

University of Sheffield
Thesis



Improving the operation and maintenance of CSO
structures

Dissertation submitted as part requirement for the Degree of PhD

By

Nan Guo

Supervisor:
Professor Adrian Saul

The University of Sheffield
Department of Civil and Structural Engineering
August 2011



IMAGING SERVICES NORTH

Boston Spa, Wetherby

West Yorkshire, LS23 7BQ

www.bl.uk

The following has been
excluded at the request of
the university

Pages 12, 14 and 15

Abstract

Combined Sewer Overflow (CSO) structures are commonly used in combined sewer systems and serve as “safety valves” for the pipe system in that they act as a hydraulic control to prevent an overload of the sewer system to prevent surcharge and flooding. They also act to retain the pollution within the sewer system and to retain such pollution, particularly aesthetic solids it has been common practice to incorporate screens into CSO chambers. However, the UK water industry is faced with an insufficient understanding of the way in which these assets perform and of the way in which they may best be managed. To better understand such performance the UK industry has installed a large number of monitoring systems that provide data on the hydraulic performance of the CSO chambers and CSO chambers with screens. This data is currently being used to develop simulation tools with a view to better understanding and providing a more efficient operational strategy, especially in respect of the frequency of maintenance visits.

The main objective of this research is to develop and validate novel mathematical techniques based on this hydraulic performance data to simulate, predict and provide a decision support system for CSO asset operation and maintenance.

To achieve this objective, three steps were completed. Firstly, data was collected on the types of structure in common use (both CSO's and screens), their monitored hydraulic performance (chamber water depth), rainfall information and their maintenance requirements (number of pro-active and reactive visits and associated costs). Secondly to use this data to develop and validate a mathematical model, using artificial intelligence techniques in the form of an adaptive linear neural network approach, to predict the hydraulic performance of chambers, which installed with different types of screens in response to rainfall. Thirdly, based on predicted CSO hydraulic performance to utilise a fuzzy logic approach to describe the operational and pro-active maintenance requirements of the different types of CSO structures and screen arrangements.

The models were successfully developed using data from one catchment and subsequently applied to a second catchment, again successfully, to test their validity and transferability. The final section of the thesis attempts to describe how the methodologies developed may be incorporated into industry standard and practical CSO asset management.

Glossary and Abbreviations

Term	Definition
ADALINE	Adaptive Linear Neuron
AI	Artificial Intelligence
AMP	Asset Management Period
ANN	Artificial Neural Networks
Autocorr	Auto correlation (serial correlation)
BDN	Binary Decision Units
CAPEX	Capital Expenditure
Ce	Correlation coefficient
Corr	Correlation analysis
Cov	Covariance
CSO	Combine Sewer Overflow
Defra	Department for Environment Food and Rural Affairs
DWF	Dry Weather Flow
EA	Environment Agency of England and Wales
FIS	Fuzzy Inference Systems
FL	Fuzzy Logic
FST	Fuzzy Set Theory
GIS	Geographic information system
LMS	Least Mean Square
LR	Linear Regression
NN	Neural Networks
O&M	Operation and Maintenance
OPEX	Operational Expenditure
PPM	Performance Predictor Model
RMSE	Root Mean Squared Error
RTS	Remote Telemetry System
SSE	Sum of Squares Due to Error
TES	Total Mean Error Square
TPE	Total Predicted Error Square
TTE	Total Trained Error Square
UCSO	Unsatisfactory CSOs
WRc	Water Research Centre
WWTP	Waste Water Treatment Plant
Xcorr	Cross correlation
YWS	Yorkshire Water service Ltd.

Contents

1. Introduction.....	1
1.1. Project Aims.....	4
2. Literature Review.....	5
2.1. Combined Sewer Overflow (CSO).....	5
2.1.1. Construction chamber types.....	7
2.1.2. Screen Types.....	12
2.1.3. Subheading –super/subcritical conditions.....	16
2.1.4. Heading – depth discharge relationships.....	17
2.1.5. Monitoring of CSO.....	19
2.1.6. Asset operation and maintenance.....	19
2.2. Artificial Neural Network.....	20
2.2.1. General concept and development.....	20
2.2.2. ADALINE.....	22
2.2.3. Least Mean Square.....	25
2.2.4. Application in water industry.....	27
2.3. Fuzzy Logic expert system.....	31
2.3.1. Fuzzy set.....	31
2.3.2. Membership function.....	32
2.3.3. Fuzzification.....	34
2.3.4. Fuzzy logic operators.....	38
2.3.5. Fuzzy logic If-Then rule.....	38
2.3.6. Inference Mechanism.....	39
2.3.7. Fuzzy logic application in Engineering.....	39
3. Data Collection.....	42
3.1. Introduction.....	42
3.2. Chamber and screen types.....	45
3.3. CSO chamber water depth.....	50
3.4. Rainfall information selection.....	53
3.5. CSO performance failures.....	60
3.6. Operation and Maintenance records.....	65
4. CSO hydraulic performance prediction model.....	68
4.1. ADALINE Methodology.....	68

4.1.1.	Linear relationship function	69
4.1.2.	Model Input	71
4.1.3.	Learning rule	78
4.2.	Model framework:	81
4.3.	Model training:	83
4.4.	Model predicting:	88
4.5.	Model sensitivity testing:.....	93
4.5.1.	Avoidance of over fitting	94
4.5.2.	Test learned system	97
4.5.3.	Prediction range	101
4.6.	Summary	105
5.	CSO pro-active operation and maintenance decision support model ...	106
5.1.	Methodology	107
5.2.	CSO O&M Fuzzy System	109
5.3.	Fuzzy Sets	111
5.4.	Model Input Parameters.....	114
5.4.1.	Input 1: Chamber Weir Height.....	119
5.4.2.	Input 2: Flow rate at 1st spill.....	122
5.4.3.	Input 3: spill rate at 1 in 5yr rainfall.....	126
5.4.4.	Input 4: Total spill duration over inspection period	129
5.4.5.	Input 5: Total spill volume over the duration of the inspection period	134
5.5.	Output Parameter: Determine O&M action	141
5.6.	Rules of FL system	145
5.7.	Model Surface View	149
5.8.	Model Testing	154
5.8.1.	Comparison of different screen devices	158
5.9.	FL model Summary	161
5.10.	Conclusion:.....	163
6.	Practical application.....	165
6.1.	Case Study Introduction.....	165
6.2.	Asset management approach	166
6.3.	Outputs	169
6.3.1.	Hydraulic performance prediction.....	169
6.3.2.	Proactive O&M strategy	177

6.4. Project Achievement.....	181
7. Conclusion.....	183
7.1. Project summary.....	183
7.2. Future works.....	185
7.2.1. Rainfall radar data application.....	185
8. References.....	188

Figures

Figure 1-1: Serious (category 1 and 2) pollution incidents where source is from sewage and water industry(www.environment-agency.gov.uk/research/library/publications/34019.aspx).....	3
Figure 2-1: High-side weir chamber (Balmforth and Henderson, 1988).....	8
Figure 2-2: End weir stilling pond chamber (Saul et al., 1997).....	9
Figure 2-3: Hydrodynamic separator (here: Storm King™, Saul et al., 1997) 11	
Figure 2-4: Hydro-international Hydro-Static Screen (Hydro-International™) 12	
Figure 2-5: Huber ROATAMA Mechanical Storm screen (Huber™)	13
Figure 2-6 Heliscreen by Hydro-International	14
Figure 2-7: Hydro-international Rotary Jet™ Screens (Hydro-International™)	15
Figure 2-8: Possible flow condition at inlet pipe to CSO chamber (WaPUG, 2006).....	17
Figure 2-9: Relationship between inflow rate and CSO chamber water depth (Fach et al 2008).....	18
Figure 2-10: Monitoring device installation in CSO tank (Weiss et al. 2006)..	19
Figure 2-11: Adaptive Linear Network Architecture (Widrow, 1985)	23
Figure 2-12: Membership function gallery indicated in MATLAB.....	34
Figure 2-13: Straight line membership function	35
Figure 2-14: Trapezoidal membership function.....	36
Figure 2-15: Gaussian membership function	37
Figure 2-16: Triangular membership function	37
Figure 3-1: Data collection details.....	43
Figure 3-2: CSO database interface	44
Figure 3-3: Terry Avenue CSO upstream and downstream networks (Source from Yorkshire Water GIS system 'Odyssey'). The red arrow indicates the flow direction from upstream to downstream and the numbers adjacent to the pipelines are the diameters of each pipe; the full black rectangle represents the CSO asset and full black circles represent the manholes.	48
Figure 3-4: Example of as built drawing of Terry Avenue CSO Chamber (Source from Yorkshire Water EDMS, scale 1:25, produced by MWH)	49
Figure 3-5: Example of Terry Avenue CSO screen general arrangement (Source from Yorkshire Water EDMS, scale 1:50, produced by Hydro International UK Ltd)	50
Figure 3-6: CSO chamber dry weather flow. Grids were defined as 288 time steps which represented the time length of 24 hours.....	51

Figure 3-7: An example of CSO chamber water depth performance and rainfall intensity (5 minutes interval).....	52
Figure 3-8: CSO chamber water depth performance during rainfall event.	53
Figure 3-9: Example of Geo-locate relation between Terry Avenue CSO and nearby rain gauge stations (Source from Yorkshire Water GIS system 'Odyssey' with a set of 'rain gauge layout')	56
Figure 3-10: Correlation analysis example: two rain gauge value during storm No1	57
Figure 3-11: Correlation analysis example: two rain gauge values during storm No2	59
Figure 3-12: CSO performance with maintenance actions.....	66
Figure 3-13: Maintenance records document.....	67
Figure 4-1: Example chart of model linear relationship explanation.....	71
Figure 4-2: Model inputs selection	72
Figure 4-3: Delay response of chamber water depth to rainfall ()	73
Figure 4-4: Cross correlation analysis of rainfall event example No. 1	75
Figure 4-5: Cross correlation analysis of rainfall event example No. 2	76
Figure 4-6: water depth serial correlation analysis.....	76
Figure 4-7: Model input selection and model predicted duration.....	77
Figure 4-8: ADALINE network system for CSO hydraulic performance prediction model (Developed from Widrow, 1985)	80
Figure 4-9: CSO hydraulic performance prediction model framework.....	82
Figure 4-10: Training dataset and testing dataset.....	84
Figure 4-11 Model training process.....	85
Figure 4-12: Example of model training output	86
Figure 4-13: Example of correlation analysis between training outputs and original value.....	87
Figure 4-14: Model prediction process.....	89
Figure 4-15: Model prediction output samples.	90
Figure 4-16: Correlation analysis of sample in Figure 4-15.....	91
Figure 4-17: comparison between predicted and actual value of model testing. Chamber water depth from CSO asset with telemetry device errors	92
Figure 4-18: Error example of correlation coefficient analysis in prediction testing.	93
Figure 4-19: Example 1 of model over fitting	95
Figure 4-20: Example2 of model over fitting	96
Figure 4-21: Example of output by using 1200 values for training and 3000 values for testing.....	98

Figure 4-22: 'corrcoef' analysis of model by using 1200 values for training and 3000 values for testing.....	98
Figure 4-23: Example of output by using 1200 values for training and 6000 values for testing.....	99
Figure 4-24: 'corrcoef' analysis of model by using 1200 values for training and 6000 values for testing.....	99
Figure 4-25: Example of output by using 1200 values for training and 9000 values for testing.....	100
Figure 4-26: 'corrcoef' analysis of model by using 1200 values for training and 9000 values for testing.....	100
Figure 4-27: Demonstration of model prediction (from T+1 time step) Wednesday.....	102
Figure 4-28: Model total mean error square compared with length of prediction period.....	104
Figure 5-1: Predicted performance of system for proactive maintenance	107
Figure 5-2: Proposed consideration of CSO O&M system.....	109
Figure 5-3: Fuzzy system in FL tool box	111
Figure 5-4: Fuzzy input linguistic description	112
Figure 5-5: Fitting RMSE comparison between different type of membership functions	117
Figure 5-6: Demonstration of model input development	118
Figure 5-7: Curve fitting result of chamber weir height.....	122
Figure 5-8: Curve fitting result of flow rate at 1 st spill	125
Figure 5-9: Curve fitting result of spill rate at 1 in 5yr rainfall	129
Figure 5-10: Example of spill duration of a CSO chamber water depth performance.....	131
Figure 5-11: Membership function representation of spill duration.....	134
Figure 5-12: Demonstration graph of Trapezoid rules.....	136
Figure 5-13: Water head on chamber weir.....	137
Figure 5-14: Membership function figure of CSO total spill volume	141
Figure 5-15: Curve fitting result of model output – requirement of CSO O&M action	144
Figure 5-16: Process of system rules development.....	146
Figure 5-17: Graphic expression of defined rules	148
Figure 5-18: Model viewer - Weir height Vs. Flow rate at 1 st spill	150
Figure 5-19: Model viewer - Weir height Vs. Spill rate at 1 in 5 year rainfall	151
Figure 5-20: Model viewer – Spill rate at 1 in 5 year rainfall Vs. Spill duration	152

Figure 5-21: Model viewer – CSO weir height Vs. Spill duration.....	153
Figure 5-22: Model viewer - Spill volume Vs. weir height.....	153
Figure 5-23: Different defuzzification outputs comparison	156
Figure 5-24: Output viewer of CSO with mechanic screen and non-screen example 1	159
Figure 5-25: Output viewer of CSO with mechanic screen and non-screen example 2	160
Figure 5-26: Prediction accuracy of FL CSO O&M model.....	163
Figure 6-1: Asset management case study frameworks	167
Figure 6-2: Example of NN model output of CRESCENT TERRACE CSO .	170
Figure 6-3: Predicted spill event No.1 for CRESCENT TERRACE CSO	172
Figure 6-4: Correlation analysis of predicted of spill No 1 (CRESCENT TERRACE CSO).....	173
Figure 6-5: Predicted spill event No.2 for CRESCENT TERRACE CSO	173
Figure 6-6: Correlation analysis of predicted of spill No 2 (CRESCENT TERRACE CSO).....	174
Figure 6-7: Predicted spill event No.3 for CRESCENT TERRACE CSO	174
Figure 6-8: Correlation analysis of predicted of spill No 3 (CRESCENT TERRACE CSO).....	175
Figure 6-9: Prediction error comparison.....	176
Figure 6-10: FL expert system application process.....	178
Figure 7-1: Comparison of measured and modelled flows (Shepherd et al. 2010).....	186

Tables

Table 1.1: Definition of pollution incidents categories(www.statistics.gov.uk/STATBASE/Expodata/Spreadsheets/D7828.xls)	2
Table 2.1: Aesthetic Control Requirements	6
Table 3-1: Summarized CSO information list.....	46
Table 3.2: Level of each rainfall event	54
Table 3-3: List of CSO structural information and causes of performance failures	61
Table 3-4: Pollution incidents by CSO chamber and screen type	65
Table 5.1: Concerning chamber feature and hydraulic parameters of FL model	107
Table 5.2: Routine maintenance action List.....	113
Table 5.3: Emergency incident maintenance action List.....	114
Table 5.4: Examples of curve fitting RMSE analysis.....	116
Table 5.5: Summary of chamber weir height (mm)	119
Table 5.6: Membership function definition of weir height	120
Table 5.7: Summary of asset flow rate at 1 st spill event (l/s).....	123
Table 5.8: Membership function definition of flow rate at 1 st spill	124
Table 5.9: Summary of asset spill rate at 1 in 5 year rainfall event (l/s)	126
Table 5.10: Membership function definition of spill rate	127
Table 5.11: Summary of asset spill duration (mins)	131
Table 5.12: Membership function definition of spill duration	132
Table 5.13: Summary of asset spill volume (m ³).....	137
Table 5.14: Membership function definition of spill volume	139
Table 5.15: Membership function definition of model output.....	142
Table 5.16: FL system rules.....	147
Table 5.17: Test outputs of different operators	155
Table 5.18: Test outputs of different Defuzzification approach	157
Table 6.1: 20 CSO asset list for case study	165
Table 6.2: 20 Rainfall return level and duration	168
Table 6.3: Prediction correlation coefficient summary.....	170
Table 6.4: Predicted spill event numbers	171
Table 6.5: Summary of predicted overflow events	176
Table 6.7: List of predicted O&M actions	179
Table 6.7: Evaluation results of applying FL model.....	180

Acknowledgement

I am heartily thankful to my supervisor, Professor Adrian Saul, whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. The financial support provided by Yorkshire Water Service Ltd. is greatly appreciated. Special thanks go to Mr David Hanson from YWS on the supporting of project data collection works.

