, consumer perception of when water can be deemed to be different from "normal" is inherently subjective (WHO, 2011). Slaats et al., (2003) found this perception threshold was on average around 8 - 10 FTU.

2.3.2 Discolouration Events in the UK

While UK water companies have high drinking water standards, unfortunately discoloured water is one thing many of their customers would have experienced at some time or another (Cook, 2007; Vreeburg and Boxall, 2007). On average over a third of all customer contacts each year are due to discolouration (Cook, 2007; DWI, 2014a).

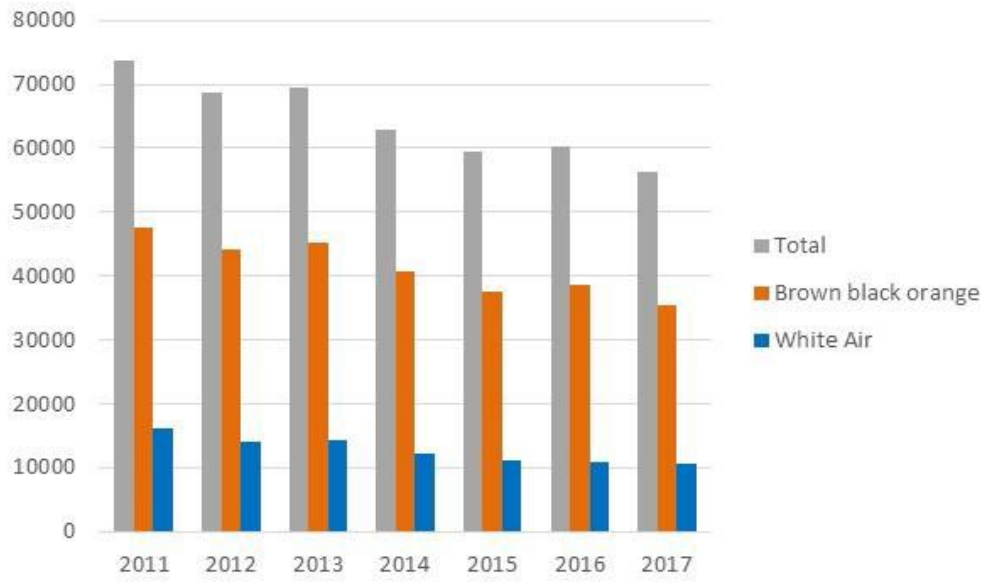


Figure 2.1. Number of contacts to companies reporting appearance issues 2011-2017 in England. Taken from (DWI, 2018a).

While Figure 2.1 shows steady progress in reducing the number of appearance related contacts, in 2017 alone there were still over 50 thousand contacts. While this is already a high number, this is only the number of people who directly contacted their water provider. It is estimated that over 2 million UK customers were actually affected by discolouration in 2013 and over 1.7 million effected in 2017 (DWI, 2014b, 2018a).

Water companies are required to report any "event" where there is a potential negative impact on public confidence in the water supply. There are 5 categories for rating the severity of the event, they are: Not Significant, Minor, Significant, Serious and Major. For the latter three categories a warranted Inspector has to make a detailed investigation of the event (DWI, 2005).

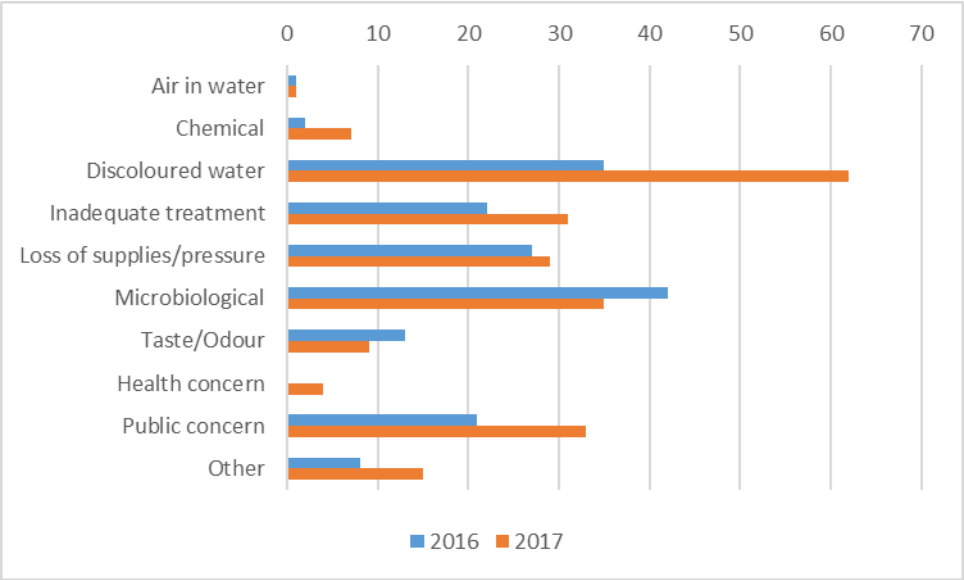


Figure 2.2. Number of Significant, Serious and Major Events in 2016 and 2017.

Figure 2.2 shows that over 20% of the 171 Significant, Serious and Major Events in 2016 were due to discoloured water. This percentage increased in 2017, where discoloured water accounted for over 27% of the 226 Significant, Serious and Major Events.

Discolouration events are hard to predict due to complex hydraulic processes in a distribution network, especially as elements such as pipe bursts and use of hydrants for firefighting can cause discolouration events. However, of the discolouration events classed as Significant, Serious or Major in 2013, 43% of those events were caused by planned works (DWI, 2014a, 2014b). With advanced knowledge of how the underlying hydraulics in the distribution system will be changed, such as disconnecting a pipe or closing of a valve, a discolouration risk model could simulate what potential discolouration may exist and how to prevent it from causing customer contacts. Even if preventative methods were not available or too costly, customers could still be informed in advance as so not to be alarmed. With such a high number of discolouration events caused by planned work, correct use of discolouration risk models could drastically reduce the number of customer contacts and in turn save a water company a lot of money.

2.3.3 Other Impacts of Discolouration

While discoloured water has been linked to several health risks (DWI, 2005; LeChevallier et al., 2003; Mann et al., 2007; WHO, 2011), the health concerns are generally only applicable through regularly drinking said discoloured water. However, even putting

aside the validity of public health concerns, discoloured water can still undermine consumer confidence and negatively impact a water utility's reputation.

Discoloured water can stain laundry and with more household appliances connecting directly to the water supply, such as fridges that automatically dispense ice, the potential damage discoloured water could cause to the contents and appliances themselves has increased.

If a customer experiences a discolouration event and believes that they have been exceptionally inconvenienced they may seek remuneration. If they can show that drinking water supplied was unfit for human consumption and the water company did not do enough to prevent the discoloured water within reason, the DWI will prosecute on a customer's behalf. While many cases are settled before going to court, the cases that do go to court have been recorded by the DWI since 1995. Of all water quality claims that went to court, from 1995 to January 2014, over 74% of these are attributed to discolouration (DWI, 2018b). And in this period, UK water companies have had to spend over 1 million pounds in claims to customers due to discolouration alone, this is separate from any additional fines they may have occurred from DWI.

2.3.4 Regulatory Standards and Guidance

The WHO recommends that turbidity in distribution networks should not exceed 1 NTU so as not to affect the effectiveness of the disinfection (WHO, 2011). The DWI similarly state that the UK regulatory limit for water leaving water treatment works, but prior to disinfection, should be no more than 1 NTU (DWI, 2014a). The DWI also states that turbidity should not exceed 4 NTU at customers' taps.

With increasing water quality standards and improving technology available to help address discolouration, in 2009 as part of the PR09 process, OFWAT introduced two new serviceability indicators for drinking water quality, one of which is 'Customer Contacts of Discolouration' and is collected by the DWI (OFWAT, 2009). Water companies that perform above average are rewarded and those that do not are financially heavily penalised.

For PR14, South West Water (SWW) performed many customer service surveys in order to identify what customers were willing to pay for in their water bill and if they were willing for a deterioration in some service areas if it decreased their bills. Out of all the service improvements offered, customers were willing to pay the most to reduce the

number of discolouration events. This was 34% more than the second highest service improvements offered, which was reducing leakage (South West Water, 2014).

2.3.5 Discolouration Risk Management

Traditionally water companies deal with discolouration issues in a reactive way. It is only upon receiving a customer complaint that a water company will know of a discolouration event and then potentially take action. Although customer complaints represent a measure of the discolouration issue in a network, the obvious disadvantage from a water company perspective is that this requires a customer complaint to be received, which makes it too late to mitigate the damage of that discolouration event.

Cleaning a trunk main can result in the improvement of bulk water quality leaving the trunk main which has been shown to reduce the likelihood and magnitude of downstream discolouration events (E. J. M. Blokker and Schaap, 2015; Cook et al., 2015; Sunny et al., 2017). This in turn also reduces the frequency of maintenance and cleaning required in downstream distribution pipes. This is further evidenced by a study in the UK showing that between 30% and 50% of discolouration events seen in the reported District Metered Areas (DMAs) could be linked to imported discolouration material from upstream trunk mains (Cook et al., 2015). The study also found that depending on the DMA, between 0% and 51% of discolouration contacts could be linked to upstream trunk mains. While the mean was only 9%, this shows that some WDSs are significantly more susceptible to discolouration events from trunk mains than others. The benefits of cleaning trunk mains are easy to observe in extreme examples, such as when a trunk main supplying 1.75 million customers saw an associated 62% reduction in overall customer contacts after being cleaned (Husband, 2010).

As discolouration material has been shown to continually accumulate ubiquitously on pipe walls, one-off capital and operational measures should be coupled to an ongoing operational programme to keep discolouration risk low.

2.4 Discolouration Modelling

2.4.1 Introduction

This section starts with a brief introduction and comparison between a physically-driven modelling approach and data-driven modelling approach. This section then looks at a

number of discolouration models and assesses their application domains, inner workings, merits and deficiencies.

2.4.2 Physically Driven and Empirically Driven Models

All hydrological modelling approaches can be placed somewhere on a spectrum between fully physically driven and fully empirically driven (Romano et al., 2012; Solomatine et al., 2009; Tao et al., 2014).

Physically driven modelling approaches use explicit knowledge of the system being modelled to correctly capture the behaviour of the system (Blokker et al., 2009; Skipworth et al., 2002). Physically driven models can simulate and explain the relationships between variables, however this can often be non-trivial as combinatorial variables and interdependent processes can exponentially increase the model's complexity. Furthermore, if the underlying processes being modelled are not fully understood, the resulting model may have limited performance or produce misleading results.

Conversely, an empirically driven modelling approach is based on analysing the data about a system. This is also why empirically driven modelling is synonymous with data-driven modelling and the methods involved in these types of models are aptly named data-driven methods. Data-driven methods work by finding relationships between the input, output and internal variables from the data given (Furnass et al., 2011). It does this without requiring explicit knowledge of the underlying processes of the system.

Machine learning methods can be considered as a subset of data-driven methods and are used to learn the relationship between a system's inputs and outputs (Breiman, 2001; Dietterich, 2002; He et al., 2014; Rashidi et al., 2016). The model is trained on a training dataset that is representative of the system's behaviour. After training, the model's performance can be checked using an independent test dataset to determine how well it can generalise to unseen data. If the model performs well on the test dataset, then it is considered to have captured the underlying processes of the system and not just the relationship between values in the training dataset.

While data-driven models do not need to be given knowledge of the underlying processes, this knowledge may still be leveraged in the model's construction to improve performance or speed up training (Arad et al., 2013; Furnass et al., 2014; Machell et al., 2010; Romano and Kapelan, 2014). In this way, the design of data-driven models is typically influenced to some extent by existing theories of the system. A caveat of data-

driven models is that the learnt relationship between a model's inputs and outputs can be difficult to understand. For this reason, even if a data-driven model performs well, extracting new insights into the underlying processes of a system may be non-trivial (Gevrey et al., 2003; Shojaeefard et al., 2013). This is why data-driven models are also sometimes referred to as being a "black box".

While data-driven approaches may have some strong advantages over physically driven approaches, an inherent and significant limitation of data-driven approaches is that they typically need substantial amounts of reliable data. Lack of sufficient reliable data for a model to train on can mean the model does not accurately learn the underlying processes of the system and thus is likely to perform poorly on unseen data.

Physically driven models that are calibrated from observed data may also be able to benefit from having large amounts reliable data available. However, the ability to easily leverage large amounts of data to improve the performance of the model is not inherent to physically driven models and thus impact of large amounts of available data will likely be very model specific.

2.4.3 Resuspension Potential Method (RPM)

The Resuspension Potential Method (RPM) was developed by KWR Water Recycle Institute in the Netherlands and is a method for planning and implementing mains cleaning (Vreeburg, 2007). Reportedly most Dutch water companies use this method, as well as some Australian water companies. RPM is based on the principle that a hydraulic disturbance causes the resuspension of particulates already present in the system. RPM compares the presence and mobility of discolouration deposits before and after a hydraulic disturbance in the system.

The RPM works by isolating a pipe and then increasing the flow in the pipe to determine the base amount of discolouration material in that pipe. Primarily based on the turbidity distribution, points are given for five criteria based on the discolouration reaction. Each criterion can score up to 3 points, for a total of 15 points, with 0 being no resuspension potential and 15 being maximum risk. As the corresponding score only ranges from 0 to 15, a broad priority rank is formed.

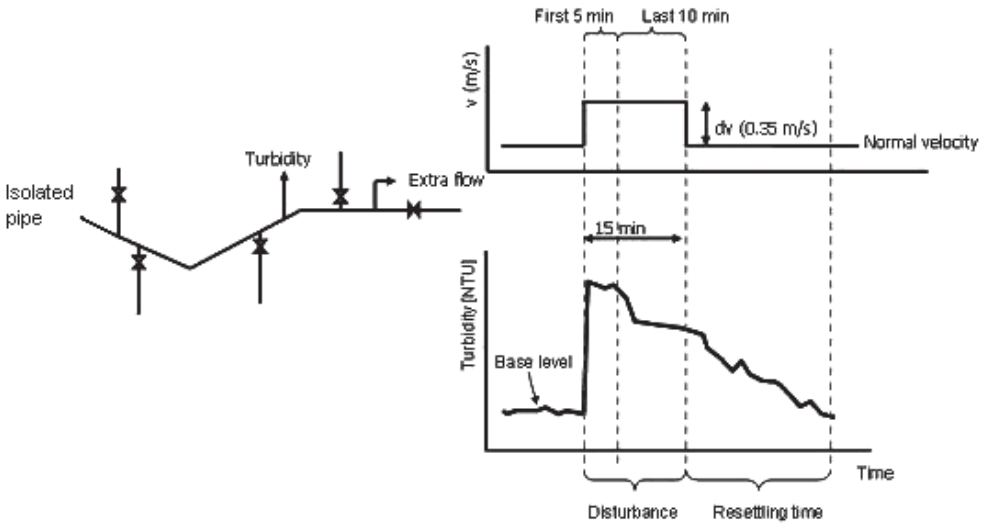


Figure 2.3. Concept of RPM procedure (Vreeburg, 2007), taken from (Kjellberg, 2007).

Due to the simplistic principal of RPM measuring and then ranking discolouration, it allows for multiple applications of RPM (Vreeburg, 2010). One common use of RPM is being able to compare conditions before and after a mains flushing, and thus the effectiveness of the mains flushing program can be evaluated and conclusions about the regeneration rate of discolouration in the network drawn (Vreeburg, 2010; Vreeburg and Tankerville, 2011).

RPM is simple and does not require a calibrated hydraulic model of the network or specific understanding of the network, yet as direct measurements are taken from the water distribution system it can still produce accurate results from the observed hydraulic profile. RPM does however have an inherent flaw when used to measure turbidity of the same location multiple times without flushing between each measurement. This is because when RPM takes a measurement, it is to some degree cleaning the pipe. Thus, repeatedly performing RPM in the same location straight after each other would lead to vastly different results.

Another downside to this method is that it can be extremely time consuming and expensive when used on larger networks as turbidity measurements need to be taken on site for most pipe lengths. RPM was not developed with whole pipe networks in mind as even with a cut down and time limited variation of RPM, only around 4 samples can be feasibly collected by one team in an 8-hour working day. However, once again, due to RPM's simplistic nature, an unskilled workforce could be trained relatively quickly to implement this method.

2.4.4 Particle Sediment Model (PSM)

Not much is known about the Particle Sediment Model (PSM) due to its limited appearance in academic literature, with Ryan et al. (2008) being the only source of information about this model. The PSM appears to be an empirical concentration-based model that was developed to model sedimentation and resuspension from different points along a pipe and material accumulation along the pipe's circumferences based on Van der Waals forces. Due to the severely limited information and model validation, the validity and application of the PSM is questionable.

2.4.5 Cohesive Transport Model (CTM)

The Cohesive Transport Model (CTM) is based on the assumption that discolouration material is held in place by forces greater than gravity settling. Thus, contrary to prior models based on traditional sediment transport theory, the CTM instead models discolouration as stable cohesive layers attached to the pipe walls (Boxall and Dewis, 2005; Boxall and Saul, 2005).

The strengths of these cohesive layers are modelled as a proportional inverse of the discolouration potential, meaning an increase in discolouration potential implies a decrease in layer strength. The initial strength of the cohesive layers is conditioned by the maximum daily shear stresses experienced within the pipes. Therefore, pipes that only experience low shear stresses on a daily basis, like the end of a network branch, will have high discolouration potential as discolouration material could sufficiently accumulate there over time. If a sufficiently large hydraulic disturbance appears (e.g. a pipe burst) that increases the shear stress in a pipe to exceed its maximum daily shear stress, then any potential material stored there may be mobilised to create a discolouration event.

The core of the CTM approach can be represented in four equations (Cook, 2007; Husband, 2010). The first of these is the shear stress equation describing the force acting on the internal pipe wall in the direction of water flow:

$$\tau = \rho g R_h S_0 \qquad \text{(Eq. 2.1)}$$

where τ is the actual shear stress (N/m^2), ρ is water density (kg/m^3), g is gravitational acceleration (m/s^2), R_h is hydraulic radius (m) and S_0 is hydraulic gradient (Unitless).

The relationship between the discolouration potential and layer strength is given as:

$$\tau_s' = \frac{C^b - C_{max}}{k} \quad \text{(Eq. 2.2)}$$

where τ_s' is the corrosion layer yield strength (N/m^2), C is the stored turbidity volume of the layer (NTU/m^3), C_{max} is the maximum turbidity that could ever be stored in a pipe (NTU/m^3), k is the gradient of the potential as a function of layer strength and b sets the first order relationship. The effect that k and b have on the relationship between stored turbidity and layer strength is shown in Figure 2.4.

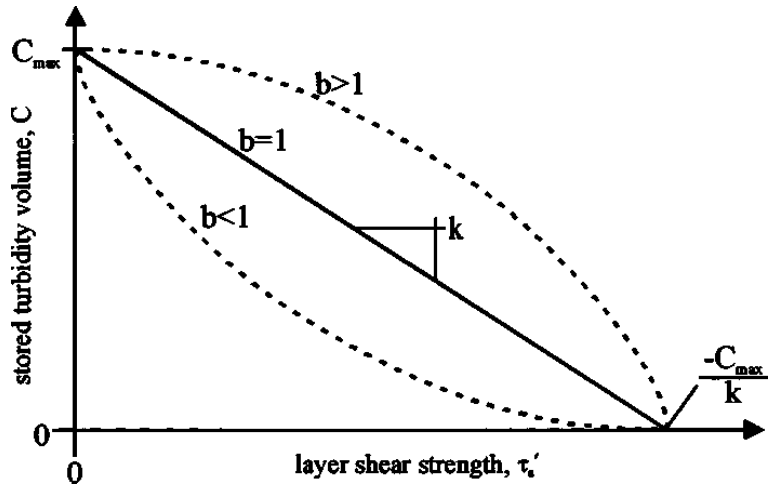


Figure 2.4. Turbidity potential versus layer strength relationship, taken from (Boxall and Saul, 2005).

The hydraulic force that mobilises the discolouration material is represented as:

$$R = P(\tau - \tau_s')^n \quad \text{(Eq. 2.3)}$$

where R is the rate of supply, $(\tau - \tau_s')$ gives the excess shear stress and the two terms P and n are empirical parameters that define the amount of material released as a function of excess shear stress. The final equation expresses the change in turbidity over time from the erosion of the discolouration layer:

$$\Delta N = \frac{RA_s}{Q} \quad \text{(Eq. 2.4)}$$

where ΔN is the change in turbidity, A_s is the pipe surface area and Q is the flow rate. The advantage of CTM is that the complex and uncertain processes involved in the formation of discolouration deposits do not need to be explicitly stated in the model in order to use it.

2.4.6 Prediction of Discolouration in Distribution Systems (PODDS)

The Prediction Of Discolouration in Distribution Systems (PODDS) model is the original implementation of the Cohesive Transport Model and was developed by the Pennine Water Group at the University of Sheffield (Vreeburg and Boxall, 2007). It is arguably the widest used discolouration model in the UK as many of the major UK Water Companies have at least once used a version of PODDS implementation. PODDS was developed for the cleaning of single pipe stretches with minimal invasive action required. It does this by determining a hydraulic stepping plan to gradually clean the water distribution pipes, one at a time, without crossing the estimated threshold where a customer can detect discolouration (Boxall and Dewis, 2005; Husband et al., 2011; Husband and Boxall, 2016). An alternative objective of the PODDS model is in using it to predict the regeneration rate of discolouration material in pipes, however due to the complex nature of the regeneration process, this research has yet to be validated (Husband et al., 2012; Machell et al., 2014).

Due to unknown pipe conditions and discolouration material build up rates, the PODDS model requires a hydraulic model with onsite model calibration before each use (Boxall and Saul, 2005; Husband and Boxall, 2011, 2009). This is especially so because the PODDS model parameters have been shown to be highly sensitive to pipe diameter, material and source water (Boxall and Saul, 2005). For this calibration, the PODDS model requires high resolution pressure, flow and turbidity profiles of a turbidity event in order to gauge the current discolouration response of the system. In order to get this calibration data, a small controlled hydraulic disturbance is created while all ends of the pipe monitored with flow and turbidity sensors. This hydraulic disturbance is created by sufficiently increasing the pipe flow. The calibration process uses the controlled hydraulic disturbance to minimisation of the error between initial turbidity predictions of the PODDS model and the controlled hydraulic disturbance observations.

As the PODDS model cannot model the continuous material erosion and regeneration of discolouration material in pipes, it can only be produce reliable predictions of turbidity levels straight after calibration. If a second turbidity prediction is needed for same pipe at a later date, a new hydraulic disturbance and recalibration is required first for reliable predictions. This limitation of the PODDS model to be incapable of modelling discolouration processes other than short-term erosion is acknowledged by the then developer of the PODDS model (Furnass, 2015).

The PODDS model was implemented on top of the open-source modelling software EPANET (Boxall and Saul, 2005; Rossman, 2000). The PODDS formulae is used to calculate the amount of discolouration material mobilised before the existing EPANET code then calculates the advection and mixing of suspended material in the bulk water. Because EPANET assumes perfect mixing at junctions and steady-state hydraulic conditions, the PODDS model as it is currently implemented is fundamentally unable to account for transient flow in pipe networks.

2.4.7 The Variable Condition Discolouration Model (VCDM)

The Variable Condition Discolouration Model (VCDM) was developed to build upon the PODDS model in order to emulate material erosion and regeneration in pipes over time, for the purpose of tracking discolouration potential over an extended period (Furnass et al., 2014). The VCDM does this by representing the latent material quantity versus shear strength relationship of wall-bound discolouration material as a relative quantity per each strength band, where each band can erode or regenerate irrespective of other bands.

For model parameter calibration, the VCDM still requires repeated pipe-specific turbidity events to capture the local condition in the pipe and a calibrated hydraulic model to track the turbidity responses. The sole paper on the VCDM showed only historical discolouration potential being simulated due to the fact that all available turbidity was used for calibration (Furnass et al., 2014). Furthermore, only a single pipe was assessed here, highlighting that while the VCDM added additional functionality to the PODDS model, it was unable to be used for the prediction of turbidity events and on a network scale.

2.4.8 Discolouration Propensity Model

Mouchel's Discolouration Propensity Model (DPM) was developed to aid in mid-term rehabilitation prioritization and long-term investment plans on whole network scales (Dewis and Randall-Smith, 2005; Randall-Smith et al., 2011). DPM is a risk-based model that produces a relative risk scale for entire networks to aid in the prioritization of pipe rehabilitation and replacement schemes.

The first component of DPM implements a simplified CTM which enables the estimation of the volume of accumulated material in a network (McClymont et al., 2013; Randall-Smith et al., 2011). This simplified CTM sticks to the core principles of CTM, which is: a) the impact of mobilized material is proportional to and limited by the material stored

in the network; b) by reducing the potential material in the network the associated discolouration risk will also reduce.

DPM requires a hydraulic model but does not require the specific calibration data required by a full CTM model (i.e. site operational history and pressure, flow and turbidity profiles of a discolouration event). It circumvents this problem by not modelling the turbidity response profile, instead it just estimates the total of discolouration potential that could be released given a hydraulic event. Because of this DPM only requires two of the core CTM calculations, those being the shear stress calculation and the mobilisation equation describing the relationship between the shear stress and the amount of material that is stored in a pipe.

Once the daily conditioning shear stress has been calculated by the hydraulic model, the pipe's maximum potential for stored material is calculated. Then for each pipe in the network, DPM runs two hydraulic simulations of the network:

- With the pipe closed off to the rest of the network, simulating the pipe being closed for maintenance work.
- With an exceptional demand applied to simulate a burst in that pipe.

A matrix of all hydraulic results through the CTM calculations gives two relative ratings on: a) the impact a pipe will have on the discolouration performance on the rest of the network; b) the amount of discolouration generated by a pipe given bursts elsewhere in the network.

Because the required processing time of Mouchel's DPM is $O(n^2)$ where the n is the number of pipes in a network, the simplicity of the underlying calculations is somewhat important. While runs of DPM have shown that a network of approximately 10,000 pipe spans require tens of hours on a current high-end desktop computer to process, this is well within the operational timescales that DPM results would be used for.

The simplicity of the CTM calculations that makes this model possible is also the biggest limiting factor of DPM. This is because CTM does not model turbidity transportation through the network and its results are limited to a relative estimation of turbidity. The only actual factors a pipe has in directly calculating its discolouration potential is its own diameter and roughness factor. This is obviously a simplification in comparison to the many factors that could significantly affect a pipe.

The aggregated results of the CTM component is used as an estimated consequence for risks, while the likelihood of risk given by pipe assessment tools evolved from the once called Discolouration Risk Management (DRM) tool (Dewis and Randall-Smith, 2005). The DRM tool bases component weights on many factors like age, material type and soil fracture index to give the likelihood of pipe failure. These factors were inferred from a panel of experts and arranged into a hierarchical tree structure.

The asset management tools and the implementation of CTM work together to produce weighted risk-based results. DPM results can recommend specific intervention options for individual pipes, for example if a particular pipe has a high discolouration potential, high sensitivity to the rest of the network but a low failure probability then regular flushing of that pipe is recommended. However, if for instance a pipe had a low discolouration potential itself but high failure probability and the network has high enough sensitivity to that pipe, then replacement is recommended instead.

Although DPM can only provide an approximation of discolouration in relation to other pipes in the network and does not produce an actual turbidity value for any pipes, the only real requirement for DPM is a semi-calibrated hydraulic model. This means that any company that already has hydraulic models for other reasons, like pressure management or future demand estimation, can easily incorporate DPM. Another advantage of DPM not being based on complex calibrations is it is fast and robust to changes in the distribution network and thus can be easily used to compare multiple water distribution network designs for ones that minimize discolouration risk as a multi-objective function (McClymont et al., 2013).

2.4.9 Comparisons between Data-Driven and Physically Based Models

For physically based discolouration models, a calibrated hydraulic network model is usually required before additional discolouration specific aspects are modelled. Thus even for simple WDS networks, some initial time and capital expense to construct and maintain a hydraulic model is required.

Adapting an existing water quality engine, like PODDS did, can limit the applicability of the model to single pipe spans. The reason for this is that discolouration mobilisation is not limited to any one point in the WDS that then simply travels downstream. The mobilisation of material leading to discolouration can occur anywhere along the flow path and the actual quantity mobilised depends on a number of factors. These factors include the quantity of material available locally (sediments deposited or material available for

'peeling off' the pipe wall) and the local hydraulic forces, both of which change spatially and temporally throughout the WDS. As a consequence, continually modelling turbidity at some point in a WDS is more complex than just accounting for travel time at an upstream point in the system. This highlights why physically based models, such as the PODDS model, require multiple additional discolouration related parameters in addition to a calibrated hydraulic model.

A data-driven approach does not require a hydraulic model, only sufficient flow and turbidity observation data is needed. This enables the model to have the potential to be transferable to any WDS that has sufficient data available. As data-driven models are based on establishing the relationship between datasets, the data used in training a data-driven discolouration model will need to include a range of observed discolouration events for the model to forecast accurately across a range of discolouration events. Unfortunately, due to the difficulty in interpreting the internal workings of data-driven models, little new knowledge about the accumulation and mobilisation processes of discolouration is likely to be generated by using these models.

2.4.10 Summary

Water companies have a selection of physically driven methods and models to quantify discolouration risk in mains. Direct comparisons of these approaches to quantify discolouration risk with each other is not entirely possible as they have differing application domains.

RPM is a simple and easily implementable discolouration scoring method that requires no hydraulic model or substantial training to use. The trade-off however, is that the amount of information in the results it produces is equally as simple and limited in use.

The PODDS model has been substantially validated for selective pipe conditioning where short-term interventions are required on a particular pipe. PODDS can be used to make short-term turbidity predictions of how a single main will respond to an increase in hydraulic conditions. However, due to unknown pipe conditions and discolouration material build up rates, PODDS requires a hydraulic model with onsite model calibration before each use, making it unsuitable in the context of continuous (rather than individual event based) turbidity prediction.

The VCDM builds upon the PODDS model and incorporates a mechanism for modelling the accumulation of cohesive layers in a pipe over time. However, the VCDM requires

repeated site-specific turbidity events for model parameter calibration and a calibrated hydraulic model to track the turbidity response. This twin modelling constraint also increases the complexity and therefore potential for error when applied to operational applications.

DPM produces a relative discolouration risk scale for entire pipe networks to aid in the prioritization of pipe rehabilitation and replacement schemes. While DPM requires a hydraulic model like PODDS and VCDM, no onsite discolouration related calibration is required for the DPM to produce relative discolouration risks of every pipe in a network. However, it is worth noting that typical calibrations of hydraulic models are usually based on one off short-term data that may not capture the system's true states (Machell et al., 2010; Preis et al., 2010).

All these aforementioned approaches to quantify discolouration risk have been primarily physically driven and use explicit knowledge of the system being modelled to correctly capture the behaviour of the system.

2.5 Real-Time Modelling of Water Distribution Systems

2.5.1 Introduction

Water distribution network modelling has been a field of increasing activity in the past few years whereby medium and long-term retrofit and design plans are made with the assistance of WDS models (Husband et al., 2012; Keedwell and Khu, 2005; Kleiner et al., 1998; Savic and Walters, 1997). While these medium and long-term management plans are effective, ever stricter regulations by OFWAT push water companies to the limits of traditional techniques for managing issues like leakage, low pressure and discolouration (Giustolisi et al., 2008; Jowitt and Xu, 1990; Martinez et al., 1999; Su et al., 1987). New techniques and methodologies are required to improve service and meet the regulatory requirements effectively and efficiently.

Traditionally water companies only monitored Water Distribution Systems (WDS) and then reacted to changes, recently however, real-time modelling techniques are becoming more common place (Bougadis et al., 2005; He et al., 2014; Okeya et al., 2014; Romano et al., 2012). This is mainly due to better and cheaper technology where now even small Water Treatment Works (WTW) can afford Supervisory Control And Data Acquisition (SCADA) equipment. Real-time modelling techniques can take advantage of the

improvements in sensor technology and their increasingly wide spread use in water distribution systems in the UK and worldwide.

2.5.2 Real-Time Monitoring

Traditionally when a water company wants additional information about the current state of a WDS beyond the existing SCADA systems, such as to create and calibrate a hydraulic model of a WDS, then personnel are sent to install temporary monitors such as flow, pressure and turbidity meters (Gitelson et al., 1993; Glasgow et al., 2004; Herkelrath et al., 1991; Lambrou et al., 2014; Ong et al., 2008). These monitors only collect data for a set amount of time (typically days/weeks) before being recovered. This installation and collection process is usually time consuming and expensive. However, as communication technology and sensor reliability has improved, the use of permanent real-time monitors in WDSs is becoming more cost effective. Once installed, the monitor usually requires only infrequent cleaning and recalibration of the sensor and battery replacement if it does not use mains power.

Advances in solar power and radio transmitting technologies like General Packet Radio Service (GPRS) now allows for permanent monitor placement even in remote areas previously impossible. Additionally, data sampling and data collection rates have improved to where a data sample from the monitor to the control centre can be automatically transmitted, typically every 15/30 minutes (more frequently possible too, if deemed necessary) (O'Flynn et al., 2010; Sunkpho and Ootamakorn, 2011). The increased collection rate of these permanent monitors allows for the constant monitoring of important infrastructure points such as reservoirs and near real-time responses that previously may have gone unnoticed until a customer complaint (Chung et al., 2006; Dehua et al., 2012; Gutiérrez et al., 2014, 2014; Nasirudin et al., 2011).

While monitoring technologies like sensors have decreased in cost, they are still considered a non-trivial expense and many WDS will only have handful of permanent sensors, most of which are flow meters and pressure sensors. This means that optimal placement of these sensors is crucial and plays a role in determining a District Metered Area's (DMA's) boundaries (Alvisi and Franchini, 2014; Di Nardo and Di Natale, 2011; Diao et al., 2012; Gomes et al., 2012; Izquierdo et al., 2009; Ulanicki et al., 2008). A DMA is a sectored part of the pipe network where all water in and out of the DMA is measured and balanced in order to help water companies to estimate demand and leakage for that area.

A wide range of techniques have been proposed for automatically finding optimal locations for various types of sensors based on minimizing specific objectives (Aral et al., 2009; Berry et al., 2006; Dorini et al., 2008; Hart and Murray, 2010; Huang et al., 2008; Krause et al., 2008; Shastri and Diwekar, 2006). Dorini et al. (2010) presented a multi-objective method aimed at placing sensors to detect contaminant intrusion in WDS. Likewise, Weickgenannt et al. (2010) presented a method for optimised sensor placement to detect contaminant intrusion, but used a risk-based objective instead to minimize the number of people affected, rather than just minimising the time it takes to detect the contamination.

2.5.3 On-Line Modelling

While off-line hydraulic models have become more common in water companies to aid in operational investigations and network design, on-line hydraulic models have yet to be mass adopted in the water industry (Preis et al., 2010; Sanders, 2007; Storey et al., 2011). A model can be classed as 'on-line' when data is continuously fed to the model and internal parameters are updated that effect that model's outputs. An on-line model can not only adapt to the changes in the system over time (e.g. water demand), but it also allows for the model to incorporate new data that was not available when the model was originally calibrated. In contrast traditional off-line models, on-line models have been shown to improve performance over time, however this is usually at the expense of increased computational processing times (Hutton and Kapelan, 2015; Machell et al., 2014; Okeya et al., 2014). On-line modelling can enable improvements in associated technologies such as real-time control strategies, where the advantages of reduction in energy consumption and reduction of unnecessary water supplying can be realised (Bakker et al., 2012; Banna et al., 2014; Wei et al., 2004).

Instead of just displaying the correct and up-to-date information, on-line modelling of WDSs typically builds upon real-time monitoring systems in two ways. The first is by making estimations of unmeasured parameters and parts of the WDS, this is usually called State Estimation. The second is by predicting the future states of the measured and unmeasured data to a degree of certainty, this is usually called Forecasting.

State estimation is a process of combining measurement data and mathematical models to allow the calculation of variables of interest that are not directly measured (Bargiela and Hainsworth, 1989; Leendertse and Gritton, 1971; Moradkhani et al., 2005). In WDSs the measurement data is usually generated by flow, pressure and turbidity monitors, and

the model is a hydraulic network. State estimation is usually separated from a similar set of methods called parameter estimation by the type of variables being estimated. State estimation generally focuses on variables that can change with each time step (e.g. pressure, flow rates and water quality), while parameter estimation generally focuses on time invariant parameters that, within the scope of the model, are not usually expected to change (e.g. pipe roughness) (Arbués et al., 2003; Kang and Lansey, 2008)

Real-time state estimation combines real-time measurement data with a WDS model to aid operators in making time sensitive decisions that may otherwise affect customers. Two critical elements of state estimation are the provision of sufficient amounts of reliable data and the accuracy of the model that relates the known values to the unknown variables. If there is insufficient data or the data has too many inaccuracies then the model may perform equivalent to or possibly worse than a similarly calibrated offline model.

The forecasting techniques used for WDSs vary depending on application and time scale and have been extensively researched and compared, from using Artificial Neural Networks and Genetic Algorithms to Kalman Filters and Bayesian Inference (Bai et al., 2014; Bakker et al., 2012; Bougadis et al., 2005, 2005; Caiado, 2009; Donkor et al., 2012; Herrera et al., 2010; Kurunç et al., 2005). While no specific time frames have been commonly agreed upon, Billings and Jones (2011) define short-term forecasting as being between an hour and 3 months, medium-term forecasting as 3 months to 2 years and long-term forecasting as greater than 2 years. However, when talking about on-line modelling the time frame being forecast is typically less than 24 hours.

2.5.4 Real-Time Event Detection

Event detection methods are used to alert operators of abnormal (or failure) events in the system, such as a burst or equipment failure (Mounce et al., 2009; Mounce and Machell, 2006; Oliker and Ostfeld, 2015; Palau et al., 2011; Ye and Fenner, 2013). Currently most in use systems have simple, fixed threshold alarms that go off when some observed value goes above or below that threshold. While this is useful for detecting large events that effect the whole DMA, smaller events can go undetected. To detect these smaller events with a single value threshold, the threshold values need to be adjusted to be more sensitive to the network changes (Mounce et al., 2009). The trade off with this is an increase in false alarms that could be caused by simply low water consumption demand.

While there will always be a trade-off between false alarms and the non-detection of smaller events, this is less significant for calibrated offline systems that use multiple time

and state variant thresholds and can simulate the probable state of the system to allow early warnings. False alarms are even less significant for on-line event detection systems, as while able to do all that a calibrated offline system can do, on-line systems can also adjust these thresholds to more accurately model the WDS limits as it changes over time (Arad et al., 2013; Romano and Kapelan, 2014).

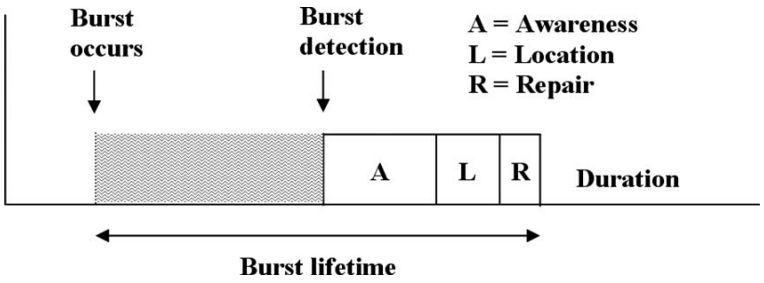


Figure 2.5. The life cycle of a pipe burst, taken from (Mounce et al., 2009).

Resolving bursts and leaks is a large issue that can cause large customer issues and affect a company's performance indicators set by OFWAT. Minimizing the time to detect and locate the burst or leak is important in reducing the time to resolving the event, consequently, a wide range of research about event detection in a WDS has been focused around bursts and leakage (Ahadi and Bakhtiar, 2010; Doumit and Lynch Jr, 2003; Fuchs and Riehle, 1991; Gao et al., 2005; Muggleton et al., 2002; Vítkovskỳ et al., 2000; Wu et al., 2009). For example, Palau et al. (2011) uses a multivariate statistical technique called Principal Component Analysis to analyse inflows and outflow of DMAs to determine the network's underlying behaviours, thus allowing for a change to be detected easily. Mounce et al. (2009) used the combination of an artificial neural network and a fuzzy inference system to detect abnormal flows in a WDS. Romano and Kapelan (2014) used the combination of an artificial neural network, a statistical process control model and a Bayesian inference system to infer the probability of a pipe burst occurring in a DMA. Even an artificial immune network has been used to detect and locate bursts in WDS (Tao et al., 2014).

While the majority of research on event detection is for bursts and leakage, there is also much research on contamination events and generic anomalies in WDS (Arad et al., 2013; Dorini et al., 2010; Shang et al., 2008).

2.6 Chapter Summary

While UK water companies have high drinking water standards, unfortunately decolorated water is one thing many of their customers would have experienced at some time or another (Cook, 2007; Vreeburg and Boxall, 2007). Due to the complex nature of discolouration in water distribution systems, it is difficult to model and predict.

Traditionally water companies deal with discolouration issues in a reactive way. It is only upon receiving a customer complaint that a water company will know of a discolouration event and then potentially take action. Although these customer complaints represent a measure of the discolouration issue in a network, the obvious disadvantage from a water company perspective is that this still requires customer complaints to be received. Once a customer complaint about a discolouration event is received, it is too late to mitigate the damage of that discolouration event.

One of the leading ways to reduce the risk of discolouration is through the use of discolouration models. Discolouration models have greatly helped in improving the accuracy, resolution and understanding of discolouration risk assessments. Insight gained from these models has allowed for a strategic approach in targeting sources of discolouration for regular cleaning or rehabilitation. Thus, allowing for informed and accurate financial investments where needed.

Currently these models are physically based and typically require specialised calibrations in order to accurately model discolouration. However, with the increasing fidelity and decreasing cost of real-time sensors for pressure, flow and turbidity, there are new opportunities for real-time data-driven models (Machell et al., 2010; Okeya et al., 2014).

Modelling discolouration in a WDS is a non-trivial task because the underlying interdependent processes of discolouration formation and mobilisation are still not fully understood. Add this to the aforementioned substantial increase of real-time data in WDSs and a data-driven approach that does not require explicit knowledge of the underlying processes of discolouration to quantify discolouration risk appears to not only be an emerging opportunity, but also presently be a gap in the current academic literature. A data-driven approach applied correctly could potentially allow for the real-time detection of discolouration events and enable proactive interventions that could mitigate the impact of a discolouration event.

Chapter 3 Case Study Data

3.1 Introduction

This chapter presents detailed information on the real trunk mains and their associated hydraulic and turbidity data used in this thesis. The data is presented before the subsequent chapters for the sake of conciseness as each of the following methodologies in the subsequent chapters use different aspects or subsets of the data presented here.

Section 3.2 provides details about the sites where data has been captured, such as hydraulic, network topography and meter placement. Section 3.3 describes the characteristics of the captured flow and turbidity observations. Section 3.4 defines what is considered a turbidity event here and provides information on the number and scale of turbidity events across the sites examined here.

3.2 Description of Sites

Flow and turbidity measurements were recorded over 2 years and 11 months from three hydraulically distinct parts (i.e. sites) of a real Water Resource Zone (WRZ) in the UK, starting from 1 September 2013 until 1 August 2016. A summary of the three sites is shown in Table 3.1.

Table 3.1. Summary of Sites.

Site	Total Length	Pipe Diameters	Inlets	Turbidity Meters (TM)	Flow Meters (FM)
Site 1	16.64 km	300 to 700 mm	1 Inlet	2 TM	7 FM
Site 2	6.45 km	450 mm	1 Inlet	1 TM	3 FM
Site 3	23.1 km	400 mm	2 Inlets	1 TM	4 FM

As it can be seen from Table 3.1, the three sites range from 6 km to 23 km in network length, 300 mm to 700 mm in pipe diameter size and are each primarily comprised of Ductile Iron (DI). While Site 1 has two turbidity meters, Sites 2 and 3 both have only a single turbidity meter.

Site 1 shown in Figure 3.1 is a trunk main network with one import and six exports. Aside from a flow meter placed directly after the upstream service reservoir, which is the sole inlet for the trunk main network, the other six flow meters were each placed at an exporting branch. Site 1 can be broken down into six distinct pipe lengths which will be denoted as pipes A, B, C, D, E and F. Pipes A, B, C and D are located upstream of Turbidity Meter (TM) A, and pipes A, B, E and F are located upstream of TM B. There is a small pipe that joins pipe B and pipe C but is excluded from being marked as a distinct pipe length due to its relatively insignificant length (< 10 m).

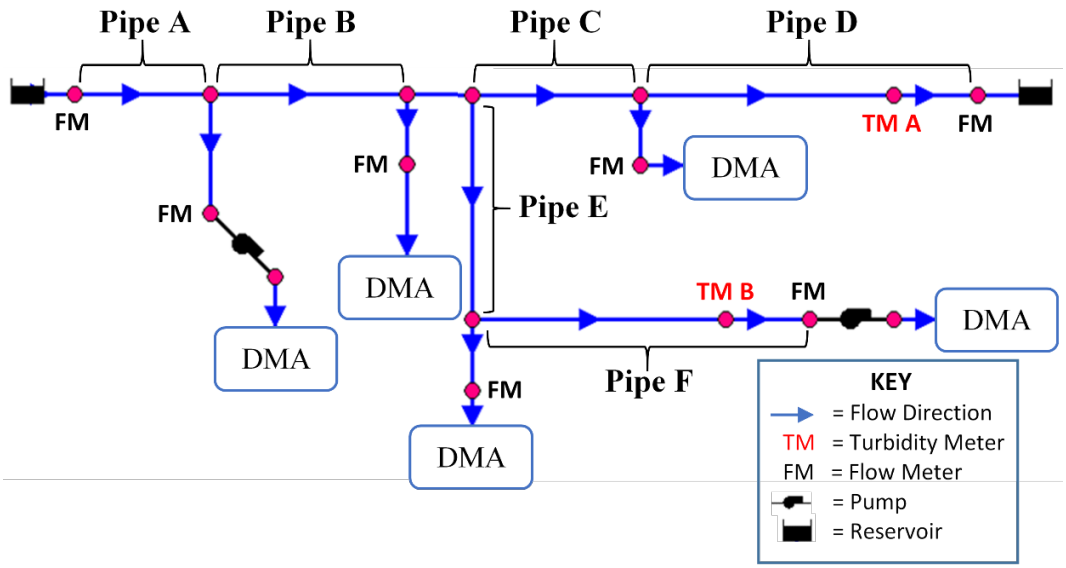


Figure 3.1. Schematic of Site 1. DMA, District Metered Area.

Site 2 shown in Figure 3.2 is a 6.5 km trunk main with one import and two exports. The flow rate in this trunk main is primarily controlled by two pumps at the downstream end of the main. Site 2 can be broken down into two distinct pipe lengths that are upstream of TM C and will be denoted as Pipes G and H. As the demand between pipes G and H is relatively small, there is only a minor variation in the hydraulic profile between the two pipes.

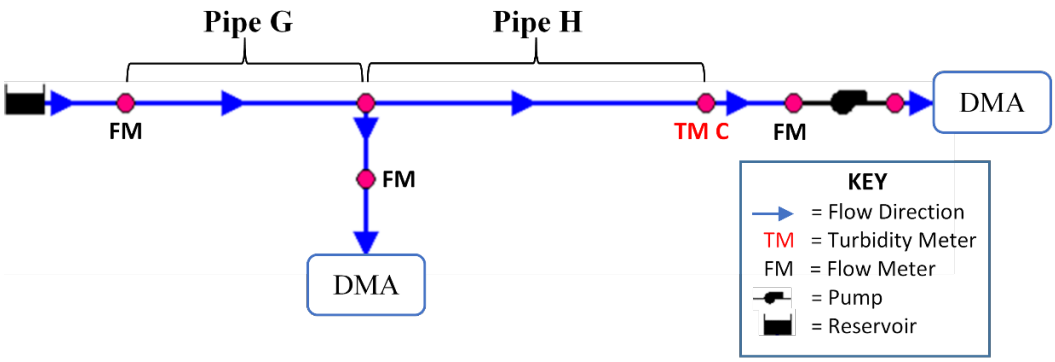


Figure 3.2. Schematic of Site 2. DMA, District Metered Area.

Unlike Sites 1 and 2, Site 3 shown in Figure 3.3, has two flow imports from two separate water sources. Site 3 can be broken down into three distinct pipe lengths that are upstream of TM D and will be denoted as Pipes I, J and K. As the water from the further downstream import, between Pipes J and K, is less expensive, only a small flow is typically seen across the almost 20 km combined length of Pipes I and J. However, when the downstream reservoir's water level is low, the upstream pumps are engaged to supply additional water.

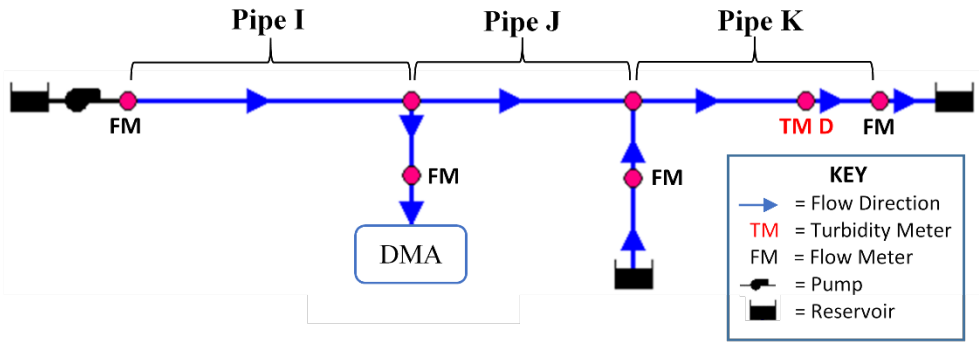


Figure 3.3. Schematic of Site 3. DMA, District Metered Area.

A summary of each pipe upstream of a turbidity meter is shown in Table 3.2.

Table 3.2. Pipe characteristics and the 99th velocity percentile over all observed data for each site.

Site	Pipe	Length	Diameter	Upstream of Turbidity Meter	99th Velocity Percentile
1	A	1.8 km	700 mm	TM A, TM B	0.94 m/s
1	B	1.6 km	700 mm	TM A, TM B	0.92 m/s
1	C	1.9 km	600 mm	TM A	0.17 m/s
1	D	5.1 km	300 mm	TM A	0.65 m/s
1	E	1.8 km	400 mm	TM B	0.86 m/s
1	F	4.4 km	400 mm	TM B	0.80 m/s
2	G	1.9 km	450 mm	TM C	0.83 m/s
2	H	4.6 km	450 mm	TM C	0.80 m/s
3	I	11 km	400 mm	TM D	0.79 m/s
3	J	8.5 km	400 mm	TM D	0.70 m/s
3	K	3.6 km	400 mm	TM D	0.73 m/s

The velocities for each pipe is derived from the measured flow meter data with the 99th velocity percentile for each pipe being shown in Table 3.2. The 99th velocity percentile is of interest as it shows the highest expected velocity in a pipe which if significantly low enough is indicative of a higher chance of material accumulation in the pipe. The low 99th velocity percentile of Pipe C thus indicates Pipe C has a high potential for discolouration material build up.

Aside from the aforementioned, little information about the operations of these sites was provided. It was claimed by the water company managing the above sites that except for a few insignificantly small water consumptions taken directly off some trunk mains, every inlet and outlet of each site was hydraulically metered by a flow meter. However on examination of the sites' hydraulic data, some significant discrepancies appeared and pulled into question the reliability of the hydraulic data. Of the three sites presented above, Site 3 had the most the most reliable hydraulic data, record of bursts and list of planned and unplanned maintenance work. Figure 3.4 shows the daily mass-balances for 16 months of hydraulic data for Site 3.

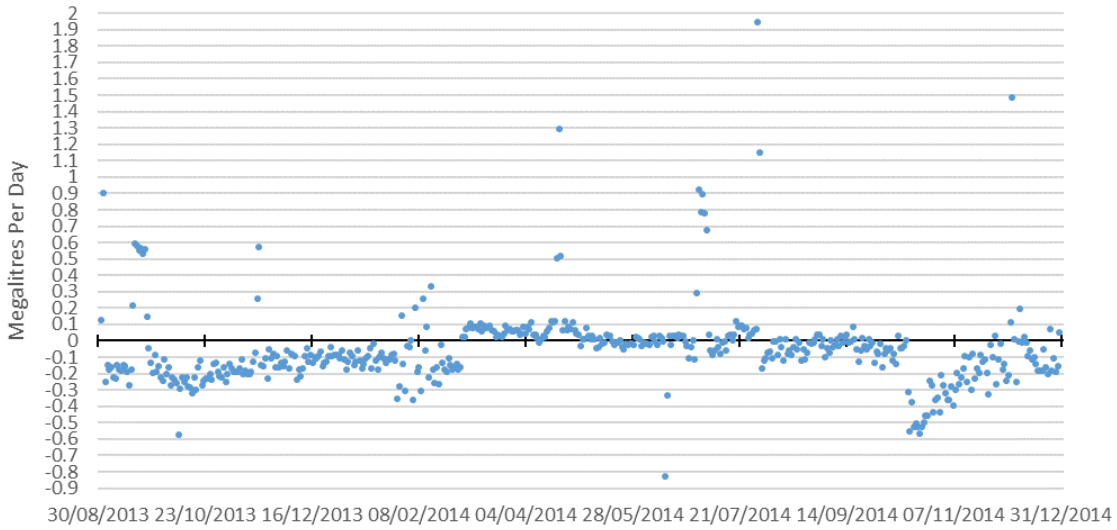


Figure 3.4. The daily mass-balances for 16 months of Site 3.

The significant positive and negative fluctuations in Figure 3.4 were cross-checked against the record of bursts and list of planned and unplanned maintenance work, however this only explained a very small portion of the variance. Further inquiries were raise with the managing water company, however no further information explaining this was received.

3.3 Flow and Turbidity Data

All turbidity meters were placed at the downstream end of each site and just upstream of a flow meter so that each turbidity measurement has an associated flow measurement. All flow and turbidity observations were logged at 15 minute intervals. Flow meters recorded the cumulative volume of water passing through the meter per 15 minute interval (i.e. m^3/15 min), and turbidity meters recorded the current turbidity value at the interval in NTU. All turbidity observations exceeding 10 NTU were caped at 10 NTU by the turbidity meters.

While all turbidity meters captured data over the same time period, TM D in Site 3 was offline for a total of five months, from July 2014 to November 2014 and then from June 2016 onwards. This is a factor in why TM D has fewer turbidity observations (including those over 1 NTU) than the other turbidity meters. A summary of the turbidity data for each turbidity meter is shown in Table 3.3.

Table 3.3. Summary of turbidity observations for each site.

Turbidity Meter	Duration Monitored	50th percentile (NTU)	99th percentile (NTU)	Observations > 1 NTU
TM A (Site 1)	2 years, 11 months	0.08	0.41	265
TM B (Site 1)	2 years, 11 months	0.07	0.42	328
TM C (Site 2)	2 years, 11 months	0.07	0.36	290
TM D (Site 3)	2 years, 6 months	0.08	0.46	204

Each turbidity meter was installed near a flow meter at a service reservoir, allowing for a constant power supply, connection to the SCADA system and ease of access for regular maintenance. All turbidity meters were serviced on a monthly basis and as a result, no drift was seen in the recorded turbidity data.

3.4 Turbidity Events

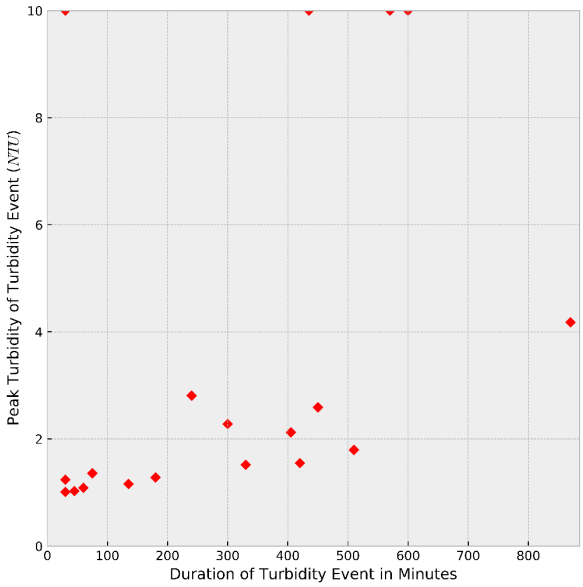
A turbidity event is defined here as occurring when the turbidity level exceeds 1 NTU, continues for at least one timestep (i.e. 15 minutes) and ends when the turbidity level subsequently drops below 1 NTU. If a new turbidity event occurs within four hours of the preceding turbidity event ending, then the preceding event, new event and all observations between the two events are all relabelled as a single turbidity event. The turbidity threshold of 1 NTU was chosen as it is a clear measurable response above typical background turbidity levels of less than 0.2 NTU and is also the UK regulatory limit for water leaving water treatment works (DWI, 2014a). This means it is unlikely for turbidity events originating from upstream treatment works to be a significant factor in the total number of turbidity events observed. Table 3.4 shows the number of turbidity events that were recorded by each turbidity meter along with the mean and standard deviation of the duration of those events.

Table 3.4. The number of turbidity events and mean and standard deviation of the duration of turbidity events recorded by each turbidity meter.

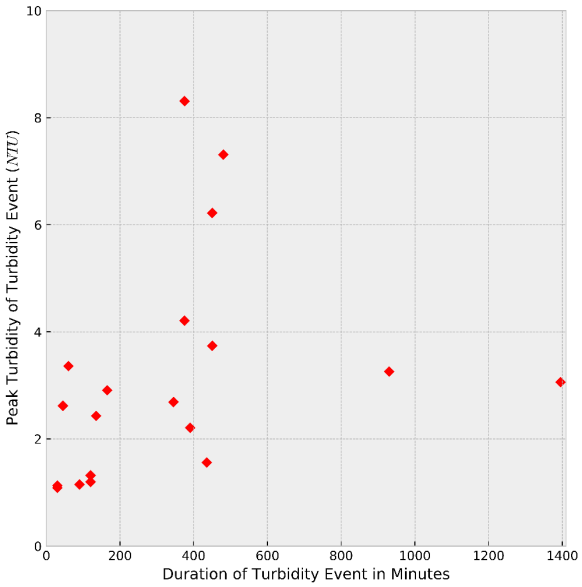
Turbidity Meter	Turbidity Events	Mean Duration of Events	Standard Deviation
TM A (Site 1)	19	4 Hours, 2 Min	2 Hours, 32 Min
TM B (Site 1)	19	2 Hours, 41 Min	1 Hours, 24 Min
TM C (Site 2)	25	4 Hours, 4 Min	2 Hours, 45 Min
TM D (Site 3)	28	7 Hour, 54 Min	4 Hours, 4 Min

From the number of turbidity events and the mean and standard deviations of event durations shown in Table 3.4, it can be seen that not only do events significantly vary between each site, but they can also significantly vary even within the same site.

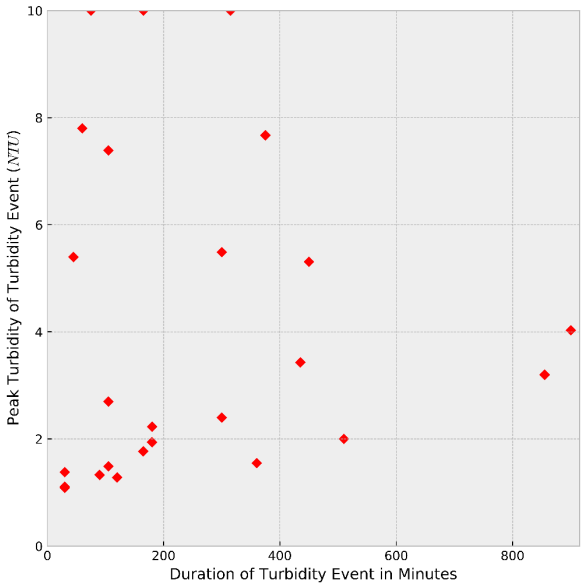
A finer breakdown is given by Figure 3.5, which shows the peak turbidity value and duration of all turbidity events for each turbidity meter. The highest turbidity observations are 10 NTU because as mentioned above, 10 NTU was the upper limit that the turbidity meters were set to record.



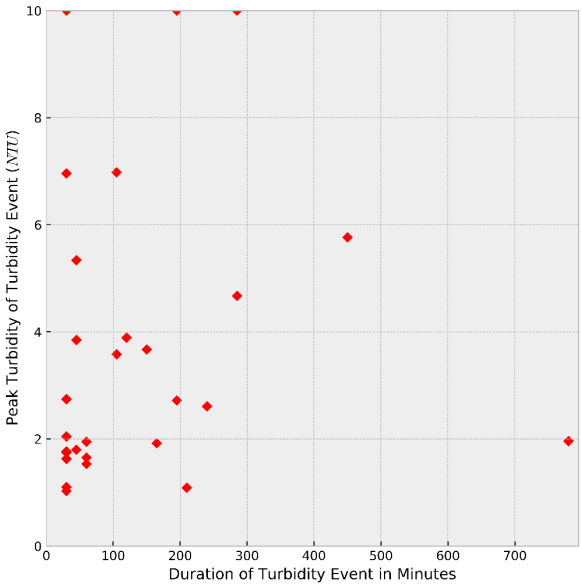
(a) Turbidity Meter A (Site 1)



(b) Turbidity Meter B (Site 1)



(c) Turbidity Meter C (Site 2)



(d) Turbidity Meter D (Site 3)

Figure 3.5. The peak turbidity value and duration of each turbidity event for each of the four turbidity meters. NTU: Nephelometric Turbidity Units.

In order to easily refer to a specific subset of events, the prefix of an NTU value to an event (ex. 4 NTU turbidity event) means that the peak turbidity value of that event exceeds the prefix NTU value. For example, if the peak value of a turbidity event is 7.5 NTU, then this event can be referred to as a 2 NTU turbidity event, a 7 NTU turbidity event but not an 8 NTU turbidity event. This means that the number of higher NTU events will always be either equal to or less than lower NTU events. Note that by the definitions given here, the set of turbidity events is identical to the set of 1 NTU turbidity events.

Chapter 4 Short-Term Turbidity Forecasting

4.1 Introduction

Given the significant amounts of on-line data being readily available in Water Distribution Systems (WDSs), a continuous turbidity forecasting methodology is developed to detect discolouration mobilisation in trunk mains and when that discolouration would reach the downstream District Metered Area (DMA) or service reservoir.

This methodology makes use of machine learning methods and does not require a hydraulic or any other type of physically based simulation model. This makes the methodology free from the costs associated with developing and maintaining calibrated hydraulic models and ensures application to almost any WDS that has suitable meters installed. A comparison of the methodology using different machine learning methods is made and forecasts up to 8 hours ahead are assessed.

After this introduction, Section 4.2 presents regression and classification approaches to forecasting turbidity along with the associated model inputs and performance metrics. Section 4.2 also presents a range of machine learning methods to be examined. Section 4.3 shows what data was used to train and test the models in this chapter. Section 4.4 reports the test results of the methodology on a real trunk main network. The test results and their implications are discussed in Section 4.5. Finally, a summary of the chapter and its main conclusions are presented in Section 4.6.

4.2 Short-Term Turbidity Forecasting Methodology

4.2.1 Overview

There are three significantly complex processes that need to be captured by the machine learning methods presented later in this section in order for the model to correctly forecast turbidity at the downstream turbidity meter.

The first process required is that the model must detect if any hydraulic force capable of mobilising discolouration material has occurred. This is a significant challenge to the model because while there are numerous studies showing that velocities significantly

above the daily conditioning values are able to mobilise discolouration material (Husband et al., 2010; Vreeburg and Boxall, 2007), the scale and extent to which typical velocities cause naturally occurring discolouration events has not been investigated. Additionally, when looking at the model results it is important to take into consideration that the model only uses velocity and turbidity measurements as inputs, thus any turbidity mobilised by other factors like temperature or bio-film shedding may not be detected.

Linked to the first process, the second required process of the model is to estimate how much turbidity the mobilisation of discolouration material has resulted in. This second process is based on the concept that if there was a significant discolouration event recently then there may be less discolouration material that can be mobilised presently. This process is especially important if the goal is to predict the exact turbidity value.

Finally, the third process also linked to the first process, is estimating where the discolouration material was mobilised from (locally and/or upstream) and then the amount of time it will take for that material to reach the downstream turbidity meter. This is important because after the discolouration material has been mobilised it can still require many hours to reach the downstream turbidity meter. Determining which general section of the network the discolouration material is being mobilised from is a complex and difficult problem (Blokker et al., 2011), however the model also has to predict the expected travel time of the discolouration material. This can mean that the model could accurately detect and estimate the size of a discolouration event but not correctly forecast when that increased turbidity will arrive at the downstream meter, resulting in a false alarm and poor overall model performance.

4.2.2 Turbidity Forecasting Models

In order to address the significant complexities required by the model, two separate modelling approaches were taken; a regression approach and a classification approach. The regression approach requires the model to predict the actual turbidity measurement at a specified period of time in the future. The classification approach requires the model to predict if turbidity will be above or below a preselected turbidity threshold at a specific period of time in the future.

For the classification approach, all the turbidity measurements are placed into one of two classes, Positive or Negative. If the turbidity measurement is above the preselected threshold then it is labelled as being in the positive class, conversely if it is below the threshold then it is labelled as part of the negative class. The classification model then

only predicts if the turbidity at a prespecified future time is positive or negative and not what the actual turbidity will be.

Once the model has made its prediction of positive or negative, those predictions are further divided into True or False depending if the prediction made was correct or not. A True Positive (TP) and True Negative (TN) are predictions made by the model that were correct for their respective classes, while conversely a False Positive (FP) and False Negative (FN) were incorrect predictions. This is illustrated in Figure 4.1.

	Predicted Event	Predicted No Event
Actual Event	True Positive (TP)	False Negative (FN)
Actual No Event	False Positive (FP)	True Negative (TN)

Figure 4.1. A confusion matrix diagram used to describe the performance of classification based models.

Three classification based models with different preselected turbidity threshold values of 1 NTU, 2 NTU and 4 NTU are trained and tested. The 1 NTU threshold was chosen as it is a clear measurable response above the background turbidity levels and is the regulatory limit for water leaving water treatment works. The 2 NTU threshold was also chosen as it is indicative of a more serious discolouration event soon occurring, and finally the 4 NTU threshold was chosen as it is the UK regulatory turbidity limit at customers' taps. Together these models could be seen as a three-level warning system where operational staff could decide at what level they wish to take action.

4.2.3 Model Inputs

The data presented as inputs to the above models is measurement data from the flow and turbidity meters. Recent historical data of turbidity and velocity is required for the model to make accurate predictions about the amount of discolouration material that has been mobilised and how long it will take that material to reach the downstream turbidity meter. However, the more inputs given to the model makes the model more likely to overfit on the training data. Thus, only a limited number of lagged meter measurements are given to the model as inputs at any one time. This is what is known as time-delay embedding

or the sliding window method (Dietterich, 2002; Gershenfeld et al., 1993), as shown in Figure 4.2.

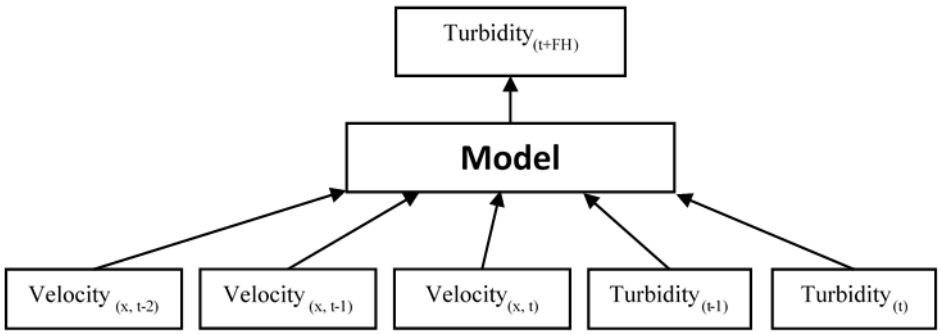


Figure 4.2. A simplified diagram of where current and previous velocity and turbidity measurements are used as inputs to a model predicting the turbidity value at a future time step. In the diagram "t" denotes the current time step, "x" denotes the flow meter and "FH" is the forecast horizon (i.e. how far into the future the predictions are made).

When choosing the amount of previous measurements to present to a model, also known as the size of the sliding window, it is important to select the optimum size. A window size too small will not provide the models with enough information to accurately reproduce the modelled system's dynamics and thus resulting in poor prediction performance. Conversely, a window size too large can result in the model overfitting on the training data and requiring increased training times (Abarbanel et al., 1993). The False Nearest Neighbours (FNN) algorithm has been shown to be relatively accurate in approximating the correct embedding dimension of dynamic systems and ANNs (Frank et al., 2001; Rhodes and Morari, 1997). The FNN algorithm with minor adjustments was used here to choose the optimal window size for each meter.

Previous work showed that providing the peak velocity and turbidity values for each meter in the previous days and weeks as separate inputs improves model accuracy with only a relatively minor increase in model complexity (Meyers et al., 2016). Thus, in addition to the time lagged meter measurements, three other types of velocity inputs were presented to the models. These inputs are functions of current and previous meter readings and are named the *peak x day velocity,* the *peak velocity difference* and the *peak turbidity level*. The peak x day velocity is the maximum measurement seen by a specific meter in the last x days. This is in order to allow the models to determine if current velocities exceed the thresholds of pipe conditioning levels or if a recent high velocity has flushed

away any potential discolouration material. The peak velocity difference is the difference between the current velocity and the daily peak velocity. While similar to the peak x day velocity input that acts as a static threshold measurement that is unlikely to change for many days at a time, this input acts as a dynamic threshold measurement that also provides information to the model about how long ago the discolouration material was likely to have been mobilised. This should aid the models in determining when the discolouration material will reach the downstream turbidity meter. The peak turbidity level aids the models in determining if recent hydrualic events have removed discolouration material or if there is significant potential material build up that could be mobilised.

A total of 95 inputs was presented to each model which comprised of; a) previous velocity and turbidity measurements for each meter used in the turbidity prediction models with separate window sizes, ranging from three to eight timesteps of 15 minutes; b) daily and weekly versions of the peak x day velocity; c) daily and weekly versions of the peak velocity difference; d) daily and weekly versions of the peak turbidity level inputs.

4.2.4 Model Performance Metrics

Performance metrics are an essential part of evaluating and comparing forecasting models and while a visual inspection of a graphical plot of a model's forecast is recommended (Green and Stephenson, 1986; Martinec and Rango, 1989), this can become impractical when comparing and/or presenting multiple models with different parameters and over multiple forecast horizons. As performance metrics inherently tend to have a bias toward particular aspects of one model over another, multiple performance metrics have been used here (ASCE, 1993).

4.2.4.1 Regression Model Metrics

The *Nash-Sutcliffe Model Efficiency Coefficient (NSE)* and *Naïve Relative Squared Error (NRSE)* metrics were chosen to assess the performance of the regression based models. These were chosen because they provide a numerical value of performance and useful information on how well the forecasts fitted with the observed data (Armstrong, 2001; Hyndman and Koehler, 2006).

The Nash-Sutcliffe Model Efficiency Coefficient is widely adopted by the Hydroinformatics community in measuring model performance. NSE indicates the model performance relative to the mean of all observational data as a forecast at every time step. In a similar way to NSE, the Naïve Relative Squared Error (NRSE) indicates performance

relative to the naïve method. The naïve method forecasts each time step is the same value as the last observed value - this is the optimal forecast method if there is no additional information on what the next observation will be (Gershenfeld et al., 1993). The formulas for NSE and NRSE are as follows:

$$NSE = 1 - \frac{\sum_t (F_t - A_t)^2}{\sum_t (M - A_t)^2} \quad \text{(Eq. 4.1)}$$

$$1 - NRSE = 1 - \frac{\sum_{t=1} (F_t - A_t)^2}{\sum_{t=2} (A_t - A_{t-1})^2} \quad \text{(Eq. 4.2)}$$

where F_t is the forecast at time step t, A_t is the actual value observed at time step t and M is the mean of all observational data. Note that NRSE is subtracted from one so that a value of 1 indicates perfect model performance for both NSE and NRSE metrics.

4.2.4.2 Classification Model Metrics

Due to the infrequent nature of discolouration where little or no discolouration can occur for multiple weeks at a time, less than 1% of the logged data points are above 1 NTU. Because of this significantly disproportionate number of negatives (i.e. non-discoloration events), most error metrics that use the number of negatives as a factor is likely to be misleading. An example of this would be the *Accuracy* metric that is the percentage of observations that are correctly classified. Using this measure, a model that simply always predicted negative would have greater than 99% accuracy even though the model predicts no discolouration events. For this reason, the Accuracy metric was not used. In addition, metrics such as the *False Positive Rate (FPR)* and *Receiver Operator Characteristic Area Under the Curve (ROC AUC)* were avoided for the same reason. The *Matthews Correlation Coefficient* (MCC) was used instead as it is regarded as a good measure of performance even when there is an extreme skew in class sizes (Baldi et al., 2000).

$$MCC = \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad \text{(Eq. 4.3)}$$

While MCC may give a performance value that can be compared with other models, just like the Root Mean Square Error metric, it is difficult to determine at what performance value a model is considered useful in real world application. Given this, the True Positive Rate and False Detection Rate metrics were additionally used here.

The True Positive Rate (TPR), also known as the probability of detection, is the probability that the model will correctly predict positive class values (i.e. turbidity above pre-specified threshold). For example, a TPR of 0.8 for a model means that of the positive class values, the model will detect 80% and miss 20% of actual events.

$$TPR = \frac{\mathrm{TP}}{TP + FN} \quad \text{(Eq. 4.4)}$$

The False Discovery Rate (FDR) is the probability that the model predicts a positive (i.e. a discolouration event, as defined above) when in reality no such event occurred (also known as a false alarm). It is important to keep FDR low to ensure that operational staff maintain confidence in the system when an alarm sounds.

$$FDR = \frac{\mathrm{FP}}{FP + TP} \quad \text{(Eq. 4.5)}$$

4.2.5 Machine Learning Methods

Building the turbidity prediction models (both regression and classification based) requires the use of a machine learning method. Three mainstream machine learning methods tested here are an Artificial Neural Network (ANN), a Random Forest (RF) and a Support Vector Machine (SVM) (Breiman, 2001; Gevrey et al., 2003; Ghiassi et al., 2008; Glorot and Bengio, 2010; Govindaraju and Rao, 2013; Mounce and Machell, 2006; Smola and Schölkopf, 2004; Vicente et al., 2012). These methods were chosen based on three factors: (a) a significant, proven ability to generalise complex non-linear relationships between potentially noisy inputs and outputs; (b) each method uses a significantly different learning mechanism to the other two methods; (c) that a regression and classification version of the method exists.

ANNs were developed based on models of human brain function and are able to solve complex problems. The ANN used here was a feed-forward multilayer perceptron ANN with hidden layers using the hyperbolic tangent sigmoid transfer function and the output layer employing a linear transfer function. The ANN was trained with a stochastic gradient-based optimizer (Kingma and Ba, 2014). The number of hidden layers and the number of neurons in each hidden layer was set via a hyperparameter tuning process that

utilised 5-fold cross-validation on the training dataset only (Golub et al., 1979; Kohavi, 1995; Silverman, 1984).

RFs can essentially be considered an ensemble of decision trees where each individual decision tree is a weak classifier but where good performance, scalability and generalisation can be achieved through the combination of numerous trees into an ensemble via the 'bootstrapping' of data (Archer and Kimes, 2008; Pal, 2005; Rodriguez-Galiano et al., 2012; Svetnik et al., 2003). The RF used here was the Extra Trees RF variant with 1000 weak classifiers and the entropy splitting criterion derived empirically. The Extra Trees RF variant adds an additional layer of randomness to the model by adding a number of randomized decision trees (a.k.a. extra-trees) to further decrease the chance of overfitting (Geurts et al., 2006). The percentage of features to consider when looking for the best tree split and the maximum depth of trees was set via hyperparameter tuning using 5-fold cross-validation on the training dataset only.

SVMs aim to find the maximum margin between decision boundaries in training data with the aim to be less prone to overfitting and thus have a lower generalisation error when tested (Hearst et al., 1998; Khan and Coulibaly, 2006; Tong and Chang, 2001; Widodo and Yang, 2007). The kernel, gamma and the regularisation penalty used by the SVM was set via hyperparameter tuning using 5-fold cross-validation on the training dataset only.

4.3 Experimental Setup

4.3.1 Data

Only a subset of the data shown in Chapter 3 was initially used here. This is because of the vast amount of data presented in Chapter 3, the cause of the majority of turbidity events in the subset of data chosen here is better understood by the water utility. This enabled better validation and clearer analysis of the model.

The velocity and turbidity measurements used here is from Site 1 (shown in Figure 3.1) and was taken over 11 months, starting from 01 September 2013 to 01 August 2014. The hydraulic data was taken from seven flow meters, the one import flow meter and six export flow meters. Turbidity data was taken from a single turbidity meter, TM A.

As each meter logged once every 15 minutes, the 11 months of measurement data resulted in 32,160 15 minute timesteps with seven velocity and one turbidity measurements per

timestep. The first seven months (01 September 2013 to 31 March 2014) of data was used for training and the following four months (01 April 2014 to 01 August 2014) of data was used for testing. This split resulted in 20,353 timesteps in the training dataset and 11,807 timesteps in the test dataset.