The author would also like to acknowledge the help of Dr. Andy Swan, Dr. Will Shepherd and Dr. Alma Schellart for all their help and guidance during this whole work.

Lastly, I offer my regards and thankful to my parents and all of those who supported me in any respect during the completion of the project.

Declaration

Except where specific reference has been made to the work of others, this thesis is the result of my own work. No part of it has been submitted to any university for a degree, diploma or other qualification.

1. Introduction

CSO structures are commonly used in combined sewer systems and serve as “safety valves” for the pipe system by limiting the quantities of flow passed forward to treatment to a level that the downstream sewer and sewage treatment system can practically and economically accommodate. However, in 2007, the Environment (EA) reported that 30% of all CSOs installed in England and Wales performed in an unsatisfactory manner when judged against a range of environmentally based criteria (EA, 2007). Potential CSO asset performance failures were considered to cause both urban flooding and to seriously impact the quality of receiving water courses (Defra, 2006).

As the regulatory authority for all waste management issues the EA require that complaints by the public and reported incidents of pollution are logged and categorised according to their severity (see Table 1.1). The category describes the impact of each incident on water, land and air. The impact of a single incident on each medium is considered and reported separately. In Table 1.1, the impact of incidents is classified into different categories, for example: the incident is classified as Category 1, if it causes persistent and extensive effects on water quality, major damage to the ecosystem, closure of a potable abstraction, major impact on property, or major damage to agriculture and/or commerce. The classification of Category 2 and Category 3 incidents are shown in Table 1. If no impact occurred for particular media, the incident is reported as a Category 4.

Table 1.1: Definition of pollution incidents
categories(www.statistics.gov.uk/STATBASE/Expodata/Spreadsheets/D7828.xls)

	Water
Category 1 - the most serious	
persistent and extensive effects on quality	*
major damage to the ecosystem	*
closure of a potable abstraction	*
major impact upon amenity value	*
major damage to agriculture and/or commerce	*
serious impact upon man	*
Category 2 - significant but less severe	
significant effect on quality	*
significant damage to the ecosystem	*
non-routine notification of abstractors	*
reduction in amenity value	*
significant damage to agriculture and/or commerce	*
impact on man	*
Category 3 - relatively minor	
minimal effect on quality	*
minor damage to local ecosystems	*
marginal effect on amenity value	*
minimal impact to agriculture and/or commerce	*

For Category 1 and Category 2 incidents in 2006, the causes of pollution were summarized by the Environment Agency and these are shown in Figure 1.1. This highlights that 14% of such serious pollution incidents were derived from CSO.

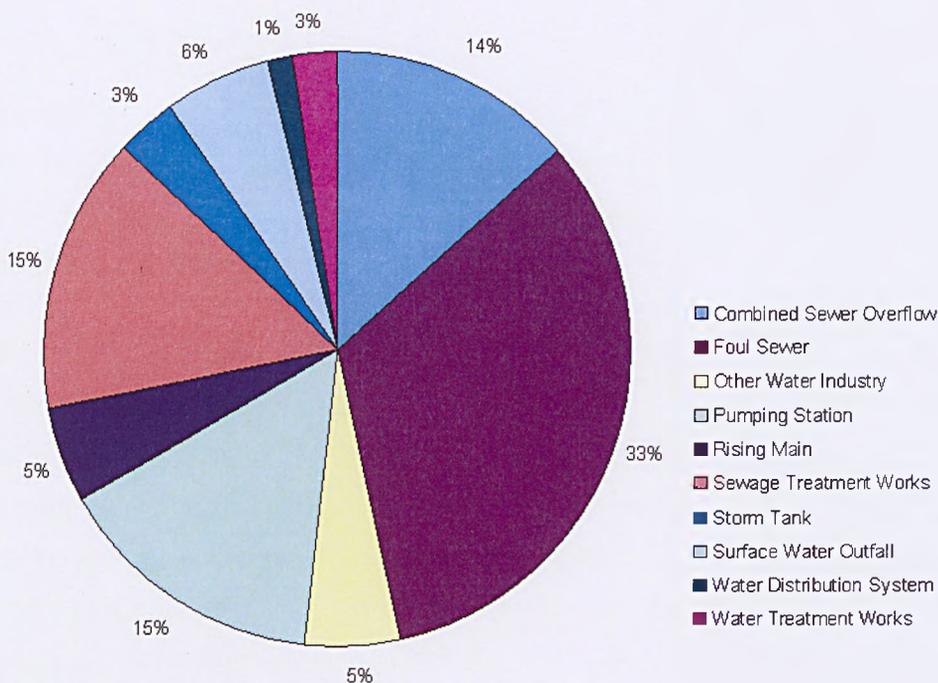


Figure 1.1: Serious (category 1 and 2) pollution incidents where source is from sewage and water industry(www.environment-agency.gov.uk/research/library/publications/34019.aspx)

Within the 2000-2005 Asset Management Plan (AMP3) the water industry has built and upgraded between 2,500 and 3,000 CSOs at a total cost of about £1 billion. During this CSO upgrading, some thousands of new or rehabilitated CSOs were installed with screens, usually based on the WaPUG CSO design guide (WaPUG, 2001), to control and retain aesthetic solids from release in an overflow spill event. In many cases, aesthetic pollution was considered as the main cause of serious pollution incidents (FWR, 2005).

However, the long-term hydraulic performance of CSO chambers and particularly those that incorporate screens is relatively unknown. There is no generic solution for summarising the hydraulic performance of CSO's and CSO's with screens, as there are a large number of different CSO chamber geometries and screen combinations that serve catchments with a wide range of characteristics (Andoh, 1999, Saul, 2000, Burt, 2002).

1.1. Project Aims

The principle aim proposed of this thesis is to build a better understanding of the performance of currently operated CSO assets and subsequently to develop more efficient CSO asset management approaches. The research has been completed in four phases:

Firstly, background literature has been reviewed on the function and design of combined sewer overflows and screen arrangements, together with details of performance problems and current asset management approaches. This has been used to set the scene and to identify the gaps in knowledge that form the basis of this thesis. In addition introduce the initial concept and development of two individual models: An Adaptive Linear Neural Network (ADALINE) CSO chambers hydraulic performance model and a Fuzzy Logic (FL) CSO asset pro-active operation and maintenance decision support model.

Secondly, to interrogate the databases of a UK Water (Yorkshire Water) company and to extract details of the types of CSO structure and screen arrangements used in practice and to determine data on CSO hydraulic performance and their operational and maintenance management. The data used in this research is presented in this section.

Thirdly, to explain and demonstrate how ADALINE CSO chambers hydraulic performance model and a FL CSO asset pro-active operation and maintenance decision support model were developed, tested and applied to water company datasets and to examine the model sensitivity.

Fourthly, to apply the ADALINE performance prediction model and FL pro-active decision support model to a field scale case study. The aim was to demonstrate that the approach gave a better understanding of the CSO operation and need for maintenance, leading to an asset management strategy with recommendations for its implementation.

2. Literature Review

The Literature review for this research included three primary areas:

- ✓ CSO design and current asset operation status
- ✓ Artificial neural network methodology and, in particular, the ADALINE algorithm
- ✓ FL theory and FL expert systems as decision support tools

2.1. Combined Sewer Overflow (CSO)

Generally, CSO design capacity is based on the definition of CSO setting: The CSO setting is the flow the CSO retains in the sewer system for downstream treatment and may be defined by Formula A, equation 2.1. (Ministry of Housing and Local Government, 1970):

$$CSO_{\text{setting}} = DWF + 1360P + 2E \text{ (litres/day)}$$

Equation 2.1

For which $DWF = PG + I + E$

P population

G water consumption per person (litres/day)

I pipe infiltration rate from ground water (litres/day)

E average industrial effluent (litres/day)

DWF dry weather flow

Formula A was considered appropriate for a CSO setting and was based on dry weather flow (DWF) plus some storage allowance and an additional allowance for industrial effluents. During storm events the foul water (dry weather flow) is mixed with rainwater in the combined sewer and this mixes with the waste water

pollutants. It is very important that CSOs only operate (provide spill flow) during heavier rainfall events so that the receiving watercourse and the environment in general are not polluted with foul water. However, combined sewer overflows may also have to meet other water quality objectives with respect to dissolved and finely suspended pollutants and bacteria.

Aesthetic control requirements were developed for all new and existing unsatisfactory discharges to inland and tidal waters in England and Wales and these requirements are based on the combined criteria of the amenity use of the receiving water and the spill frequency, shown as Table 2.1 (WaPUG, 2001).

Table 2.1: Aesthetic Control Requirements

Amenity Classification	Spill Frequency	Aesthetic Control Requirement
High Amenity		
i) Receiving water passes through formal public park	> 1 spill per annum	6 mm solids separation ⁽¹⁾
ii) Formal picnic site		
iii) Influences area where bathing and water contact sport (immersion) is regularly practised (wind surfing sports canoeing)	≤ 1 spill per annum	10 mm solids separation ⁽²⁾
iv) Shellfish waters		
Moderate Amenity		
i) Boating on receiving water	> 30 spills per annum	6 mm solids separation ⁽¹⁾
ii) Popular footpath adjacent to watercourse		
iii) Watercourse passes through housing or frequented town centre area (bridge, pedestrian/shopping area)		
iv) Recreation and contact sport (non-immersion) area	≤ 30 spills per annum	10 mm solids separation
Low Amenity		
i) Basic amenity use only	Not applicable	Solids separation achieved through "best engineering design" of CSO chamber (high side weir, stilling pond, vortex)
ii) Casual riverside access on a limited/infrequent basis (bridge in rural area, footpath adjacent to watercourse)		
Non-Amenity		
i) Seldom or never used for amenity purposes		
ii) Remote or inaccessible area		

Before the WaPUG Guide was published in 2001 (WaPUG, 2001), CSOs were designed according to the industry standard guideline FR0488 (Balmforth et al., 1994) and did not contain screening devices in the CSO structure. This was due

to the fact that if designed properly, the FR0488 CSOs were found to provide significant retention of aesthetic solids and other finer settleable or floating materials. However, a significant proportion of aesthetic solids are neutrally buoyant and do not therefore lend themselves to separation and retention in a conventional CSO chamber (WaPUG, 2001). Hence to meet the aesthetic criteria the UK water companies have adopted an approach based on the retention of all solids of dimension 6mm in 2 dimensions. A number of different types of mesh screens have subsequently been developed to meet the retention of these aesthetic solids and these are now widely implemented in practice. Therefore, currently most CSOs incorporate screens that ensure suspended solids and other solid matters are kept within the sewer network to be removed at the WWTP rather than to be released directly to the watercourse (WaPUG 2001).

2.1.1. Construction chamber types

Generally there are three main construction types of CSOs.

High-side weir chamber:

The high-side weir chamber (see Figure 2.1) consists of a stilling zone upstream of the weir and a storage zone downstream of the weir. In the stilling zone the gross solids are supposed to settle to the chamber bottom, whereas the floatables are retained within the surface waters of the storage zone. In front of the weir a scumboard is placed to protect the floatables from spilling over. A high-side weir chamber may have single or double weirs (Saul et al., 1997).

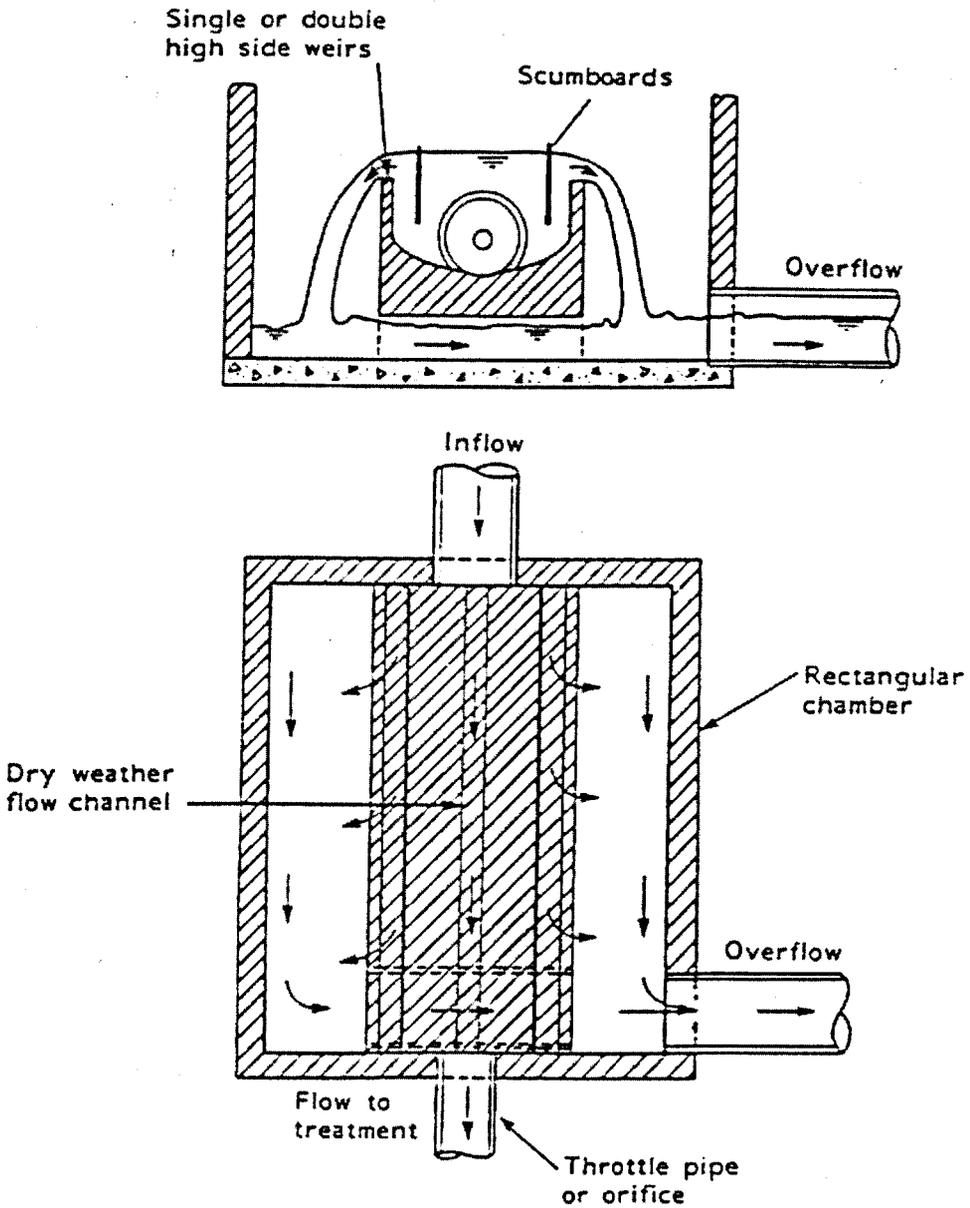


Figure 2.1: High-side weir chamber (Balmforth and Henderson, 1988)

Stilling Pond:

A stilling pond chamber has a transverse weir at the downstream end of its rectangular plan shape (see Figure 2.2). Similar to the high-side weir chamber,

the inlet and the continuation pipe lie on the longitudinal centre line. Particles with a fall (slow) velocity settle to the base of the chamber and are discharged from the continuation flow to downstream network and treatment works. With chamber flow velocity increasing, particles with a rise velocity are prevented from passing over the chamber weir and retained by a scumboard. These particles are then transported to the upstream of the chamber, because of the recirculatory of the flow. These particles are retained until the end of the overflow event and discharged in continuation flow as the flow in the chamber subsided (Saul et al., 1997).