Velocity and turbidity measurements were standardised based on the mean and standard deviation calculated from the training dataset only.

4.3.2 Model Setup

The results of the hyperparameter tuning using 5-fold cross-validation for each model lead to optimal parameter settings as follows: (a) the a single hidden layer with 60 neurons in the hidden layer for ANN classification and regression models; (b) 10% of features to be considered at each tree split and a maximum depth of each tree limited to 2 for the RF classification and regression models; (c) a radial basis function (RBF) kernel with a gamma of 0.01 and 0.1 for the regularisation penalty for the SVM regression models; (d) a RBF kernel with a gamma of 0.001 and 1.25 for the regularisation penalty for the SVM classification models.

Each machine learning method required less than five hours to find its optimal parameter settings on an Intel Core i7-5600U CPU and 16GB of RAM. Once trained, each machine learning method was able to produce forecasts for the four months of unseen test data in less than five seconds. This indicates that regular retraining to include new data is feasible.

4.4 Results

All results and figures shown below have been calculated from the unseen test datasets by using models trained on seen (i.e. training) datasets.

4.4.1 Regression Based Turbidity Forecasting Models

Figure 4.3 shows the performances of each of the regression based turbidity machine learning prediction models for different forecasting horizons up to 2 hours. While it is not clear from this figure if the ANN or RF model performs the best overall, it is clear that they both outperform the SVM model. The SVM model's poor performance indicates the SVM was unable to generalise well on the unseen (i.e. test) data. The ANN and RF model perform well initially before rapidly degrading in performance as the forecast horizon

increases. For example, the models exhibit questionable reliability when approaching the 20-35 minute forecast horizon. Forecasts made after 50 minute forecast horizon are only slightly better than forecasting the mean of the measurement.

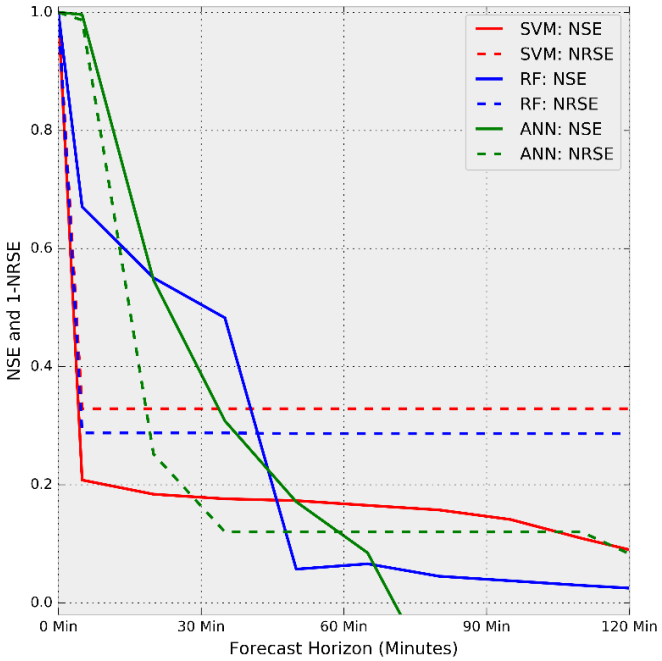


Figure 4.3. Nash-Sutcliffe Model Efficiency (NSE) and the 1-Naïve Relative Squared Error (1-NRSE) performance metrics (lower values indicate worse model performance in both cases).

Figure 4.4 shows the forecasts of the regression based ANN turbidity prediction model forecasting 20 minutes ahead against the actual turbidity measurements of a typical event. While the model can be seen to accurately forecast the start and shape of the turbidity event, it under-predicts the magnitude of the event. This is the typical behaviour of the ANN model across all events.

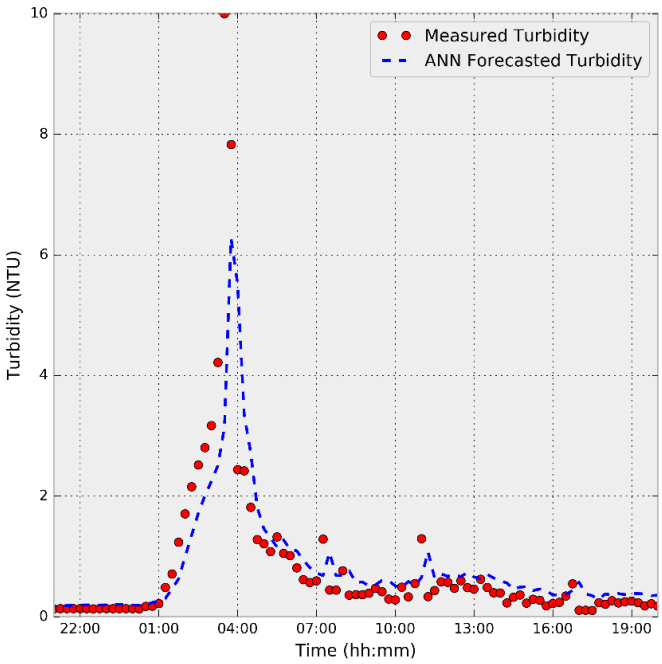


Figure 4.4. The regression based ANN turbidity prediction model forecasting 20 minutes ahead over a typical event in the test dataset.

The regression based ANN and RF turbidity prediction models have potential to aid operational staff in the immediate future with relatively accurate forecasts up to approximately 20 minutes, however that is unlikely to be sufficient lead time for operational staff to act upon in most cases.

4.4.2 Classification Based Turbidity Forecasting Models

Figure 4.5 shows the mean Matthew's Correlation Coefficient (MCC) calculated from the classification based turbidity prediction models for each of the three machine learning methods over a range of forecast horizons up to 8 hours ahead. As it can be seen from this figure, the ANN and RF based models clearly outperform the SVM based model. While the ANN based model initially outperforms the RF based model, it quickly becomes worse performing than the RF based model as the forecast horizon increases.

The sharp drop in the mean MCC value seen from forecast horizons less than 30 minutes across all models is primarily due to an increasing number of false positives which have a significantly stronger influence on the MCC because of the large class skew (significantly fewer actual event measurements limits the number of possible true positives but not false positives). The particularly poor performance of the SVM based model indicates that the SVM was unable to generalise well and produced a greater

number of false predictions as a result. From the results in Figure 4.5 it is determined that of the models presented here, the RF based model has the best performance overall.

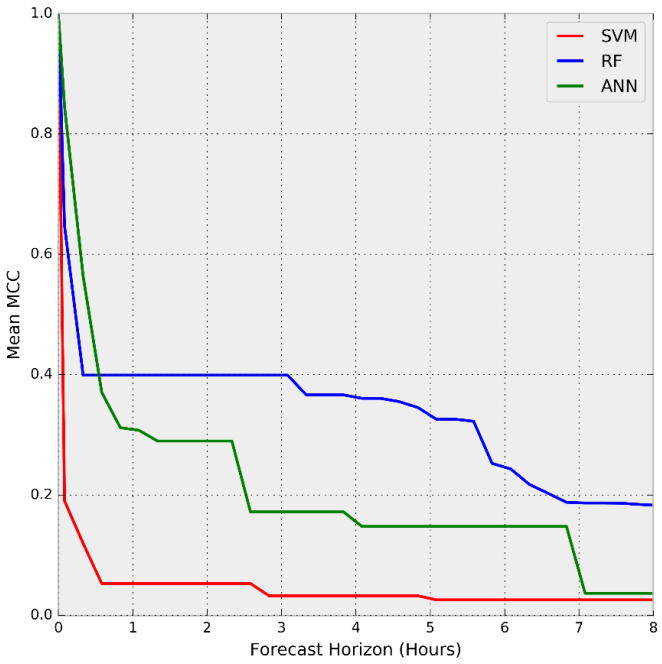


Figure 4.5. The mean Matthews correlation coefficient (MCC) is shown for each of the classification based turbidity prediction models over an 8 hour forecast horizon. A lower value indicates worse performance.

Based on these results a closer examination of the RF based turbidity model is shown in Figure 4.6. This figure shows the values of True Positive Rates (TPRs) and False Detection Rates (FDRs) for the RF model at the three different threshold values of 1 NTU, 2 NTU and 4 NTU and for different forecasting horizons ranging from 5 minutes to 8 hours. From Figure 4.6 it can be observed that it is possible to obtain high TPR values (i.e. values above 0.9) even for longer forecast horizons of over 5 hours. The TPR values are high initially for all models however the 2 and 4 NTU threshold models quickly drop to a mid-region values (0.3 to 0.45) for lead times around 30 minutes. TPRs for the 2 and 4 NTU threshold models then again increase to high values of above 0.9 for lead times of over 5 hours before the 2 NTU threshold model once again drops. The increases in each model's TPR are explained by the model's sensitivity and trade-off between detecting real turbidity events (i.e. the TPR) and raising false alarms (i.e. the FPR). This is because each forecast horizon of the classification based model was trained with more weight placed on keeping the FPR low, and thus also keeping the FDR low, than was placed on

keeping the TPR high. This extra weight was removed if the TPR drops to 0.3 or lower during training.

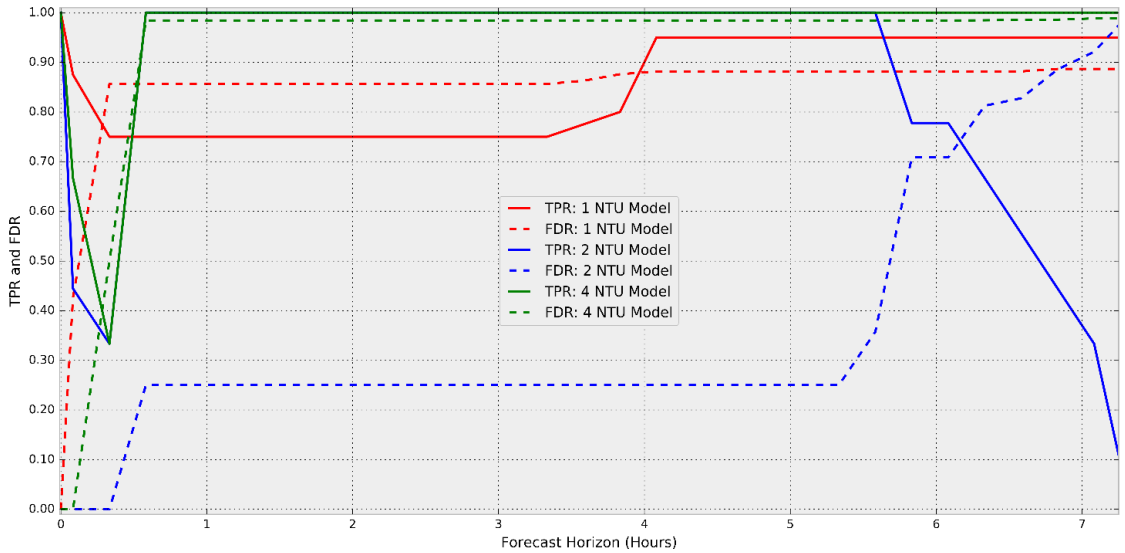
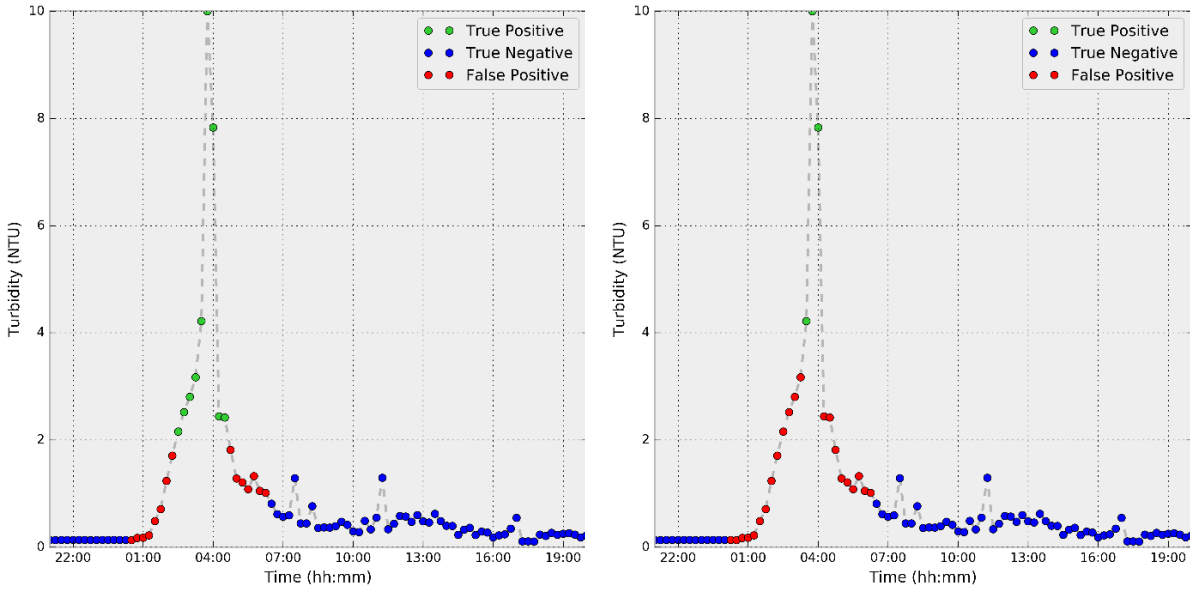


Figure 4.6. The performances of the 1, 2, 4 NTU threshold classification based models across a range of forecasting horizons. The True Positive Rate (TPR) shows how well a model can detect real events and the False Detection Rate (FDR) gives the likelihood of a real event that has been detected turning out to be a false positive. Perfect results would show a TPR=1 and FDR=0.

The TPR and FDR of the 1 NTU threshold model increases at approximately 3 hours and 30 minutes because during training the TPR would have dropped below 0.3, thus the model made the trade-off of being more sensitive and accepting more false positives in order to keep the TPR higher. This is also why the FDR for all models in Figure 4.6 is seen to only monotonically increase.

It can also be observed from Figure 4.6 that unlike the 2 and 4 NTU threshold prediction models, the 1 NTU threshold prediction model never reaches a TPR of 1. A closer examination of the model showed that this is primarily because the 1 NTU threshold prediction model struggled to forecast the many sudden transient turbidity spikes that slightly exceed 1 NTU for a single time step. Some of these are accurately forecasted at very short lead times (less than 20 minutes) by autocorrelation with previous turbidity but are quickly lost when forecasting further into the future. However, events of this magnitude and duration are highly unlikely to be detected by customers. Furthermore, their short duration means that it is not clear what remedial action an operator could take to mitigate this.

Note that the FDRs shown in Figure 4.6 can be misleadingly high as the majority of FPs that make up the FDRs are not randomly spread out, but instead clustered around actual turbidity events. This is shown in Figure 4.7 where the 2 and 4 NTU threshold RF turbidity prediction models are shown forecasting 5 hours and 20 minutes ahead. As it can be seen from this figure, both models correctly predict that turbidity will go above the corresponding thresholds hours before any rise in turbidity is observed. This clearly shows that the model is able to detect the mobilisation of discolouration material via flow meter measurements alone. However, the high number of FPs that surround the event show that the models struggle to predict exactly when the discolouration material would reach the downstream turbidity meter. Because the models detected the mobilisation of discolouration material but struggled to estimate its travel time to the downstream turbidity meter, the models learned to predict positive for all values close to when the discolouration material was estimated to arrive.



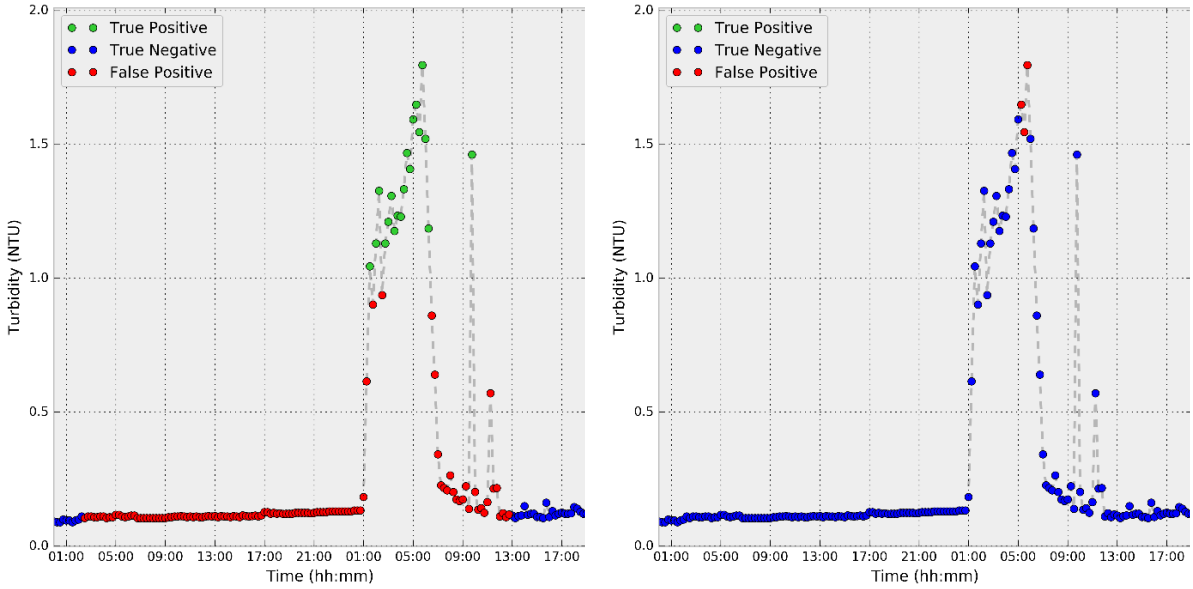
(a) 2 NTU Threshold Prediction Model (b) 4 NTU Threshold Prediction Model

Figure 4.7. The 2 and 4 NTU threshold RF turbidity models forecasting 5 hours and 20 minutes ahead over a typical turbidity event.

While this behaviour ensures that TPs are not missed and enables the models to have high TPRs for long lead times, it also means many FPs are generated in the process which raises the FDRs. For the 2 and 4 NTU threshold RF turbidity models forecasting 5 hours and 20 minutes ahead, the FPs that occur 6 hours before and after all actual events account for 79% and 88% of FPs respectfully.

It is worth noting that as the FDR is normalised by the number of true and false positives, the FDR for the 4 NTU threshold prediction model was expected to be higher than the other models because there are fewer 4 NTU or greater observations and thus fewer positives for the model can accurately predict. This is shown to some extent in Figure 4.7 where the unequal number of TPs per threshold prediction model results in this individual event having an FDR of 63% for the 2 NTU threshold prediction model but an FDR of 88% for the 4 NTU threshold prediction model.

While the 2 and 4 NTU threshold prediction models may be more appealing to detecting larger impact discolouration events, the high sensitivity required by the 1 NTU threshold prediction model should not be overlooked. Figure 4.8 shows the 1 and 2 NTU threshold RF turbidity prediction model forecasting over a low NTU event 5 hours and 20 minutes ahead. The event seen in Figure 4.8 is a non-typical turbidity event caused by trunk mains cleaning in an upstream branch of the network.



(a) 1 NTU Threshold Prediction Model (b) 2 NTU Threshold Prediction Model

Figure 4.8. The 1 and 2 NTU threshold RF turbidity models forecasting 5 hours and 20 minutes ahead over a non-typical low concentration turbidity event.

It can be observed here that the 1 NTU threshold prediction model picked up a change in the WDS over a day before the resulting turbidity event. The change that was detected was due to operational staff filling up the upstream and downstream service reservoirs in preparation for the cleaning works. Figure 4.8 (b) shows that because the 2 NTU threshold

prediction model was trained to be less sensitive to smaller changes in measurement data, it did not raise the same alarm.

4.5 Discussion

The ANN and RF regression based turbidity prediction models were able to produce relatively accurate forecasts up to approximately 20 minutes ahead. This is almost certainly an insufficient amount of lead time for a water company's operational staff to act upon.

The RF classification based turbidity prediction models performed very well at predicting turbidity events with a lead time of several hours, as shown by the high TPR values obtained. However, as lead time increases the model struggles to predict when the increased turbidity will arrive at the downstream turbidity meter, thus there is a significant increase in the number of false positives around the forecasted turbidity events. Ultimately, at the expense of the FDR, the largest reliable, increased turbidity forecasting time horizon for all classification models in this trunk main system is just over 5 hours. While lead times of the order of several hours does not allow the system operators to do anything more substantial in nature (e.g. flush the relevant pipe(s)), this should provide water company operational staff with a sufficient lead time to act on limiting or mitigating the discolouration event (e.g. to reduce flow rates to prevent further discolouration mobilisation). This can be considered more proactive compared to the commonly adopted reactive approach of waiting for customers to report occurrences of discoloured water and then cleaning the trunk main only after a sufficient number of these complaints have been received. Therefore, of the two modelling approaches, the classification based approach is deemed to be more practical than making direct predictions of turbidity values (i.e. regression based approach).

In the classification based approach a maximum turbidity threshold of 4 NTU was assessed which is the UK regulatory limit for discoloured water at customers' taps. However the visibility threshold where water can be deemed to be different from normal by a customer it is estimated to be approximately 8 - 10 NTU (Slaats et al., 2003). While this means a 4 NTU prediction may not result in a customer contact, immediate reduction of the WDS flow rates could prevent higher concentrations of discolouration material from being mobilised that would be visible. While this is limited to WDS where the flow rates can be automatically controlled, the immediate reduction of WDS flow rates after a

high turbidity forecast could also significantly increase the travel time of already mobilised discolouration material and thus significantly increase the lead time for operational staff to act.

The discovery of the optimal trade-off between TPR, FDR and lead time will heavily depend on the purpose and requirements of the resulting proactive management strategy. For example, a relatively high FDR may not be an issue if the flow rate can be immediately reduced to prevent further discolouration material from mobilising. Similarly, a long forecasting lead time may not be required if a downstream actuated valve can automatically divert discoloured water (e.g. into a separate reservoir cell or elsewhere) and a very low FDR would be important to not needlessly spam customers who are warned of potential discolouration in advance via text message.

Most of these proactive management strategies should be possible to implement providing a modern SCADA system is already in use. These strategies could not only directly reduce the number of discolouration contacts but could also aid indirectly by reducing discolouration material build up in downstream distribution pipes. This further emphasises the cost-benefit trade-off in comparison to the expensive alternative of regular invasive mains cleaning and their associated risks.

While the trunk main network used here is arguably simple, constructing and maintaining a hydraulic model for this network would still require some time and capital expense. Furthermore, building and calibrating a hydraulic model alone is not sufficient to accurately predict turbidity, even in a relatively simple system configuration like the trunk main system shown here. The reason for this is that discolouration mobilisation is not limited to any one point in the WDS which then simply travels downstream. The mobilisation of material leading to discolouration can occur anywhere along the flow path and the actual quantity mobilised depends on a number of factors. These factors include the quantity of material available locally (sediments deposited or material available for ‘peeling off’ the pipe wall) and the local hydraulic forces, both of which change spatially and temporally throughout the pipe system. As a consequence, continually predicting turbidity at the downstream end of a WDS, even in the case of pipes in series, is more complex than just accounting for travel time at an upstream point in the system. This explains why physically based models such as the PODDS model (Husband and Boxall, 2016) require multiple additional discolouration related parameters, in addition to a calibrated hydraulic model. Furthermore, these discolouration related parameters have to be calibrated with onsite turbidity observations for each trunk main being modelled.

The turbidity forecasting methodology presented here is data-driven and hence does not require a hydraulic model. Only readily available velocity and turbidity observation data is needed from the trunk main network to forecast turbidity. Therefore, the methodology has the potential to be transferable to WDS that have suitable meters installed and sufficient historical data already captured. However, the collected historical data will need to include a range of observed turbidity events within the WDS for the data-driven models to forecast accurately. Thus, while there is conceptually no issue with transferring the methodology to other WDS, the accuracy and leads times that could be achieved in other WDS is currently unknown and may vary significantly between different WDS.

Once a model is trained for a specific water system/site, the turbidity forecasting model would require retraining if that water system's configuration is altered in some significant way (e.g. by doing some rehabilitation type work or otherwise). Depending on the nature of the alteration, it is possible that no human interaction would be required as the model's parameters can be automatically retrained. This will likely require new observed data to be collected for a period of time which would be dependent on the requisite range of discolouration events taking place in the new system. This, however, is not seen as a major limitation of the methodology proposed here as most water supply and distribution systems are rather well developed and with network configurations that are not subject to regular, significant change.

Finally, due to the difficulty in interpreting the internal workings of data-driven models, little new knowledge about the accumulation and mobilisation processes of discolouration has been generated by using these models, although this remains an avenue for further investigation.

4.6 Chapter Summary

The turbidity forecasting methodology developed and presented in this chapter is data-driven and hence does not require hydraulic or water quality (turbidity) type network models. The hydraulic model is expensive and time consuming to build and maintain and, more importantly, a reliable, physically based, discolouration model capable of accurately predicting turbidity continually in a pipe network does not currently exist.

Three data-driven turbidity forecasting models based on Artificial Neural Networks, Random Forests and Support Vector Machines were developed and presented. Each of these models takes current and past velocity and turbidity measurements at a number of

upstream locations in the system to either directly predict turbidity (regression based model) or classify the turbidity at the downstream location as being above (or not) a pre-specified threshold (classification based model). Three threshold values of 1, 2 and 4 NTU were used to define the occurrence of different discolouration events. Different forecasting horizons ranging from 5 minutes to 8 hours ahead were used in the analyses. The methodology was tested and verified on a real water system in the UK with 11 months of velocity and turbidity measurement data.

The results obtained lead to the following conclusions:

(a) It is possible to reasonably reliably forecast occurrence of turbidity above some pre-specified threshold and hence detect the corresponding discoloration event in a real system by using a data-driven (i.e. non-physically based) methodology only;

(b) A classification based turbidity model is more reliable than direct turbidity predictions made by a regression type model. In a real water system analysed here, reasonable forecasts of discolouration events with increased turbidity can be made up to approximately 5 hours ahead using the classification based approach, as opposed to only 20 minutes when using the regression based approach.

(c) The classification based turbidity model results show that this methodology was capable of detecting the mobilisation of discolouration material, estimating if sufficient turbidity will be generated to exceed the pre-specified threshold and approximating the travel time required for that material to reach the downstream turbidity meter. A deeper look at the classification based model results show that the poor False Discovery Rate can be misleading as most of the false positives generated are clustered around a legitimate event. This further showed that the high number of false positives are from the model struggling to predict turbidity travel time to the downstream meter and not whether discolouration has been mobilized;

(d) Random Forests based turbidity forecast models performed best out of the three data-driven methods used;

(e) Despite being data-driven, the methodology shown here is generic and therefore transferable to other water systems.

Chapter 5 Data-Driven Study of Discolouration Mobilisation

5.1 Introduction

Due to the potential consequences associated with water supply in trunk mains, research on trunk mains has been primarily limited to areas where the benefits are clearly evident to water companies. In particular, significant research has been carried out on incrementally increasing the flow rate in a trunk main to remove discolouration material from the trunk main (Husband and Boxall, 2009; Sunny et al., 2017; Vreeburg and Beverloo, 2011). This is also known as flushing, and it has been shown that significant enough increases to the applied hydraulic force on pipe walls are able to mobilise discolouration material in pipes (Boxall and Saul, 2005; Husband and Boxall, 2016; Verberk et al., 2006).

Thus while the hydraulic mobilisation of iron and manganese deposits has been long known to result in discoloured water, the presence of discoloured water can also be due to other non-hydraulic processes, such as some biofilm sloughing or chemical interactions between pipe materials and water acidity (Fish et al., 2017; Husband et al., 2010; Husband and Boxall, 2009; Liu et al., 2017; Makris et al., 2014). Furthermore, only a small number of studies with often conflicting results describe the scale and frequency of hydraulically mobilised discolouration events under usual WDS operating conditions (Cook et al., 2015; Gaffney and Boult, 2012).

Water companies primarily deal with discolouration by cleaning, i.e. flushing WDS mains. Once a sufficient number of discolouration complaints have been reported in the area, the company may decide to reline (or replace) old mains believed to be the cause of significant discolouration (Husband et al., 2010; Vreeburg et al., 2008), particularly if this is going to help address additional issues (e.g., leakage). However, cleaning and especially rehabilitating WDS mains is expensive and can still potentially only have a limited effect if the discolouration material was mobilised from a different section of the network (Cook et al., 2015). Thus, determining where the significant causes of discolouration are in a WDS is important to efficiently reduce discolouration risk.

This chapter presents a methodology to determine the amount of discolouration that can be linked to preceding flow changes in upstream pipes. This methodology further aims to

identify the origin of mobilised discolouration material, thus identifying areas or pipes in the network that discolouration material accumulates in. This could aid in targeted proactive cleaning or replacement of problem pipes.

This chapter is organised as follows. Section 5.2 presents the methodology that determines the amount of discolouration that can be linked to preceding upstream hydraulic forces. Section 5.3 presents the experimental setup and a statistical test to determine the statistical significance of the results found by the methodology. Section 5.4 reports the results of the methodology. The results and their implications are discussed in Section 5.5. Finally, a summary of the chapter and its main conclusions are presented in Section 5.6.

5.2 Methodology for Assessing Discolouration Mobilisation

The methodology presented here evaluates the percentage of turbidity observations downstream that can be linked to preceding hydraulic events in an upstream pipe and thus identifying where discolouration material is likely to have originated, i.e. accumulated in the WDS. This in turn can enable targeted trunk main rehabilitation and cleaning operations.

The methodology is formed from three principles: (a) the hydraulic force that mobilised the discolouration material which resulted in a high turbidity observation occurred just prior (i.e. minutes or hours, not weeks) to the high turbidity observation; (b) a stronger hydraulic force (i.e. flow, velocity or shear stress) would result in more discolouration material being mobilised, provided that there is available material to mobilise (Boxall and Saul, 2005; Husband et al., 2011; Slaats et al., 2003); (c) discolouration material is constantly being regenerated/built up in all pipes (Furnass et al., 2014; Vreeburg and Beverloo, 2011). Based on these three principles, a downstream turbidity observation is thought to be the result of a hydraulically-based mobilisation process if a hydraulic force in an upstream pipe preceding the turbidity observation exceeds the recent prior hydraulic forces experienced in that pipe.

The percentage of selected turbidity observations that can be linked to preceding hydraulic events in an upstream pipe is given by the Hydraulically Mobilised Turbidity Percentage (HMTP) introduced here as follows:

$$HMTP(\varepsilon, T, x, y) = \frac{\sum_{\tau \in T} \left[\beta^{\varepsilon}_{y,\tau} > \alpha^{\varepsilon}_{x,\tau}\right]}{|T|} \times 100 \quad \text{(Eq. 5.1)}$$

where ε is the upstream pipe being assessed, T is the set of turbidity observations τ given in Nephelometric Turbidity Units (NTU) at a downstream location, x and y are periods of time, $\beta^{\epsilon}_{y,\tau}$ is the recent peak velocity (m/s) in pipe ε during a period of time of y duration preceding the turbidity observation τ and $\alpha^{\epsilon}_{x,\tau}$ is the peak velocity (m/s) in pipe ε during a period of time of x length that precedes the period of time for $\beta^{\epsilon}_{y,\tau}$.

From the perspective of a turbidity observation, the recent preceding peak velocity $\beta^{\varepsilon}_{y,\tau}$ of pipe ε is only assumed to have caused that turbidity observation if it has exceeded the prior peak velocity $\alpha^{\varepsilon}_{x,\tau}$ of that pipe. This is because the prior velocities in that pipe should have mobilised all the discolouration material that they could, and only a higher velocity should be able to mobilise additional material. The prior peak velocity $\alpha^{\varepsilon}_{x,\tau}$ will be called the preconditioned hydraulic threshold, and the recent preceding peak velocity $\beta^{\varepsilon}_{y,\tau}$ will be called the peak mobilising velocity. Thus, the y parameter determines how far back in time the HMTP should look for hydraulic mobilisation, and the x parameter determines the minimum size of hydraulic events being considered.

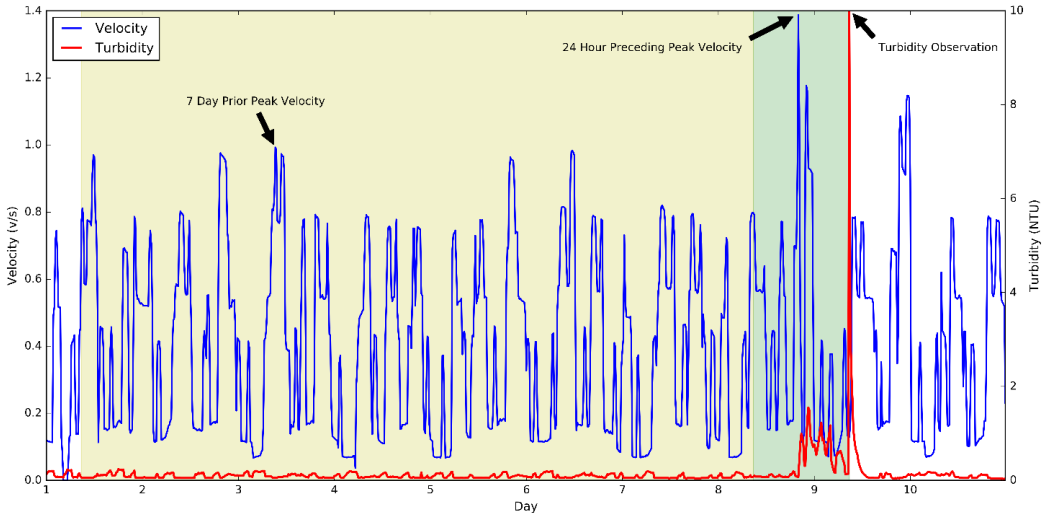


Figure 5.1. An example of the methodology showing that the turbidity observation could be linked to the preceding upstream hydraulic mobilisation of discolouration material. The 24 hours preceding the turbidity observation are highlighted green, and the 7 prior days are highlighted yellow. The preceding 24-h peak velocity is shown to be the cause of the high turbidity observation as it exceeds the prior 7-day peak velocity. NTU: Nephelometric Turbidity Units.

An example of the methodology is shown in Figure 5.1 where real velocity and turbidity observations are displayed. For the sake of brevity, the methodology is visualised for a single turbidity observation and for which the peak turbidity observation is chosen. The length of time set for y is 24 h because it was sufficiently long for all potentially mobilised material from the furthest upstream point to reach the downstream turbidity meter. The length of time set for x in this example is 7 days and was chosen solely for the ease of visualizing this example.

The velocity profile 24 h preceding the turbidity observation (highlighted green) is where the peak mobilising velocity is calculated. The 7 days prior (highlighted yellow) is where the preconditioned hydraulic threshold is calculated. The peak mobilising velocity (i.e. $\beta_{y,\tau}^{\varepsilon}$) of this upstream pipe is greater than its preconditioned hydraulic threshold (i.e. $\alpha_{x,\tau}^{\varepsilon}$), and thus, the turbidity observation is determined to have resulted from the hydraulic mobilisation of discolouration material in this pipe.

From the velocity and turbidity measurements shown in Figure 5.1, it can also be seen that the velocity just before the start of day 10 also exceeds the peak velocity indicated on day 3. However, no subsequent turbidity response is seen on day 10 because that pipe is now reconditioned to the new preconditioned hydraulic threshold at the end of day 8

(i.e. all discolouration material that could have been mobilised by this new velocity was already mobilised by the peak velocity at the end of day 8).

5.3 Experimental Setup

5.3.1 Chi-Square Test for Independence

While the high turbidity observation examined in Figure 5.1 is determined to have been caused by the hydraulic mobilisation of discolouration material, it is possible that the velocity profile and preceding turbidity response were coincidental. Thus, the chi-square test for independence will be used to determine the statistical significance of the results.

All turbidity observations will be divided into two turbidity sets of over 1 NTU observations and under 1 NTU observation, and then, each set of turbidity observations will be examined separately. $HMTP_{>1NTU}$ will show the percentage of turbidity observations above 1 NTU that are deemed to be caused by hydraulic mobilisation upstream, and likewise, $HMTP_{<1NTU}$ will show the percentage of turbidity observations below 1 NTU that are deemed to be caused by hydraulic mobilisation upstream. The turbidity threshold of 1 NTU was chosen as it is a clear quantifiable response above background turbidity levels and is the UK regulatory limit for water leaving water treatment works (DWI, 2014a). Therefore, a turbidity observation over 1 NTU can be considered as part of a turbidity event, and turbidity observations under 1 NTU can be considered as the absence of a turbidity event.

The proposed null hypothesis is that the turbidity level (i.e. over 1 NTU or under 1 NTU) is independent of an upstream pipe's preceding peak velocity that exceeds the preconditioned hydraulic threshold. The proposed alternative hypothesis is that higher turbidity levels (i.e. over 1 NTU) are dependent on an upstream pipe's preceding peak velocity that exceeds the preconditioned hydraulic threshold. Thus, for a statistically-significant result where the null hypothesis can be rejected, a $HMTP_{>1NTU}$ significantly greater than a corresponding $HMTP_{<1NTU}$ is expected to be seen. The significance level chosen is 0.01, and the chi-square test statistic with 1 degree of freedom is used to calculate the statistical significance.

5.3.2 Pipes and Pipes in Series

The methodology examines each pipe upstream of the turbidity meter, where a pipe is determined here by stretches of piping where the velocity remains the same. This means

an import and export branch or change in diameter determines the boundaries of a pipe. The individual pipes of each site are shown in Figure 3.1, Figure 3.2 and Figure 3.3.

While each pipe can be examined individually to estimate the amount of discolouration material linked to that pipe, the preconditioned hydraulic thresholds of multiple pipes can be simultaneously exceeded and discolouration material mobilised from multiple pipes simultaneously. This would mean that some turbidity observations are counted as originating from more than one pipe.

Therefore, to accurately assess the total amount of turbidity observations that can be linked to hydraulic mobilisation, all pipes upstream of the turbidity meter are also jointly assessed. This is done by separately assessing if any pipe upstream of the turbidity meter experienced a velocity that exceeded their preconditioning hydraulic threshold. The multiple pipes that are jointly assessed will be called pipe sets.

5.3.3 Data

All available data for each turbidity meter in each site presented in Chapter 3 was used here.

5.4 Results

5.4.1 Hydraulically Mobilised Turbidity Percentage

The results of the HMTP using an x of 1 day and y of 1 day applied to each turbidity meter and its corresponding jointly assessed pipe set are shown in Table 5.1.

Table 5.1. Hydraulically Mobilised Turbidity Percentage (HMTP) carried out on each pipe in series between the upstream sources and downstream service reservoirs to assess the amount of turbidity observations that can be linked to hydraulic mobilisation. The x and y parameters of HMTP were set to 1 day.