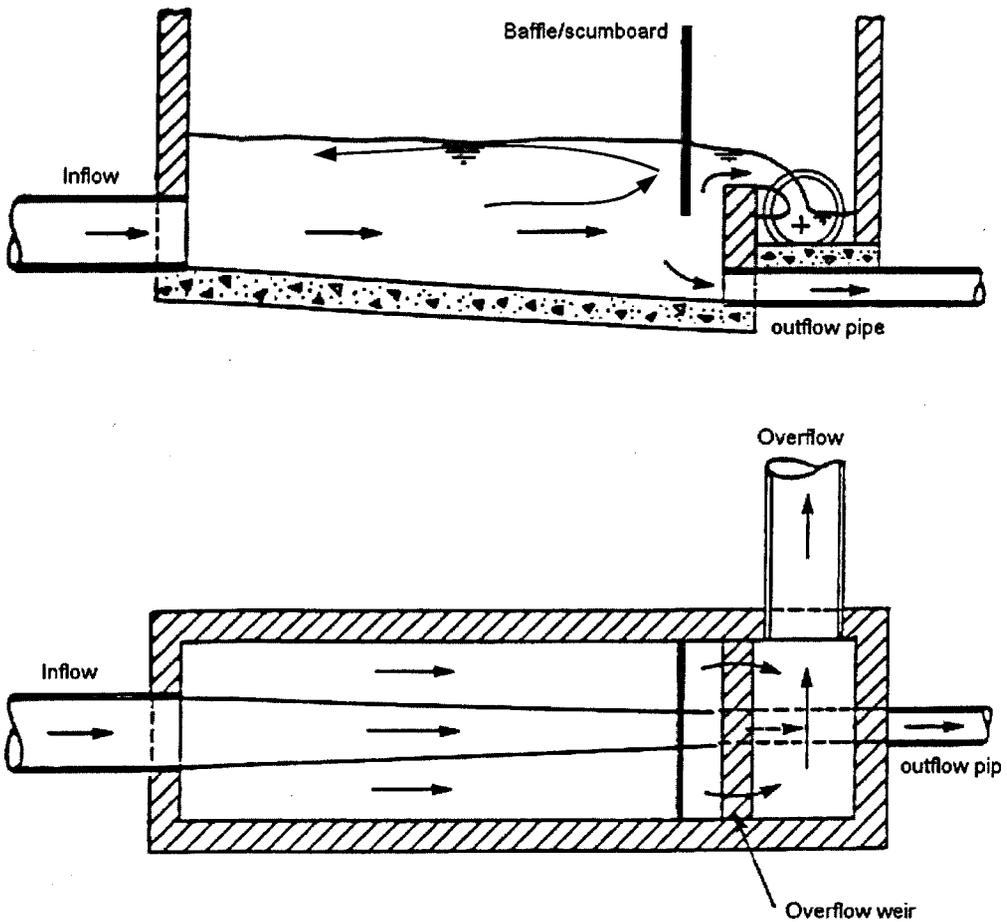


Figure 2.2: End weir stilling pond chamber (Saul et al., 1997)

Hydrodynamic separators (i.e.: Storm King TM):

Hydrodynamic separator is an enhanced rotary flow field and velocity distribution within the chamber achieved by specifically designed internal components. Downward helical flow in the outer region of the chamber and an upward helical flow near the centre are created. With the underflow outlet pipe the separated gross solids and settled material is discharged in the continuation flow. To improve the retention of floatables, several baffles and a dip plate are installed within the chamber to keep the floatables between the dip plate and the outer wall of the structure. The spilled flow will be discharged over a circular (in plan) overflow weir from where it is conveyed via a spillway channel to the overflow pipe (Saul et al., 1997).

Further examples of hydrodynamic separators are the US EPA Swirl Concentrator Regulator and the German-designed Fluid separating device (Saul et al., 1997).

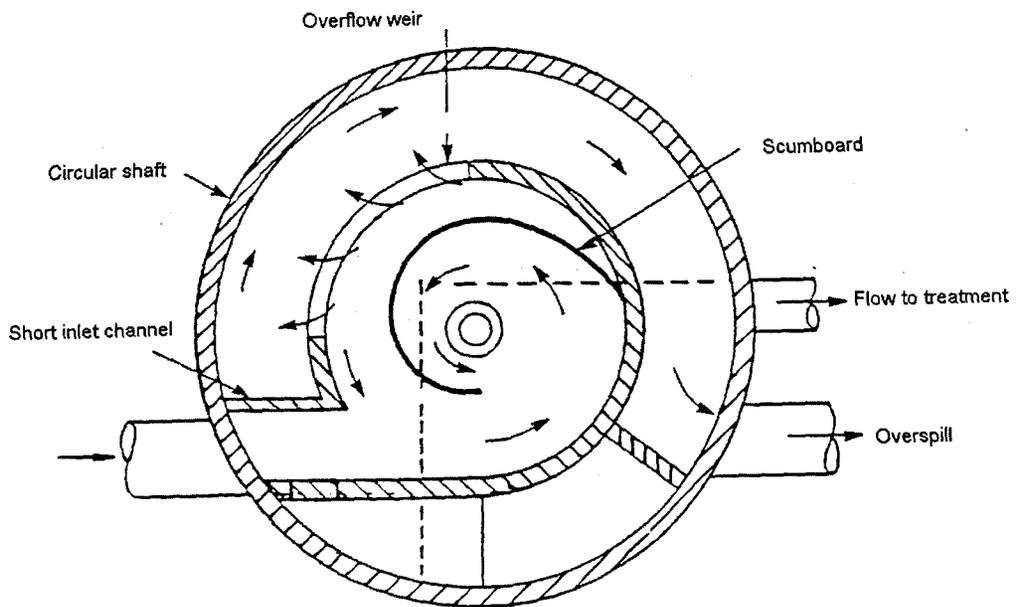
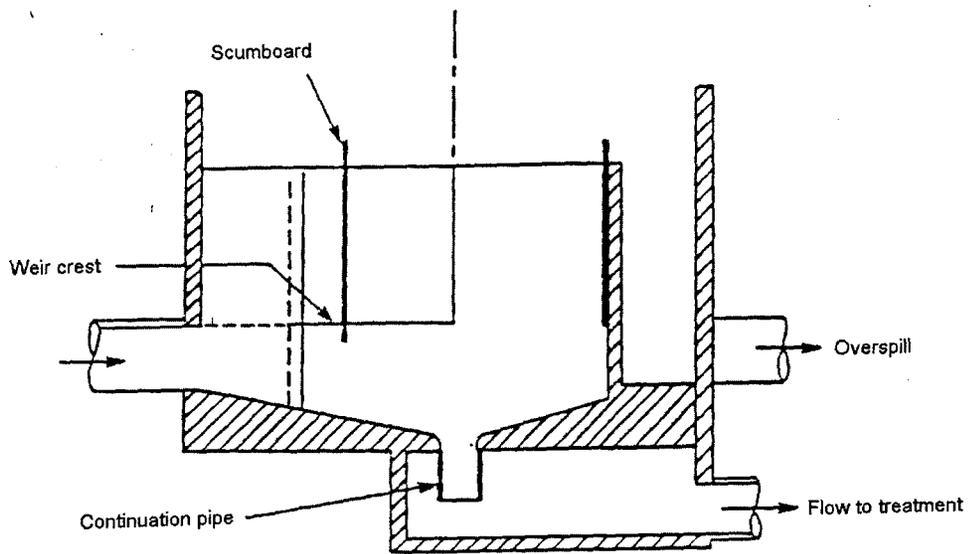


Figure 2.3: Hydrodynamic separator (here: Storm King™, Saul et al., 1997)

Generally, such static screens are simply designed non-power screening technology suitable for sites with infrequent overflows. Static screens normally require a low capital cost but may incur relatively high operational costs compared with other types of screens due to the requirement for regular cleaning.

Mechanical Screens:

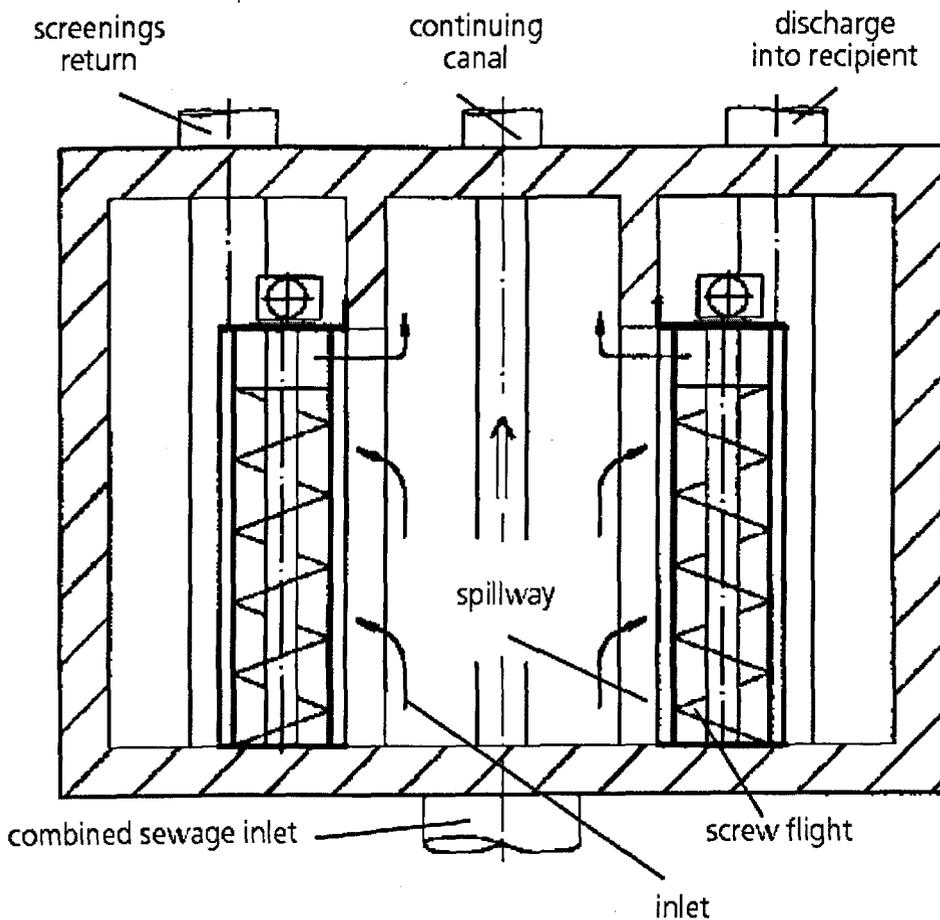


Figure 2.5: Huber ROATAMA Mechanical Storm screen (Huber™)

ROTAMAT mechanical screen produced by HUBER tech was introduced (Figure 2.5). The characteristics were described as:

- ✓ Hydrodynamic separation and integral storage with Storm King® Overflows
- ✓ 6 mm two dimensional screening
- ✓ Rotary Jet™ Screens
- ✓ No power required
- ✓ Self-cleansing and self-activating

2.1.3. Subheading –super/subcritical conditions

According to the CSO design guide (WaPUG, 2006), it is desirable that the hydraulic performance of flow inside CSO chamber is suitable for the effective operation of a screen device. The initial requirement is that the flow throughout the CSO chamber is subcritical, i.e. the Froude Number of the flow is less than 1. The Froude Number is a ratio of the flows inertia to its gravitational forces (Equation 2.2):

$$Fr = \frac{V}{\sqrt{g \frac{A}{W}}}$$

Equation 2.2

V = velocity of flow, *m/s*

g = gravitational acceleration = $9.81m/s^2$

A = cross section area of flow, m^2

W = width of water surface,

Conversely, any flow with a Froude Number is larger than 1 called supercritical flow. Typical subcritical and supercritical flow conditions are shown in Figure 2.8:

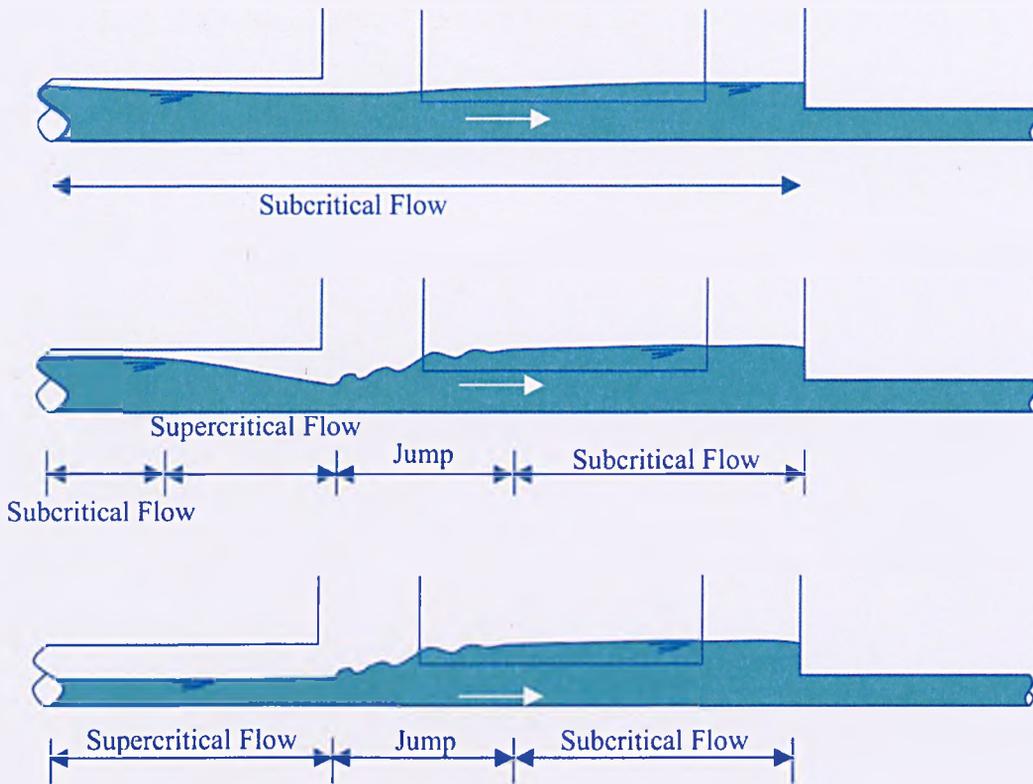


Figure 2.8: Possible flow condition at inlet pipe to CSO chamber (WaPUG, 2006)

As seen from Figure 2.8, the transition from super to sub is achieved with a hydraulic jump, which results in significant turbulence. The turbulence flow in CSO chamber has a significantly effect on the performance of screen devices (WaPUG, 2006).With the purpose of avoiding such a hydraulic jump, the WaPUG CSO Design Guide provides guidance on the required length of the overflow weir and the diameter of the inlet pipes to eliminate supercritical flow inside CSO chamber.