Turbidity Meter	Pipes in Set	$HMTP_{<1NTU}$	$HMTP_{>1NTU}$	p-Value
TM A (Site 1)	A, B, C, D	81%	100%	$p < 10^{-9}$
TM B (Site 1)	A, B, E, F	77%	91%	$p \approx 10^{-8}$
TM C (Site 2)	G, H	53%	93%	$p < 10^{-9}$
TM D (Site 3)	I, J, K	66%	84%	$p \approx 10^{-6}$

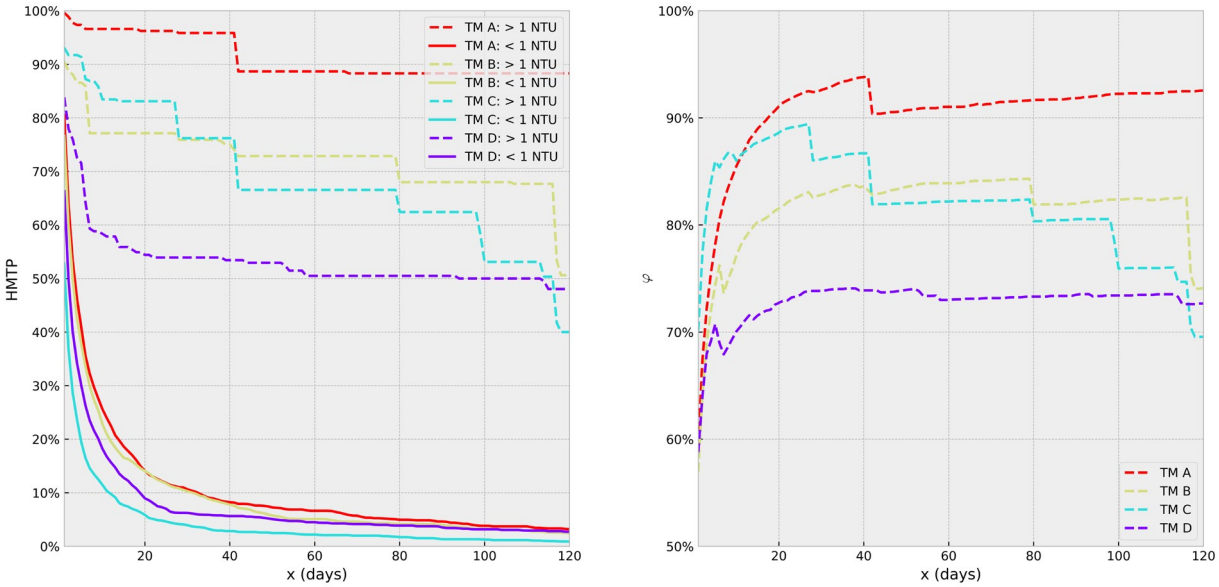
A length of 1 day was chosen for the *y* parameter because it was sufficiently long enough for material mobilised from the furthest upstream points of each site to reach their respective downstream turbidity meter. The *x* parameter was set to 1 day to show the maximum amount of turbidity in each site that can be linked to preceding upstream hydraulic events.

The $HMTP_{>1NTU}$ results in Table 5.1 range from 84% - 100%, thus showing that the majority of turbidity can be linked to preceding hydraulic events upstream. However, because the requirements for the methodology to determine if there was a hydraulic event preceding a turbidity observation are quite low with only an *x* of 1 day (i.e. the peak velocity in the 0-24 hour time period prior to the turbidity observation exceeds the peak velocity in the 24-48 hour time period prior to the turbidity observation), the $HMTP_{<1NTU}$ results in Table 5.1 are also substantially high. While the *p*-values show that the null hypothesis can be rejected at a 0.01 level of significance, a significantly larger gap between the $HMTP_{<1NTU}$ and $HMTP_{>1NTU}$ results would indicate greater confidence in the methodology and results.

The effect of increasing the *x* parameter for $HMTP_{<1NTU}$ and $HMTP_{>1NTU}$ can be seen as plotted in Figure 5.2a. This figure shows that the $HMTP_{<1NTU}$ of each TM exponentially decays while the $HMTP_{>1NTU}$ decreases at a substantially slower rate. A function calculating the trade-off between the $HMTP_{<1NTU}$ and $HMTP_{>1NTU}$ is given by the formula shown below:

$$\varphi = \frac{\mathrm{HMTP}_{>1\mathrm{NTU}} + (1 - \mathrm{HMTP}_{<1\mathrm{NTU}})}{2} \qquad \text{(Eq. 5.2)}$$

Figure 5.2b shows φ plotted for each TM over increasing values of x. A higher φ indicates a better trade-off between a low $HMTP_{<1NTU}$ and a high $HMTP_{>1NTU}$.



(**a**) (**b**)

Figure 5.2. (**a**) The $HMTP_{<1NTU}$ and $HMTP_{>1NTU}$ are shown for each TM over increasing *x* parameters; (**b**) the objective formula φ is shown for each TM over increasing *x* parameters.

From the results shown in Figure 5.2a and Figure 5.2b, a 30-day length for *x* was selected to be shown here in Table 5.2. This length was chosen as it is deemed a reasonable trade-off between minimizing $HMTP_{<1NTU}$ and maximizing $HMTP_{>1NTU}$ across all four TMs while keeping $HMTP_{<1NTU}$ and $HMTP_{>1NTU}$ results relatively insensitive to an increase or decrease of a couple of days. The corresponding results for each jointly assessed pipe set followed by the results for each individual pipe that makes up that pipe set are shown in Table 5.2. As it can be seen from this table, the percentage of turbidity observations under 1 NTU that is deemed to be hydraulically mobilised ranges from 3% to 5% for individual pipes and 4% to 11% for the grouped pipe sets. At the same time, the percentage of turbidity observations over 1 NTU ranges from 0% to 84% for individual pipes and 54% to 96% for the grouped pipe sets.

Table 5.2. Hydraulically Mobilised Turbidity Percentage (HMTP) results with the *x* parameter set to 30 days and the *y* parameter set to 1 day.

Turbidity Meter	Pipes	$HMTP_{<1NTU}$	$HMTP_{>1NTU}$	*p*-Value
TM A (Site 1)	A, B, C, D	11%	96%	$p < 10^{-9}$
	A	4%	26%	$p < 10^{-9}$
	B	4%	33%	$p < 10^{-9}$
	C	5%	84%	$p < 10^{-9}$
	D	5%	81%	$p < 10^{-9}$
TM B (Site 1)	A, B, E, F	10%	76%	$p < 10^{-9}$
	A	4%	75%	$p < 10^{-9}$
	B	4%	75%	$p < 10^{-9}$
	E	5%	1%	$p = 1$
	F	5%	0%	$p = 1$
TM C (Site 2)	G, H	4%	76%	$p < 10^{-9}$
	G	4%	76%	$p < 10^{-9}$
	H	3%	76%	$p < 10^{-9}$
TM D (Site 3)	I, J, K	6%	54%	$p < 10^{-9}$
	I	3%	41%	$p < 10^{-9}$
	J	3%	39%	$p < 10^{-9}$
	K	4%	52%	$p < 10^{-9}$

Note that the sum of $HMTP_{>1NTU}$ for individual pipes belonging to each turbidity meter exceeds 100%. This was expected because, as mentioned above, a turbidity observation can be linked to multiple pipes if a hydraulic event occurs in both pipes simultaneously.

For the results of TM A, a high $HMTP_{>1NTU}$ of 84% is given for Pipe C. This is important to note because as can be seen from Table 1, Pipe C had a low 99th velocity percentile of 0.17 m/s, which indicates a high potential for material build up. The 12% difference between the $HMTP_{>1NTU}$ of 84% for Pipe C and the $HMTP_{>1NTU}$ of 96% for the pipe set is assumed to come from Pipes A and B and not Pipe D because Pipes C and D have very similar velocity profiles (when compared to Pipes A and B).

Comparing the TM A and TM B cases for Site 1, Table 4.2 shows two very different sets of $HMTP_{>1NTU}$ results for Pipes A and B, even though they are both located on the same

site. This indicates that significant discolouration material is being mobilised from Pipes A and B. However, the material that travels towards TM B reaches it, whilst the material that travels towards TM A ends up settling/re-attaching in Pipe C and then remobilises at a later time.

5.4.2 Turbidity and Velocity Relationships

Figure 5.3 shows the peak velocity in the 24 h preceding a turbidity observation plotted against the peak velocity in the 30 days prior to the 24 h preceding the turbidity observation for the four pipes that had the highest $HMTP_{>1NTU}$ for their respective turbidity meters as shown in Table 5.2. The size of each data point in a plot is relative to the turbidity measurement where a higher turbidity value results in a bigger data point. Because observations from the same turbidity event are likely to have the identical 24-h preceding peak velocity and prior 30-day peak velocity, most data points are overlapping and thus only the largest turbidity observation for a turbidity event is visible in Figure 5.3. The dashed line, which is also the identity line, shows where the preconditioned hydraulic threshold for data points along the *y* axis is. Hence, data points above the identity line are considered to have exceeded their preconditioned hydraulic thresholds for that pipe because they have experienced a velocity in the preceding 24 hours that is higher than all velocities experienced in the prior 30 days. Although data points below the identity line of a specific pipe cannot be linked to hydraulic mobilisation from that pipe, it does not mean that they cannot be linked to hydraulic mobilisation from a different pipe upstream of the turbidity meter.

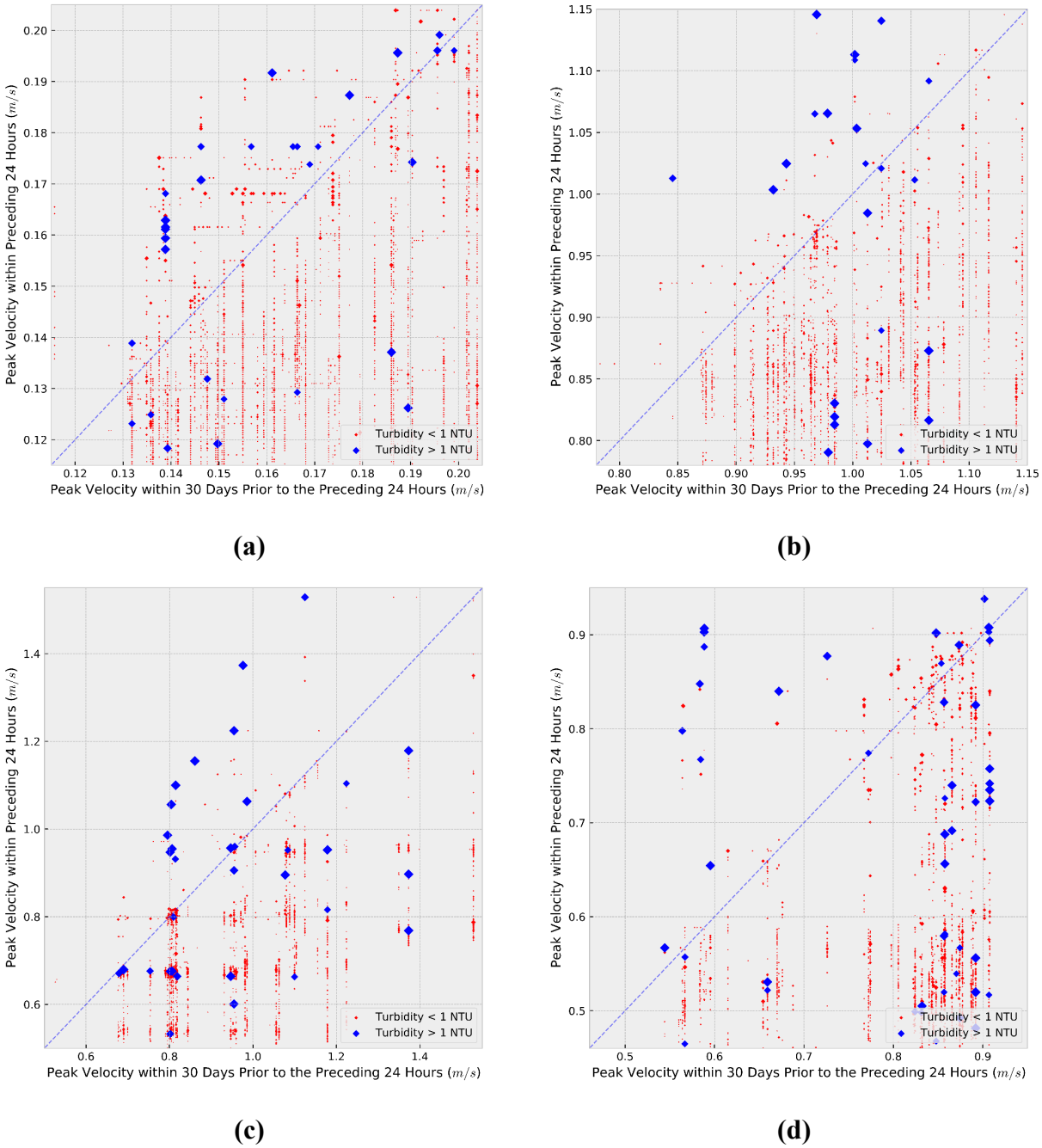


Figure 5.3. For each turbidity observation, the 24-h preceding peak velocity was plotted against the peak velocity of the 30 days prior to the 24 hours preceding the turbidity observation. (**a**) Site 1, Turbidity Meter A, Pipe D; (**b**) Site 1, Turbidity Meter B, Pipe B. (**c**) Site 2, Turbidity Meter C, Pipe H; (**d**) Site 3, Turbidity Meter D, Pipe K. NTU: Nephelometric Turbidity Units.

If a cleaning velocity (i.e. velocity at which all material is removed from the pipe) existed in any pipe, a significant number of low and only low turbidity observations would be seen above the identity line after a sufficiently high prior peak velocity (i.e. the *x* axis of Figure 5.3). However, this is not observed in any pipes, and thus, no clear mobilisation limit is seen, as expected. Similarly, by looking at only the peak velocities in pipes during

the 24 h preceding the arrival of turbidity (i.e. *y* axis of Figure 5.3), a clear minimum mobilising velocity was not seen in any pipes, again, as expected. Instead, it can be clearly seen that most of the turbidity observations under 1 NTU are ubiquitously spread below the identity line. The turbidity observations over 1 NTU that are below the identity line could not be linked mobilisation forces in that specific pipe, however the group pipe set $HMTP_{>1NTU}$ results presented in Table 5.2 show that the majority of turbidity observations can be linked to mobilisation forces in at least one pipe upstream of each turbidity meter.

Table 5.3 shows Spearman's rank correlation coefficients for three different sets of velocity calculations correlated with their associated turbidity sets. Spearman's rank correlation coefficient measures the dependence of two parameters as described as a monotonic function. Linearity between the two parameters is not assumed in Spearman correlation, which is not the case for Pearson's correlation coefficient. A Spearman correlation coefficient of 1 or −1 indicates a perfect monotonic relationship. For each correlation coefficient, an associated *p*-value is derived from a statistical *t*-test, which indicates the probability of an uncorrelated system generating datasets that have a correlation at least as extreme.

The first correlation shown in Table 4.3 is the *24h Peak Velocity*, which is the 24-h preceding peak velocities of each turbidity observation, regardless of turbidity value, correlated with respective turbidity observations. The very weak to weak positive correlation across all pipes shows that higher preceding velocities alone rarely indicate the appearance of higher turbidity concentrations downstream. However, these correlations are predominantly driven by the many low flow turbidity observations and hence tell little about the correlation between peak velocities preceding a turbidity event and the amount of turbidity mobilised in that event. To further analyse this, the second correlation set is shown in Table 5.3, denoted as the *24h Peak Event Velocity.* In this set, the 24-h preceding peak velocities of turbidity events are correlated with the downstream turbidity observations of those turbidity events. Interestingly, there are a few negative correlations, but because they are very weak correlations, not much can be inferred.

Table 5.3. Spearman's rank correlation coefficients and associated *p*-values for three sets of correlations: (a) 24-h preceding peak velocities of each turbidity observation correlated with those turbidity observations; (b) 24-h preceding peak velocities of each turbidity event correlated with the turbidity observations of that event; (c) the difference between the 24-h preceding peak velocity and the 30-day preconditioned threshold of each turbidity event correlated with the turbidity observations of that event.

Turbidity Meter	**Pipe**	**(a) 24-h Peak Vel.**	***p*-Value**	**(b) 24-h Peak Event Vel.**	***p*-Value**	**(c) Exceeded Vel. Difference**	***p*-Value**
TM A	A	0.30	$p < 10^{-40}$	−0.16	$p \approx 10^{-23}$	0.55	$p < 10^{-40}$
(Site 1)	B	0.28	$p < 10^{-40}$	−0.13	$p \approx 10^{-15}$	0.38	$p \approx 10^{-21}$
	C	0.29	$p < 10^{-40}$	0.40	$p < 10^{-40}$	0.13	$p \approx 10^{-08}$
	D	0.29	$p < 10^{-40}$	0.41	$p < 10^{-40}$	0.16	$p \approx 10^{-10}$
TM B	A	0.38	$p < 10^{-40}$	0.22	$p < 10^{-40}$	0.42	$p < 10^{-40}$
(Site 1)	B	0.38	$p < 10^{-40}$	0.22	$p < 10^{-40}$	0.48	$p < 10^{-40}$
	E	0.06	$p < 10^{-40}$	−0.18	$p \approx 10^{-32}$	−0.30	$p \approx 10^{-14}$
	F	0.02	$p \approx 10^{-10}$	−0.17	$p \approx 10^{-31}$	−0.26	$p \approx 10^{-08}$
TM C	G	0.32	$p < 10^{-40}$	0.40	$p < 10^{-40}$	0.41	$p < 10^{-40}$
(Site 2)	H	0.30	$p < 10^{-40}$	0.40	$p < 10^{-40}$	0.41	$p < 10^{-40}$
TM D	I	0.13	$p < 10^{-40}$	0.15	$p \approx 10^{-22}$	0.44	$p < 10^{-40}$
(Site 3)	J	0.27	$p < 10^{-40}$	0.14	$p \approx 10^{-20}$	0.52	$p < 10^{-40}$
	K	0.31	$p < 10^{-40}$	0.22	$p < 10^{-40}$	0.24	$p \approx 10^{-20}$

The third correlation set shown in Table 5.3 is the *Exceeded Velocity Difference*, where the difference between the 24-h preceding peak velocity and the 30-day preconditioned hydraulic threshold (i.e. the identity lines shown in Figure 5.3) of each turbidity event is correlated with the turbidity observations of those events. While the largest correlation is only 0.55, the majority of correlations are moderately positive, which is significantly stronger compared to the *24h Peak Velocity* and *24h Peak Event Velocity* correlations. This shows that the exceeded velocity difference is a better indicator of the resulting turbidity event size than a preceding increase in velocity alone.

5.5 Discussion

Discolouration material clearly does build up in the trunk mains observed in the three sites analysed and this is despite the relatively high velocity percentiles shown in Table 1 that vastly exceed the 'self-cleaning' velocity ranges (0.25 m/s – 0.4 m/s) associated with smaller diameter distribution pipes (Blokker et al., 2010; Vreeburg and Boxall, 2007). This agrees with the findings of other authors conducted on large diameter pipes (i.e. over 200 mm), which find no evidence for self-cleaning velocities or shear stresses (Husband et al., 2011; Seth et al., 2009; Vreeburg and Beverloo, 2011).

The majority of *p*-values in Table 5.1 and Table 5.2 show the null hypothesis being overwhelmingly rejected at a 0.01 level of significance and thus conclude that higher turbidity levels (i.e. over 1 NTU) can be explained by an upstream pipe's preceding peak velocity exceeding its preconditioned hydraulic threshold. These *p*-values here are particularly low due to the high number of turbidity observations considered (i.e. over 100,000), thus making it very unlikely that these values would be seen in uncorrelated results.

The only exceptions in results seen in Table 5.2 and Table 5.3 are Pipes E and F for TM B, which are distinctly different from all other results as, conversely, only a few high turbidity observations can be linked to preceding hydraulic events in these pipes. This may be indicative of an underlying process that is either preventing discolouration material from sufficiently accumulating in these pipes or the methodology from accurately modelling them.

It is important to note that the majority of velocity peaks that precede turbidity observations are relatively small increases in comparison to their average daily peak velocity, these velocity peaks prior to turbidity events are seen in the datasets used here as typically being less than 110% of the average daily peak velocity. This emphasises how sensitive discolouration material can be to mobilising velocities and indicates why discolouration events are so often attributed to maintenance or refurbishment works that alter the velocities in WDS.

Table 5.1 shows that a maximum amount of 84% to 100% of turbidity observations over 1 NTU in the trunk mains examined here can be linked to hydraulic mobilisation. This also conversely shows that between 0% and 16% of turbidity observations cannot be linked to the hydraulically-driven mobilisation process outlined in this section. This

leaves a few possibilities about the mobilisation of the remaining turbidity: (a) some of the discolouration material could have been mobilised from further upstream (i.e. reservoirs or treatment works); (b) a hydraulic process not accounting for such as a transient event or flow reversal caused some mobilisation; (c) a non-hydraulic process caused some mobilisation (e.g., biofilm detachment/sloughing that can sometimes occur without an increase in hydraulic force).

The methodology presented here does not make any assumptions about what the discolouration material consists of (e.g., manganese, biofilms), what form the discolouration material takes inside pipes (e.g., sediment, cohesive layers), nor does it assume a rate (e.g., linear, exponential) at which discolouration material is mobilised. Additionally, because the mobilisation condition has been reduced to a simple "greater than prior" condition, as long as the hydraulic force has a monotonic relationship to the flow rate, it also does not matter what the hydraulic force is (e.g., velocity, shear stress, laminar boundary layer size). This means, in theory, that the application of the methodology to a WDS is not limited by the material composition, layout and range of flow rates of that WDS.

As flow meters are already ubiquitous in WDS, this methodology also shows the potential information gain from installing even a single turbidity meter. As the accuracy of the methodology to identify the primary sources of discolouration increases with more data, installing a turbidity meter at an existing downstream facility (e.g. flow meter, service reservoir, etc) where regular maintenance is easily achievable is advised. If possible, further turbidity meters should be installed at the downstream ends of different network branches. This would enable the correlation of methodology results to further identify high discolouration risk pipes, as was shown done in this section for Site 1.

Regarding the frequency of flow and turbidity observations, while a 15 minute sampling frequency was deemed sufficient for the sites examined here, a higher frequency may be required for WDSs that can experience sharp, but short-lived velocity spikes. This is because a significant, but short-lived velocity spike that could cause discolouration may only present as a minor increase in the cumulative flow over a 15 minute period.

5.6 Chapter Summary

This section presents a long-term continuous study of discolouration mobilisation and a methodology to determine the approximate amount and origin of hydraulically mobilised

turbidity in trunk mains. The results of the methodology are presented from real turbidity and flow observations collected from three UK trunk main networks over a period of two years and 11 months. The following conclusions are made based on the case studies results obtained:

(a) The methodology shows that for the four turbidity meters used in this study, a maximum of 84%, 91%, 93% and 100% of downstream turbidity observations over 1 NTU could be linked to preceding hydraulic forces that exceeded an upstream pipe's hydraulically preconditioned state. This shows that the mobilisation of discolouration material is predominantly determined by hydraulic forces, which, in turn, indicates significant potential for modelling and predicting discolouration events.

(b) The methodology showed that even without a calibrated hydraulic model, it is possible to determine the approximate origin of discolouration material that had been hydraulically mobilised within each site analysed. This can be used as an aid in the prioritisation of cleaning trunk mains and targeted mains rehabilitation.

(c) The level of turbidity is shown to be significantly dependent on preceding upstream velocities that exceed a pipe's preconditioned state. Furthermore, discolouration material is shown to accumulate regardless of the velocity magnitude, thus indicating that controlling the variability of the hydraulic profile where possible is significantly more vital in effectively managing discolouration risk.

Chapter 6 On-Line Turbidity Event Forecasting

6.1 Introduction

The accuracy and lead forecast times of the short-term turbidity forecasting methodology shown in Chapter 4 were limited. One of the factors in its limited accuracy was that while the model was able to detect the mobilisation of discolouration material sufficiently well, it struggled to predict when the resulting material would arrive at the downstream turbidity meter. This was especially true for longer forecasting horizons (i.e. above 5 hours).

Furthermore, in Chapter 5 it was shown that not all high turbidity observations (i.e. above 1 NTU) could be linked to preceding upstream hydraulic events. This means that the short-term turbidity forecasting methodology presented in Chapter 4 was, to some degree, trained using turbidity observations that had no corresponding preceding hydraulic response. This would have very likely contributed somewhat to the model's limited accuracy.

In this chapter, a novel on-line turbidity event forecasting methodology is presented that addresses the two aforementioned issues. This methodology detects the start of a hydraulic event in a WDS and predicts if that hydraulic event will mobilise a sufficient amount of discolouration material to cause a downstream turbidity event within the subsequent 24 hours.

When the methodology forecasts that a hydraulic event will indeed cause a downstream turbidity event, customers could be warned via text or email of the discolouration issue. Furthermore, if the methodology is employed in a network where the flowrates can be somewhat controlled (i.e. through pumps, automated valves, etc), then proactive prevention of the discolouration event occurrence is possible. This would be done by reducing or even limiting the flowrate from further increasing and thus limiting the magnitude of the hydraulic event which in turn could prevent the turbidity event from occurring. Even if the turbidity event is not prevented from occurring, by limiting the magnitude of the hydraulic event, the magnitude of the resulting turbidity event may still be prevented from reaching visible turbidity levels (i.e. > 8-10 NTU) downstream. Hence, not resulting in customer complaints for discoloured water.

This chapter is organised as follows. Section 6.2 presents an on-line turbidity event forecasting methodology along with the associated model inputs and performance metrics. Section 6.3 shows what data is used by the models. Section 6.4 reports and discusses the results of training and testing the methodology. Finally, a summary of the chapter and its main conclusions are presented in Section 6.5.

6.2 On-Line Turbidity Event Forecasting Methodology

The methodology presented here analyses the velocity in WDS pipes to detect the start of a hydraulic event and then predict if that hydraulic event will cause a downstream turbidity event in the next 24 hours that at least exceeds 1 NTU in peak magnitude. As the optimal use of the methodology is to prevent (or at the very least limit) the forecasted turbidity event by way of reducing the magnitude of the hydraulic event, the forecast has to be made at the start of the hydraulic event. This is to provide as much time as possible for the magnitude of the hydraulic event to be limited by closing actuated values, turning off pumps, etc. Even if the methodology is used in a WDS where the flow cannot be controlled, forecasting a turbidity event as early as possible provides more time for operational staff to take any other preventative actions and warn likely affected customers.

6.2.1 Hydraulic Event Detection

Because the turbidity event forecast is made from assessing the start of a hydraulic event, the hydraulic event detection needs to be simple and needs to include as broad a range of hydraulic events as possible. Therefore, the start of a hydraulic event is detected when the velocity in a pipe exceeds a threshold value. However, in order to account for the variation in velocity between different WDSs and individual pipes, the threshold value for a specific pipe is set to be the peak velocity from a predefined number of trailing prior days and is recalculated (i.e. updated) for each time step (i.e. every 15 minutes). The length of this trailing time period, in which the peak velocity is taken from, will be denoted as λ. This can alternatively be interpreted as the length of time after a hydraulic event in which velocities of equal or lower magnitudes to the ones seen in the hydraulic event will not be recognised as the start of a new hydraulic event. The value λ needs to be chosen carefully, as explained below.

Laboratory experiments by Husband et al. (2008) showed that turbidity observations measuring over 1 NTU could be seen from mobilising a week's worth of accumulated

discolouration material in a 79 mm polyethylene pipe. Therefore, it is thought somewhat likely that a λ bigger than 7 days would mean that some hydraulic events that cause turbidity events would not be recognised. However, having a too small λ such as 1 or 2 days would mean that even typical daily peak velocities are likely to be incorrectly recognised as the start of a hydraulic event. The value finally chosen for λ was 7 days as this would also clearly link to a WDS's weekly water consumption pattern. Consequently, a hydraulic event can now be more clearly defined as occurring if the velocity exceeds the previous week's peak velocity. Additionally, it is worth noting that as the hydraulic threshold is set from the preceding 7 days, variation due to seasonality is also taken into account.

As multiple hydraulic events can occur in relatively quick succession (i.e. < 24 hours), defining the end of a hydraulic event is important. A hydraulic event is defined here as having ended when the velocity in the pipe has remained below the hydraulic threshold for a continuous period of two hours. This time period of two hours was chosen after conducting a limited sensitivity analysis. As the hydraulic threshold is recalculated for each timestep, a new hydraulic event can only be triggered shortly after the old hydraulic event if the velocity in that pipe exceeds the previous hydraulic event's peak velocity. An additional hydraulic event occurring in the same pipe within 24 hours of the previous hydraulic event will be called here a 'subsequent hydraulic event'.

Because velocities can vary between different pipes in a network and thus will be conditioned to different levels, each distinct pipe section (i.e. pipe length where the velocity remains constant for a given flow rate) upstream of the turbidity meter is given its own separate hydraulic threshold that is recalculated at each timestep. However, if there were two distinct pipes in the WDS then a hydraulic event will be detected if either pipes' hydraulic threshold is exceeded by velocities occurring in their respective pipes. A hydraulic event occurring at the same time as another hydraulic event but in a different pipe will be called here a 'simultaneous hydraulic event'.

6.2.2 Classification Based Turbidity Event Forecasting Model

Unlike the turbidity forecasting methodology presented in Chapter 4, the methodology here forecasts the arrival of downstream turbidity events to be within the subsequent 24 hours. A 24-hour forecast period was chosen as it was deemed to be sufficiently long enough for material mobilised from the furthest upstream points, of the sites examined here, to travel to their respective downstream turbidity meters.

A binary classification approach is used in this methodology meaning that the model only predicts 'Positive' (i.e. predicts a turbidity event will occur in the next 24 hours) or 'Negative' (i.e. predicts a turbidity event will not occur in the next 24 hours). If the model prediction matches the subsequent turbidity observations then it is prefixed with 'True' (i.e. True Positive or True Negative), otherwise the prediction is prefixed with 'False' (i.e. False Positive or False Negative). For example, if the model predicts that there will be a turbidity event and a turbidity event is observed at the downstream turbidity meter within the next 24 hours, then the prediction is labelled as a 'True Positive'. Conversely, if the model predicts there will not be a turbidity event and a turbidity event actually is observed at the downstream turbidity meter within the next 24 hours, then the prediction is labelled as a 'False Negative'. This relationship is illustrated by Figure 4.1.

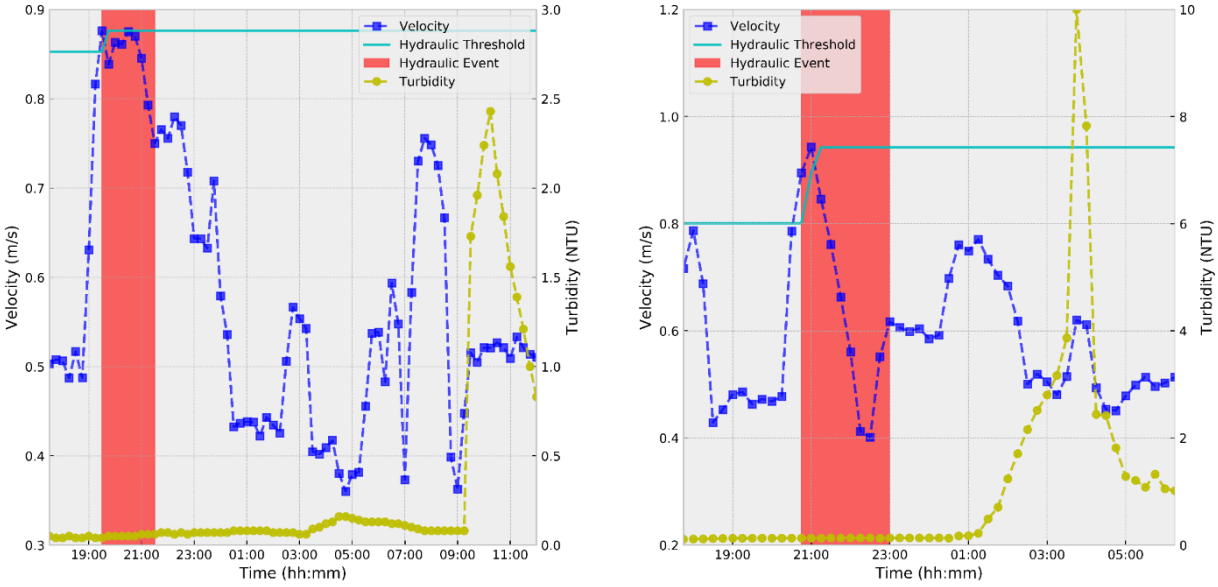


Figure 6.1. Two hydraulic events and their respective subsequent downstream turbidity events are shown from (a) Site 1, Pipe B, TM B and (b) Site 1, Pipe B, TM A. The start of the hydraulic event is triggered by exceeding the 7-day hydraulic threshold with the whole hydraulic event highlighted red. NTU: Nephelometric Turbidity Units.

Figure 6.1 shows two real hydraulic events observed in pipes and their respective subsequent downstream turbidity events. As defined by the methodology, the hydraulic event highlighted in red occurs when the velocity in a pipe exceeds the hydraulic threshold and ends after remaining below the hydraulic threshold for a continuous period of two hours. As the positive turbidity event forecast is made right at the start of hydraulic events, the turbidity event forecast for the hydraulic event shown in Figure 6.1a is made

at 9:30 and thus is forecasted 14 hours before the turbidity level exceeds 1 NTU. For the hydraulic event shown in Figure 6.1b, the turbidity event forecast is made at 20:45 and thus is forecasted 5 hours and 15 minutes before the turbidity level exceeds 1 NTU.

6.2.3 Model Performance Metrics

The True Positive Rate (TPR), also known as the probability of detection, is the probability that the model will correctly predict positive class values (i.e. discolouration material will be hydraulically mobilised resulting in a turbidity event). The formula for the TPR is given in Section 4.2.4.2.

The False Positive Rate (FPR), also known as the false alarm rate, is the probability that the model will incorrectly predict negative class values (i.e. no turbidity event due to hydraulic mobilisation). It is important to keep FPR low to ensure that confidence in the system remains when an alarm is given. The formula for the FPR is given below:

$$FPR = \frac{\mathrm{FP}}{FP + TN} \quad \text{(Eq. 6.1)}$$

6.2.4 Random Forest

Building the forecasting model requires the use of a machine learning method. A RF classifier was chosen to be used here due to its good performance shown in Section 4.2. The RF classifier comprises of 200 weak binary tree classifiers, each using the entropy splitting criterion. The percentage of features to consider when looking for the best tree split was set to 20%.

The optimal maximum depth of trees was set via hyperparameter tuning using 2-fold cross-validation on the training dataset only. The hyperparameter tuning of the optimal maximum depth of trees was limited to depths between 3 to 7.

6.2.5 Model Inputs

The range and number of inputs given to the RF model here has been significantly reduced compared to the methodology used in chapter 4. This is in part based on the findings of work done in chapters 4 and 5, and in part due to this model not likely requiring the large amounts of recent historical velocity inputs that the models in chapter 4 required. This is because the model here does not need to make predictions on exactly how long discolouration material will take to reach the downstream turbidity meter.

Thus, here only one type of novel feature is extracted from current and historical data and that is the *Current Exceeded Velocity* (CEV) which is defined as follows:

$$CEV(x) = v - Max\ Vel(x) \quad \text{(Eq. 6.2)}$$

where v is the current velocity and *Max Vel(x)* gives the maximum velocity from the period of the prior x time but excluding the current timestep. The CEV provides the model with information on the potential amount of discolouration material accumulation in a pipe.

Six inputs are given to the model per each distinct pipe upstream of the turbidity meter. These six inputs are the current velocity for that pipe and the CEV of a 7-day period, 14-day period, 30-day period, 60-day period and 120-day period. For example, a site with only two pipes upstream of the turbidity meter would have a total of 12 inputs presented to the model and these are listed out in Table 6.1.

Table 6.1. Example list of inputs given to a model for a site with only two distinct pipes upstream of the turbidity meter.

Input Number	Source	Input Type
1	Pipe A	Current Velocity
2	Pipe A	CEV(7 days)
3	Pipe A	CEV(14 days)
4	Pipe A	CEV(30 days)
5	Pipe A	CEV(60 days)
6	Pipe A	CEV(120 days)
7	Pipe B	Current Velocity
8	Pipe B	CEV(7 days)
9	Pipe B	CEV(14 days)
10	Pipe B	CEV(30 days)
11	Pipe B	CEV(60 days)
12	Pipe B	CEV(120 days)

6.2.6 Model Training

The training dataset is made up of positive samples (i.e. hydraulic events where a downstream turbidity event is observed within the next 24 hours) and negative samples

(i.e. hydraulic events where no downstream turbidity event is observed within the next 24 hours). Each positive and negative sample used in the training of the model can be assigned a Sample Weight. The weight of a sample can be considered as the relative amount of importance the RF classifier places on a sample to correctly learn that sample. For example, a model trained using only 3 samples with a weighting of 1, 1, and 3 respectfully, would prioritise the model to correctly predict sample 3 over samples 1 and 2 combined.

The weight for each sample in the training dataset is given by the formula shown below:

$$Sample\ Weight = Class\ Weight \times Reliability\ Weight \times Age\ Weight \quad \text{(Eq. 6.3)}$$

6.2.6.1 Class Weight

The Class Weight for a sample is set to be inversely proportional to the size of the class that the sample belongs to. This way, the Class Weight addresses the imbalance between the number of positive samples and negative samples in the training dataset.

6.2.6.2 Reliability Weight

As each pipe upstream of the turbidity meter has its own hydraulic threshold, it is possible to have multiple hydraulic thresholds exceeded in different pipes within a relatively short time span of each other. This then raises a potential problem if there is indeed a subsequent downstream turbidity event (i.e. it is a positive sample) because it is not possible to automatically determine if the downstream turbidity event was actually caused by only certain preceding hydraulic events and not others. This means that there is a chance that a negative sample could actually be miss-labelled as positive if that specific hydraulic event did not cause any discolouration mobilisation but was in fact solely caused instead by a subsequent or preceding hydraulic event. While it could be that all those hydraulic events do mobilise discolouration material that contributes to the turbidity event, the uncertainty makes these positive samples somewhat less reliable.

Note that this is not an issue if there is no actual resulting downstream turbidity event as all the samples can be reliably labelled negative. Likewise, this is not a problem if multiple pipes exceed their respective thresholds within the same time step because these are counted as a single positive sample.

Using these unreliable positive samples in training could significantly reduce the model's prediction performance. This is especially true when very few positive samples are used to train the model, which is expected to be the case here as turbidity events in general are

rare. While removing these unreliable positive samples is an option, this would further reduce an already low number of positive training samples.