2.1.4. Heading – depth discharge relationships

The CSO chamber hydraulic performance during a rainfall event was discussed by Fach et al (2008). A CSO with single side weir chamber was used to test the hydraulic performance with purpose of determining relationship between CSO

inflow rate and chamber water level performance (Fach et al 2008). See Figure 2.9:

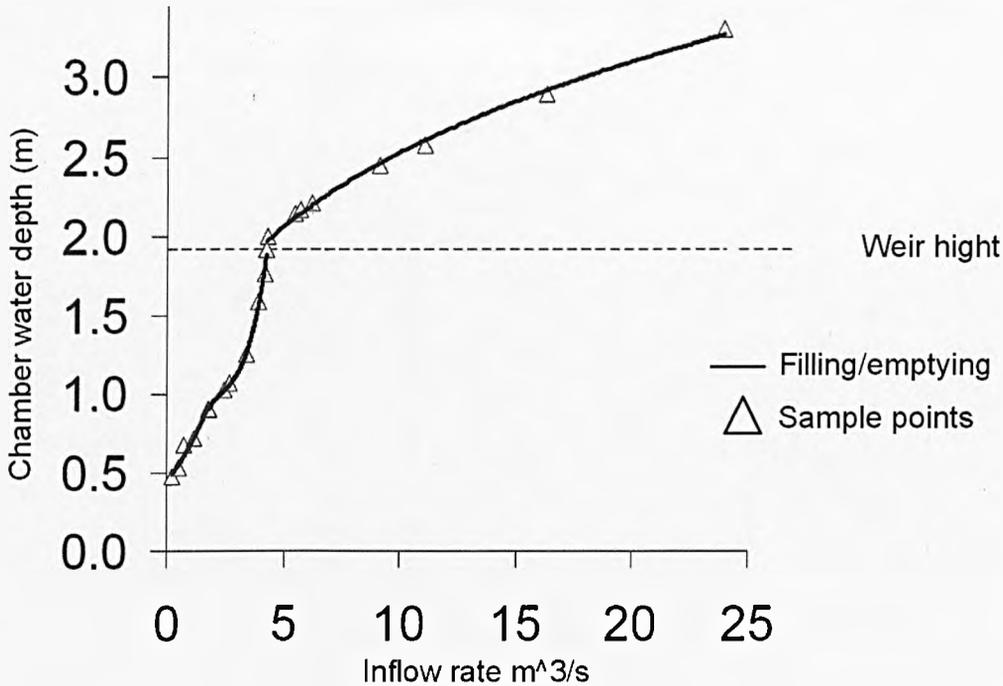


Figure 2.9: Relationship between inflow rate and CSO chamber water depth (Fach et al 2008)

Fach et al (2008) analysed concerning the relation between rate of inflow and corresponding water depth, the result is showing in Figure 2.9. The increasing pattern of CSO chamber water depth can be split into two phases

1. The chamber water depth rises sharply with inflow velocity is increasing until the water depth reaches the weir crest; this phase represents a filling of the chamber storage with outflow via the continuation pipe.
2. When spill occurs, the rising rate of water depth becomes lower due to the change in the governing equations (both orifice and weir flow).

2.1.5. Monitoring of CSO

With the purpose of obtaining CSO hydraulic performance data, monitoring devices had been installed in a large number of CSO chambers. An example of a CSO monitoring system is shown in Figure 2.10:

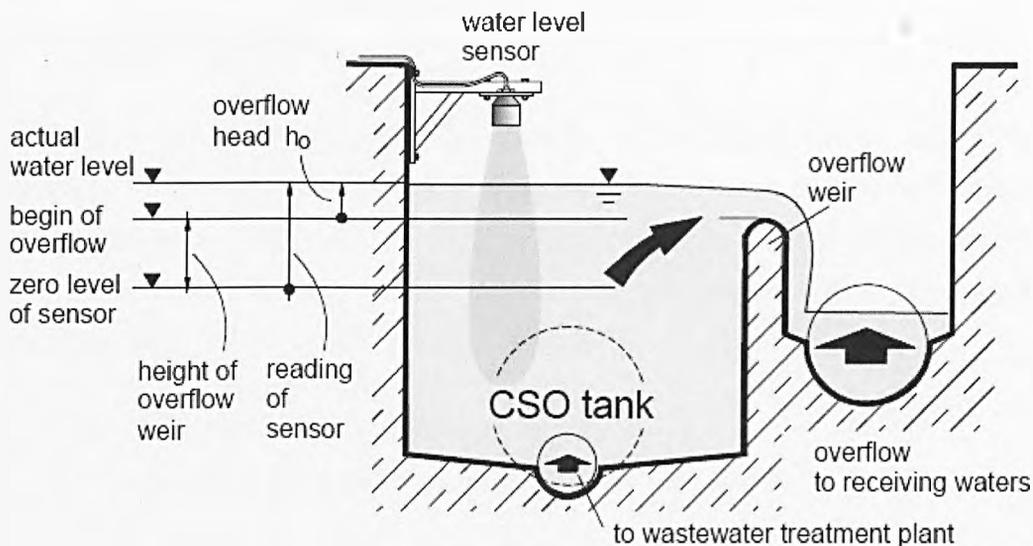


Figure 2.10: Monitoring device installation in CSO tank (Weiss et al. 2006)

As shown as Figure 2.10, the water level sensor that is used to measure the CSO chamber water depth is installed above the main body of flow within the chamber. The data from the depth monitor is generally transmitted, remotely, to a database held by the operating company. The data is then analysed to detect anomalies or unexpected changes in performance, for example, spill flow operation in dry weather Weiss et al. (2006).

2.1.6. Asset operation and maintenance

Guidance on the operation and maintenance of CSO's is given by the Capital Maintenance Planning Common Framework developed by UKWIR (UKWIR, 2002). With the guidance of this common framework, water companies have

developed their operation and maintenance plans, which were assessed by Ofwat using a four stage process (Ofwat, 2004):

1. Review historic expenditure and serviceability as a baseline
2. Identify differences in future requirements
3. Make assumptions about future efficiency
4. Take account of interactions with schemes to improve quality or the supply
– demand balance

From an engineering viewpoint, an iterative and comprehensive risk based approach to management of sewer system assets compatible with Common Framework was introduced in Sewer Rehabilitation Manual (WRc, 1994). Specifics on CSO asset operation and maintenance were introduced generally as two categories (BS EN 752-7, 1998) (Butler, 2008):

- ✓ Routine operation and maintenance actions
- ✓ Reactive Operation and Maintenance (O&M)actions

Maintenance actions are carried out to prevent performance failure or in response to a reported failure or due to an alarm triggered from the measured data.

2.2. Artificial Neural Network

2.2.1. General concept and development

Generally, an Artificial Neural Network (ANN) is an information processing system that is inspired by the way biological nervous systems, such as the brain, process information. The key element of this system is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working together to solve specific problems. ANNs, like human beings, learn from examples. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments

to the synaptic connections that exist between the neurons, which is true for ANNs as well (typically called weights).

Neural networks rely on training data to initialise and update the system. Thus, an ANN requires an appropriate training set that allows the system to learn and generalise on future input data. The combination of inputs is very similar to previous training data. They are recognised and result in a similar output, while new data (or incomplete and/or noisy data) can be matched as closely as possible to patterns previously learned by the system (Medsker et al., 1999).

McCulloch and Pitts (1943) devised simple electrical networks composed of Binary Decision Units (BDN). Each BDN has a number of inputs and is constructed to emit a unit pulse if the total activity coming to them from similar units is greater than a certain (threshold) value, otherwise they are silent. It was shown that such a network could perform any logical function on its inputs. What made this approach very interesting was the fact that the BDN is a very simple model, operationally very similar to the human nerve cell used for thinking in the brain.

Some years later, Donald Hebb (1949) wrote about early theories of neural learning and Rosenblatt (1958, 1962) developed the perceptron model. Hebb proposed a learning mechanism called Hebbian learning. Hebb based his rule on real neuronal observations. It involved reinforcing active connections only. Rosenblatt further developed this technique into a method of training a network of model neurons (perceptron). Active connections could be strengthened or weakened. Rosenblatt demonstrated that they could correctly categorise some kinds of patterns, even with noise at the inputs.

Rosenblatt studied single-layer and two-layer perceptron, but was only able to prove that his single-layer perceptron (the inputs are fed directly to the outputs via a series of weights) could separate inputs into two classes if the two classes were linearly separable which means completely separated by a single line. The

only way out of this theoretical impasse was to move onto networks with multiple layers in which the outputs of one network are used as the inputs to the next layer. This provided the necessary complexity to solve linearly inseparable problems, but the delta rule could not be applied to a multi-layer perceptron.

Minsky and Papert (1969) proved that the single-layer perceptron could not solve a linearly inseparable problem (a large class of problems) and interest in neural network research suffered a period of decline.

It was not until 1986 when Rumelhart, McClelland and Williams published their paper detailing a method for training a Multi-Layer Perceptron (MLP) that the problems posed by Minsky and Papert were solved. MLPs trained with back propagation are, in theory and given sufficient training data, universal computing machines capable of arbitrary function approximation.

2.2.2. ADALINE

ADALINE is a single layer neural network. As one of the pioneer neural network models, it was developed by Professor Bernard Widrow and a graduate student, Ted Hoff, at Stanford University in 1960. ADALINE is based on the McCulloch-Pitts neuron that consists of a weight, a bias and a summation function. However, the difference between ADALINE and standard McCulloch-Pitts perceptron is indicated as that during the learning phase the weights of neurons are adjusted according to the weighted sum of net inputs. In the sense of standard perceptron, the net is processed to a certain transfer function and output of this function is applied to adjust the weights.

Generally, ANNs can be classified by their architecture and their learning algorithm. There are two main types of neural network architectures, the feedforward and feedback structures. Similarly, there are two main mechanisms for learning: supervised and unsupervised. ADALINE basically followed the feedforward learning structure with supervised learning mechanism and

implemented least mean square error (LMS) learning algorithm (Widrow, 1962). ADALINE architecture is shown as Figure 2.11.

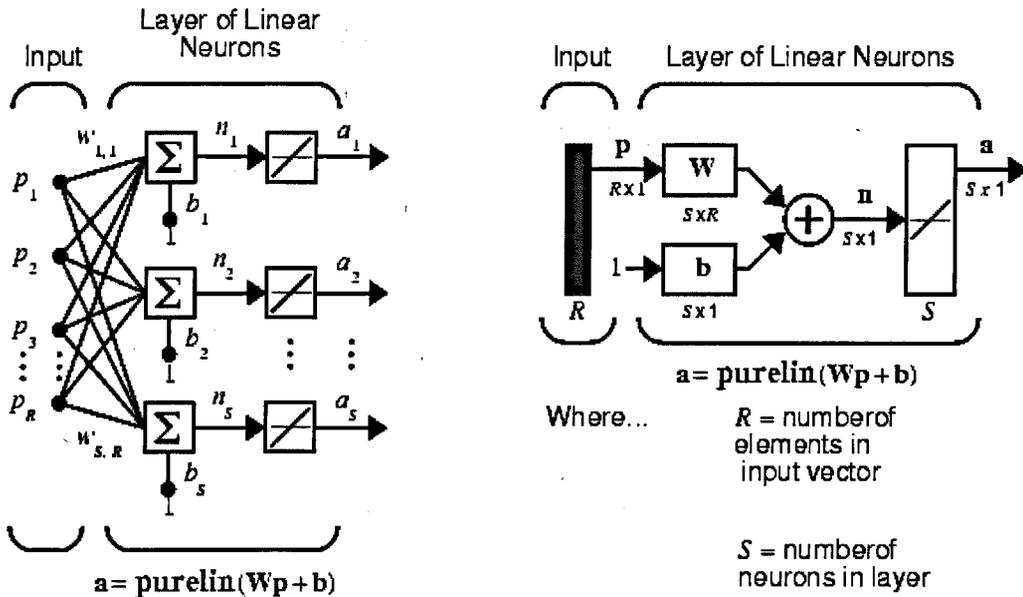


Figure 2.11: Adaptive Linear Network Architecture (Widrow, 1985)

- ✓ p_1, p_2, \dots, p_R : Elements in input vector
- ✓ $W_{11}, \dots, W_{S,R}$: Weight of each input element in transfer function (linear)
- ✓ b_1, b_2, \dots, b_S : Bias each transfer function
- ✓ a_1, a_2, \dots, a_S : Output of linear neural layer
- ✓ R : Number of elements in input vector
- ✓ S : Number of neurons in layer

Feedforward ANNs allow signals to travel one way only; from input to output. There are no feedbacks (loops) i.e. the output of any layer does not affect that same layer. Feedforward ANNs tend to be straight forward networks that associate inputs with outputs. The MLP (Multi Layer Perceptron) is an example of a feedforward neural network.

Most ANN solutions have been trained with supervised learning. With supervised learning the actual output of a neural network is compared with the desired output. Weights, which are usually randomly set to begin with, are then adjusted by the network training algorithm so that the next iteration, or cycle, will produce a closer match between the desired and the actual output. The learning method tries to minimise the current errors of all processing elements. This global error reduction is created over time by continuously modifying the input weights until the acceptable network accuracy is reached.

The least mean square error algorithm is an example of the supervised training approach, in which the learning rule is provided with a set of samples of desired network behaviour. The mathematical expression shown as:

$$[P_1, t_1], [P_2, t_2], \dots, [P_n, t_n]$$

Where

P_n is an input to the network, and t_n is the corresponding target (the real value). As each updated input is applied into the network, the related network output is assessed by comparing with the target. Therefore, the error is captured and calculated as the difference between the target output and network output. The initial purpose is to minimize the average of the sum of these errors, shown as the Equation 2.3:

$$mse = \frac{1}{n} \sum_{k=1}^n e(k)^2 = \frac{1}{n} \sum_{k=1}^n [t(k) - a(l)]^2$$

Equation 2.3

The LMS learning algorithm adjusted the weights and biases of ADALINE, to minimise the mean square error (Widrow, 1995).

2.2.3. Least Mean Square

Least mean square algorithm was also applied initially in another popular mathematical prediction method –linear regression (LR).

According to the introduction, the adaptive linear prediction approach considered the whole system and inputs developing mechanism appeared to be highly similar with linear regressions (Draper, 1998).

Linear regression:

Considering a given a data set as $\{y_i, x_{i1}, \dots, x_{ip}\}_i^n$ of n units, a linear regression model assumes that the relationship between the dependent variable y_i and the p -vector of regressions x_i is approximately linear. This approximation relationship is modelled through a 'disturbance term'- bias $-\varepsilon$, the model takes the form as Equation 2.4:

$$\beta_1 x_{i1} + \dots + \beta_p x_{ip} + \varepsilon_i = y_i = x_i' \beta + \varepsilon_i \quad i = 1, \dots, n,$$

Equation 2.4

Where $x_i' \beta$ is the inner product between vector x_i and β which may be expressed in the form of Equation 2.5:

$$y = X\beta + \varepsilon$$

Equation 2.5

When expended this may be described by Equation 2.6:

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, X = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{np} \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}, \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

Equation 2.6

Least mean square application in linear regression:

Assume that:

- ✓ W: weight
- ✓ X: input vector
- ✓ Y: output vector
- ✓ ε : bias
- ✓ 1,2,3,...i means W, X, Y are all i dimensional vectors

$$Y = [X]^T W + \varepsilon$$

Equation 2.7

ε in Equation 2.7 represents bias which may be expressed as:

$$\varepsilon = Y - [X]^T W$$

Equation 2.8

Defining the bias in Equation 2.9:

$$\text{LMS} = \min\left(\frac{\varepsilon^2}{W}\right) \Leftrightarrow \frac{\partial \varepsilon^2}{\partial W} = 0$$

Equation 2.9

The calculation of Least Mean Square may be calculated using Equation 2.10:

$$\frac{\partial \varepsilon^2}{\partial W} = 0 \Leftrightarrow \frac{\partial}{\partial W} \{[Y - XW]^2\} \Leftrightarrow \frac{\partial}{\partial W} \{[Y^2 - 2YXW + XWXW]\} \Leftrightarrow 0$$

Equation 2.10

In integral form Equation 2.10 becomes

$$W = [X^T \cdot X]^{-1} \cdot X^T \cdot Y$$

Equation 2.11

In Equation 2.11, where X and W are vectors and hence the display sequence cannot be changed in the equation. T means the inverse dimension of the matrix X .

Compared with the application of adaptive learning algorithms in an ADALINE, the linear regression model requires a re-calculation from the first input data in the case when either new data is updated into this system or the data is based on a prediction model based on a dynamic system.

Explained as Equation 2.12:

$$y = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \cdot w_\varepsilon$$

Equation 2.12

The matrix equation was re-calculated every time when a change happened in the system or new data was updated in a dynamic system and hence the working load was massive compared with the application of neural network model that contained learning and memorising approaches. The weight parameters of each input were adjusted every time step moving forward. As a consequence the ADALINE approach was used in this study.

2.2.4. Application in water industry

ANNs have become very popular for application in the water industry in recent years. The following section provides a brief overview of recent ANN applications in areas related to this research project such as rainfall-runoff modelling, river level, flow and flood forecasting. Specific to CSO hydraulic performance, papers were also reviewed, which studied the applicability of ANN methodology to predict CSO discharges.

Shamseldin (1997) applied a multilayer feed forward neural network with one hidden layer for rainfall runoff modelling comparing the performance of the technique with models using similar input information such as simple linear model (SLM), seasonally based linear perturbation model (LPM) and the nearest neighbour linear perturbation model (NNLPM). Using the data of six real catchments in Nepal, Ireland, USA, Australia and two in China, the approach shows promising, but variable results. In some catchments the results are good, in others quite poor. Therefore, it was suggested considering different network architectures or different transfer functions as the chosen logistic activation function.

Tokar et al. (2000) compared ANN models with traditional conceptual models in predicting watershed runoff as a function of rainfall, snow water equivalent, and temperature. The ANN technique was applied to model watershed runoff in three basins with different climatic and physiographic characteristics. Back-propagation was used as supervised training algorithm in the chosen MLP by developing the prediction model for each catchment, and the outputs were compared with different conventional water balance model. It was concluded that ANN approach appeared to be more efficient than traditional conceptual models in predicting watershed runoff with the knowledge of local rainfall information. Also, they indicated the potential of applying ANN techniques to forecast CSO discharge performance, which is a direct function of precipitation-runoff.

Solomatine et al. (2003) compared a model tree approach with a multi-layered perception (MLP) network trained with the back-propagation algorithm to learn the rainfall runoff relationship. A nonlinear activation function was used in the hidden layer, whereas a linear activation function was used in the output layer because it is unbounded and able, to a certain extent, to extrapolate beyond the range of the training data. For the model tree (MT), the same sets of input-output, and training and verification data were considered as for the ANN. Both ANNs and MTs produced excellent results for 1 hour forward prediction, acceptable results for 3 hour forward prediction and conditionally acceptable

results for 6 hour forward prediction under a 1 in 10 year return rainfall. Both techniques have almost similar performance for 1 hour forward prediction of runoff, but the result of the ANN is slightly better than the MT for longer lead times. However, it was emphasised, that the results of the MT were easier to understand and that the ANN is not the only data-driven model that can be used in hydrology or elsewhere. Attention to ANNs is without any doubt justifiable, but other models deserve attention as well.

Hessami et al. (2004) conducted a comparison based on ANNs including multilayer feed forward networks and radial basis functions for the post-calibration of weather radar rainfall estimation. The multilayer feed-forward training algorithms consisted of four variants of the gradient descent methods; four variants of the conjugate gradient method were compared: Quasi-Newton, One Step Secant, Resilient back-propagation, Levenberg-Marquardt method and Levenberg-Marquardt method using Bayesian regularisation. In general, results showed that the Levenberg-Marquardt algorithm using Bayesian regularisation can be introduced as a robust and reliable algorithm for post-calibration of weather radar rainfall estimation. The author believed that the proposed artificial neural network may be used as a general data integration and data calibration.

Schulze et al. (2005) published a paper to determine the possibility of using artificial neural networks in integrated water management by illustrating some samples of applications of ANN on several critical aspects of integrated water management such as prediction of water quality parameters, which indicated that, a relatively simple flood warning or protection system can be developed based on ANN techniques. Mostly, the exact causes and relations are not clear, but potential significant variables are known. An ANN was very suitable to fit the best relations and trace all possible relations based on information in the past. Therefore, the author came up with a conclusion of developing a new integrated water management methodology with the purpose of achieving real time control of the entire catchment.

River flow, level and flood forecasting research based on rainfall and runoff parameters by applying ANN techniques were also successfully carried out in the last few years. Aqil et al., (2007), developed the three layers feed forward neural network model with two types of neural network architectures and three types of training algorithm to predict flood level in a river basin by using the collected rainfall value from local rain gauge. The author announced that, compared with six different neural network models, the one with Levenberg-Marquardt algorithm model is able to forecast the flood level up to 5 hours in advance with reasonable prediction accuracy.

Similar research on ungauged catchment in the UK was done by Dawson et al., (2006). ANN had been applied to develop prediction model to predict T- year flood events and the index flood (the median of the annual maximum series). This model was developed across the scale of small catchment data provided by the Centre for Ecology and Hydrology's Flood Estimation Handbook (FEH). The conclusion from this research demonstrated the feasibility of using ANNs to model flood events in ungauged catchment, however, the Neural network approach was recognised as heavily data dependent, which can be hardly applied into a system with limited data set. In addition physical processes existed in the system which cannot be accounted by ANN model, which reduced confidence in model prediction.

Sumer et al., (2007) carried out research, specific to sewer systems, on Sanitary Sewer Overflow (SSO) water depth and flow prediction models by applying a multilayer perceptron neural network. Historical observation of SSO flow information from multiple sites over system-wide upstream to downstream were collected and used as the model input dataset in the Sumer et al., (2007). Model and came up with positive output on SSO flow prediction. Predicted alarm mechanisms were also mentioned in this research, mathematical models can be used in conjunction with control theory to detect these disruptions that presented the detected abnormal asset performance.

A feed forward multilayer perceptron with three hidden layers model was developed by Kurth et al., (2008) to predict CSO chamber flow by taking into account of both recorded CSO flow value and rainfall information. This study also applied the prediction model to 2 case studies where rainfall and CSO depth had been recorded for a one year period. The model was able to predict three time steps (5 minutes per step) in advance. Hence the research completed by Kurth, proved the feasibility of applying an ANN approach on CSO hydraulic performance prediction. However, Kurth stressed that the trained ANN model was only fit for purpose within the boundaries of the data set used to train the model, in this case where the rainfall event is lower than return level of 1 in 5 years. Outside this range of data, the application of the model is less certain.

Kurth also reported that the model training process of multilayer ANN models was time-consuming, compared to the actual prediction time when applied to complex systems. In contrast to the multilayer ANN model, the ADALINE approach takes much less time to train the model but that the results are less accurate when applied to complex system performance prediction.

In summary, it is clear that there is significant potential to apply ANN's to predict the performance of CSO chambers and in this study the ADALINE model was adopted. This decision was based on the speed of application and the data requirements used to train the model, relative to the more complex models reported above.

2.3. Fuzzy Logic expert system

2.3.1. Fuzzy set

'Fuzzy logic' concepts were first merged and developed with the theory of fuzzy sets by Lotfi Zadeh in 1965 (Hajek, 1998). Compared with a crisp dataset, a fuzzy set is a set with elements that have degrees of membership. Based on classical 'set' theory, the membership of elements included in this set is assessed following a common rule: if this element belongs to the set or not,

which is called a 'bivalent' set. However, fuzzy set theory permits a gradual assessment instead of pure 'either-or' rule. There is a membership function that is valued in a real unit interval of 0 to 1. With the special cases of membership function in fuzzy sets, the membership function boundary was defined by an indicator function that takes a; value of 0 or 1 (Zadeh, 1965).

To grade the parameters in a fuzzy set, the concept of degree of truth has to be explained:

A fuzzy set A of a crisp set X is characterised by evaluating each individual data x included in X with the degree of membership of x in A . If X is defined as a set of propositions then element x may be presented as their degree of truth, which may be "absolutely true", which means the degree of truth value is 1 or "absolutely false", which means the degree of truth value is 0 or some intermediate truth degree such as a proposition may be presented as 'more true than another proposition'.

A fuzzy set (A, m) in which A is the set and: $m: A \rightarrow [0,1]$

For each $x \in A$, $m(x)$ is called the grade of membership of x in the set A .

if $m(x) = 0$, then x called absolutely false to crisp set X

if $m(x) = 1$, then x called absolutely truth to crisp set X

The usual membership function with values from 0 to 1 are called [0, 1]-valued membership function (Goguen, 1967).

2.3.2. Membership function

The membership function of a fuzzy set is developed with the purpose of presenting the degree of truth as an extended valuation. The conceptual differences between degrees of truth and probabilities, is that the fuzzy truth represents a membership based on a fuzzy defined set. However, their

probabilities are a description of some event or condition (Zadeh, 1965). There are many types of fuzzy logic membership function and Figure 2.12 highlights the distributions that may be used in MATLAB. Here, the terms of each function's definition are explained:

- ✓ Trapmf: Trapezoidal-shaped built-in membership function
- ✓ Gbellmf: Generalized bell-shaped built-in membership function
- ✓ Trimf: Triangular-shaped built-in membership function
- ✓ Gaussmf: Gaussian curve built-in membership function
- ✓ Gauss2mf: Gaussian combination membership function
- ✓ Smf: S-shaped built-in membership function
- ✓ Zmf: Z-shaped built-in membership function
- ✓ Psigmf: Built-in membership function composed of the product of two sigmoidally-shaped membership functions
- ✓ Dsigmf: Built-in membership function composed of the difference between two sigmoidal membership functions
- ✓ Pimf: Π -shaped built-in membership function
- ✓ Sigmf: sigmoidally-shaped membership function

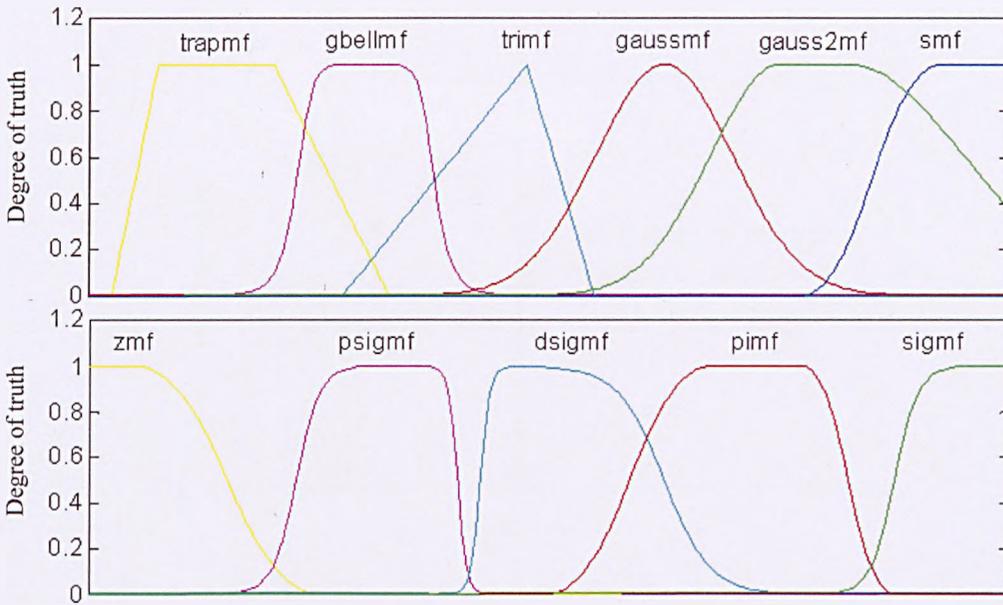


Figure 2.12: Membership function gallery indicated in MATLAB

The process of defining membership functions is the most subjective part of the process of FL modelling. Each variable must have membership functions, which usually are represented by linguistic terms, which define the entire range of possible values. The linguistic terms normally describe a concept that is related to the value of the variable, such as low, average and high. These linguistic membership functions define the degree to which a particular numerical value of a variable fits the concept expressed by the linguistic term. The value of r ranges from zero (not part of the set) to one (perfectly represents the linguistic concept).

2.3.3. Fuzzification

The membership function is used to associate a grade, which is a actually value of input dataset, to each linguistic term. The process of transforming crisp values into grades of membership for linguistic terms of fuzzy sets is called Fuzzification. Four examples of Fuzzification process based on commonly used types of membership function and applied to CSO chamber water depth are introduced below (see Figure 2.13 to Figure 2.16):

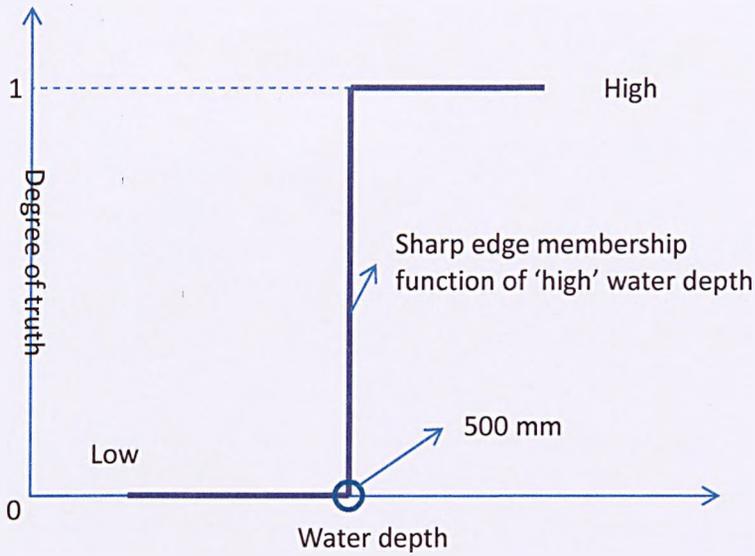


Figure 2.13: Straight line membership function

Figure 2.13 shows a straight line membership function of the CSO chamber water depth, which contains a sharp edge membership function of 'high', is the simplest type. As an example that was shown in Figure 2.13, any water depth that is higher than the value of 500 mm is defined as 'high'.

Water depth is presented as a line with a certain slope in trapezoidal type membership function, as shown in Figure 2.14. The example of trapezoidal membership function indicates that the definition between "low" water depth and 'high' water depth is gradually changing. As can be seen from Figure 2.14, when the water depth is 450mm, the truth degree of 'high' is 0.95, which is identified as more likely to be a 'high' water depth. However, when water depth is 200mm the degree of truth is only 0.1, which means it is unlikely to be defined as 'high' water depth.

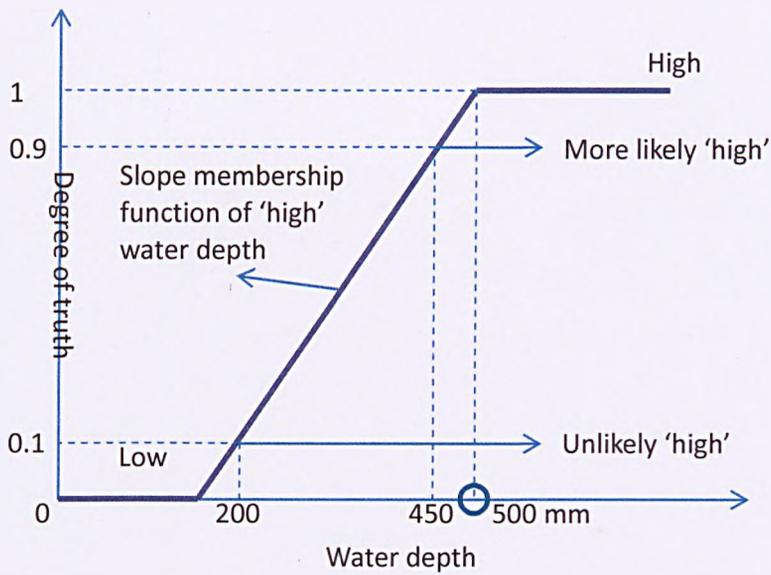


Figure 2.14: Trapezoidal membership function

In the example of Gaussian type of membership function, a 'medium' water depth is defined as the value of CSO chamber water depth is 300 mm. In Figure 2.15, the membership function, which is defined with a boundary of 'medium water depth' is defined as a Gaussian expression. Where the chamber water depth is 300mm, the truth of degree to 'medium' water depth is 1. The sample point that is close to middle line is identified as most likely to be 'medium' water depth.