To address this issue, a Reliability Weight is set for each sample. If a sample is negative (i.e. no subsequent turbidity event) then it is given a weight of 1, however if the sample is positive (i.e. there is a subsequent turbidity event) then a sample weight of 1 is divided across each sample that occurs just prior to the subsequent turbidity event. For example, if there is only one positive sample then it has a Reliability Weight of 1, if there are 2 positive samples then each sample has a Reliability Weight of 0.5 and if there are 3 positive samples then each sample has a Reliability Weight of 0.33. By doing this, higher importance is placed on training the model to samples that have higher certainty.

6.2.6.3 Age Weight

The Age Weight for a sample is given by the below formula:

$$Age\ Weight = 2 - 1.0001^{\sigma} \quad \text{(Eq. 6.4)}$$

where σ is the sample age rank. The newest sample is assigned the age rank of 0 and the oldest sample has the age rank equal to the size of the training dataset minus 1. Figure 6.2 shows the Age Weight for any given age rank between 0 and 1000.

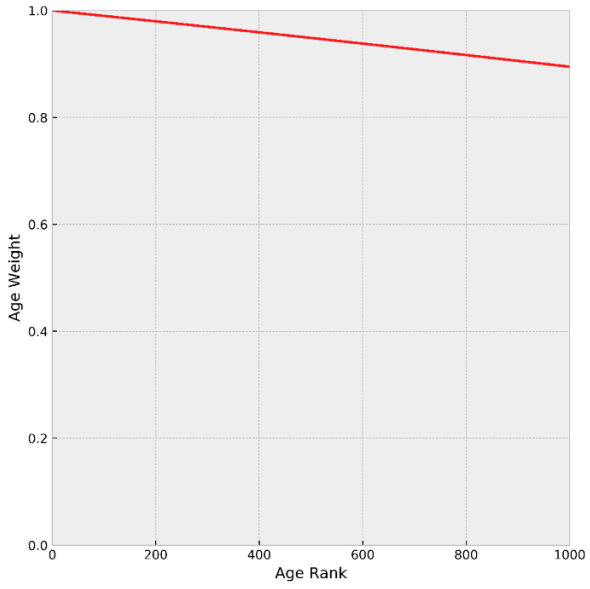


Figure 6.2. Age Weight based on the age rank of the sample.

The Age Weight ensures that the newer samples, which more closely represent the current state of the system, are prioritised in training the model against. The Age Weight equation was formulated so that exponentially less weight is given to older samples, however the

base of the exponent (i.e. 1.0001) is very small so that significant weight is still given to the oldest samples. This way, only if there is a substantial number of newer samples available will old data not be utilised by the model.

It is important to note that the weights described here are only used on the training dataset to train the model and nowhere else.

6.2.7 On-Line Retraining

6.2.7.1 Initial Model Training

When applying this model to a new WDS with no available historical data, new measurement data first needs to be collected from the WDS for the initial training of the model. The amount of time required for sufficient training data to be collected is dependent on the range of events observed taking place in the WDS. Thus, in order to initially train and use the model as soon as possible whilst retaining a minimum level of forecasting performance, the minimum number of samples required for the initial model training was chosen to be at least 2 negative samples and 2 positive samples. Furthermore, only samples occurring no closer than 24 hours of each other can contribute to this minimum sample count. This is to ensure that samples have been collected from a range of events.

6.2.7.2 Subsequent Model Retraining

After the model has been trained on the initial dataset, it can then be used to forecast turbidity events. To maintain forecasting performance on unseen data, newly observed data can be used to periodically retrain the model and thus improve its forecasting performance on unseen data. This process is illustrated in Figure 6.3.

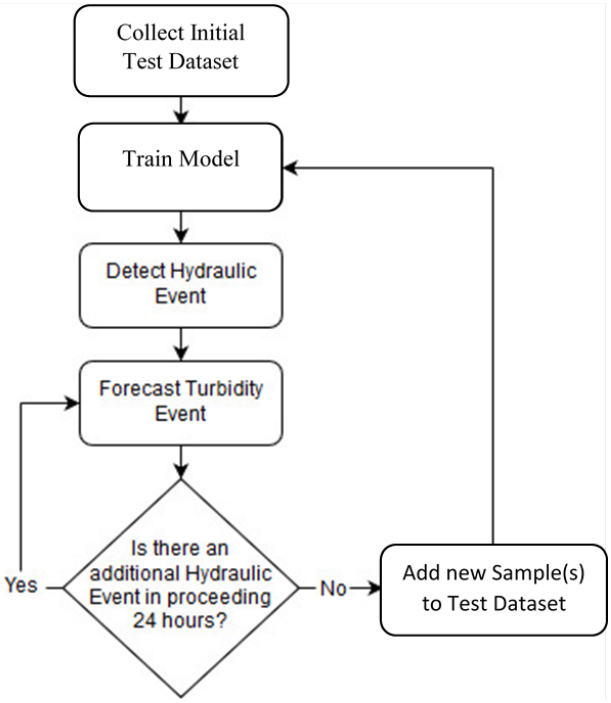


Figure 6.3. Flow diagram of retraining cycle.

The retraining cycle starts by the trained model detecting a hydraulic event and forecasting if there will or will not be a turbidity event within the subsequent 24 hours. The next 24 hours is observed to detect the presence or absence of a turbidity event and thus determine if the forecast was true or false. If any subsequent or simultaneous hydraulic events are detected during this 24-hour period, then turbidity forecasts for those events are also made and the subsequent 24 hours of each forecast is observed to determine if each respective forecast was true or false. These hydraulic events are then added to the training dataset and the model is retrained before detecting further hydraulic events. This retraining cycle repeats so that the model is regularly retrained using the latest observed hydraulic events.

6.3 Case Study Data

The methodology is applied independently to each of the four turbidity meters in the three sites presented in Chapter 3. For convenience, each model will be named after its corresponding turbidity meter.

Table 6.2. The detected number of turbidity events, hydraulic events that caused a turbidity event (i.e. positive samples) and hydraulic events that did not cause a turbidity event (i.e. negative samples) for each turbidity meter using a 7-day hydraulic threshold.

Turbidity Meter	Turbidity Events	Positive Samples	Negative Samples
TM A (Site 1)	17	31	622
TM B (Site 1)	13	18	473
TM C (Site 2)	15	18	153
TM D (Site 3)	19	39	242

From using a 7-day hydraulic threshold as stated in Section 6.2.1, the number of turbidity events, positive samples and negative samples that were detected for each turbidity meter is shown in Table 6.2. As expected, there are more turbidity-causing hydraulic events (i.e. positive samples) than turbidity events. This is because multiple subsequent and simultaneous hydraulic events are occurring before respective turbidity events.

For a turbidity-causing hydraulic event, the time difference between when the turbidity event forecast is made and when the resulting turbidity event starts (i.e. turbidity exceeds 1 NTU) is called the Forecast Lead Time. Figure 6.1 shows two turbidity-causing hydraulic events with forecast lead times of 14 hours and 5 hours and 15 minutes respectively. Figure 6.4 shows the forecast lead times for every turbidity-causing hydraulic event. It can be seen from this figure that while the forecast lead times of turbidity-causing hydraulic events vary significantly, the majority of them have shorter forecast lead times (i.e. < 6 hours). This is not surprising as hydraulic events by definition have higher velocities. The events with longer lead times are likely to have either been caused far upstream in the WDS or had the velocity in the WDS significantly drop off after the hydraulic event.

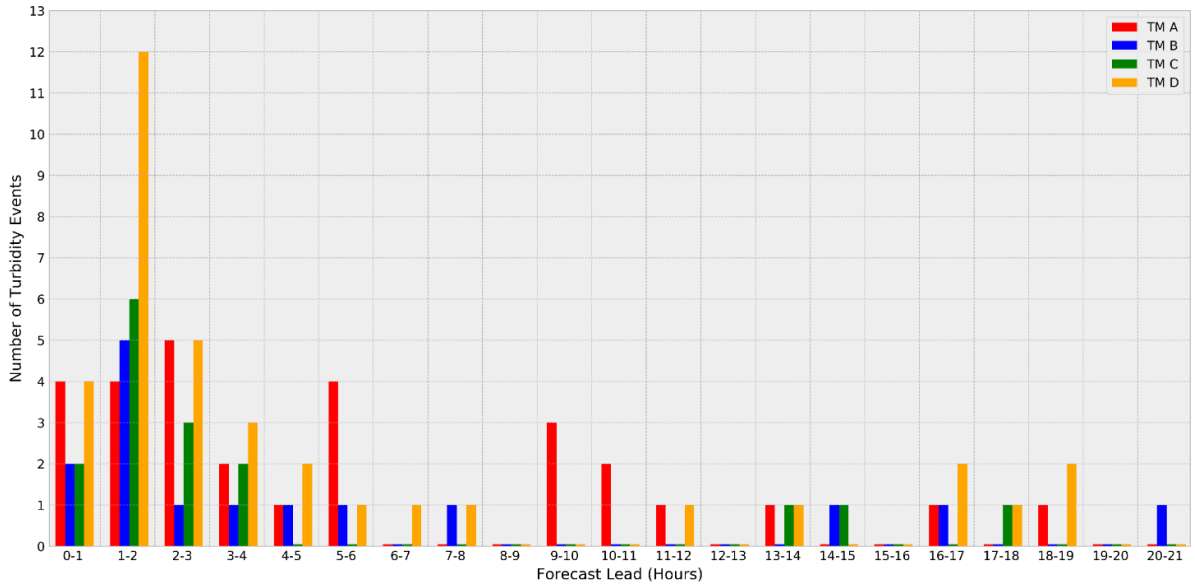


Figure 6.4. Forecast lead times for all positive samples.

Although the RF model forecasts whether or not a turbidity event will occur within 24 hours, the longest forecast lead time for all turbidity-causing hydraulic events seen across all sites is only 20 hours. This indicates that 24 hours is, for the sites analysed here, a sufficiently long enough period of time for any material mobilised from the furthest upstream points of each site to travel to their respective downstream turbidity meters.

As a true negative forecast from the model means that there will not be a turbidity event in the next 24 hours, the forecast lead times for all non-turbidity-causing hydraulic events (i.e. negative samples) can be considered to be 24 hours.

6.4 Results and Discussion

Each model is treated as if initiated on the 1st of September 2013 in the real sites described in Chapter 3 with no historical data available. Each model is continuously run until the 1st of August 2016 (i.e. 2 years and 11 months).

6.4.1 Initial Training

Table 6.3 shows the initial training dataset size and time required for the initial training dataset to be collected for each model. As shown in the table, the initial training dataset size for each model is bigger than the 2 positive and 2 negative sample requirements stated in Section 6.2.7.1. This is due to the rare occurrence of turbidity events in WDSs,

meaning that for the four turbidity meters used here, between 22 and 94 days were required for the collection of at least two positive samples from two different turbidity events.

Table 6.3. Initial training dataset summary for each turbidity event forecasting model.

Model	Initial Training Dataset		Collection Time
	Positive Samples	Negative Samples	Required
TM A (Site 1)	2	60	94 Days
TM B (Site 1)	3	6	22 Days
TM C (Site 2)	2	9	64 Days
TM D (Site 3)	3	5	48 Days

6.4.2 On-Line Forecasting

Table 6.4 shows the dates that each model started and stopped its on-line forecasting, the number of days each model was forecasting for and the number of times each model was retrained based on the all available historical data. Note that the forecasting duration for Model TM D is significantly lower than the other models, this is because the turbidity meter in Site 3 stopped capturing data between July 2014 to November 2014 and then from June 2016 onwards.

Table 6.4. The on-line forecasting start date, end date, duration and number of retraining cycles for each turbidity event forecasting model.

Model	On-Line Forecasting Time			Number of Retraining cycles
	Start Date	End Date	Duration	
TM A (Site 1)	03/12/2013	01/08/2016	973 Days	205
TM B (Site 1)	23/09/2013	01/08/2016	1044 Days	232
TM C (Site 2)	03/11/2013	01/08/2016	1003 Days	131
TM D (Site 3)	18/10/2013	01/06/2016	834 Days	175

The processing time required to retrain the forecasting models increased with the size of the training dataset, however, the final retraining for each model (i.e. the retraining with the largest dataset size) required less than a minute on an Intel Core i7-7820HQ CPU with 16GB of RAM. This shows that retraining to include new data could even be feasibly completed between flow and turbidity capture timesteps (i.e. 15 minutes).

Table 6.5 shows the forecasting models' on-line forecasting performances which is accumulated over the duration of their on-line forecasting periods that was shown in Table 6.4. From very good TPRs and FPRs shown in Table 6.5, it can be seen that all the forecasting models can very accurately forecast if a hydraulic event will cause a subsequent downstream turbidity event.

Table 6.5. The True Positive Rate (TPR) and False Positive Rate (FPR) shown for each model's on-line forecasting.

Model	**Predicted:**	**Actual:** Positive	**Actual:** Negative
TM A	Positive	29/29 **(TPR: 1.00)**	6/562 **(FPR: 0.01)**
	Negative	0/29	556/562
TM B	Positive	15/15 **(TPR: 1.00)**	7/467 **(FPR: 0.02)**
	Negative	0/15	460/467
TM C	Positive	14/16 **(TPR: 0.88)**	2/144 **(FPR: 0.01)**
	Negative	2/16	142/144
TM D	Positive	33/36 **(TPR: 0.92)**	11/237 **(FPR: 0.05)**
	Negative	3/36	226/237

The two false positives for model TM C had lead times of 2 hours and 13 hours 15 minutes, while the three false positives for model TM D had lead times of 3 hours 30 minutes, 6 hours 30 minutes and 13 hours. By comparing the spread in lead times of these false positives to the lead times for all positive samples shown in Figure 6.4, there is no indication that the reason the models incorrectly forecasted these positive samples was due to their forecast lead times. Furthermore, while the models do not predict the exact time that the turbidity event will appear at a downstream end of the network, Model TM B correct forecasted a turbidity event that occurred 20 hours after the mobilising hydraulic event occurred.

As mentioned in Section 6.2.1, turbidity events can have more than one preceding hydraulic event, and further investigation of the false positive forecasts showed that the two false positives of model TM C came from failing to forecast the same single turbidity event and the three false positives of model TM D also came from failing to forecast a single turbidity event. The number of forecasted turbidity events that were accurately forecasted by the models is shown in Table 6.6. From this it can be seen that across all models only two of the 64 forecasted turbidity events were incorrectly forecasted.

Table 6.6. The accurately of forecasted turbidity events by each model and the accuracy of all turbidity events (i.e. including turbidity events that were not forecasted due to them not having a mobilising hydraulic event) by each model.

Model	Accuracy of Forecasted Turbidity Events	Accuracy of All Turbidity Events
TM A	17/17 **(TPR: 1.00)**	17/19 **(TPR: 0.89)**
TM B	13/13 **(TPR: 1.00)**	13/18 **(TPR: 0.72)**
TM C	14/15 **(TPR: 0.93)**	14/25 **(TPR: 0.56)**
TM D	18/19 **(TPR: 0.95)**	18/28 **(TPR: 0.64)**

While the forecasting results of forecasted turbidity events show very accurate forecasting performances, note that because the models are forecasting if a hydraulic event will or will not cause a turbidity event, only turbidity events with preceding hydraulic events (i.e. hydraulically mobilised turbidity events) are actually forecasted. In other words, all non-hydraulically mobilised turbidity events were not forecasted by the models. Table 6.6 shows the adjusted TPRs that include these non-forecasted turbidity events to show the TPRs of the models for all turbidity events, regardless of whether these are hydraulically or non-hydraulically caused. Note that the FPRs from Table 6.5 remain the same.

Unsurprisingly, the TPRs for all turbidity events are significantly lower than the TPRs for forecasted turbidity events. However, an argument could be made for removing these non-hydraulically mobilised turbidity events from consideration here, as by definition it is impossible to forecast a turbidity event from hydraulic data when said turbidity event has no preceding hydraulic response.

6.4.3 Proactive Flow Management

If the flowrates in the network can be automatically controlled (i.e. through pumps, actuated valves, etc) via a central SCADA system, as is in fact the case with all three sites presented here, then proactive actions may be taken to prevent or limit the magnitude of the resulting discolouration event. This would be done by reducing the flowrate as soon as the model forecasts that a hydraulic event will cause a downstream turbidity event. Note that care may need to be taken in some situations to ensure that a flow reduction does not cause a significant water interruption to the downstream supply and result in a different type of customer complaint (ex. Low Pressure).

Figure 6.1 showed two hydraulic events and the downstream turbidity events that they caused. Both those turbidity events were not only correctly forecasted by their respective models from Section 6.4, but they also likely would have not reached the peak turbidity levels that they are shown to have reached if the magnitude of their preceding hydraulic events had been immediately limited. Figure 6.5 shows another two correctly forecasted turbidity events chosen to demonstrate this point.

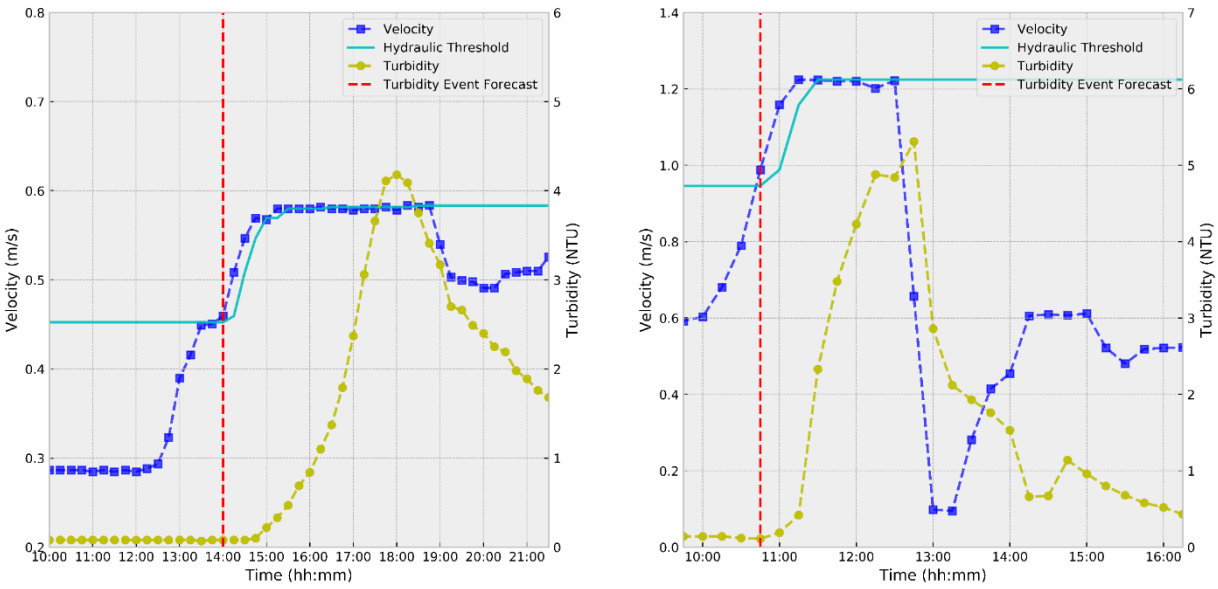


Figure 6.5. Two hydraulic events and their subsequent downstream turbidity events shown from (a) Site 1, Pipe D, TM A and (b) Site 2, Pipe H, TM C. The dotted red lines mark when the turbidity event forecasts occurred. Both forecasts were correctly made by their respective models. NTU: Nephelometric Turbidity Units.

Appendix A contains additional examples turbidity events that were correctly forecasted by their respective models.

Figure 6.5a shows that the velocity continued to rise for an hour and 30 minutes after model TM A correctly forecasted that a turbidity event will occur. Figure 6.5b likewise shows that the velocity continued to rise for 45 minutes after model TM C correctly forecasted that a turbidity event will occur. If the velocity was reduced upon or soon after the forecasts in Figure 6.5a and Figure 6.5b, it is likely that the turbidity levels for both would not have exceeded the 4 NTU UK regulatory limit for turbidity at a customers' taps.

It should be noted however that even if turbidity events are correctly forecasted by the methodology here, some turbidity events may not be preventable or limitable. Figure 6.6 shows a turbidity event which appears to be such a case.

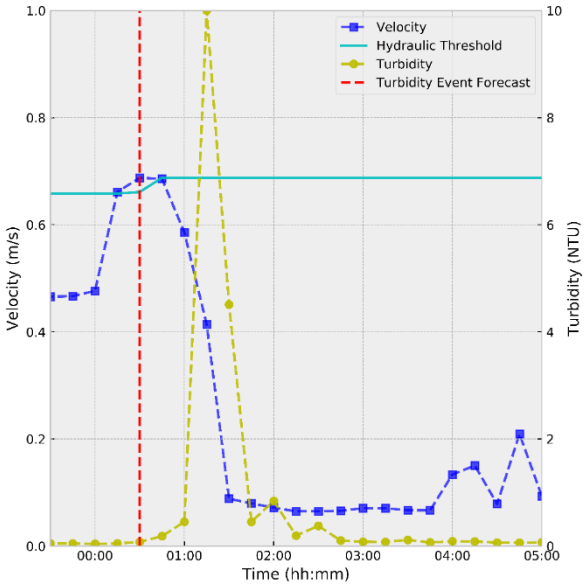


Figure 6.6. A hydraulic event and its subsequent downstream turbidity event shown from Site 3, Pipe K, TM D. The turbidity event forecast was correctly made by its respective model. NTU: Nephelometric Turbidity Units.

As the turbidity event forecast takes place at the peak of the hydraulic event, the subsequent decrease in velocity does not prevent a sudden short-lived turbidity event from exceeding 10 NTU. At which point the only option left is to warn customers of possible discolouration issues in their area.

6.4.4 Simulating Increased Velocities

While the models only forecast upon a hydraulic threshold being exceeded (i.e. hydraulic event occurring), the models can be made to continuously forecast. This allows for proactive velocity management by introducing artificially inflated velocities into the model and testing for the bounds of which velocities the model will predict a turbidity event will occur. This can be done by taking the current velocity at every timestep for a pipe and increasing it either by a fixed amount or by a percentage of the current velocity. While any amount can be chosen, Figure 6.7 shows this happening with a 20% inflated velocity.

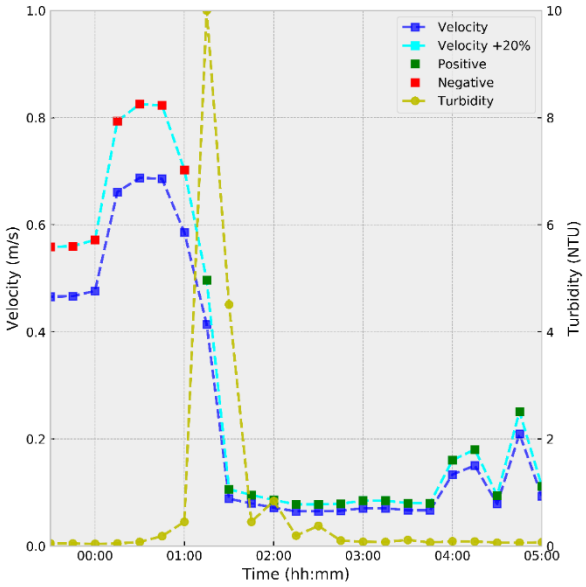


Figure 6.7. A turbidity event forecasting model continuously forecasting at each timestep using velocities from Site 3, Pipe K and artificially inflated them by 20% shown. NTU: Nephelometric Turbidity Units.

The positive labelled green squares mean that the model predicts that no turbidity event will be generated if the current velocity remains below that value. Whereas the negative labelled red squares mean that the model predicts a turbidity event would occur at the downstream turbidity meter if the current velocity reached or exceeded this value. Figure 6.7 shows an example of this in practice as the model indicated that a turbidity event would occur if the velocity increased, like it did at 00:15 and a resulting turbidity event was observed. While this could be used for a range of possible applications, such as testing potential flow rates in a network prior to maintenance/rehabilitation works, the reliability of these forecasts is unknown.

6.5 Chapter Summary

This chapter presented an on-line turbidity event forecasting methodology that was able to analyse the velocity in WDS pipes to detect the start of hydraulic events and then predict if a hydraulic event will cause a downstream turbidity event within the subsequent 24 hours that at least exceeds 1 NTU in peak magnitude. Results were tested using four turbidity meters from three real water distribution networks over the course of 2 years and 11 months.

The methodology shows accurate forecasting performance with TPRs between 1 and 0.88 and FPRs between 0.01 and 0.05. It is noted however that only hydraulically mobilised turbidity events were forecasted and thus not all turbidity events observed by the turbidity meter were forecasted. Inclusion of these non-hydraulically mobilised turbidity events significantly decreased the TPRs for all models.

If the flowrates in the network can be automatically controlled then proactive actions may be possible to prevent or limit the magnitude of a discolouration event. This is done by reducing the flowrate as soon as the model forecasts that a hydraulic event will cause a downstream turbidity event. If flowrates in a WDS are determined by water demands and cannot be controlled, customers can still be warned via text or email of the discolouration issue in their area.

The methodology presented in this chapter could enable novel proactive management strategies that can be implemented as an alternative to expensive trunk mains cleaning programs. This may not only directly reduce the number of discolouration contacts but also potentially aid indirectly by preventing discolouration build up in downstream distribution pipes.

Chapter 7 Summary, Conclusions and Future Work Recommendations

7.1 Introduction

This chapter provides a summary of the work done, conclusions derived from the resulting chapters in this thesis and discusses their significance. This chapter closes by highlighting the novel contributions of this thesis and describing several recommendations for further research.

7.2 Summary and Conclusions

Water discolouration in WDSs is an increasingly important issue in the UK and abroad. The increase in sensors and data acquisition systems installed in WDSs has led to an unprecedented amount of real-time data now produced by water companies. This substantial increase in real-time data in WDSs has opened the door to using data-driven methods to achieve what physically based models have to date been unable to.

The research presented in this thesis has resulted in the development of three novel data-driven methodologies to aid in the reduction of discolouration risk to water utility customers and further understand the underlying causes of discolouration in a WDS.

7.2.1 Short-Term Turbidity Forecasting

The first of these methodologies, presented in Chapter 4, is a data-driven continuous turbidity forecasting methodology capable of forecasting if and when turbidity will exceed a predefined threshold (i.e. if and when a turbidity event will occur). The methodology took current and past velocity and turbidity measurements at a number of locations in the network to forecast if the turbidity level at the downstream end of the network is above a pre-specified threshold. To do this, the forecasting model had to detect if discolouration material was mobilised, estimate if sufficient turbidity would be generated to exceed a preselected threshold and approximate how long the material will take to reach downstream.

In developing the methodology, three threshold values of 1, 2 and 4 NTU were used to define the occurrence of different discolouration events, and different forecasting horizons ranging from 5 minutes to 8 hours ahead were analysed. Furthermore, an

assessment of three data-driven turbidity forecasting models based on Artificial Neural Networks, Random Forests and Support Vector Machines was made. A regression based approach in directly predicting turbidity levels was also developed alongside the classification approach.

The following can be concluded from the results obtained:

- It is possible to forecast the occurrence of turbidity above some pre-specified threshold and hence detect the corresponding discoloration event in a real system by using a data-driven (i.e. non-physically based) methodology only. Forecasts of turbidity could be reliably made up to 5 hours ahead in the site examined, however forecasting further ahead was at the expense of increased false alarm rates. An examination of the results showed that the model struggled to predict turbidity travel time to the downstream meter in the site examined and not whether a discolouration event had occurred;
- A comparison of results obtained by the two prediction approaches, regression based direct prediction of turbidity and classification based prediction of turbidity exceeding a pre-specified threshold. Of which the latter approach showed substantially better performance than the former;
- Of the three data-driven methods developed, the Random Forest based models performed the best overall.

This methodology could be implemented as an alternative to expensive trunk mains cleaning programs and enable new proactive management strategies in WDSs to reduce discolouration contacts.

7.2.2 Assessing Discolouration Mobilisation in Trunk Main Networks

The second of the three novel methodologies in this thesis, and presented in Chapter 5, is a methodology for assessing discolouration mobilisation in WDSs. The methodology can estimate the percentage of downstream turbidity observations that can be linked to an upstream pipe and thus show where discolouration material is likely to have originated. This identifies network areas (i.e. pipes) that discolouration material accumulates in, and thus aids in targeted proactive cleaning or replacement of problem pipes.

The methodology does this by examining the time preceding a downstream turbidity observation for if a hydraulic force in an upstream pipe had exceeded the pipe's

hydraulically conditioned state. This methodology does not make any assumptions about what the discolouration material consists of (e.g., manganese, biofilms), what form the discolouration material takes inside pipes (e.g., sediment, cohesive layers), nor does it assume a rate at which discolouration material is mobilised (e.g., linear, exponential). This means, in theory, the application of the methodology to a real WDS is not limited by the material composition, layout and range of flow rates of that WDS, and thus transferable to any WDS with the requisite data.

This methodology was also used to investigate the extent to which typical hydraulic conditions could mobilise discolouration material in a long-term continuous study. The study used data collected from three UK trunk main networks over a period of two years and 11 months. The results of the study showed that most of downstream turbidity observations over 1 NTU could be paired with a preceding hydraulic force that exceeded an upstream pipe's hydraulically preconditioned state. This empirically showed that the mobilisation of discolouration material in the three sites examined is predominantly determined by the variation of hydraulic forces in pipes. These findings further highlight the potential benefits in proactively managing the variance of hydraulic forces to reduce discolouration risk and improve customer service.

7.2.3 On-Line Turbidity Event Forecasting

The final methodology in this thesis, presented in Chapter 6, is an on-line turbidity event forecasting methodology that built upon the two aforementioned methodologies to reliably predict if a hydraulic event occurring upstream will cause a turbidity event downstream.

This methodology analyses the historic and current hydraulic forces in WDS pipes to detect the start of a hydraulic event and predict if that hydraulic event will cause a downstream turbidity event in the subsequent 24 hours. The prediction is made at the start of the hydraulic event to provide as much lead time as possible for the necessary action(s) to be taken, i.e. so that the magnitude and duration of that hydraulic event can be limited by e.g. closing valves, turning off pumps, re-routing discoloured water elsewhere, etc. This way it may be possible to limit or even prevent the forecasted turbidity event from occurring. Even if the flows in the WDS cannot be controlled, customers can still be warned via text or email of the discolouration issue in their area before the discolouration event takes place.

A factor that limited the forecasting accuracy of the short-term turbidity forecasting models presented in Chapter 4, was that they struggled to predict the exact time discolouration material would arrive at the downstream ends of the network. By forecasting the arrival of downstream turbidity events to be within a subsequent 24-hour period, this methodology was able to significantly improve its forecasting accuracy.

The methodology for assessing discolouration mobilisation in Chapter 5 showed that not all turbidity events could be linked to a preceding upstream hydraulic event. Thus, the on-line turbidity event forecasting methodology was developed to train and forecast starting from the detection of hydraulic events and not turbidity events. This way the methodology automatically avoids training using turbidity events that have no corresponding preceding hydraulic response.

After the on-line turbidity event forecasting models were trained on the initial dataset, to maintain forecasting performance on unseen data, newly observed data was used to periodically retrain the model and thus improve its forecasting performance on unseen data.

From the three sites examined, the on-line turbidity event forecasting models showed accurate and reliable performance in forecasting hydraulically mobilised turbidity events. This strong forecasting performance came at the expense of not forecasting any non-hydraulically mobilised turbidity events. However, an argument is made for excluding non-hydraulically mobilised turbidity events from consideration, as by definition these events cannot be forecasted from hydraulic data alone.

This methodology could be used as an early warning system for discolouration events and thus enable a multitude of cost beneficial proactive management strategies to be implemented as an alternative to expensive trunk mains cleaning programs. This may not only directly reduce the number of discolouration contacts but also potentially aid indirectly by preventing discolouration build up in downstream distribution pipes.

7.3 Summary of Novel Contributions

The work carried out in this thesis is based on the application of existing machine learning and statistical methods. These methods are synergistically combined in novel methodological frameworks to forecast discolouration events and further the understanding of discolouration processes in WDSs. Therefore, this thesis has the cross-

disciplinary nature of Hydroinformatics and should be seen as an attempt to find synergies between AI methods and discolouration risk management.

The research work carried out in this thesis has contributed to the existing body of knowledge in several distinct ways, with the main contributions of the work presented in this thesis as follows:

- The development of a short-term turbidity forecasting methodology that can forecast turbidity above pre-specified thresholds and hence detect the corresponding discoloration event in a real system by using only a data-driven (i.e. non-physically based) approach. It is the first verified discolouration model capable of predicting turbidity continually in a pipe network and the first verified turbidity forecasting methodology that does not require a hydraulic model.
- The development of a novel methodology for assessing discolouration mobilisation in WDSs and approximate the percentage of downstream turbidity observations that can be linked to an upstream pipe and thus show where discolouration material is likely to have originated.
- A novel and unprecedentedly long continuous study into the extent to which typical hydraulic conditions could mobilise discolouration material from three UK trunk main networks over a period of two years and 11 months. Results of the sites examined showed empirical evidence that the mobilisation of discolouration material is predominantly determined by the variation of flows and related hydraulic conditions/forces in pipes.
- The development of a novel on-line turbidity event forecasting methodology that analyses the historic and current hydraulic forces in WDS pipes to detect the start of hydraulic events and predict if the hydraulic event will cause a downstream turbidity event in the subsequent 24 hours. This methodology could be used as an early warning system for discolouration events and through use of the methodology it may be possible to limit or even prevent forecasted turbidity events from occurring.

The research work carried out in this thesis could enable new cost beneficial proactive management strategies by water utilities to reduce discolouration risk in WDSs.

7.4 Recommendations for Further Research

The focus of future work should be primarily centred around further testing of the methodologies presented in this thesis. This should involve testing on additional real systems of varying size, layout, range of operational conditions and material composition and by incorporating hydraulic and turbidity events of significantly varying characteristics.

Alternative methods could be exploring to improve the overall accuracy of the short-term turbidity forecasting methodologies and attempting to forecast further into the future. Greater lead times in an early warning system would increase the usefulness of the system to operational staff. One possible route to doing this is by investigating the hybridisation of data-driven models and physically based models. While this may limit the application domain of the methodology to WDSs where hydraulic models are readily available, there is potential for a synergetic approach may to greatly improve turbidity event forecasting.

Currently the forecasting methodologies presented in this thesis have limited alarm capabilities. This could be improved by changing to a probability-based system whereby each forecast provides a confidence ranking of the forecast and operational staff could choose different proactive strategies based on the certainty of the event occurring. Likewise, a probability estimate of how long before the turbidity event appears occurs could be taken into consideration.

Collecting and assessing more turbidity data from numerous WDSs could led to the development of a discolouration prediction model that does not require turbidity data to make forecasts.

Other minor opportunities for future related work that have potential to significantly improve forecasting performances, some of which were discussed earlier in the thesis, are summarised below:

- Comparing the impact of using sub fifteen-minute frequency velocity and turbidity data in forecasting turbidity events and examine to what degree sharp, but short-lived velocity spikes mobilise discolouration material.
- Comparing the impact of using either instantaneous or averaged velocity and turbidity readings in forecasting turbidity events.

- Further examination of the trade-off between of the TPR and FPR in forecasting turbidity events.

APPENDIX A: Additional Examples of Forecasted Turbidity Events

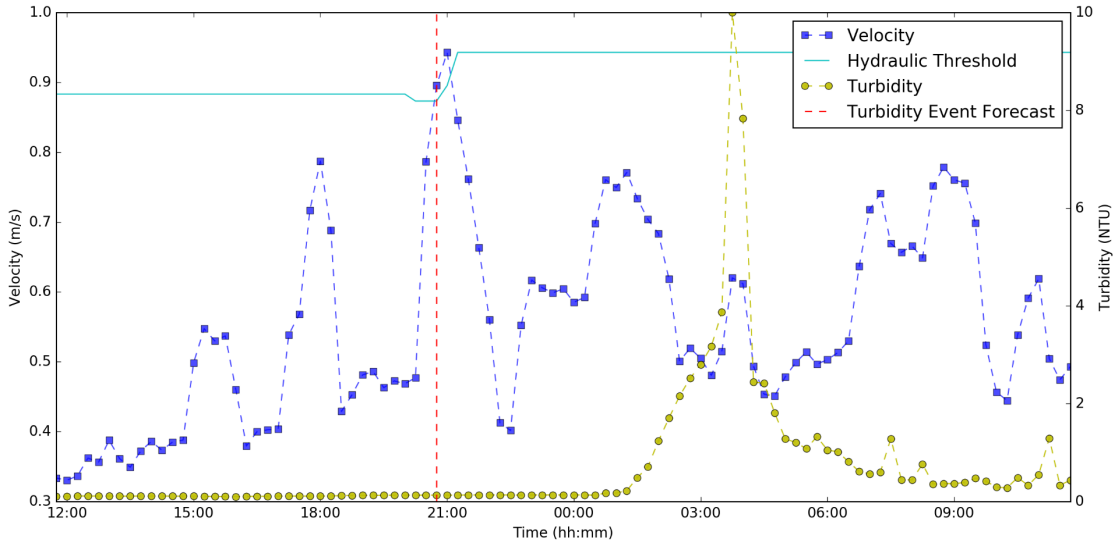
The on-line turbidity event forecasting models presented in Chapter 6 were able to forecast with good accuracy if a turbidity event would occur sometime with the proceeding 24 hours. The four forecasting models were named based on the which turbidity meter data they were using and thus were named TM A, TM B, TM C and TM D.

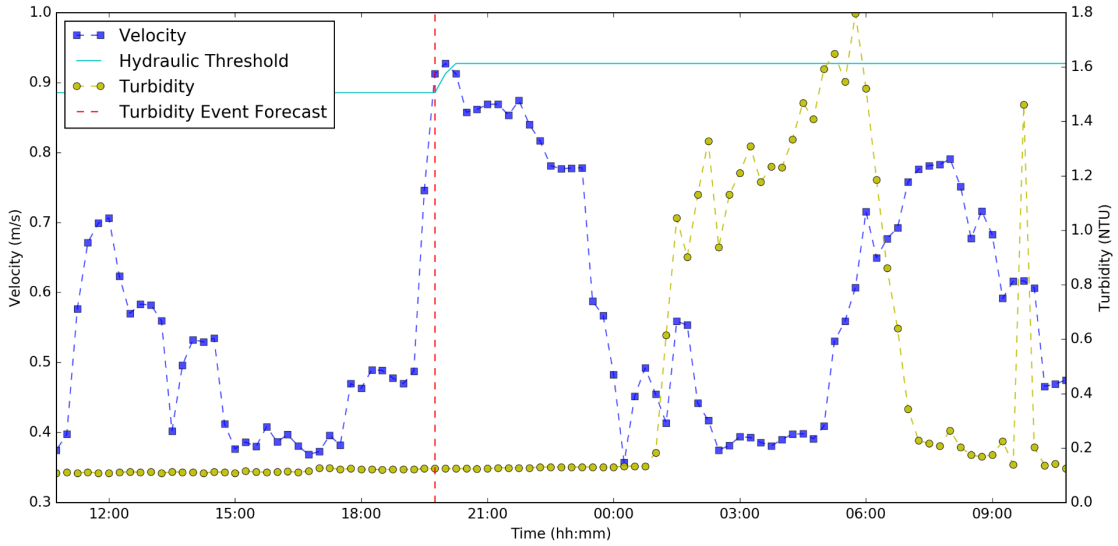
This Appendix A contains 40 turbidity events that were correctly forecasted by their respective forecasting models for each model (10 turbidity events per model).