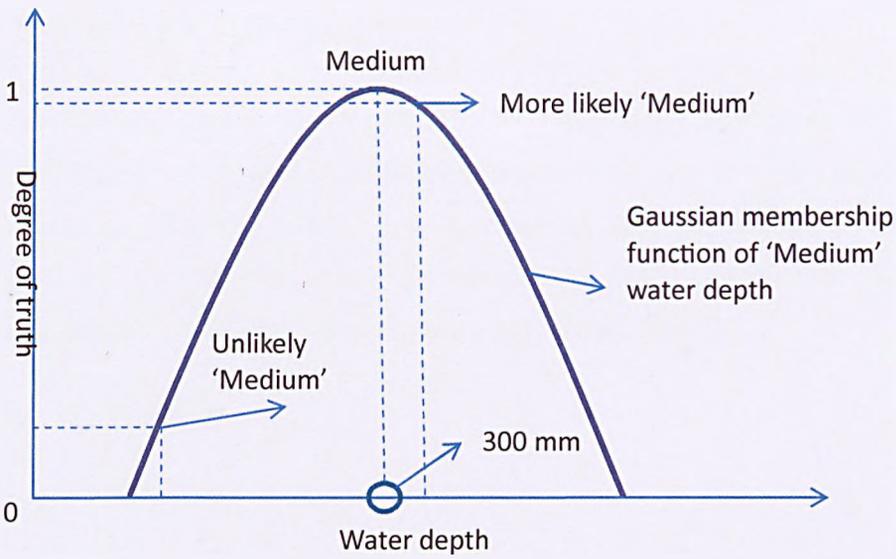


Figure 2.15: Gaussian membership function

In terms of a triangular type of membership function, a triangular boundary is used to present the 'medium' water depth definition, as shown in Figure 2.16:

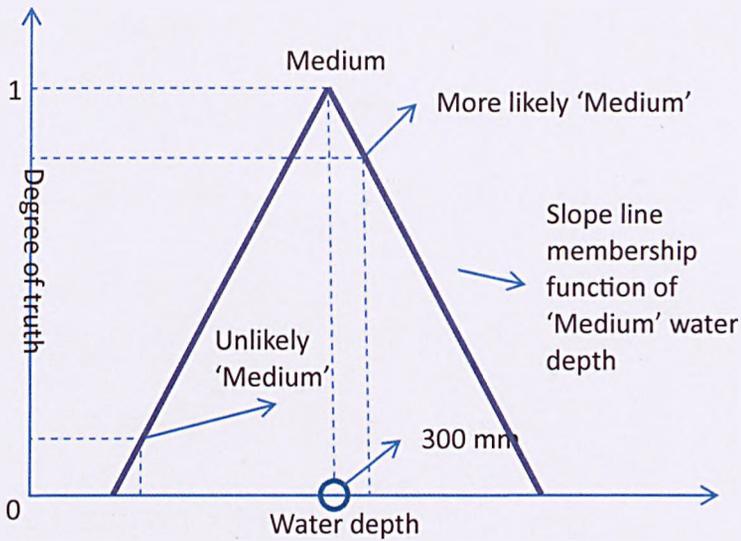


Figure 2.16: Triangular membership function

2.3.4. Fuzzy logic operators

Logic operators are used in the FL methodology. Here it is proposed to define the operations between each membership function, developed in the form of different fuzzy categories (Zadeh, 1968). The process of operator selection is also considered as an important task in developing the FL inference system. Commonly used types of logic operators are introduced below:

Max: $Max(A, B) = A \cup B$

Min : $Min(A, B) = A \cap B$

Complement: $Comp(A) = 1 - A$

Developed logic operator rules shown as:

De Morgan's low: $\overline{A \cap B} = \bar{A} \cap \bar{B}, \overline{A \cup B} = \bar{A} \cap \bar{B}$

Associativity: $(A \cap B) \cap C = A \cap (B \cap C), (A \cup B) \cup C = A \cup (B \cup C)$

Commutativity: $A \cap B = B \cap A, A \cup B = B \cup A$

Distributivity: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C), A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

∪: And operation

∩: Or operation

A, B, C: Membership functions

2.3.5. Fuzzy logic If-Then rule

Previously introduced fuzzy sets and FL operators are recognised as the subjects and verbs in the theory of developing a FL system. To complete an effective FL expert system, If-Then rules are applied to formulate the conditional

statements between each defined fuzzy set and coordinate the selected FL operators. An example of fuzzy If-Then rule is shown as:

If x is A_1 Then y is B_2

Here A_1 and B_2 are linguistic descriptions of fuzzy sets defined by membership functions in the universe of discourse X and Y respectively. The 'If' part of the rule, ' x is A_1 ' is called the antecedent or premise, whilst the 'Then part' of the rule, ' y is B_2 ' is called consequent (Zadeh, 1968).

2.3.6. Inference Mechanism

Inference is the process of describing the general nonlinear working flow from the given input space to the output space of the whole fuzzy system. Decisions for each process of such as membership function, fuzzy operator and If-Then rules definition can then be provided.

Two types of widely applied fuzzy inferences are shown as:

- ✓ Mamdani fuzzy inference (Mamdani, 1974)
- ✓ Sugeno fuzzy inference (Sugeno, 1988)

These two fuzzy inferences are similar in many respects: The first two parts of the fuzzy inference process, inputs fuzzification and applying the fuzzy operator, are exactly the same. The main difference between Mamdani (1974) and Sugeno (1968) is that the Sugeno (1968) output membership functions are either linear or constant. In this research, the outputs from each type of fuzzy inference are compared and introduced in Chapter 5.

2.3.7. Fuzzy logic application in Engineering

FL approach has been successfully applied in several fields, where the relationship between 'cause' and 'effect' (variable and results) are vague. Fuzzy

variables are used to organise knowledge that is expressed 'linguistically' into a formal analysis. Based on fuzzy set theory (FST), a FL approach has been applied in many areas where empirical relationships are not well defined or impractical for model development. The foundations of FST, to deal specifically with non-statistical uncertainties, were first developed by Zadeh (1965). Since that time, other researchers had explored the applicability of FL to a variety of problems, including engineering applications (e.g. Siskos, 1982; Seo and Sakawa, 1985) Despite the subjectivity of establishing the descriptive variables, FL model applications had been widely successful in civil engineering, particularly in situations where there were many uncertainties in the relationship between the input variables and the output results.

The development of knowledge based FL approaches has been widely applied to engineering research, and, in respect of the water industry many new applications have been developed. For example, the evaluation model of the risk of water quality failures in a distribution network was developed by Sadiq et al., (2007), whilst an FL methodology was used to develop a rainfall runoff model by Jacquin et al (2008). In this study Mamdani-type Fuzzy Inference Systems (FIS) were applied to the development of rainfall–runoff models operating on a daily basis. The model proposed using a Rainfall Index, obtained from the weighted sum of the recently observed rainfall values, as input information. The model output was designed as the daily discharge amount. Membership function parameters were calibrated using a two-stage constrained optimisation procedure, involving the use of a global and a local search method. Mean squared error and the coefficient of efficiency were used to assess the performance of the fuzzy model. Compared with three other rainfall runoff models that used the same input information as the fuzzy model, overall, the results of this study indicated that Mamdani-type FIS was a suitable alternative for modelling the rainfall runoff relationship.

River seasonal runoff prediction has been published by Mahabir et al., (2003). This research indicated the applicability of FL modelling techniques for

forecasting water supply. By applying FL approach, a water supply forecast model was created that classified potential runoff into three forecast zones: 'low', 'average' and 'high'. Spring runoff forecasts from the fuzzy expert systems were found to be considerably more reliable than the regression models in forecasting the appropriate runoff zone, especially in terms of identifying low or average runoff years. Based on the modelling results in these two basins, it was concluded that FL method has a promising potential for providing reliable water supply forecasts.

Kumar et al., (2010) applied FL approaches using Matlab Simulink tools to develop a methodology of assessing groundwater quality. Eight critical parameters were considered as important model input with respect to drinking water quality criteria, and the conclusion came up with positive outputs by using FL approach in ground water quality assessment.

FL methods have also been applied into infrastructure asset management. Sameh et al., (2007), and Bairaktaris et al., (2007) developed decision support models for the Rehabilitation of Deteriorating Sewers. The models which were completed by Sameh et al 2007 and Bairaktaris et al 2007 were developed based on experienced knowledge of infrastructure assets (sewer system) management. Therefore, these FL models can be only utilised to evaluate assets' serviceability and performance in a certain period, which is dependent on the availability of recorded operation information. However, to develop effective asset operation, maintenance plans and associated maintenance strategies for pro-active decision support by using an FL approaches requires consideration of all up to date asset performance information.

3. Data Collection

In Section 1.1, it was recognised that data was a key component of the proposed research. This section describes the collection of data and its preparation for use in model development. Initially, data on 34 CSOs with pollution problems and 20 normally operation CSO's was selected for the research. The pollution problems were highlighted in Yorkshire Water Service's (YWS) CSO pollution list (Feb. 2006 – Feb. 2007), which included the numbers of pollution incident and the associated pollution category for each CSO asset. Pollution categories were introduced in chapter 1 (page 2) and only incidents of category 1 or 2 were used to identify problem CSO's in the research.

3.1. Introduction

The objectives of this data collection process were to provide all necessary information of the CSO assets and of the operation and maintenance records, including costs. Some initial data analysis has also been completed. Details of collected information that related to CSOs are shown in Figure 3.1:

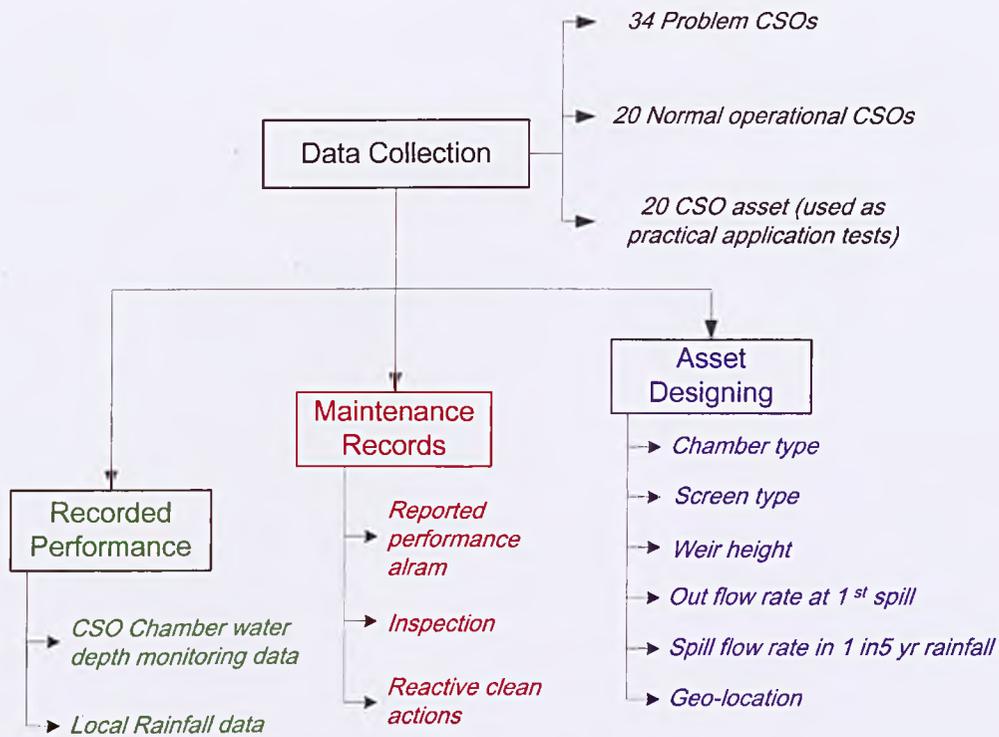


Figure 3.1: Data collection details

All data has been summarised on a DVD data disc. The DVD included the entire process of original data collection and the initial analysed outputs. The interface is shown in the Figure 3.2 below:

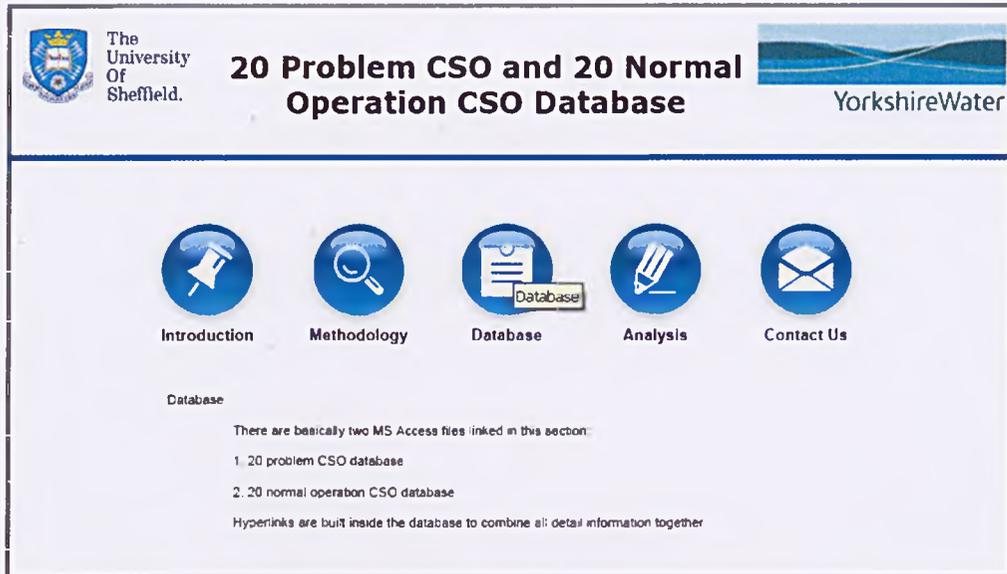


Figure 3.2: CSO database interface

The data collected was as follows:

- ✓ CSO chamber water depth

This was recorded by a telemetered Hawkeye ultrasonic depth monitor installed in each CSO chamber. Water depth values were read every 5 minutes. The CSO chamber water depth was used as one of the input variables for the CSO hydraulic performance prediction model.

- ✓ Rainfall data

Rainfall data was recorded by rain gauge devices. For some CSO chambers more than one rain gauge was located in the vicinity of the CSO. All rainfall values from each rain gauge stations were continuously recorded for each 0.2mm depth of rainfall. The rainfall data was also used as a basic input parameter to the development of the CSO hydraulic performance model.

- ✓ Maintenance records:

These were recorded by company water operators. All collected maintenance information was grouped as 'alarm response', 'field inspection' and 'reactive actions'. Summarized information was used to develop the CSO pro-active operation and maintenance decision support model.

✓ CSO chamber type and screen type

The selected CSO assets were grouped according to different types of chamber and screen.

✓ Hydraulic characteristics

Designed CSO asset hydraulic performance parameters, such as weir height, flow rate as 1st spill etc.

✓ Geo-location information

The Geo-location information indicated the location of the CSO asset and associated rain gauges.

Samples of collected data are introduced in the following section in this chapter.

3.2. Chamber and screen types

The basic information regarding all the CSOs used in this research, including chamber type and screen type, is shown in

Table 3.1:

Table 3.1: Summarized CSO information list

<u>Asset Local Name</u>	<u>Catchment</u>	<u>Overflow Type</u>	<u>Screen</u>	<u>No. of pollution incident</u>
CHANTRY BRIDGE/CSO	Sheffield Catchment	Single Sided Weir	None	8
FOULRIDGE/CSO	Bradford Catchment	Single Sided Weir	Mechanical	6
BUTCHER TERRACE/CSO	York Catchment	Single Sided Weir	Mechanical	5
THE MILL/NO 2 CSO	Chesterfield Catchment	Stilling Pond	Static	4
SHEAF BANK/CSO	Sheffield Catchment	Stilling Pond	Mechanical	3
MAYFIELD GROVE/CSO	York Catchment	Single Sided Weir	Mechanical	3
TERRY AVENUE/CSO	York Catchment	Single Sided Weir	Static	3
TERRY AVENUE/ NO 2 CSO	York Catchment	Single Sided Weir	Static	3
KEARSLEY LANE/CSO	Sheffield Catchment	Stilling Pond	Static	3
CARLETON RD SKIPTON/CSO	Bradford Catchment	VORTEX	Static	3
BROUGH GOLF COURSE/CSO	Hull Catchment	VORTEX	None	2
DEARNE HALL ROAD/CSO	Barnsley Catchment	Double Sided Weir	Mechanical	2
GREEN LANE 125/CSO	Bradford Catchment	Stilling Pond	Static	2
MYTHOLMES LANE/CSO	Bradford Catchment	Single Sided Weir	Mechanical	2
SHARLSTON/CSO	Tadcaster Catchment	Stilling Pond	Static	2
SKELDERGATE BRIDGE/CSO	York Catchment	Single Sided Weir	None	2

WOODBINE COTTAGE/CSO	Barnsley Catchment	Double Sided Weir	Mechanical	2
WORTH WAY SUN STREET/CSO	Bradford Catchment	Single Sided Weir	None	2
WYKE OLD LANE/CSO	Huddersfield Catchment	Double Sided Weir	None	2
BOROUGH BOUNDARY/CSO	Huddersfield Catchment	Double Sided Weir	None	1
CHAPEL LANE/NO 2 CSO	Dewsbury Catchment	Single Sided Weir	Mechanical	1
BEIGHTON TIP/CSO	Chesterfield Catchment	Single Sided Weir	None	1
CANAL ROAD/CSO	Leeds Catchment	Double Sided Weir	Mechanical	1
DELVES ROAD/CSO	Chesterfield Catchment	Single Sided Weir	None	1
HOLLIN DRIVE/CSO	Leeds Catchment	Single Sided Weir	None	1
SKIRLAUGH/CSO	Hull Catchment	Stilling Pond	Mechanical	1

The information about CSO chamber and screen type was collected by checking the asset construction and layout drawings, which were recorded in a YWS database. A typical layout map of is shown in Figure 3.3 (of the Terry Avenue CSO), which also highlights the upstream and downstream pipe network of the CSO.