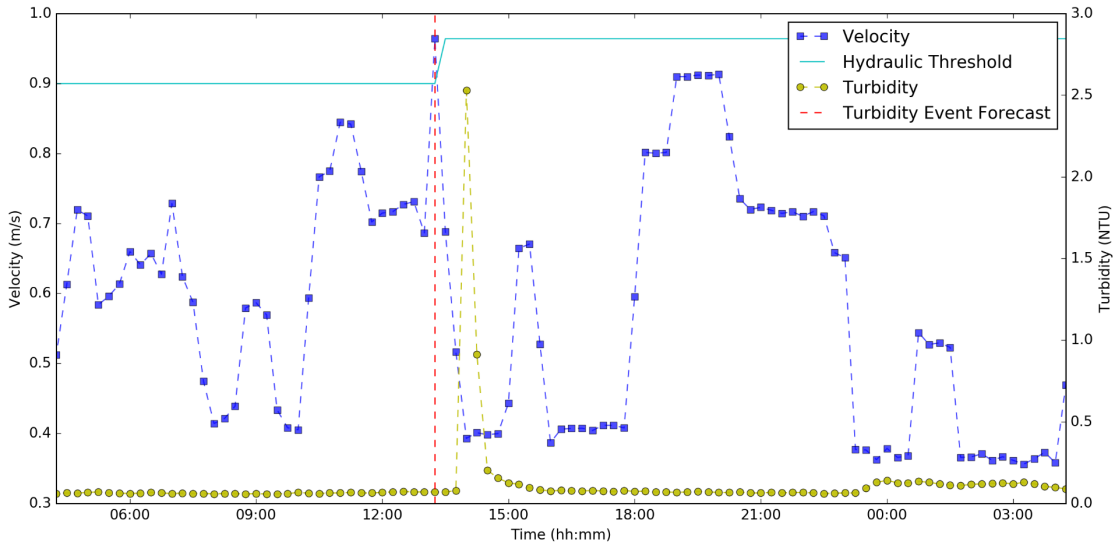
Note that just because a turbidity event was correctly forecasted by the model, this does not guaranty that the hydraulic event that triggered the model to make the forecast is the sole cause and contributor of the turbidity event.

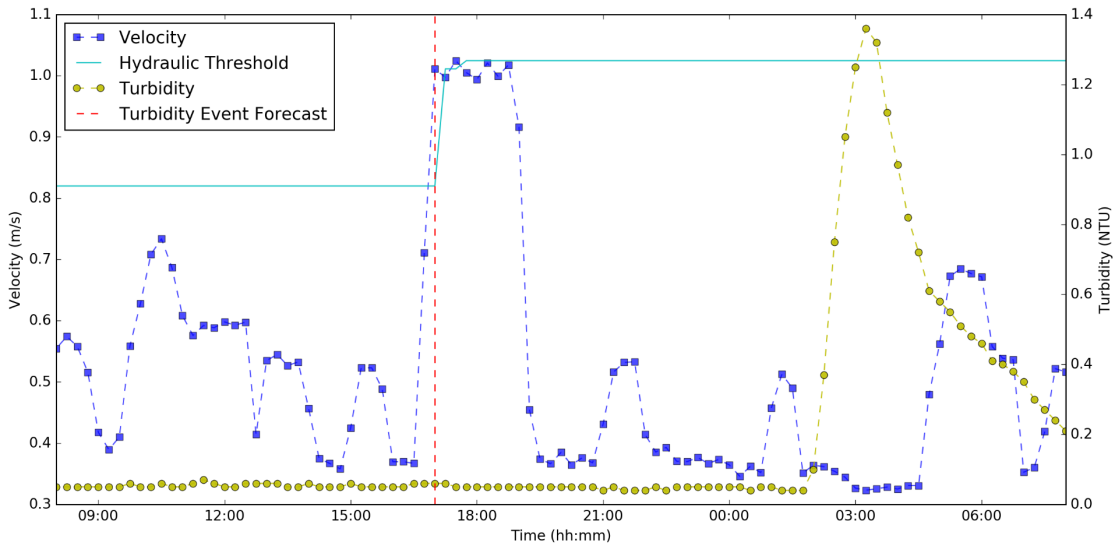
The below figures each show 24 hours focused around the forecast. This is from 9 hours prior to the forecast until 15 hours after the forecast.

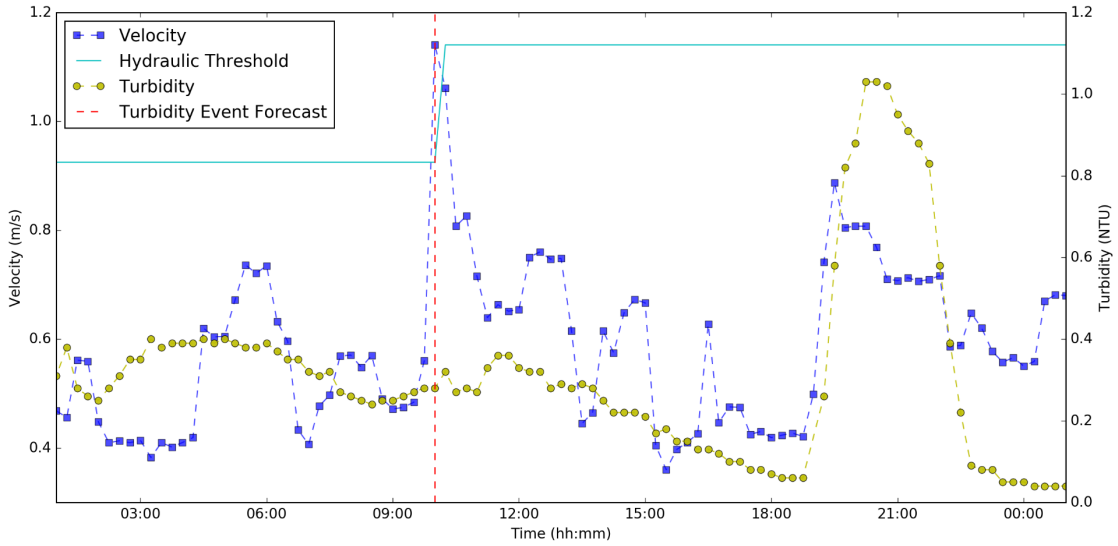
Turbidity Forecasting Model TM A

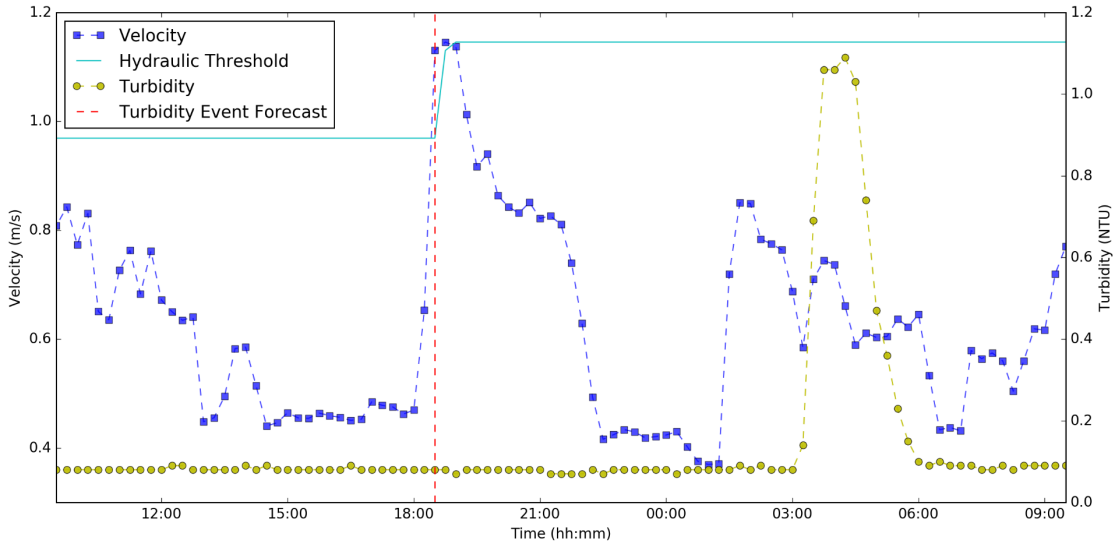


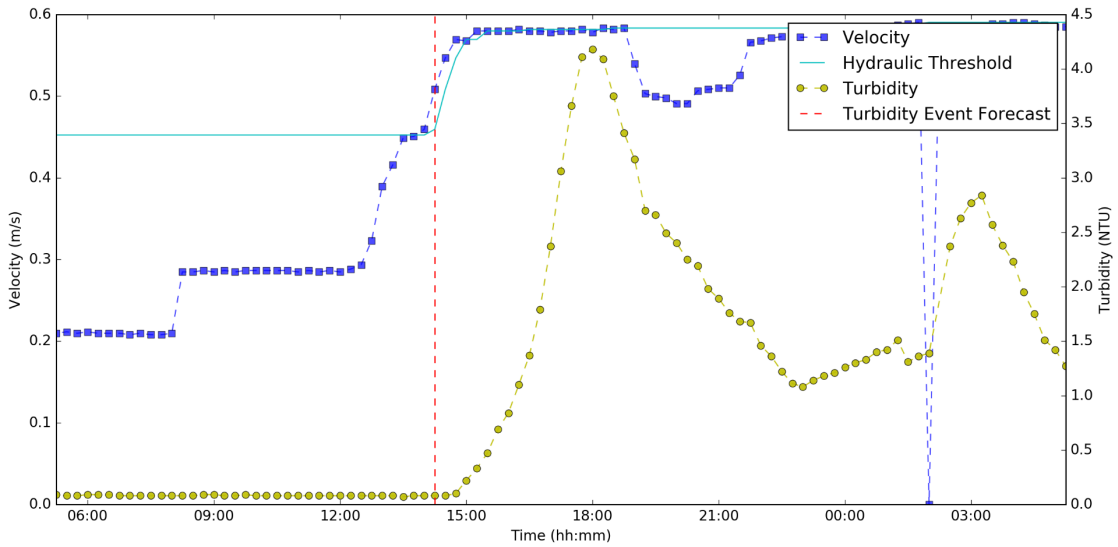

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

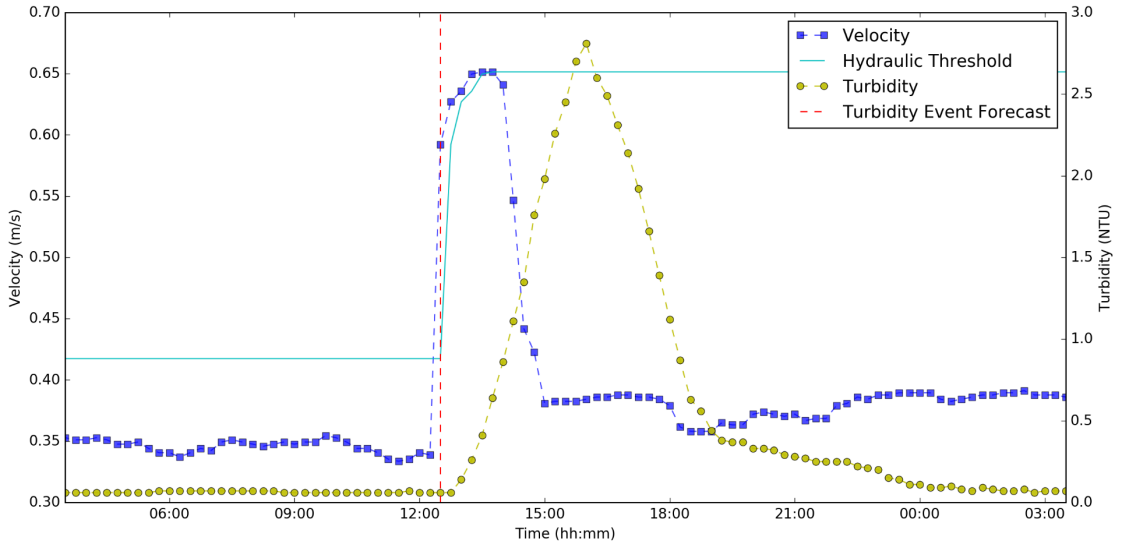

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
06:00
09:00
12:00
15:00
18:00
21:00
00:00
03:00

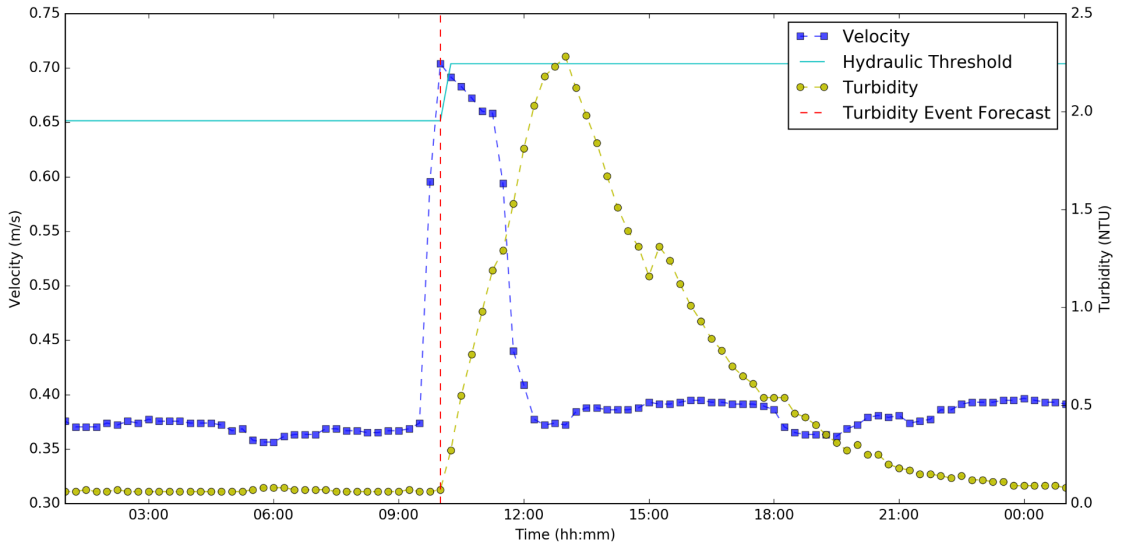

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
09:00
12:00
15:00
18:00
21:00
00:00
03:00
06:00

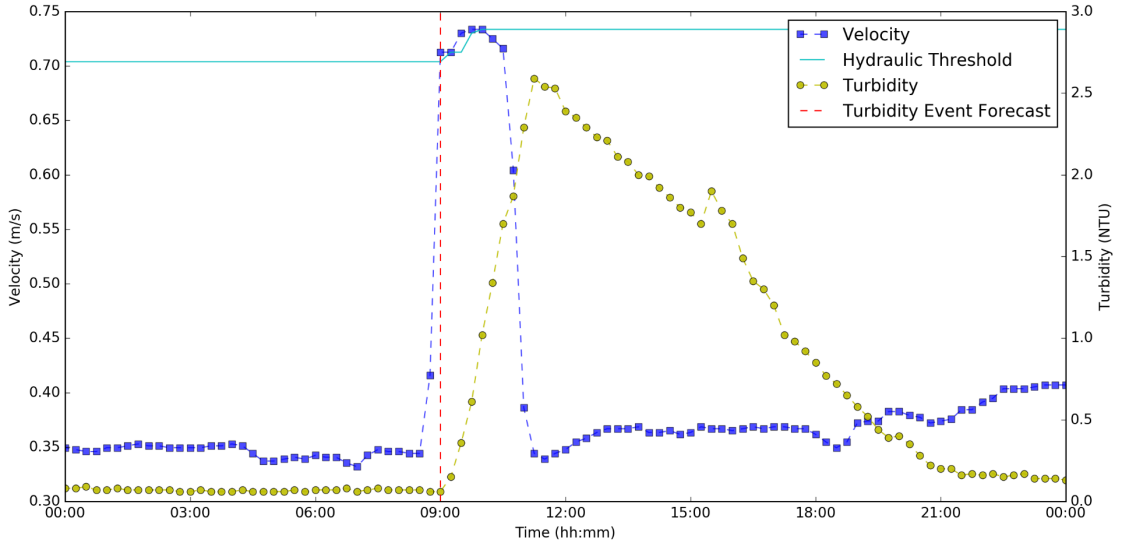

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
03:00
06:00
09:00
12:00
15:00
18:00
21:00
00:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

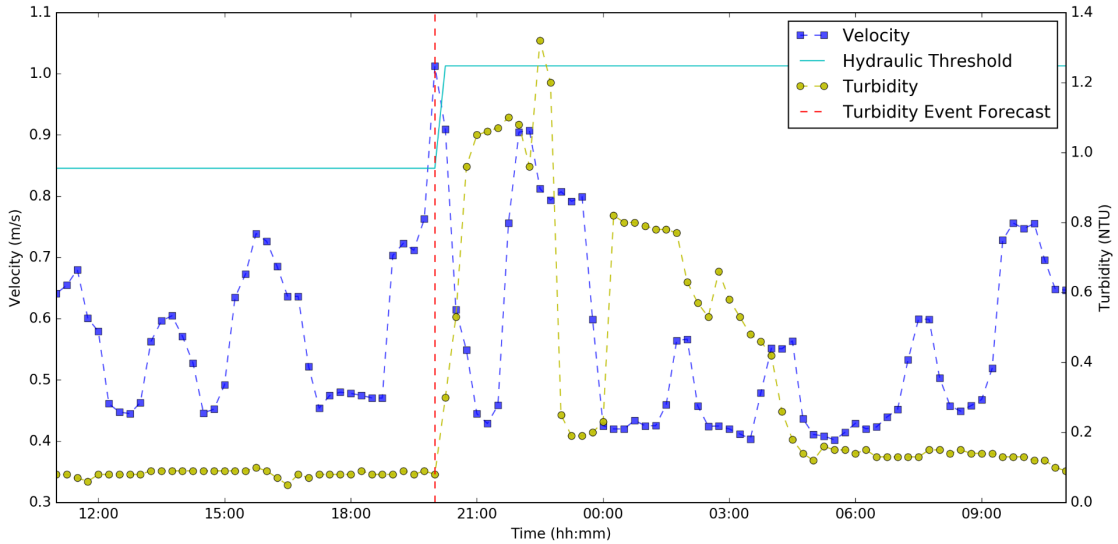

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
06:00
09:00
12:00
15:00
18:00
21:00
00:00
03:00

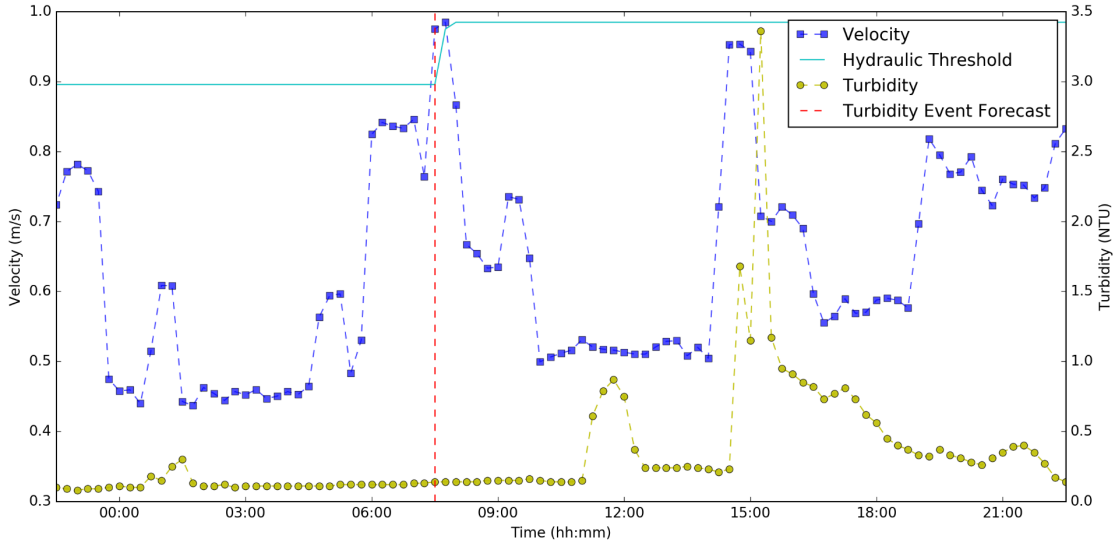

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
3.0
2.5
2.0
1.5
1.0
0.5
0.0
06:00
09:00
12:00
15:00
18:00
21:00
00:00
03:00

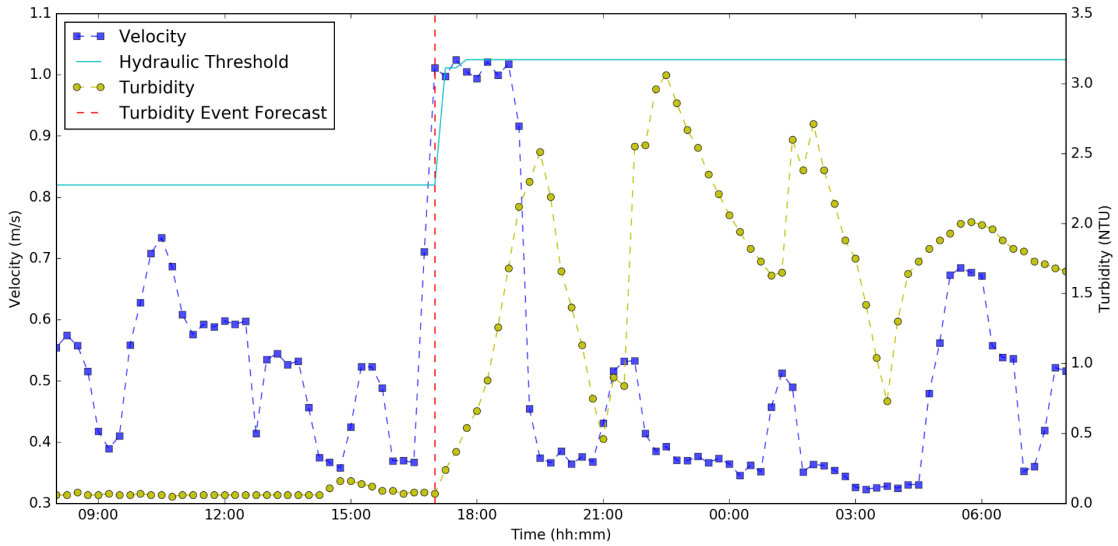

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
2.5
2.0
1.5
1.0
0.5
0.0
03:00
06:00
09:00
12:00
15:00
18:00
21:00
00:00

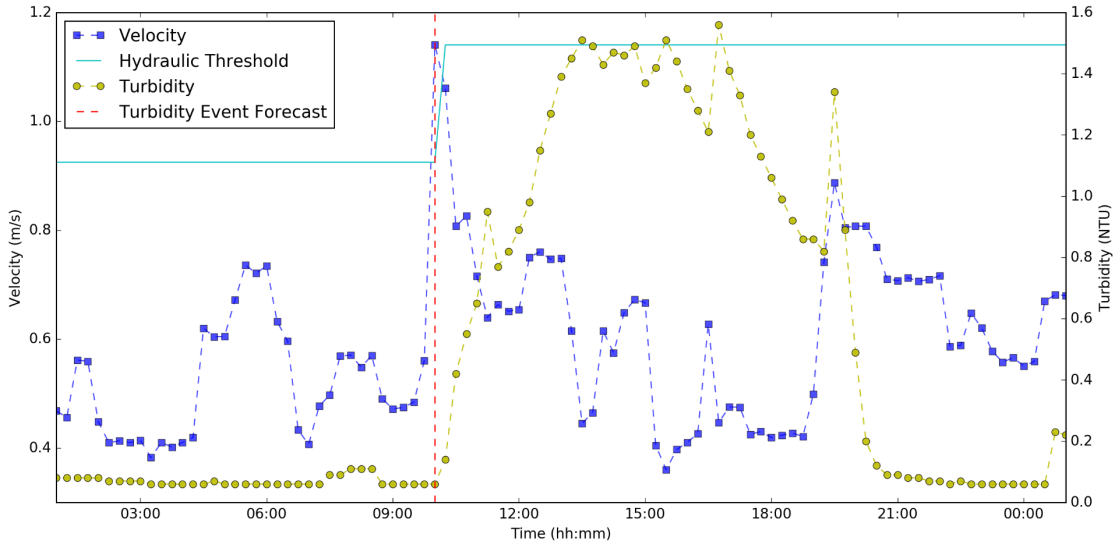

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0.30
3.0
2.5
2.0
1.5
1.0
0.5
0.0
00:00
03:00
06:00
09:00
12:00
15:00
18:00
21:00
00:00

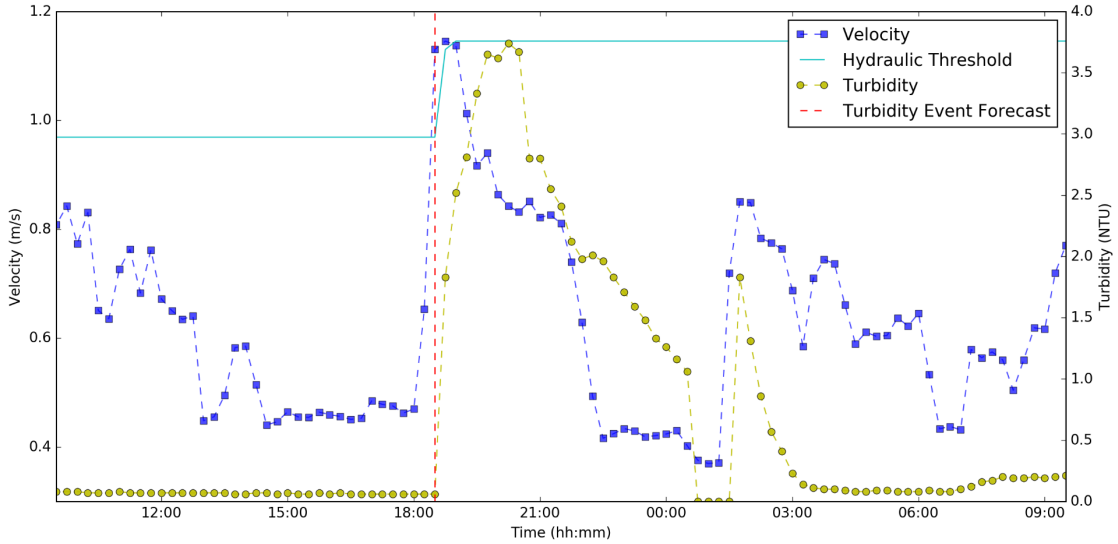
Turbidity Forecasting Model TM B

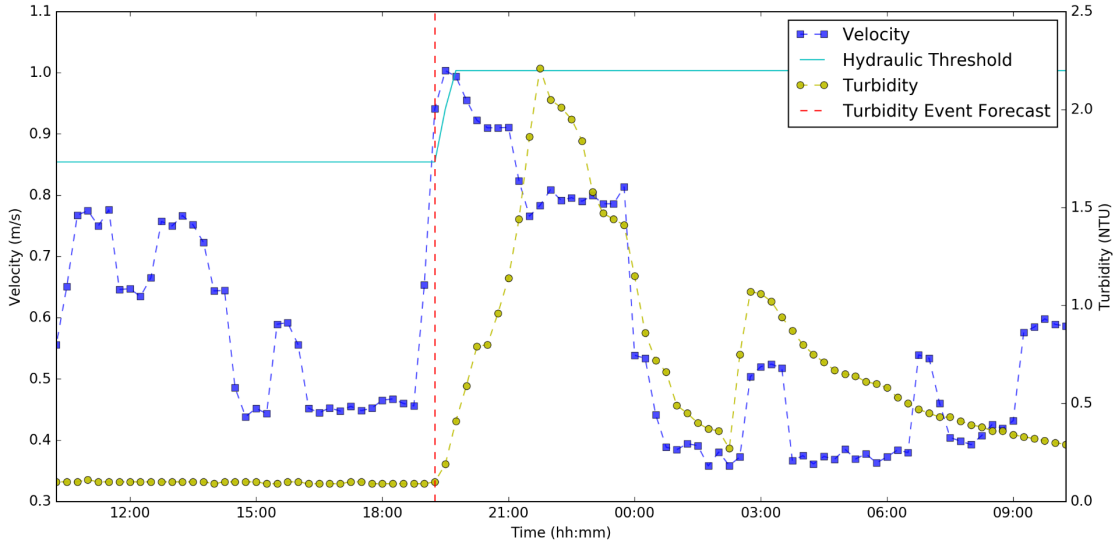


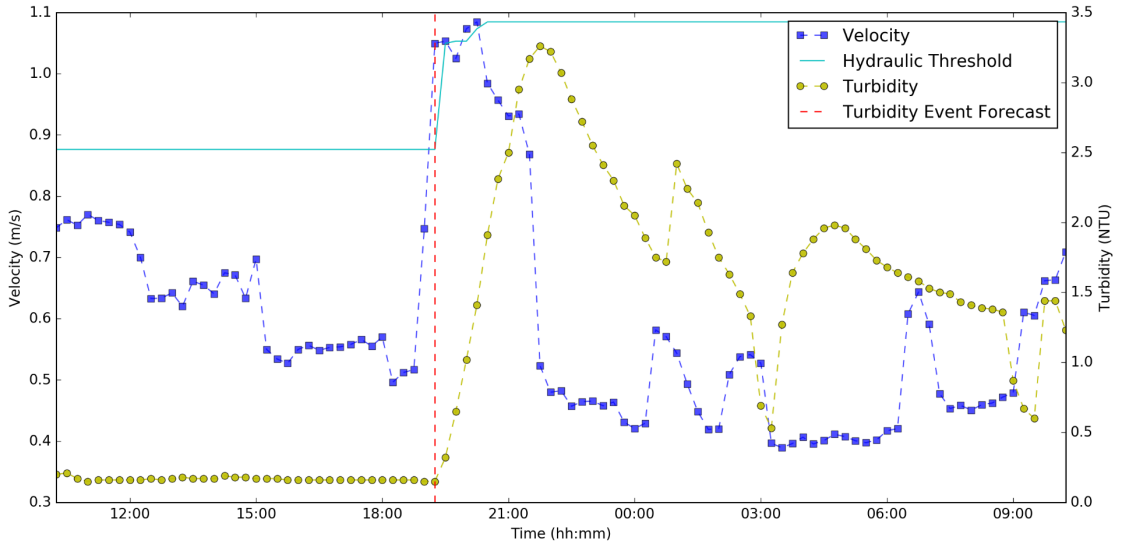


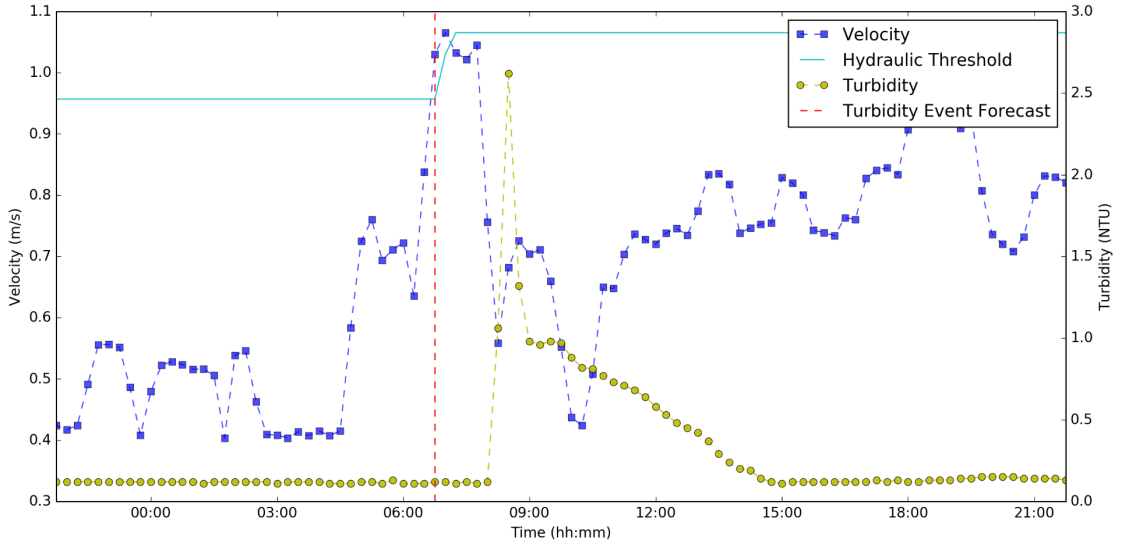


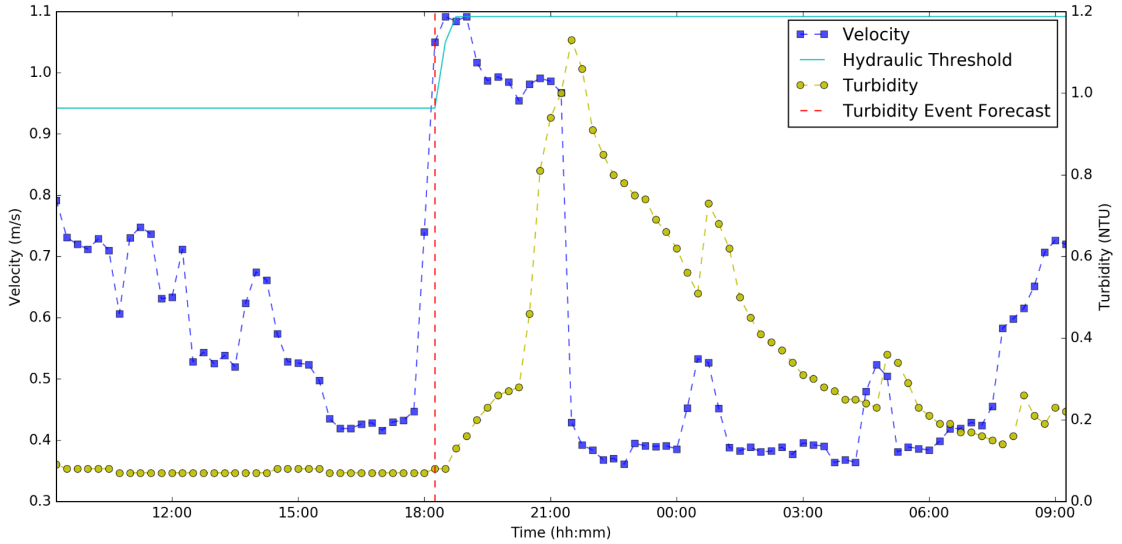

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
03:00
06:00
09:00
12:00
15:00
18:00
21:00
00:00

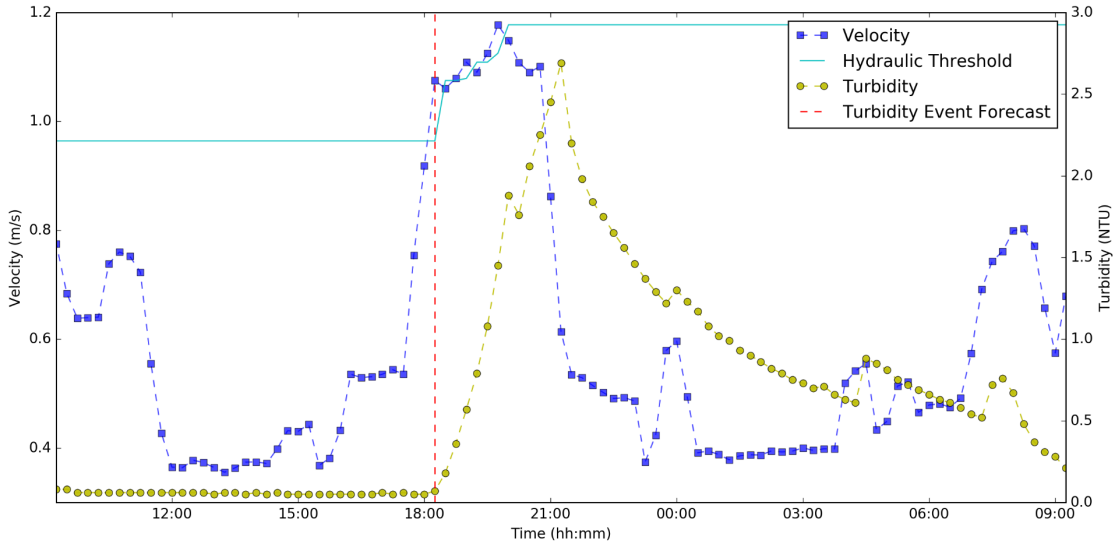

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

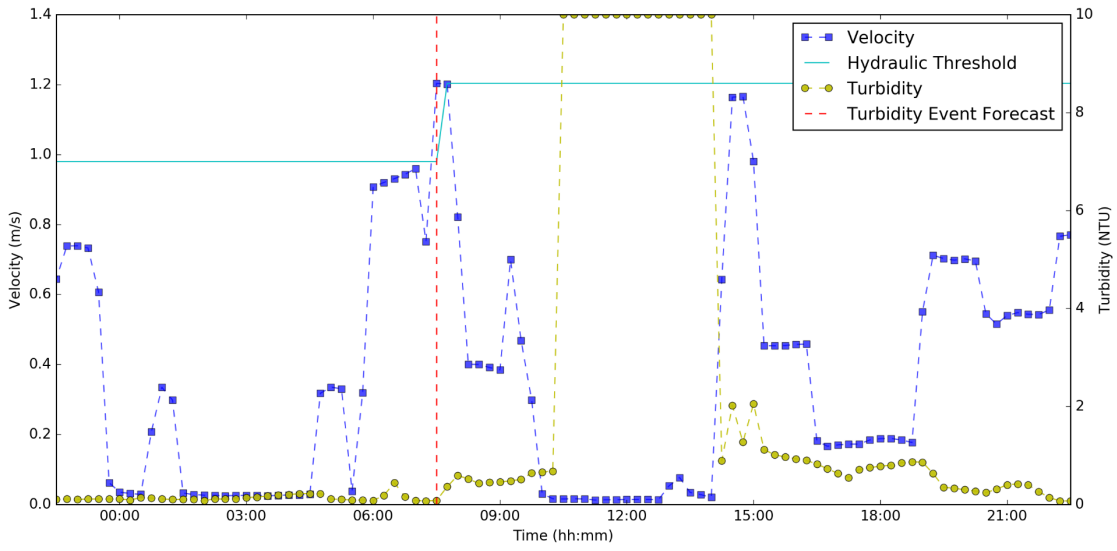

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

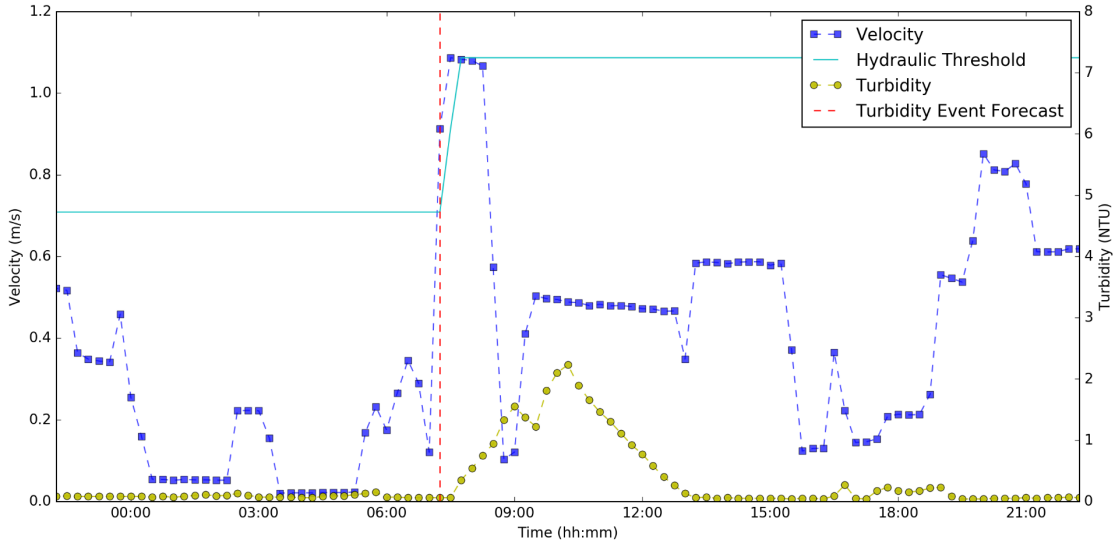

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
00:00
03:00
06:00
09:00
12:00
15:00
18:00
21:00