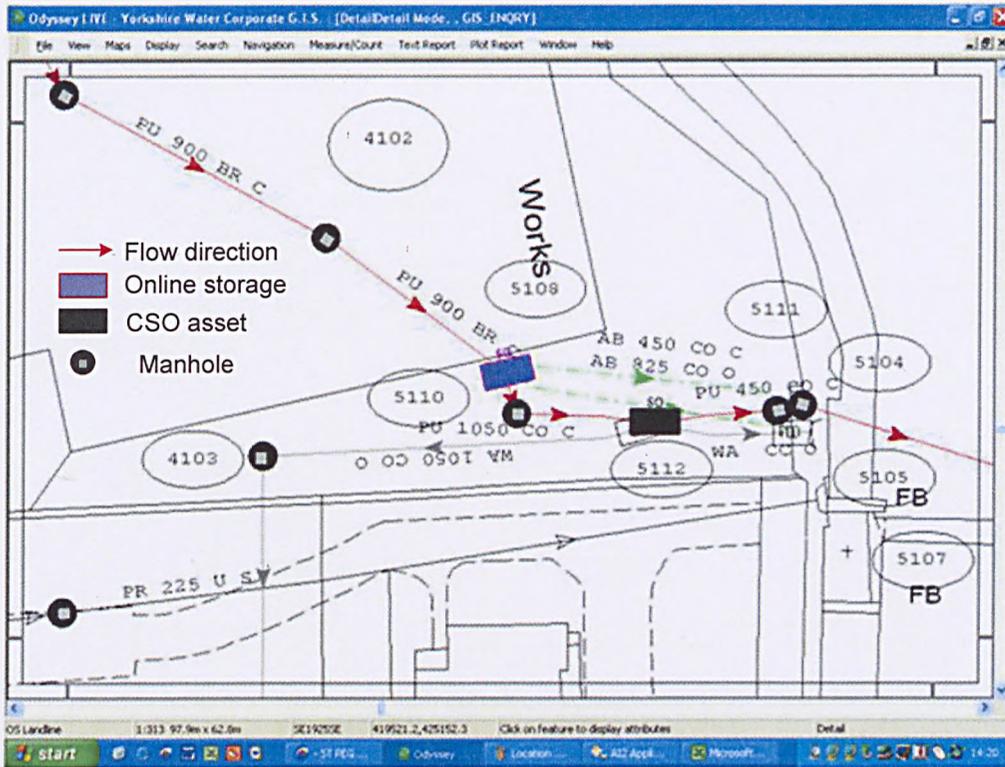


Figure 3.3: Terry Avenue CSO upstream and downstream networks (Source from Yorkshire Water GIS system 'Odyssey'). The red arrow indicates the flow direction from upstream to downstream and the numbers adjacent to the pipelines are the diameters of each pipe; the full black rectangle represents the CSO asset and full black circles represent the manholes.

In Figure 3.3, an online storage tank is shown as a purple rectangle. This storage volume was constructed at the site of an abandoned CSO and is now used to reduce the spill flow at the downstream CSO asset.

An example of CSO chamber's as built construction drawing is shown in Figure 3.4. The Terry Avenue CSO is a single side high weir chamber with a circular inflow pipe 1050 mm in diameter. The continuation pipe is 450 mm in diameter and the overflow pipe downstream of the weir is a 600mm x 1200mm box culvert. This outfall discharges the spill flow to the receiving water. The CSO chamber incorporates a screen to retain the aesthetic solids

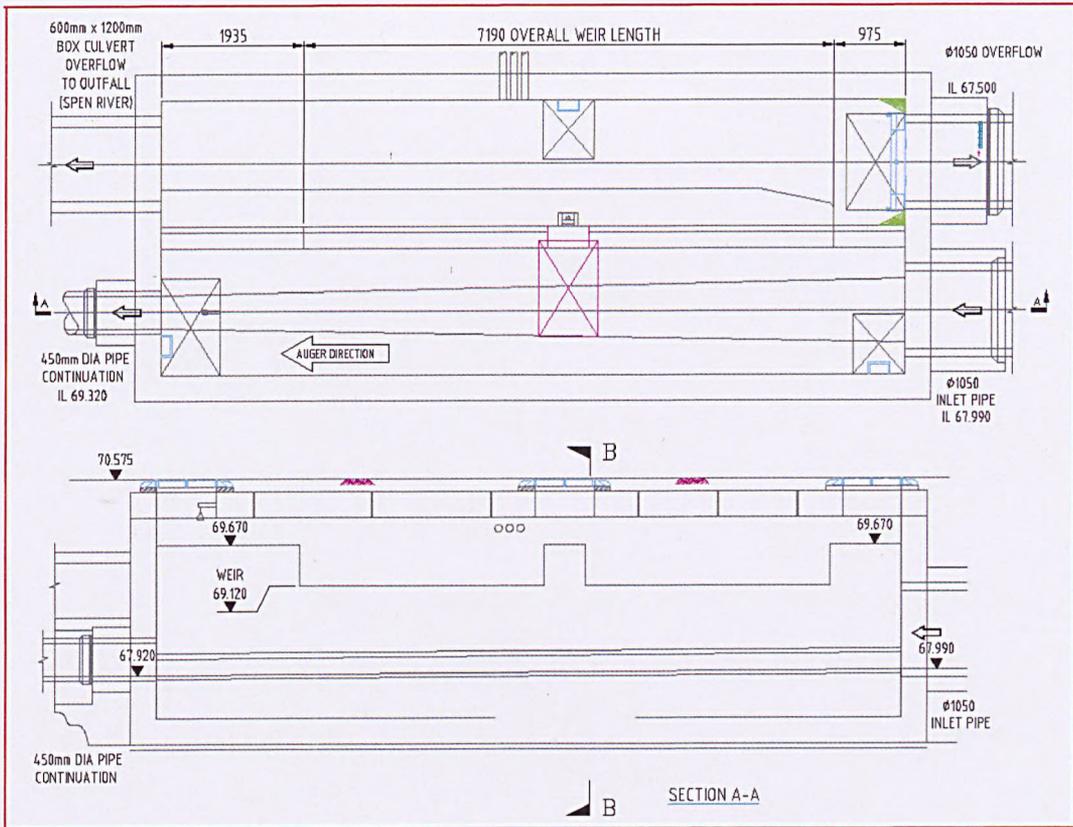


Figure 3.4: Example of as built drawing of Terry Avenue CSO Chamber(Source from Yorkshire Water EDMS, scale 1:25, produced by MWH)

Figure 3.5 shows the general arrangement of screen, which is a 500*6000 RMM Heliscreen produced by Hydro International UK Ltd.

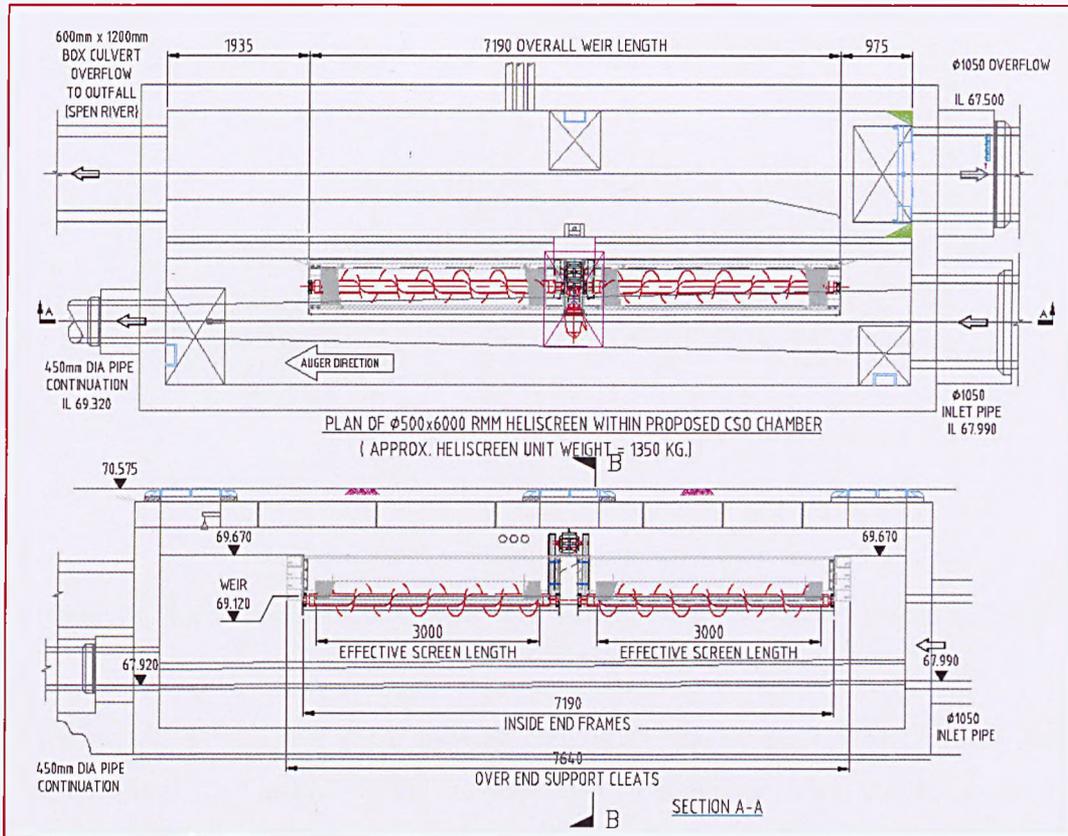


Figure 3.5: Example of Terry Avenue CSO screen general arrangement (Source from Yorkshire Water EDMS, scale 1:50, produced by Hydro International UK Ltd)

All summary details of the CSO chamber, screen information and original construction and layout drawings were recorded on the database disc that is included as an appendix to the thesis.

3.3. CSO chamber water depth

The CSO chamber water depth data was collected from the YWS CSO monitor database. The data was recorded at a standard time interval of 5 minutes.

A typical pattern of flow in dry weather condition was observed and is shown in Figure 3.6:

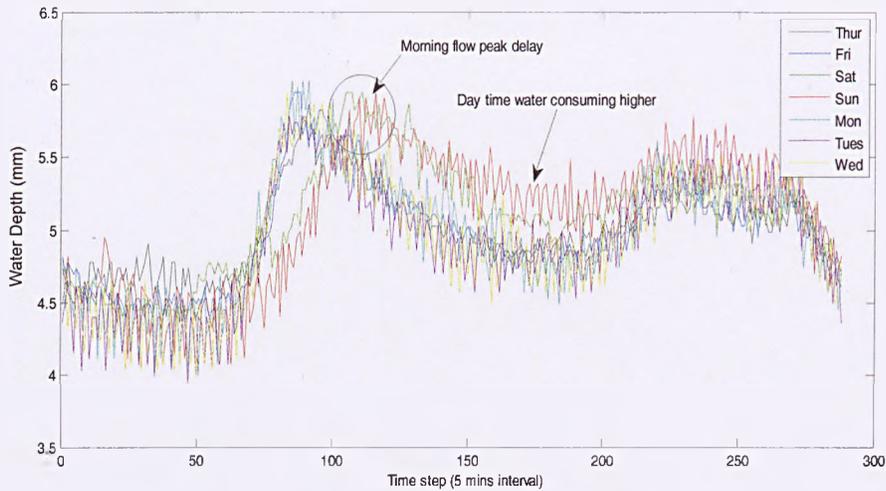


Figure 3.6: CSO chamber dry weather flow. Grids were defined as 288 time steps which represented the time length of 24 hours.

This shows a typical diurnal pattern that is similar to the daily potable water consumption. The pattern has 2 peaks corresponding to early morning and tea time activity with a minimum flow overnight. Green and red lines individually presented the water level performance in CSO chamber on Saturday and Sunday. From Figure 3.6, two significant differences between weekday and weekend patterns are highlighted. Morning peak water level appeared later than normal working days, and the overall day time water level is higher due to more water consumption during the weekend day time.

A typical record of chamber water depth over a four week period is shown in Figure 3.7:

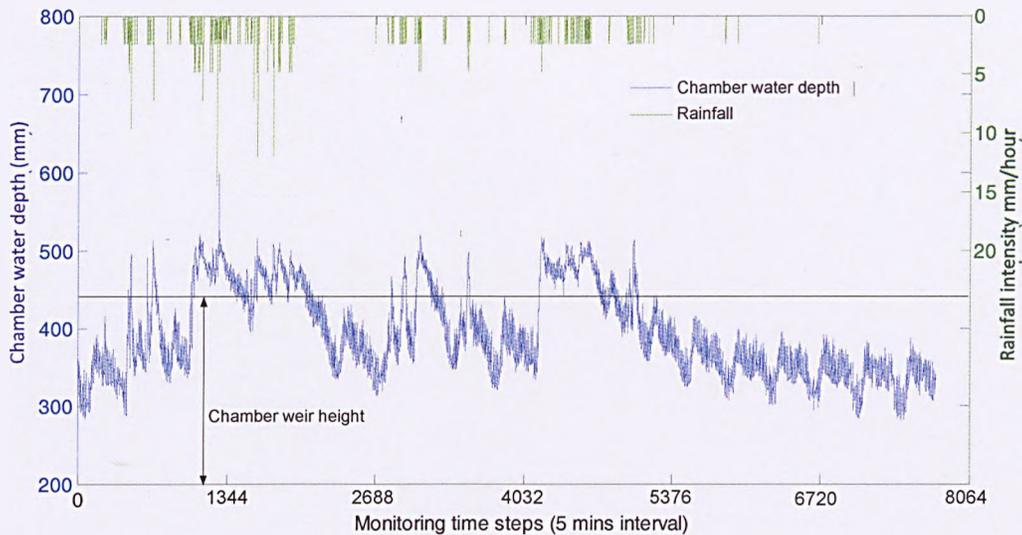


Figure 3.7: An example of CSO chamber water depth performance and rainfall intensity (5 minutes interval).

Chamber water depth values were recorded in 'mm'. In Figure 3.7; the X axis represents the time scale of data with 8064 steps, which indicates the time length of 28 days. The water depth is shown to change in response to rainfall over the catchment. Rainfall data from a rain gauge situated in the catchment is also shown in Figure 3.7.

CSO hydraulic performance under wet weather and dry weather conditions indicates that the system performance was much more complex in wet weather. Routine domestic water is no longer acting as the main volume of sewage as this is now driven by rainfall which results in much higher flow volumes. Figure 3.8 shows a close up of the response of the change in CSO depth to rainfall.