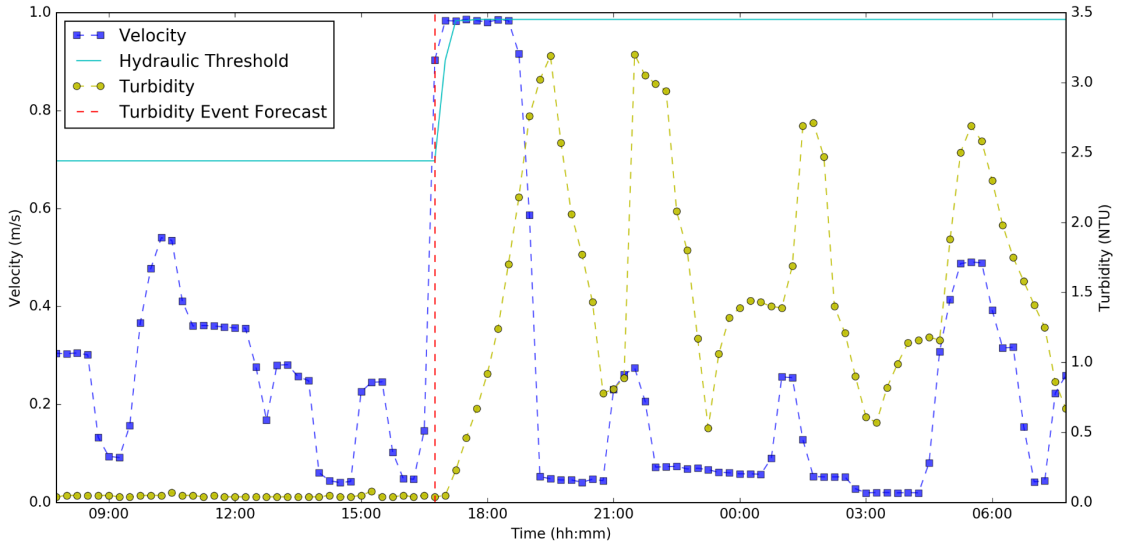

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

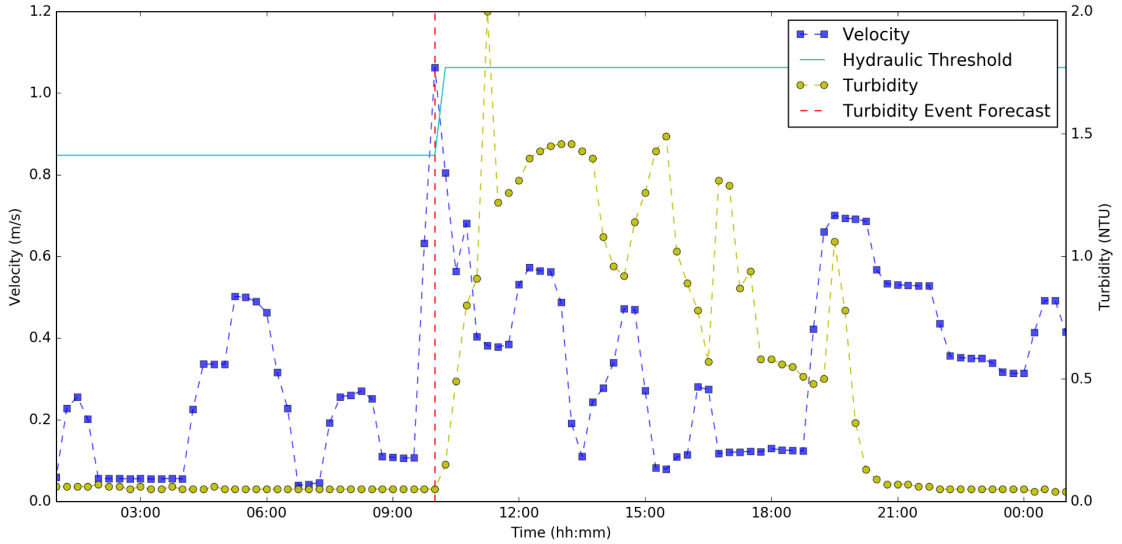


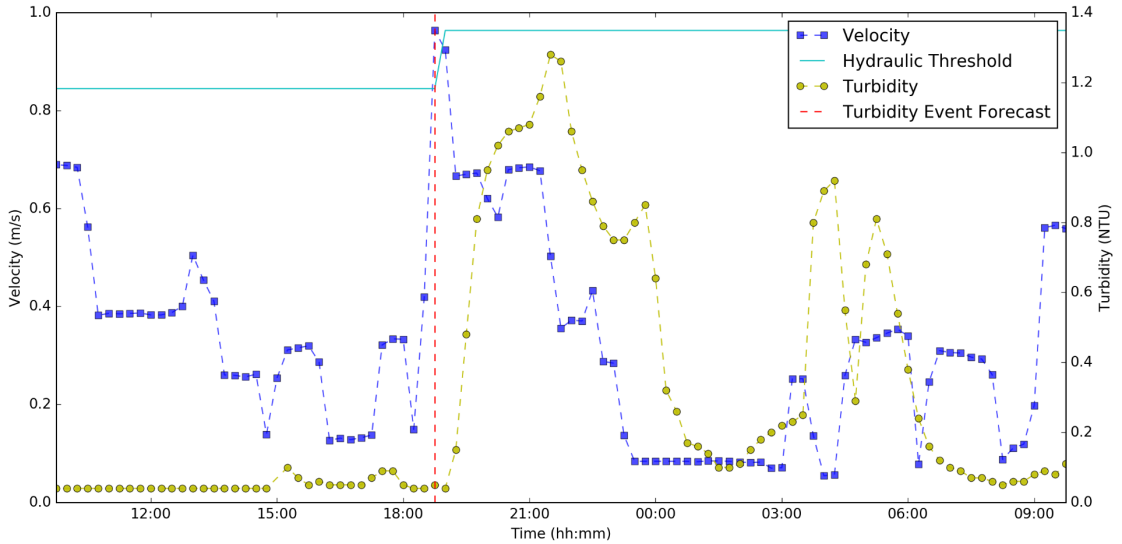
Turbidity Forecasting Model TM C

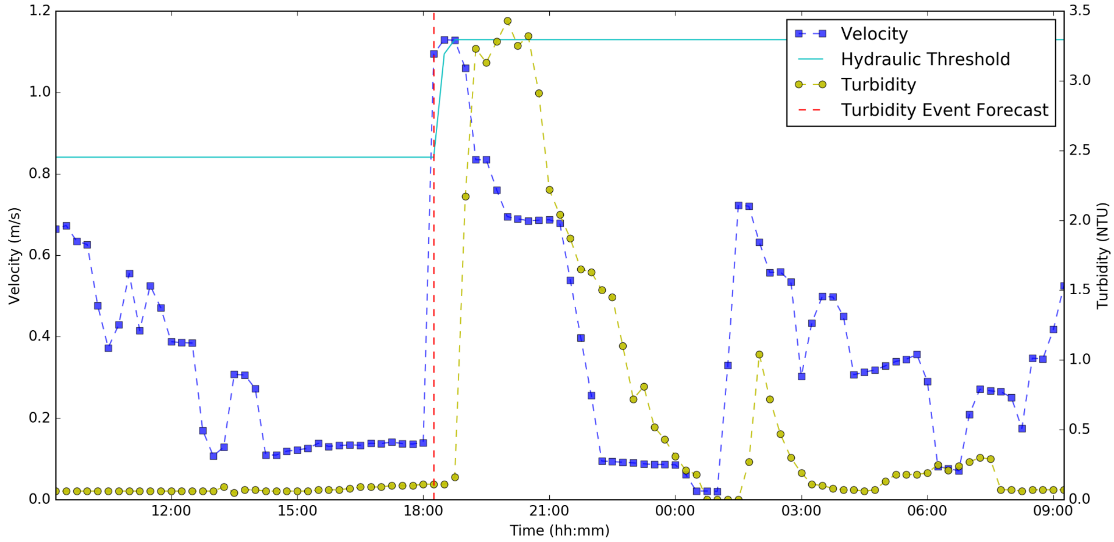


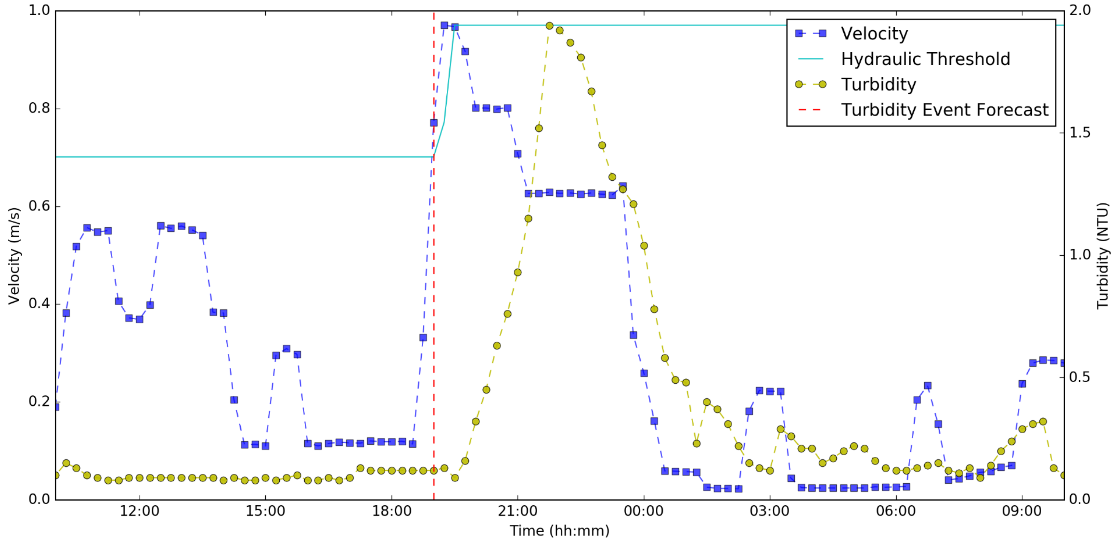


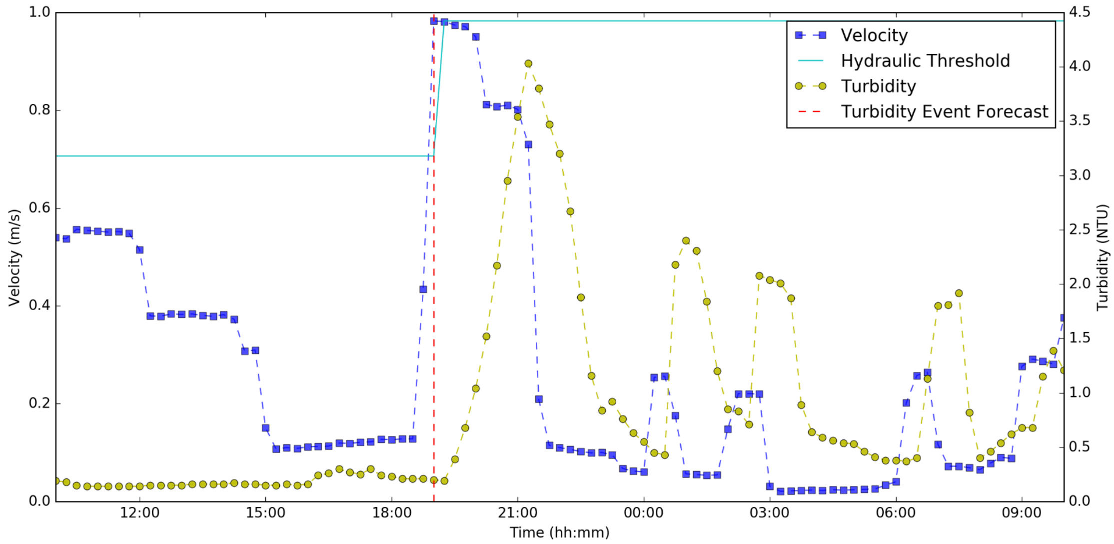

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
09:00
12:00
15:00
18:00
21:00
00:00
03:00
06:00

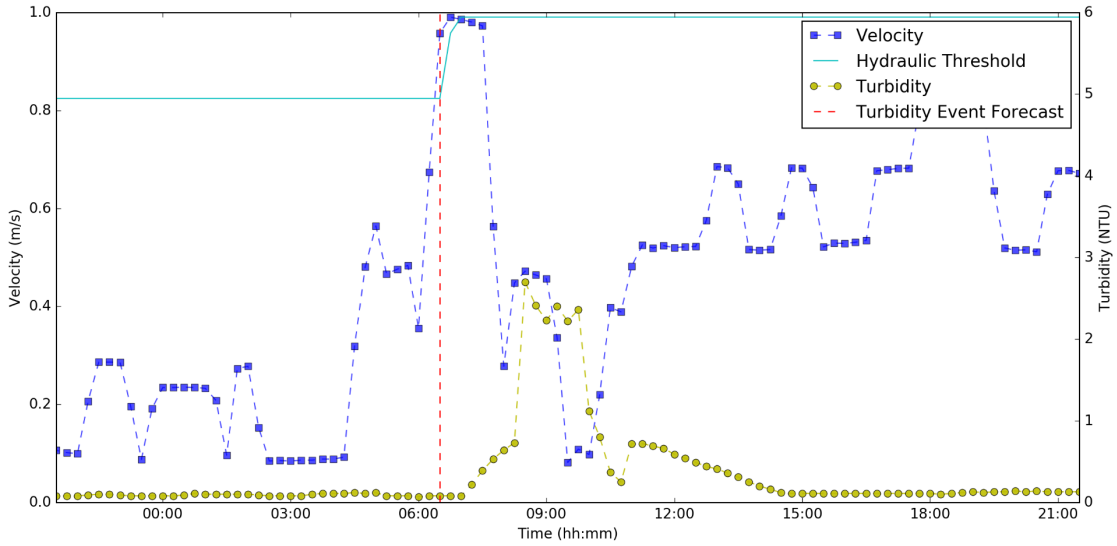

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
03:00
06:00
09:00
12:00
15:00
18:00
21:00
00:00

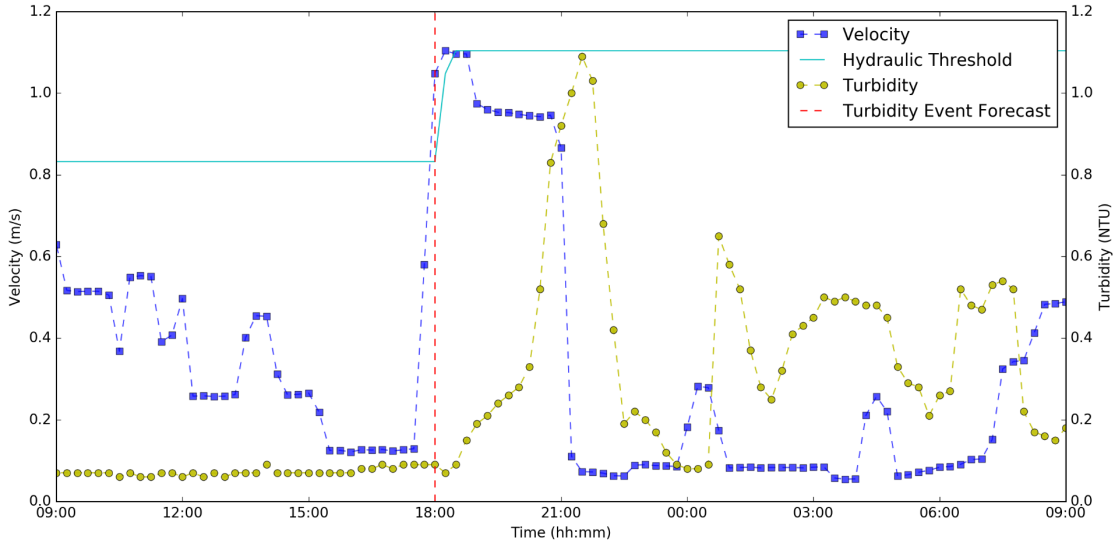

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

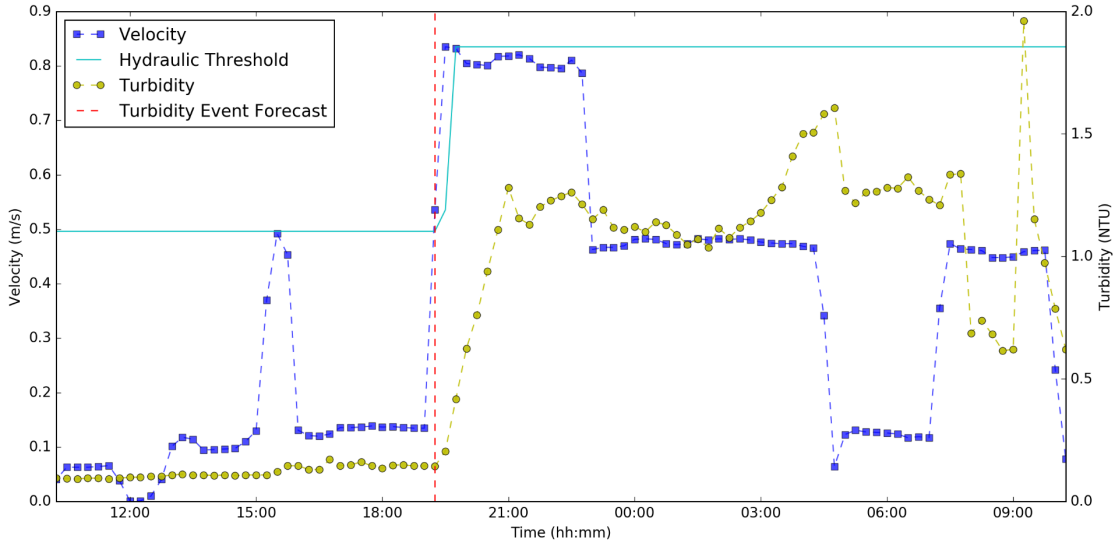

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

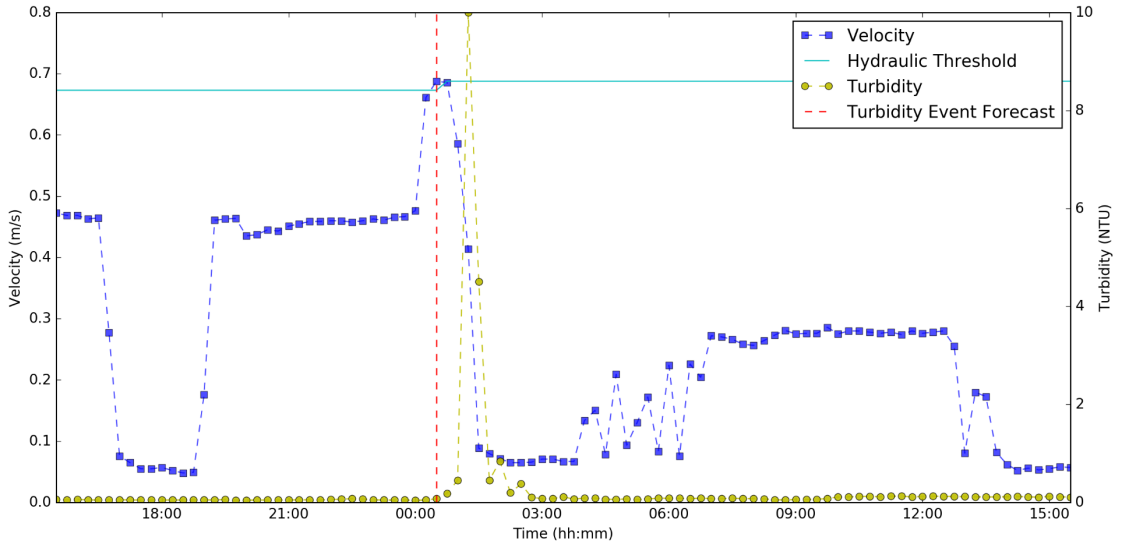

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

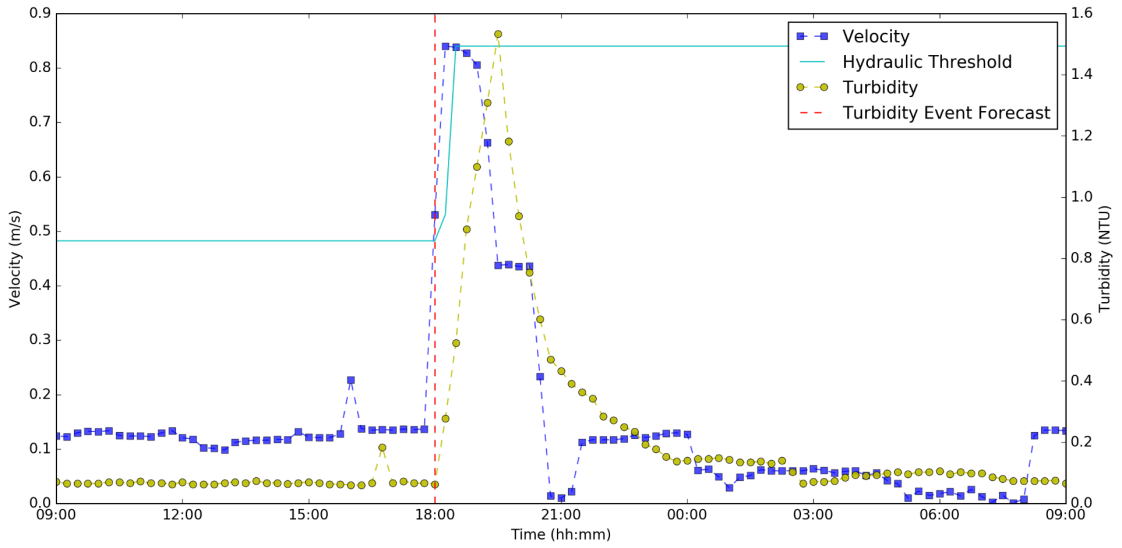


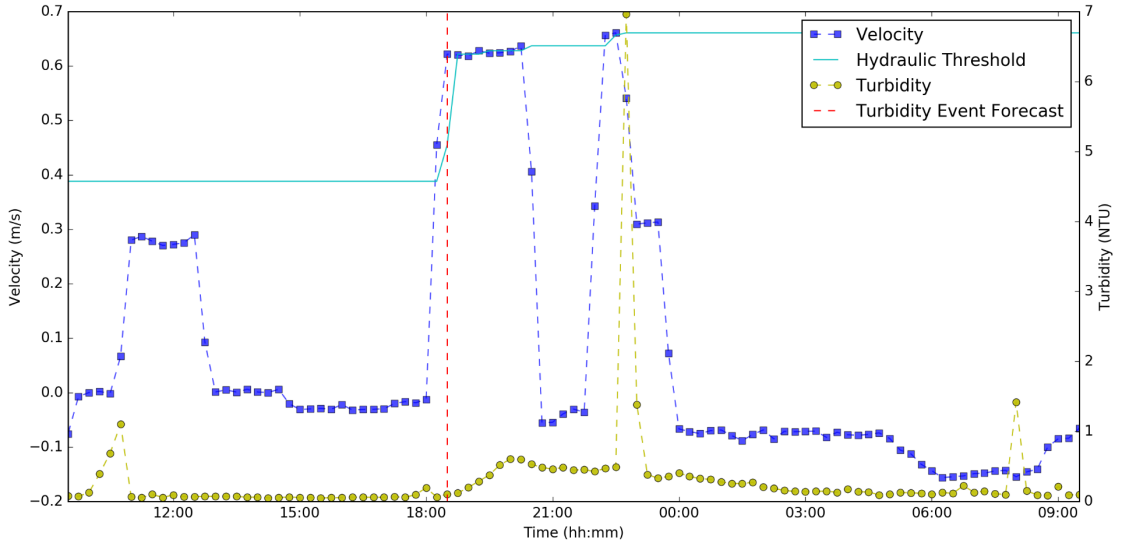


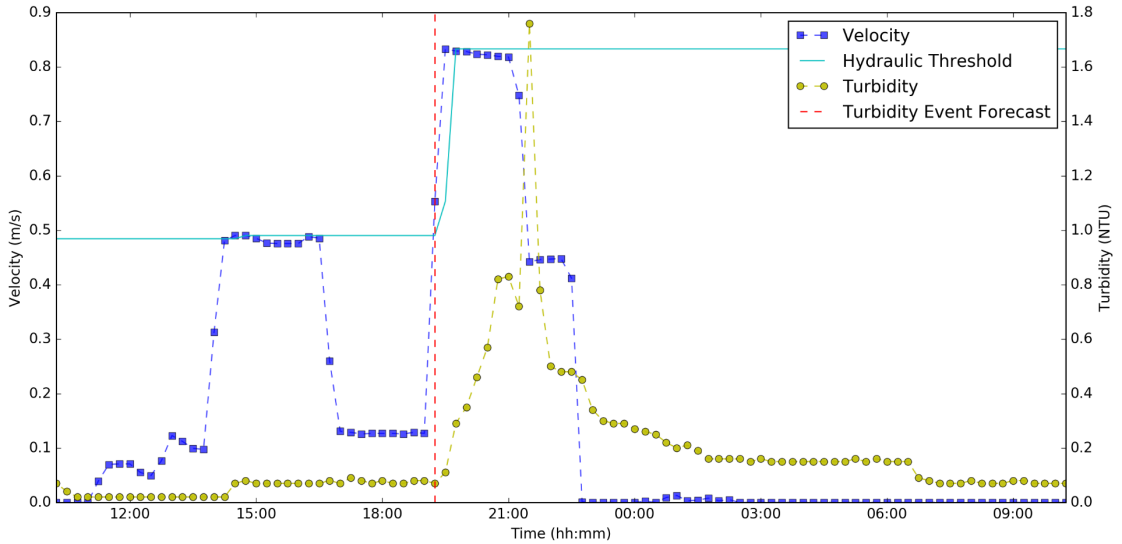
Turbidity Forecasting Model TM D

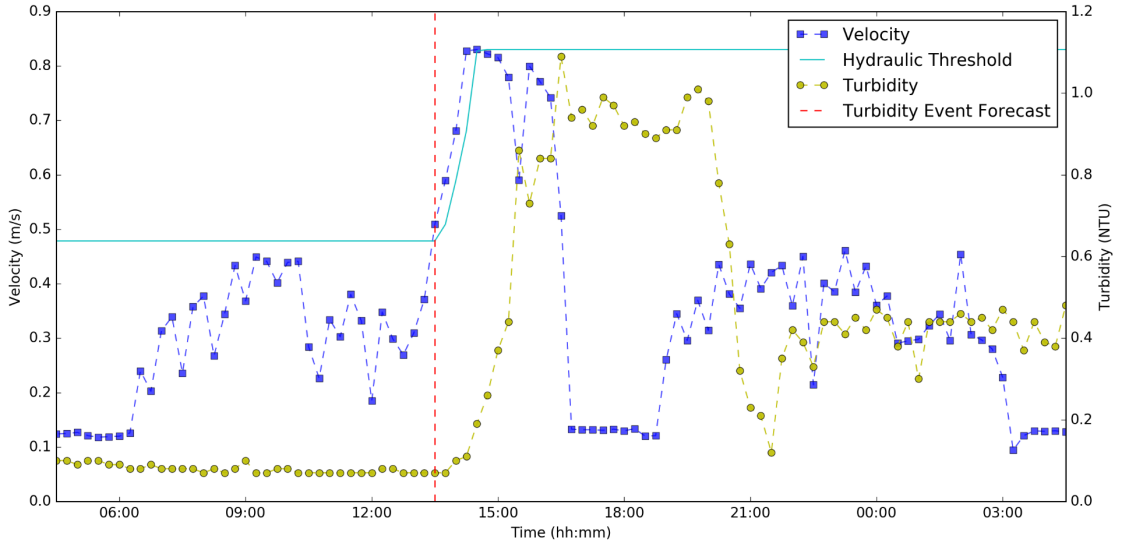


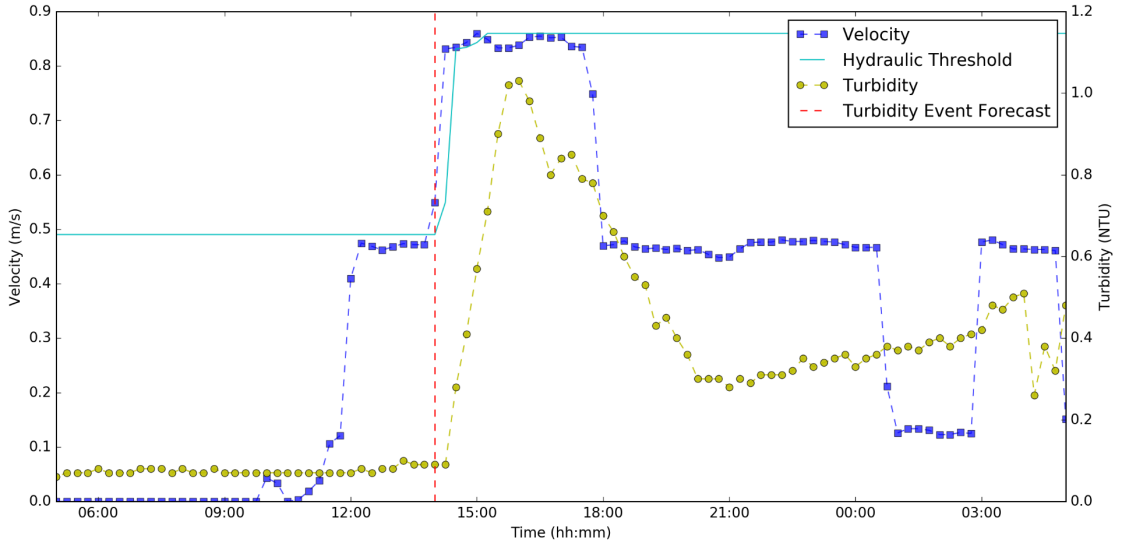

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
18:00
21:00
00:00
03:00
06:00
09:00
12:00
15:00

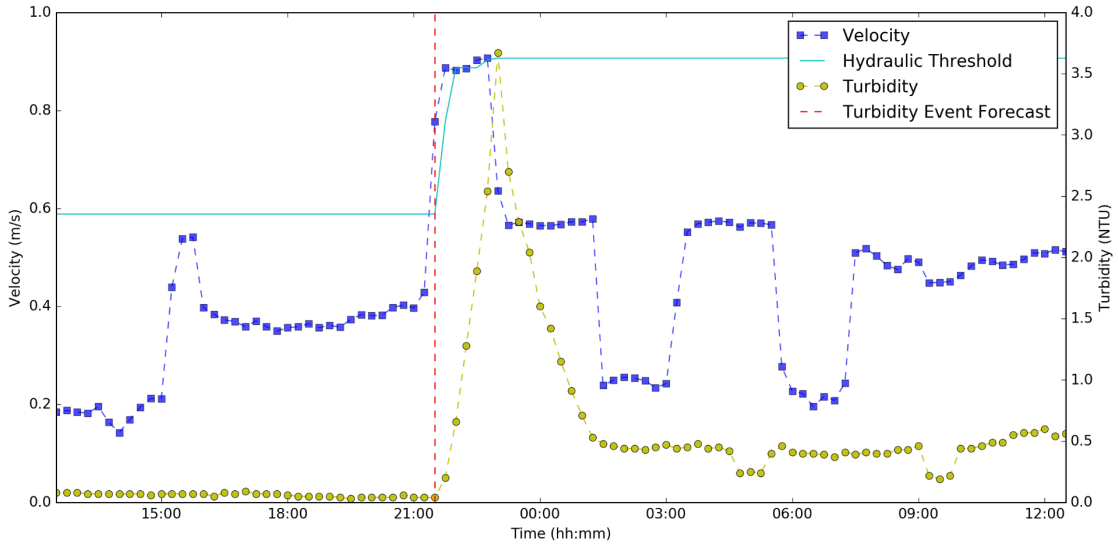

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
09:00
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

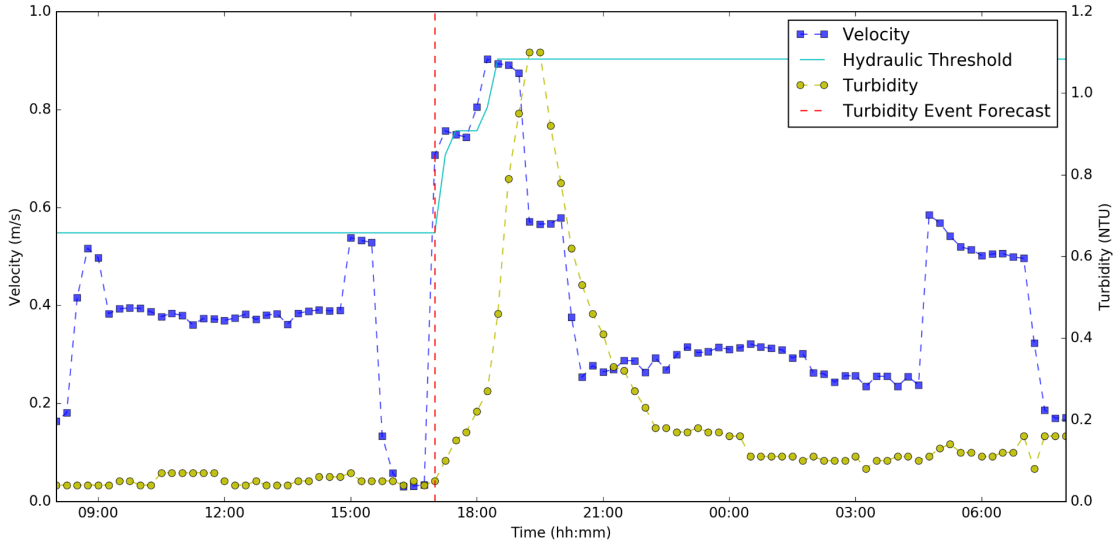

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00

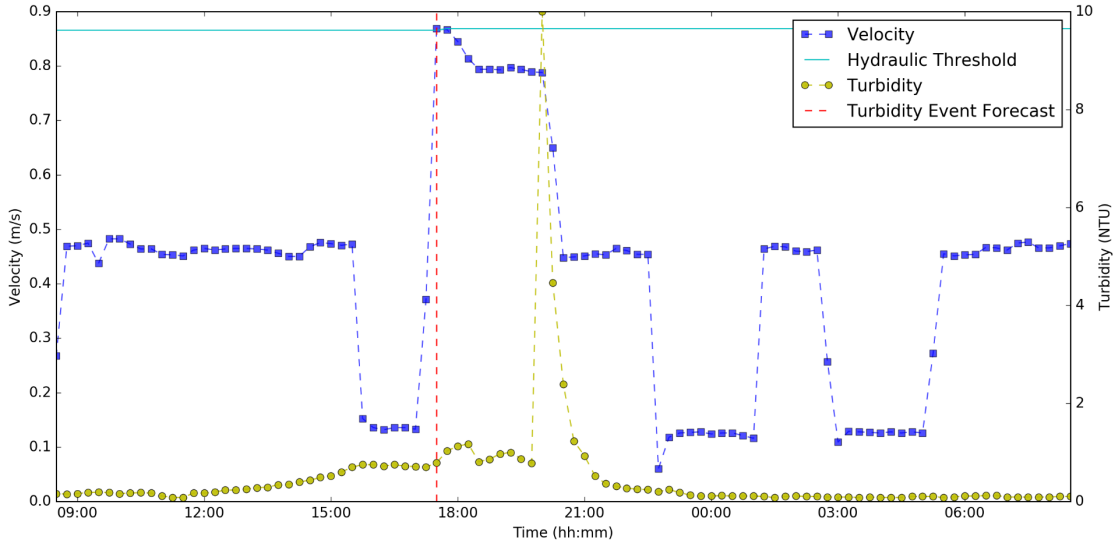

Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
12:00
15:00
18:00
21:00
00:00
03:00
06:00
09:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
06:00
09:00
12:00
15:00
18:00
21:00
00:00
03:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
06:00
09:00
12:00
15:00
18:00
21:00
00:00
03:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
15:00
18:00
21:00
00:00
03:00
06:00
09:00
12:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
09:00
12:00
15:00
18:00
21:00
00:00
03:00
06:00


Velocity
Hydraulic Threshold
Turbidity
Turbidity Event Forecast
Velocity (m/s)
Turbidity (NTU)
Time (hh:mm)
09:00
12:00
15:00
18:00
21:00
00:00
03:00
06:00

Bibliography

Abarbanel, H.D.I., Brown, R., Sidorowich, J.J., Tsimring, L.S., 1993. The analysis of observed chaotic data in physical systems. Rev. Mod. Phys. 65, 1331–1392. https://doi.org/10.1103/RevModPhys.65.1331

Abe, Y., Skali-Lami, S., Block, J.-C., Francius, G., 2012. Cohesiveness and hydrodynamic properties of young drinking water biofilms. Water Res. 46, 1155–1166. https://doi.org/10.1016/j.watres.2011.12.013

Ahadi, M., Bakhtiar, M.S., 2010. Leak detection in water-filled plastic pipes through the application of tuned wavelet transforms to acoustic emission signals. Appl. Acoust. 71, 634–639.

Aisopou, A., Stoianov, I., Arora, A., O'Hare, D., Graham, N., Boxall, J., others, 2010. Multi-parameter water quality sensors for water supply systems., in: Integrating Water Systems. Proceedings of the Tenth International Conference on Computing and Control for the Water Industry, CCWI 2009-'Integrating Water Systems', Sheffield, UK, 1-3 September 2009. CRC Press/Balkema, pp. 349–355.

Aisopou, A., Stoianov, I., Graham, N., 2011. Modelling discolouration in WDS caused by hydraulic transient events, in: Proceedings of the Water Distribution System Analysis Conference, Tucson, Arizona, ASCE Conf. Proc., Doi. p. 425.

Aisopou, A., Stoianov, I., Graham, N.J.D., 2012. In-pipe water quality monitoring in water supply systems under steady and unsteady state flow conditions: A quantitative assessment. Water Res. 46, 235–246. https://doi.org/10.1016/j.watres.2011.10.058

Alvisi, S., Franchini, M., 2014. A heuristic procedure for the automatic creation of district metered areas in water distribution systems. Urban Water J. 11, 137–159.

Arad, J., Housh, M., Perelman, L., Ostfeld, A., 2013. A dynamic thresholds scheme for contaminant event detection in water distribution systems. Water Res. 47, 1899–1908. https://doi.org/10.1016/j.watres.2013.01.017

Aral, M.M., Guan, J., Maslia, M.L., 2009. Optimal design of sensor placement in water distribution networks. J. Water Resour. Plan. Manag. 136, 5–18.

Arbués, F., Garcıa-Valiñas, M.Á., Martınez-Espiñeira, R., 2003. Estimation of residential water demand: a state-of-the-art review. J. Socio-Econ. 32, 81–102.

Archer, K.J., Kimes, R.V., 2008. Empirical characterization of random forest variable importance measures. Comput. Stat. Data Anal. 52, 2249–2260.

Armand, H., Stoianov, I., Graham, N., 2015. Investigating the impact of sectorized networks on discoloration. Procedia Eng. 119, 407–415.

Armand, H., Stoianov, I.I., Graham, N.J.D., 2017. A holistic assessment of discolouration processes in water distribution networks. Urban Water J. 14, 263–277. https://doi.org/10.1080/1573062X.2015.1111912

Armstrong, J., 2001. Evaluating Forecasting Methods. Mark. Pap.

ASCE, 1993. Criteria for Evaluation of Watershed Models. J. Irrig. Drain. Eng. 119, 429–442. https://doi.org/10.1061/(ASCE)0733-9437(1993)119:3(429)

Bai, Y., Wang, P., Li, C., Xie, J., Wang, Y., 2014. A multi-scale relevance vector regression approach for daily urban water demand forecasting. J. Hydrol. 517, 236–245. https://doi.org/10.1016/j.jhydrol.2014.05.033

Bakker, M., Vreeburg, J.H.G., Rietveld, L.C., Van der Roer, M., 2012. Reducing customer minutes lost by anomaly detection?, in: 14th Water Distribution Systems Analysis Conference, Adelaide, Australia, 24-27 September 2012.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A.F., Nielsen, H., 2000. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics 16, 412–424. https://doi.org/10.1093/bioinformatics/16.5.412

Banna, M.H., Imran, S., Francisque, A., Najjaran, H., Sadiq, R., Rodriguez, M., Hoorfar, M., 2014. Online drinking water quality monitoring: review on available and emerging technologies. Crit. Rev. Environ. Sci. Technol. 44, 1370–1421.

Bargiela, A., Hainsworth, G.D., 1989. Pressure and flow uncertainty in water systems. J. Water Resour. Plan. Manag. 115, 212–229.

Berry, J., Hart, W.E., Phillips, C.A., Uber, J.G., Watson, J.-P., 2006. Sensor placement in municipal water networks with temporal integer programming models. J. Water Resour. Plan. Manag. 132, 218–224.

Besner Marie-Claude, Ebacher Gabrielle, Lavoie Jean, Prévost Michèle, 2007. Low and Negative Pressures in Distribution Systems: Do They Actually Result in Intrusion? World Environ. Water Resour. Congr. 2007, Proceedings. https://doi.org/10.1061/40927(243)480

Billings, R.B., Jones, C.V., 2011. Forecasting urban water demand. American Water Works Association.

Blokker, E.J.M., 2010a. Stochastic water demand modelling for a better understanding of hydraulics in water distribution networks. TU Delft, Delft University of Technology.

Blokker, E.J.M., 2010b. Stochastic water demand modelling for a better understanding of hydraulics in water distribution networks. Water Management Academic Press, Delft.

Blokker, E. J. M., Schaap, P.G., 2015. Particle Accumulation Rate of Drinking Water Distribution Systems Determined by Incoming Turbidity. Procedia Eng., Computing and Control for the Water Industry (CCWI2015) Sharing the best practice in water management 119, 290–298. https://doi.org/10.1016/j.proeng.2015.08.888

Blokker, E.J.M., Schaap, P.G., 2015. Temperature Influences Discolouration risk. Procedia Eng. 119, 280–289. https://doi.org/10.1016/j.proeng.2015.08.887

Blokker, E.J.M., Vreeburg, J.H.G., Schaap, P.G., van Dijk, J.C., 2010. The self-cleaning velocity in practice, in: Proceedings of the Water Distribution System Analysis Conference. Tucson, Arizona.

Blokker, E.J.M., Vreeburg, J.H.G., Van Dijk, J.C., 2009. Simulating residential water demand with a stochastic end-use model. J. Water Resour. Plan. Manag. 136, 19–26.

Blokker, E.M., Schaap, P.G., Vreeburg, J.H., 2011. Comparing the fouling rate of a drinking water distribution system in two different configurations. Urban Water Manag. Chall. Oppor. Exeter UK.

Bougadis, J., Adamowski, K., Diduch, R., 2005. Short-term municipal water demand forecasting. Hydrol. Process. 19, 137–148.

Boxall, J.B., Dewis, N., 2005. Identification of discolouration risk through simplified modelling. ASCE ERWI World Water Environ. Water Resour.

Boxall, J.B., Saul, A.J., 2005. Modeling discoloration in potable water distribution systems. J. Environ. Eng. 131, 716–725.

Boxall, J.B., Saul, A.J., Gunstead, J.D., Dewis, N., 2003. Regeneration of discolouration in distribution systems, in: ASCE/EWRI/World Water and Environmental Resources Conference, Philadelphia.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brunone, B., Karney, B.W., Mecarelli, M., Ferrante, M., 2000. Velocity profiles and unsteady pipe friction in transient flow. J. Water Resour. Plan. Manag. 126, 236–244.

Caiado, J., 2009. Performance of combined double seasonal univariate time series models for forecasting water demand. J. Hydrol. Eng. 15, 215–222.

Chung, S.W., Yoon, S.W., Ko, I.H., 2006. Development of a Real-time Turbidity Monitoring and Modeling System for a Reservoir, in: The Proceeding of 7th International Conference on Hydroinformatics HIC. pp. 4–8.

Collins, R., Boxall, J., Besner, M.-C., Beck, S., Karney, B., 2010. Intrusion modelling and the effect of ground water conditions, in: Water Distribution Systems Analysis 2010. pp. 585–594.

Collins, R., Fox, S., Beck, S., Saul, A., Boxall, J., 2012. Intrusion and leakage through cracks and slits in plastic (MDPE) pipes, in: WDSA 2012: 14th Water Distribution Systems Analysis Conference, 24-27 September 2012 in Adelaide, South Australia. Engineers Australia, p. 807.

Cook, D., 2007. Field Investigation of Discolouration Material Accumulation Rates in Live Drinking Water Distribution Systems (Ph.D.). University of Sheffield.

Cook, D.M., Boxall, J.B., 2011. Discoloration Material Accumulation in Water Distribution Systems. J. Pipeline Syst. Eng. Pract. 2, 113–122. https://doi.org/10.1061/(ASCE)PS.1949-1204.0000083

Cook, D.M., Husband, P.S., Boxall, J.B., 2015. Operational management of trunk main discolouration risk. Urban Water J. 13, 382–395. https://doi.org/10.1080/1573062X.2014.993994

Dehua, W., Pan, L., Bo, L., Zeng, G., 2012. Water quality automatic monitoring system based on GPRS data communications. Procedia Eng. 28, 840–843.

Dewis, N., Randall-Smith, M., 2005. Discolouration risk modelling. Water Asset Manag. Int. 1, 16–18.

Di Nardo, A., Di Natale, M., 2011. A heuristic design support methodology based on graph theory for district metering of water supply networks. Eng. Optim. 43, 193–211.

Diao, K., Zhou, Y., Rauch, W., 2012. Automated creation of district metered area boundaries in water distribution systems. J. Water Resour. Plan. Manag. 139, 184–190.

Dietterich, T.G., 2002. Machine Learning for Sequential Data: A Review, in: Caelli, T., Amin, A., Duin, R.P.W., Ridder, D. de, Kamel, M. (Eds.), Structural, Syntactic, and Statistical Pattern Recognition, Lecture Notes in Computer Science. Presented at the Joint IAPR International Workshops on Statistical Techniques in Pattern Recognition (SPR) and Structural and Syntactic Pattern Recognition (SSPR), Springer Berlin Heidelberg, pp. 15–30. https://doi.org/10.1007/3-540-70659-3_2

Donkor, E.A., Mazzuchi, T.A., Soyer, R., Alan Roberson, J., 2012. Urban water demand forecasting: review of methods and models. J. Water Resour. Plan. Manag. 140, 146–159.

Dorini, G., Jonkergouw, P., Kapelan, Z., Di Pierro, F., Khu, S.T., Savic, D., 2008. An efficient algorithm for sensor placement in water distribution systems, in: Water Distribution Systems Analysis Symposium 2006. pp. 1–13.

Dorini, G., Jonkergouw, P., Kapelan, Z., Savic, D., 2010. SLOTS: Effective algorithm for sensor placement in water distribution systems. J. Water Resour. Plan. Manag. 136, 620–628.

Doumit, J., Lynch Jr, R.J., 2003. Early warning water leak detection system.

Douterelo, I., Husband, S., Loza, V., Boxall, J., 2016. Dynamics of biofilm re-growth in drinking water distributions systems. Appl. Environ. Microbiol. AEM.00109-16. https://doi.org/10.1128/AEM.00109-16

DWI, 2018a. Drinking water 2017 [WWW Document]. URL http://www.dwi.gov.uk/about/annual-report/2017/index.html (accessed 1.21.19).

DWI, 2018b. PROSECUTIONS BROUGHT BY THE DRINKING WATER INSPECTORATE [WWW Document]. URL http://dwi.defra.gov.uk/press-media/incidents-and-prosecutions/pros-caut.html (accessed 3.21.18).

DWI, 2017. Drinking water 2016 [WWW Document]. Drink. Water Insp. URL http://dwi.defra.gov.uk/about/annual-report/2016/index.html (accessed 3.20.18).

DWI, 2014a. Public water supplies in the Western region of England [WWW Document]. URL http://dwi.defra.gov.uk/about/annual-report/2014/ (accessed 3.13.16).

DWI, 2014b. Drinking water quality events in 2013 [WWW Document]. URL http://dwi.defra.gov.uk/about/annual-report/2013/dw-events.pdf (accessed 8.22.17).

DWI, 2006. Annual Provision of Information on Consumer Contacts. Drinking Water Inspectorate, London.

DWI, 2005. A Brief Guide to Drinking Water Safety Plans. Drinking Water Inspectorate, London, UK.

Fish, K., Osborn, A.M., Boxall, J.B., 2017. Biofilm structures (EPS and bacterial communities) in drinking water distribution systems are conditioned by hydraulics and influence discolouration. Sci. Total Environ. 593–594, 571–580. https://doi.org/10.1016/j.scitotenv.2017.03.176

Frank, R.J., Davey, N., Hunt, S.P., 2001. Time Series Prediction and Neural Networks. J. Intell. Robot. Syst. 31, 91–103. https://doi.org/10.1023/A:1012074215150

Fuchs, H.V., Riehle, R., 1991. Ten years of experience with leak detection by acoustic signal analysis. Appl. Acoust. 33, 1–19.

Furnass, W.R., 2015. Modelling both the continual accumulation and erosion of discolouration material in drinking water distribution systems (phd). University of Sheffield.

Furnass, W.R., Collins, R.P., Husband, P.S., Sharpe, R.L., Mounce, S.M., Boxall, J.B., 2014. Modelling both the continual erosion and regeneration of discolouration material in drinking water distribution systems. Water Sci. Technol. Water Supply 14, 81–90.

Furnass, W.R., Mounce, S.R., Boxall, J.B., 2011. A Data-Driven Methodology For Determining The Causes Of Discolouration In Distribution Networks. 11th Comput. Control Water Ind. 2011.

Gaffney, J.W., Boult, S., 2012. Need for and Use of High-Resolution Turbidity Monitoring in Managing Discoloration in Distribution. J. Environ. Eng. 138, 637–644. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000521

Gao, Y., Brennan, M., Joseph, P.F., Muggleton, J.M., Hunaidi, O., 2005. On the selection of acoustic/vibration sensors for leak detection in plastic water pipes. J. Sound Vib. 283, 927–941.

Gauthier, V., Gérard, B., Portal, J.-M., Block, J.-C., Gatel, D., 1999. Organic matter as loose deposits in a drinking water distribution system. Water Res. 33, 1014–1026.

Gershenfeld, N.A., Weigend, A.S., others, 1993. The future of time series: learning and understanding.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Mach. Learn. 63, 3–42. https://doi.org/10.1007/s10994-006-6226-1

Gevrey, M., Dimopoulos, I., Lek, S., 2003. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol. Model. 160, 249–264.

Ghiassi, M., Zimbra, D.K., Saidane, H., 2008. Urban water demand forecasting with a dynamic artificial neural network model. J. Water Resour. Plan. Manag. 134, 138–146.

Gitelson, A., Garbuzov, G., Szilagyi, F., Mittenzwey, K.H., Karnieli, A., Kaiser, A., 1993. Quantitative remote sensing methods for real-time monitoring of inland waters quality. Int. J. Remote Sens. 14, 1269–1295.

Giustolisi, O., Savic, D., Kapelan, Z., 2008. Pressure-driven demand and leakage simulation for water distribution networks. J. Hydraul. Eng. 134, 626–635.

Glasgow, H.B., Burkholder, J.M., Reed, R.E., Lewitus, A.J., Kleinman, J.E., 2004. Real-time remote monitoring of water quality: a review of current applications, and advancements in sensor, telemetry, and computing technologies. J. Exp. Mar. Biol. Ecol. 300, 409–448.

Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks, in: In Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS'10). Society for Artificial Intelligence and Statistics.

Golub, G.H., Heath, M., Wahba, G., 1979. Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics 21, 215–223.

Gomes, R., Marques, A.S., Sousa, J., 2012. Decision support system to divide a large network into suitable District Metered Areas. Water Sci. Technol. 65, 1667–1675.

Govindaraju, R.S., Rao, A.R., 2013. Artificial Neural Networks in Hydrology. Springer Science & Business Media.

Green, I.R.A., Stephenson, D., 1986. Criteria for comparison of single event models. Hydrol. Sci. J. 31, 395–411. https://doi.org/10.1080/02626668609491056

Gutiérrez, J., Villa-Medina, J.F., Nieto-Garibay, A., Porta-Gándara, M.Á., 2014. Automated irrigation system using a wireless sensor network and GPRS module. IEEE Trans. Instrum. Meas. 63, 166–176.

Hart, W.E., Murray, R., 2010. Review of sensor placement strategies for contamination warning systems in drinking water distribution systems. J. Water Resour. Plan. Manag. 136, 611–619.

He, Z., Wen, X., Liu, H., Du, J., 2014. A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region. J. Hydrol. 509, 379–386. https://doi.org/10.1016/j.jhydrol.2013.11.054

Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B., 1998. Support vector machines. IEEE Intell. Syst. Their Appl. 13, 18–28.

Herkelrath, W.N., Hamburg, S.P., Murphy, F., 1991. Automatic, real-time monitoring of soil moisture in a remote field area with time domain reflectometry. Water Resour. Res. 27, 857–864.

Herrera, M., Torgo, L., Izquierdo, J., Pérez-García, R., 2010. Predictive models for forecasting hourly urban water demand. J. Hydrol. 387, 141–150. https://doi.org/10.1016/j.jhydrol.2010.04.005

Huang, J.J., McBean, E.A., James, W., 2008. Multi-objective optimization for monitoring sensor placement in water distribution systems, in: Water Distribution Systems Analysis Symposium 2006. pp. 1–14.

Husband, P.S., 2010. Discolouration in water distribution systems: understanding, modelling and practical applications. The University of Sheffield.

Husband, P.S., Boxall, J.B., Saul, A.J., 2008. Laboratory studies investigating the processes leading to discolouration in water distribution networks. Water Res. 42, 4309–4318. https://doi.org/10.1016/j.watres.2008.07.026

Husband, S., Boxall, J., 2016. Understanding and managing discolouration risk in trunk mains. Water Res. 107, 127–140.

Husband, S., Boxall, J.B., 2011. Asset deterioration and discolouration in water distribution systems. Water Res. 45, 113–124. https://doi.org/10.1016/j.watres.2010.08.021

Husband, S., Boxall, J.B., 2009. Field studies of discoloration in water distribution systems: model verification and practical implications. J. Environ. Eng. 136, 86–94.

Husband, S., Jackson, M., Boxall, J.B., 2011. Trunk main discolouration trials and strategic planning. Urban Water Manag. Chall. Oppor. Exeter UK.

Husband, S., Whitehead, J., Boxall, J.B., 2010. The role of trunk mains in discolouration. Proc. Inst. Civ. Eng. - Water Manag. 163, 397–406. https://doi.org/10.1680/wama.900063

Husband, S., Xin, Y., Boxall, J.B., 2012. Long Term Asset Condition and Discolouration Modelling in Water Distribution Systems with Epanet MSX, in: Paper Presented at the World Environmental and Water Resources Congress 2012: Crossing Boundaries, Proceedings of the 2012 Congress. pp. 3161–3170.

Hutton, C.J., Kapelan, Z., 2015. A probabilistic methodology for quantifying, diagnosing and reducing model structural and predictive errors in short term water demand forecasting. Environ. Model. Softw. 66, 87–97. https://doi.org/10.1016/j.envsoft.2014.12.021

Hyndman, R.J., Koehler, A.B., 2006. Another look at measures of forecast accuracy. Int. J. Forecast. 22, 679–688. https://doi.org/10.1016/j.ijforecast.2006.03.001

Izquierdo, J., Herrera, M., Montalvo, I., Pérez-García, R., 2009. Division of water supply systems into district metered areas using a multi-agent based approach,

in: International Conference on Software and Data Technologies. Springer, pp. 167–180.

Jowitt, P.W., Xu, C., 1990. Optimal valve control in water-distribution networks. J. Water Resour. Plan. Manag. 116, 455–472.

Kang, D.S., Lansey, K., 2008. Real-Time Demand Estimation and Confidence Limit Analysis for Water Distribution Systems, in: Water Distribution Systems Analysis 2008. American Society of Civil Engineers, pp. 1–9.

Keedwell, E., Khu, S.-T., 2005. A hybrid genetic algorithm for the design of water distribution networks. Eng. Appl. Artif. Intell. 18, 461–472.

Khan, M.S., Coulibaly, P., 2006. Application of support vector machine in lake water level prediction. J. Hydrol. Eng. 11, 199–205.

Kingma, D., Ba, J., 2014. Adam: A Method for Stochastic Optimization. ArXiv14126980 Cs, 3rd International Conference for Learning Representations.

Kjellberg, S., 2007. Implementing Resuspension Potential Method to Optimise Mains Cleaning Program Case study: Yarra Valley Water, Melbourne, Australia. LUP Stud. Pap.

Kleiner, Y., Adams, B.J., Rogers, J.S., 1998. Long-term planning methodology for water distribution system rehabilitation. Water Resour. Res. 34, 2039–2051.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection, in: Ijcai. Montreal, Canada, pp. 1137–1145.

Krause, A., Leskovec, J., Guestrin, C., VanBriesen, J., Faloutsos, C., 2008. Efficient sensor placement optimization for securing large water distribution networks. J. Water Resour. Plan. Manag. 134, 516–526.

Kurunç, A., Yürekli, K., Cevik, O., 2005. Performance of two stochastic approaches for forecasting water quality and streamflow data from Yeşil$ıota$rmak River, Turkey. Environ. Model. Softw. 20, 1195–1200.

Lambrou, T.P., Anastasiou, C.C., Panayiotou, C.G., Polycarpou, M.M., 2014. A low-cost sensor network for real-time monitoring and contamination detection in drinking water distribution systems. IEEE Sens. J. 14, 2765–2772.

LeChevallier, M.W., Gullick, R.W., Karim, M.R., Friedman, M., Funk, J.E., 2003. The potential for health risks from intrusion of contaminants into the distribution system from pressure transients. J. Water Health 1, 3–14.

Leeder, A., Mounce, S.R., Boxall, J.B., 2012. Analysis of multi-parameter water quality data using event detection software on laboratory simulated events, in: WDSA 2012: 14th Water Distribution Systems Analysis Conference, 24-27 September 2012 in Adelaide, South Australia. Engineers Australia, p. 1018.

Leendertse, J.J., Gritton, E.C., 1971. A water-quality simulation model for well mixed estuaries and coastal seas.

Liu, G., Zhang, Y., Knibbe, W.-J., Feng, C., Liu, W., Medema, G., van der Meer, W., 2017. Potential impacts of changing supply-water quality on drinking water distribution: A review. Water Res. 116, 135–148. https://doi.org/10.1016/j.watres.2017.03.031

Machell, J., Boxall, J., Saul, A., Bramley, D., 2009. Improved representation of water age in distribution networks to inform water quality. J. Water Resour. Plan. Manag. 135, 382–391.

Machell, J., Mounce, S.R., Boxall, J.B., 2010. Online modelling of water distribution systems: a UK case study. Drink. Water Eng. Sci. 3, 21–27.

Machell, J., Mounce, S.R., Farley, B., Boxall, J.B., 2014. Online data processing for proactive UK water distribution network operation. Drink. Water Eng. Sci. 7, 23–33. https://doi.org/10.5194/dwes-7-23-2014

Makris, K.C., Andra, S.S., Botsaris, G., 2014. Pipe Scales and Biofilms in Drinking-Water Distribution Systems: Undermining Finished Water Quality. Crit. Rev. Environ. Sci. Technol. 44, 1477–1523. https://doi.org/10.1080/10643389.2013.790746

Mann, A.G., Tam, C.C., Higgins, C.D., Rodrigues, L.C., 2007. The association between drinking water turbidity and gastrointestinal illness: a systematic review. BMC Public Health 7, 256. https://doi.org/10.1186/1471-2458-7-256

Martinec, J., Rango, A., 1989. Merits of Statistical Criteria for the Performance of Hydrological Models.

Martinez, F., Conejos, P., Vercher, J., 1999. Developing an integrated model for water distribution systems considering both distributed leakage and pressure-dependent demands, in: WRPMD'99: Preparing for the 21st Century. pp. 1–14.

McClymont, K., Keedwell, E., Savić, D., Randall-Smith, M., 2013. A general multi-objective hyper-heuristic for water distribution network design with discolouration risk. J. Hydroinformatics 15, 700–716. https://doi.org/10.2166/hydro.2012.022

McNeill, L.S., Edwards, M., 2000. Phosphate inhibitors and red water in stagnant iron pipes. J. Environ. Eng. 126, 1096–1102.

Meyers, G., Kapelan, Z., Keedwell, E., Randall-Smith, M., 2016. Short-term Forecasting of Turbidity in a UK Water Distribution System. Procedia Eng., 12th International Conference on Hydroinformatics (HIC 2016) - Smart Water for the Future 154, 1140–1147. https://doi.org/10.1016/j.proeng.2016.07.534

Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R., 2005. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. Adv. Water Resour. 28, 135–147.

Mounce, S.R., Boxall, J.B., Machell, J., 2009. Development and verification of an online artificial intelligence system for detection of bursts and other abnormal flows. J. Water Resour. Plan. Manag. 136, 309–318.

Mounce, S.R., Machell, J., 2006. Burst detection using hydraulic data from water distribution systems with artificial neural networks. Urban Water J. 3, 21–31.

Muggleton, J.M., Brennan, M.J., Pinnington, R.J., 2002. Wavenumber prediction of waves in buried pipes for water leak detection. J. Sound Vib. 249, 939–954.

Mustonen, S.M., Tissari, S., Huikko, L., Kolehmainen, M., Lehtola, M.J., Hirvonen, A., 2008. Evaluating online data of water quality changes in a pilot drinking water distribution system with multivariate data exploration methods. Water Res. 42, 2421–2430. https://doi.org/10.1016/j.watres.2008.01.015

Naser, G., Karney, B.W., Boxall, J.B., 2006. Red water and discoloration in a WDS: A numerical simulation, in: 8th Annual Water Distribution Systems Analysis Symposium.

Nasirudin, M.A., Za'bah, U.N., Sidek, O., 2011. Fresh water real-time monitoring system based on wireless sensor network and GSM, in: Open Systems (ICOS), 2011 IEEE Conference On. IEEE, pp. 354–357.

Neilands, K., Bernats, M., Rubulis, J., 2012. Accumulation and modeling of particles in drinking water pipe fittings. Drink. Water Eng. Sci. 5, 47–57. https://doi.org/10.5194/dwes-5-47-2012

O'Flynn, B., Regan, F., Lawlor, A., Wallace, J., Torres, J., O'Mathuna, C., 2010. Experiences and recommendations in deploying a real-time, water quality monitoring system. Meas. Sci. Technol. 21, 124004.

OFWAT, 2009. Serviceability outputs for PR09 final determinations [WWW Document]. URL https://www.ofwat.gov.uk/wp-content/uploads/2015/11/det_pr09_finalfull.pdf (accessed 8.22.17).

Okeya, I., Kapelan, Z., Hutton, C., Naga, D., 2014. Online Modelling of Water Distribution System Using Data Assimilation. Procedia Eng., 12th International Conference on Computing and Control for the Water Industry, CCWI2013 70, 1261–1270. https://doi.org/10.1016/j.proeng.2014.02.139

Oliker, N., Ostfeld, A., 2015. Network hydraulics inclusion in water quality event detection using multiple sensor stations data. Water Res. 80, 47–58. https://doi.org/10.1016/j.watres.2015.04.036

Ong, J.B., You, Z., Mills-Beale, J., Tan, E.L., Pereles, B.D., Ong, K.G., 2008. A wireless, passive embedded sensor for real-time monitoring of water content in civil engineering materials. IEEE Sens. J. 8, 2053–2058.

Pal, M., 2005. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 26, 217–222.

Palau, C.V., Arregui, F.J., Carlos, M., 2011. Burst detection in water networks using principal component analysis. J. Water Resour. Plan. Manag. 138, 47–54.

Pothof, I.W.M., Blokker, E.J.M., 2012. Dynamic hydraulic models to study sedimentation in drinking water networks in detail. Drink. Water Eng. Sci. 5, 87–92.

Prasad, T.D., Danso-Amoako, E., 2014. Influence of chemical and biological parameters on iron and manganese accumulation in water distribution networks. Procedia Eng. 70, 1353–1361.

Preis, A., Allen, M., Whittle, A.J., 2010. On-line hydraulic modeling of a Water Distribution System in Singapore, in: Proc. WDSA10 (Water Distribution Systems Analysis Symposium), Tucson, Ariz.

Randall-Smith, M., Collingbourne, J., McClymont, K., 2011. Targeting Water Distribution Network Interventions for the Cost-Effective Mitigation of Discolouration Risk: A Case Study. Urban Water Manag. Chall. Oppor. Vol. 2, 631–633.

Rashidi, A., Sigari, M.H., Maghiar, M., Citrin, D., 2016. An analogy between various machine-learning techniques for detecting construction materials in digital images. KSCE J. Civ. Eng. 20, 1178–1188. https://doi.org/10.1007/s12205-015-0726-0

Rhodes, C., Morari, M., 1997. The false nearest neighbors algorithm: An overview. Comput. Chem. Eng. 21, S1149–S1154.

Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J.P., 2012. An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS J. Photogramm. Remote Sens. 67, 93–104.

Romano, M., 2012. Near real-time detection and approximate location of pipe bursts and other events in water distribution systems.

Romano, M., Kapelan, Z., 2014. Adaptive water demand forecasting for near real-time management of smart water distribution systems. Environ. Model. Softw. 60, 265–276. https://doi.org/10.1016/j.envsoft.2014.06.016

Romano, M., Kapelan, Z., Savić, D.A., 2012. Automated detection of pipe bursts and other events in water distribution systems. J. Water Resour. Plan. Manag. 140, 457–467.

Rossman, L., 2000. EPANET 2 Users Manual, Technical Report EPA/600/R-00/057. Water Supply Water Resour. Div. Natl. Risk Manag. Res. Lab. US Environ. Prot. Agency Cincinnati OH USA.

Ryan, G., Mathes, M., Haylock, G., Jayaratne, A., Wu, J., Noui-Mehidi, N., Grainger, C., Nguyen, B.V., 2008. Particles in water distribution systems. Coop. Res. Cent. Water Qual. Treat. Salisb. Aust.

Sanders, B.F., 2007. Evaluation of on-line DEMs for flood inundation modeling. Adv. Water Resour. 30, 1831–1843.

Sarin, P., Snoeyink, V.L., Lytle, D.A., Kriven, W.M., 2004. Iron corrosion scales: model for scale growth, iron release, and colored water formation. J. Environ. Eng. 130, 364–373.

Savic, D.A., Walters, G.A., 1997. Genetic algorithms for least-cost design of water distribution networks. J. Water Resour. Plan. Manag. 123, 67–77.

Schaap, P., Blokker, M., 2013. Zooming in on network fouling locations, in: World Environmental and Water Resources Congress 2013: Showcasing the Future. pp. 1033–1043.

Seth, A.D., Husband, P.S., Boxall, J.B., 2009. Rivelin trunk main flow test. Integrating Water Syst. Proc. 10th Comput. Control Water Ind. Boxall Maksimovic Eds Taylor Francis 431–434.

Shang, F., Uber, J., Murray, R., Janke, R., 2008. Model-based real-time detection of contamination events, in: Proceedings of the 10th Annual Water Distribution Systems Analysis Conference.

Sharpe, R., 2012. Laboratory investigations into processes causing discoloured potable water (PhD Thesis). University of Sheffield.

Shastri, Y., Diwekar, U., 2006. Sensor placement in water networks: A stochastic programming approach. J. Water Resour. Plan. Manag. 132, 192–203.

Shojaeefard, M.H., Akbari, M., Tahani, M., Farhani, F., 2013. Sensitivity Analysis of the Artificial Neural Network Outputs in Friction Stir Lap Joining of Aluminum to Brass. Adv. Mater. Sci. Eng. 2013, 1–7. https://doi.org/10.1155/2013/574914

Silverman, B.W., 1984. A fast and efficient cross-validation method for smoothing parameter choice in spline regression. J. Am. Stat. Assoc. 79, 584–589.

Skipworth, P.J., Machell, J., Saul, A.J., 2002. Empirical travel time estimation in a distribution network, in: Proceedings of the Institution of Civil Engineers-Water and Maritime Engineering. Thomas Telford Ltd, pp. 41–49.

Slaats, P.G.G., Rosenthal, L.P.M., Siegers, W.G., van den Boomen, M., Beuken, R.H.S., Vreeburg, J.H.G., 2003. Processes involved in the generation of discolored water. AWWA Research Foundation and Kiwa Water Research.

Sly, L.I., Hodgkinson, M.C., Arunpairojana, V., 1990. Deposition of manganese in a drinking water distribution system. Appl. Environ. Microbiol. 56, 628–639.

Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Solomatine, D., See, L.M., Abrahart, R.J., 2009. Data-Driven Modelling: Concepts, Approaches and Experiences, in: Practical Hydroinformatics, Water Science and Technology Library. Springer, Berlin, Heidelberg, pp. 17–30. https://doi.org/10.1007/978-3-540-79881-1_2

South West Water, 2014. C1 Customer consultation and research.

Storey, M.V., Van der Gaag, B., Burns, B.P., 2011. Advances in on-line drinking water quality monitoring and early warning systems. Water Res. 45, 741–747.

Su, Y.-C., Mays, L.W., Duan, N., Lansey, K.E., 1987. Reliability-based optimization model for water distribution systems. J. Hydraul. Eng. 113, 1539–1556.

Sunkpho, J., Ootamakorn, C., 2011. Real-time flood monitoring and warning system. Songklanakarin J. Sci. Technol. 33.

Sunny, I., Husband, S., Drake, N., Mckenzie, K., Boxall, J., 2017. Quantity and Quality Benefits of in-Service Invasive Cleaning of Trunk Mains. Drink. Water Eng. Sci. Discuss. 1–9. https://doi.org/10.5194/dwes-2017-3

Svetnik, V., Liaw, A., Tong, C., Culberson, J.C., Sheridan, R.P., Feuston, B.P., 2003. Random forest: a classification and regression tool for compound classification and QSAR modeling. J. Chem. Inf. Comput. Sci. 43, 1947–1958.

Tang, K.W., Brunone, B., Karney, B., Rossetti, A., 2000. Role and characterization of leaks under transient conditions, in: Building Partnerships. pp. 1–10.

Tao, T., Huang, H., Li, F., Xin, K., 2014. Burst Detection Using an Artificial Immune Network in Water-Distribution Systems. J. Water Resour. Plan. Manag. 140, 04014027. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000405

Tong, S., Chang, E., 2001. Support vector machine active learning for image retrieval, in: Proceedings of the Ninth ACM International Conference on Multimedia. ACM, pp. 107–118.

Tzatchkov, V.G., Alcocer-Yamanaka, V.H., Buchberger, S.G., 2008. Stochastic Demand Generated Unsteady Flow in Water Distribution Networks, in: Water Distribution Systems Analysis Symposium 2006. pp. 1–12.

Ulanicki, B., AbdelMeguid, H., Bounds, P., Patel, R., 2008. Pressure control in district metering areas with boundary and internal pressure reducing valves, in: Water Distribution Systems Analysis 2008. pp. 1–13.

van Summeren, J., Raterman, B., Vonk, E., Blokker, M., van Erp, J., Vries, D., 2015. Influence of Temperature, Network Diagnostics, and Demographic Factors on Discoloration-Related Customer Reports. Procedia Eng. 119, 416–425. https://doi.org/10.1016/j.proeng.2015.08.903

Van Thienen, P., Vreeburg, J.H.G., Blokker, E.J.M., 2011. Radial transport processes as a precursor to particle deposition in drinking water distribution systems. Water Res. 45, 1807–1817.

Verberk, J., Hamilton, L.A., O'Halloran, K.J., Van Der Horst, W., Vreeburg, J., 2006. Analysis of particle numbers, size and composition in drinking water transportation pipelines: results of online measurements. Water Sci. Technol. Water Supply 6, 35–43.

Vicente, H., Dias, S., Fernandes, A., Abelha, A., Machado, J., Neves, J., 2012. Prediction of the quality of public water supply using artificial neural networks. J. Water Supply Res. Technol. - Aqua 61, 446–459. https://doi.org/10.2166/aqua.2012.014

Vítkovskỳ, J.P., Simpson, A.R., Lambert, M.F., 2000. Leak detection and calibration using transients and genetic algorithms. J. Water Resour. Plan. Manag. 126, 262–265.

Vreeburg, I.J.H.G., Boxall, D.J.B., 2007. Discolouration in potable water distribution systems: A review. Water Res. 41, 519–529. https://doi.org/10.1016/j.watres.2006.09.028

Vreeburg, J., 2010. Discolouration in drinking water systems. IWA Publishing.

Vreeburg, J.H.G., 2007. Discolouration in drinking water systems: a particular approach. TU Delft, Delft University of Technology.

Vreeburg, J.H.G., Beverloo, H., 2011. Sediment accumulation in drinking water trunk mains, in: Urban Water Management: Challenges and Opportunities. Presented at the Computing and Control for the Water Industry, Computing and Control for the Water Industry 2011, Exeter, UK.

Vreeburg, J.H.G., Schaap, P.G., Bergmans, B., van Dijk, J.C., 2009. How Effective Is Flushing of Cast Iron Pipes?

Vreeburg, J.H.G., Schaap, P.G., Van Dijk, J.C., 2005. Particles in the drinking water system: from source to discolouration. Water Sci. Technol. Water Supply 4, 431–438.

Vreeburg, J.H.G., Schippers, D., Verberk, J.Q.J.C., van Dijk, J.C., 2008. Impact of particles on sediment accumulation in a drinking water distribution system. Water Res. 42, 4233–4242. https://doi.org/10.1016/j.watres.2008.05.024

Vreeburg, J.H.G., Tankerville, M.Z., 2011. Build-up of discolouration potential in networks as function of treatment performance, in: Reston, VA: ASCE Copyright Proceedings of the 2011 World Environmental and Water Resources Congress; May 22. 26, 2011, Palm Springs, California| d 20110000. American Society of Civil Engineers.

Wei, H., Song, W., Li, Q., 2004. A RBF network based online modeling method for realtime cost model in power plant. Proc.-Chin. Soc. Electr. Eng. 24, 246–252.

Weickgenannt, M., Kapelan, Z., Blokker, M., Savic, D.A., 2010. Risk-based sensor placement for contaminant detection in water distribution systems. J. Water Resour. Plan. Manag. 136, 629–636.

WHO, 2011. Guidelines for drinking-water quality, fourth edition.

WHO (Ed.), 2004. Guidelines for drinking-water quality, 3rd ed. ed. World Health Organization, Geneva.

Widodo, A., Yang, B.-S., 2007. Support vector machine in machine condition monitoring and fault diagnosis. Mech. Syst. Signal Process. 21, 2560–2574.

Wu, Z.Y., Sage, P., Turtle, D., 2009. Pressure-dependent leak detection model and its application to a district water system. J. Water Resour. Plan. Manag. 136, 116–128.

Ye, G., Fenner, R.A., 2013. Weighted least squares with expectation-maximization algorithm for burst detection in UK water distribution systems. J. Water Resour. Plan. Manag. 140, 417–424.