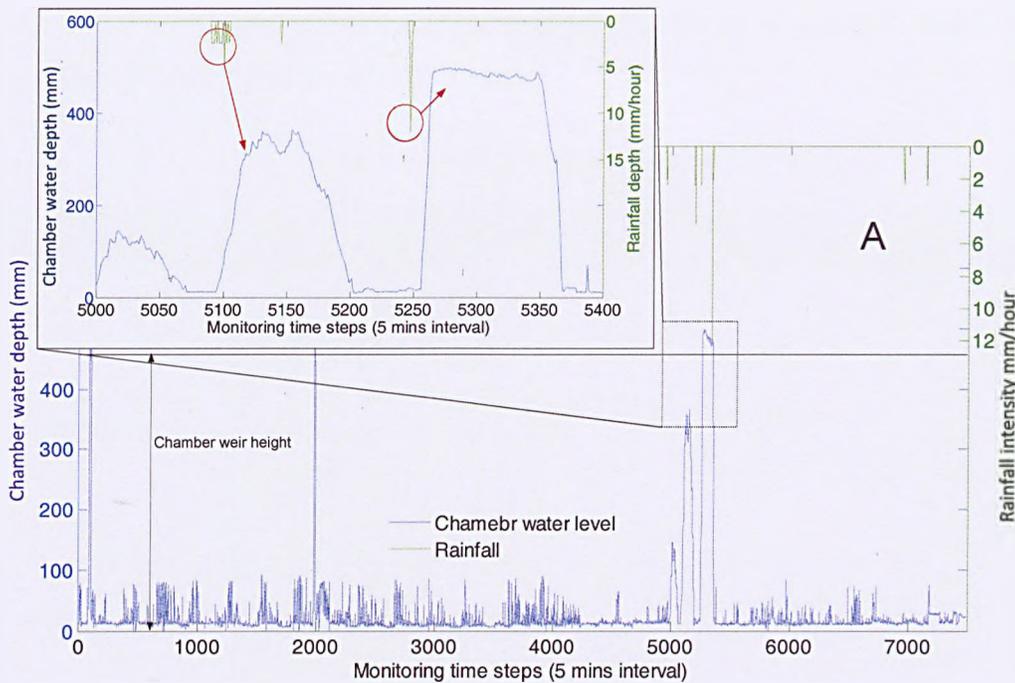


Figure 3.8: CSO chamber water depth performance during rainfall event.

Figure 3.8 highlights that there is a clear relationship between rainfall and change in water depth. This relationship is explored further in Chapter 4 of the thesis.

3.4. Rainfall information selection

As introduced in section 3.1, rainfall intensity data was collected from rain gauge records. The type of rain gauge device used in this research is tipping-bucket rain gauge. The two buckets in a tipping-bucket rain gauge rest on a pivot so that when one bucket has received 0.2 mm of rain it tips by gravity, empties the rainwater and allows the other bucket to start collection. During the tip, an electrical switch is closed and triggers a nearby autographic recorder to register each 'tilt', thus giving a fairly continuous record of precipitation and, in a more sophisticated form, even rainfall intensity.

In this research, all rainfall events recorded in the relevant catchments over the time period of 12 months (March 2006 to March 2007) smaller than the return

level of a 1 in 10 year event were considered. The return period and duration of each rainfall event was shown in Table 3.2.

Table 3.2: Level of each rainfall event

Asset Name	Rainfall Date	Chamber Type	Screen Type	Level of rainfall
CHANTRY BRIDGE/CSO	23-Mar-06	Single Sided High Weir	Mechanical	M5-40min
	09-Apr-06			M5-75min
	22-Apr-06			M2-60min
	12-May-06			M4-100min
	01-Jun-06			M2-40min
	12-Jun-06			M2-30min
	27-Jun-06			M3-100min
	31-Jul-06			M5-60min
THE MILL/NO 2 CSO	16-Mar-06	Single Sided High Weir	Static	M2-30min
	04-Apr-06			M2-75min
	13-Apr-06			M2-60min
	08-May-06			M5-40min
SHEAF BANK/CSO	15-Mar-06	Stilling Pond	Mechanical	M3-100min
	11-May-06			M5-30min
	05-Jul-06			M4-40min
KEARSLEY LANE/CSO	23-Apr-06	Stilling Pond	Static	M2-70min
	03-Jul-06			M2-60min
	03-Nov-06			M2-30min
CARLETON RD SKIPTON/CSO	15-Mar-06	VORTEX	Static	M2-60min
	23-Nov-06			M2-75min
	16-Dec-06			M3-40min
DEARNE HALL ROAD/CSO	16-Mar-06	Double Sided Low Weir	Mechanical	M5-30min
	07-Apr-06			M3-100min
TERRY AVENUE/CSO	05-Oct-06	Single Sided High Weir	Static	M2-60min
	12-Nov-2006			M3-25min
GREEN LANE 125/CSO	19-Apr-06	Stilling Pond	Static	M2-60min
	06-Jun-06			M3-40min
MYTHOLMES LANE/CSO	13-May-06	Single Sided High Weir	Mechanical	M3-60min
	18-May-06			M5-20min
SHARLSTON/CSO	17-Mar-06	Stilling Pond	Static	M2-25min
	23-May-06			M2-75min
WOODBINE COTTAGE/CSO	07-Mar-06	Double Sided Low Weir	Static	M3-40min
	10-Oct-06			M2-60min
WORTH WAY SUN STREET/CSO	21-Aug-06	Single Sided Low Weir	Static	M2-60min
	28-Aug-06			M2-60min
WYKE OLD	19-Jun-06	Double Sided Low	Static	M3-60min

LANE/CSO	14-Aug-06	Weir		M2-30min
CHAPEL LANE/NO 2 CSO	26-Jun-06	Single Sided Low Weir	Mechanical	M5-30min
CANAL ROAD/CSO	03-Oct-06	Double Sided Low Weir	Mechanical	M4-60min
BIRLEY MANSFIELD	19-Oct-06	Single Sided High Weir	Mechanical	M5-40min

The calculation of rainfall return level and duration followed the Rational Method (WaPUG, 1983).

To examine the effect of rainfall on CSO chamber water depth, the rain gauge closest to the CSO was generally used in the analysis. However, in some catchments there were 2 or 3 rain gauges in the vicinity of the CSO chamber. Due to the spatial distribution, rainfall intensities and depths at various locations in a drainage catchment are not equal for the same event, and hence, where rainfall intensity values were available from more than one rain gauge in the same operational catchment it was possible to look the CSO performance in response to the different values of measured rainfall from each of the gauges. For example, in Figure 3.9: there are two rain gauge stations in the same catchment as the Terry Avenue CSO. The location of the CSO is highlighted by the as (+) whilst the rain gauges are presented as (+).

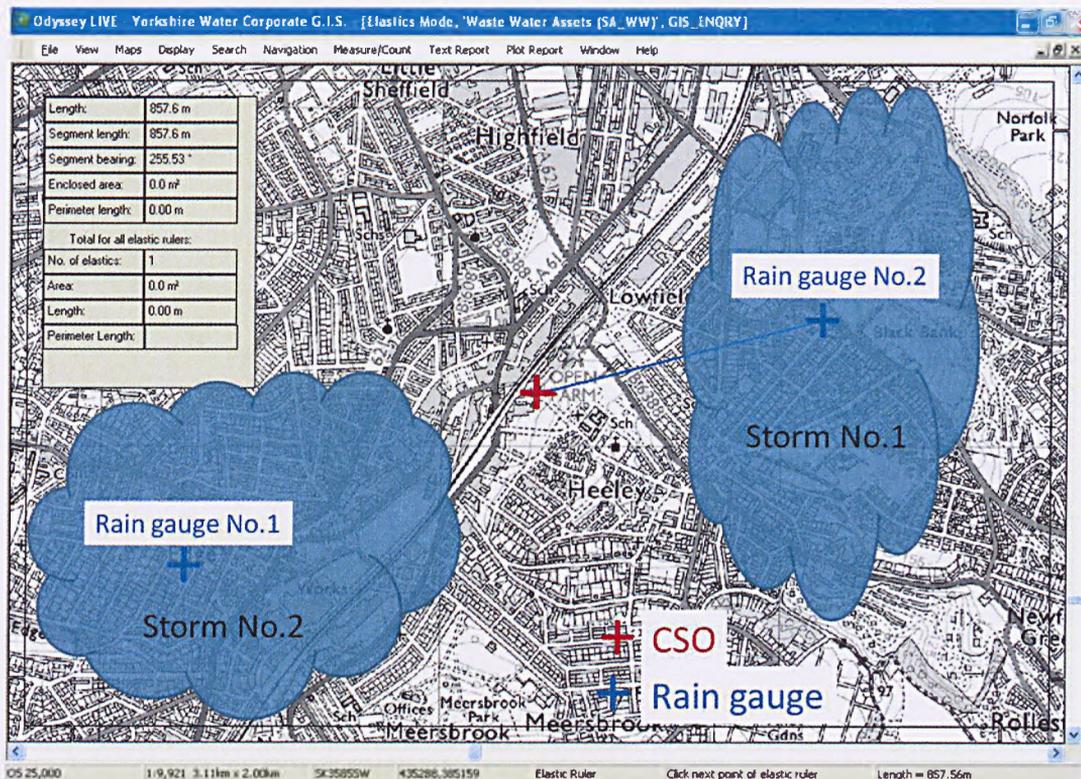


Figure 3.9: Example of Geo-locate relation between Terry Avenue CSO and nearby rain gauge stations (Source from Yorkshire Water GIS system 'Odyssey' with a set of 'rain gauge layout')

To highlight the differences in performance in the utilisation of different rain gauge data reference has been made to two rainfall events. Figure 3.9 displays the pattern of each storm is identified with following correlation analysis shown Figure 3.10. For rainfall event No. 1 (during 11th Dec. to 13th Dec. 2006), the CSO hydraulic performance chart is shown in from Figure 3.10, which is considers with rainfall intensity data from rain gauge No. 1, and 2 during the No.1 storm event:

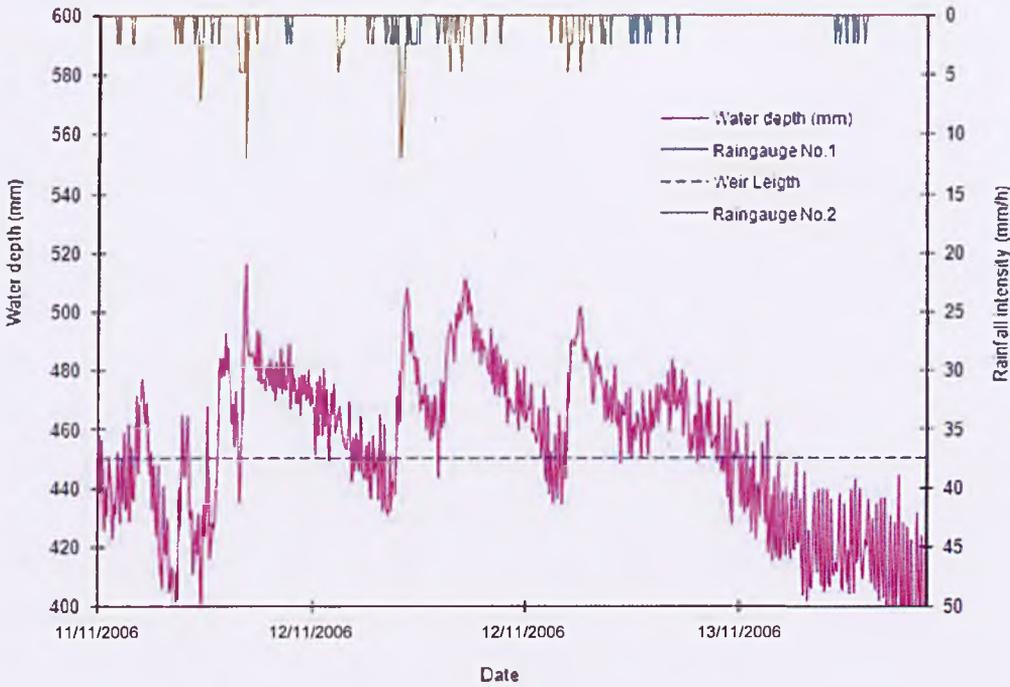


Figure 3.10: Correlation analysis example: two rain gauge value during storm No1

The correlation analysis of rainfall intensity value and CSO chamber water depth performance is following the Equation 3.1:

$$Corr(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Equation 3.1

n: Number of data point for both rainfall intensity and water depth value

X: Water depth variable

x_i : Value of water depth for sample point

\bar{x} : Sample means of X

Y: Rainfall intensity variable

y_i : Value of rainfall intensity for sample point

\bar{y} : Sample means of Y

s_x, s_y : Sample standard deviations of X and Y

During the same storm event, the correlation analysis between rainfall intensity value provided by both rain gauge No. 1 and No 2 and CSO chamber water depth performance were also carried out and are shown:

The analysis outputs of rain gauge No.1: $Corr(X, Y)_{Rain\ gauge\ No.1} = 0.223$.

The analysis outputs of rain gauge No.2: $Corr(X, Y)_{Rain\ gauge\ No.2} = 0.538$.

The correlation coefficient of CSO chamber water depth and rainfall intensity values, shows that the CSO chamber water depth performance gives better correlation to the rainfall that was measured by rain gauge No.2 than rain gauge No.1. This result indicates that, due to the storm pattern (in Figure 3.9); this storm event (No.1) did not pass over the area where rain gauge No.1 was located.

Figure 3.10 provides representative example of the relationship between local rainfall and the resultant change in flow depth. As expected, there is a lag time between the peak of rainfall intensity and peak of chamber water depth. This time lag is discussed more fully in the ADALINE prediction model sensitivity testing, Section 4.5.

For rainfall event No. 2 (during 4thOct. to 6thOct. 2006), the CSO hydraulic performance is shown in

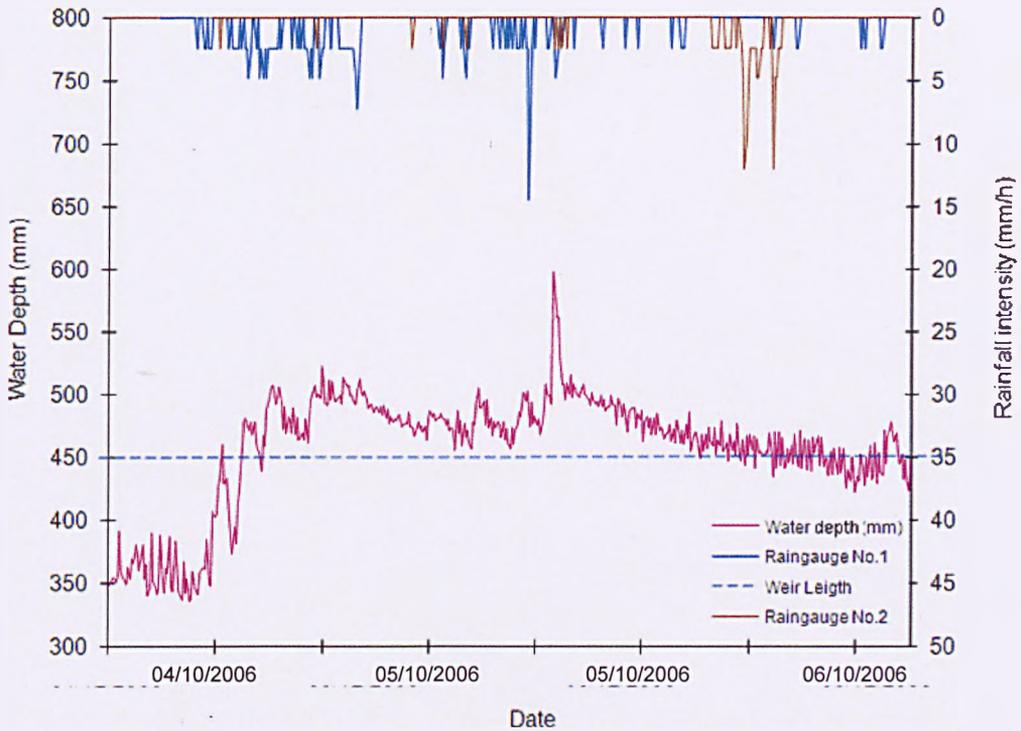


Figure 3.11. In this Figure, rainfall intensity data from both rain gauge No. 1 and No. 2 are also considered:

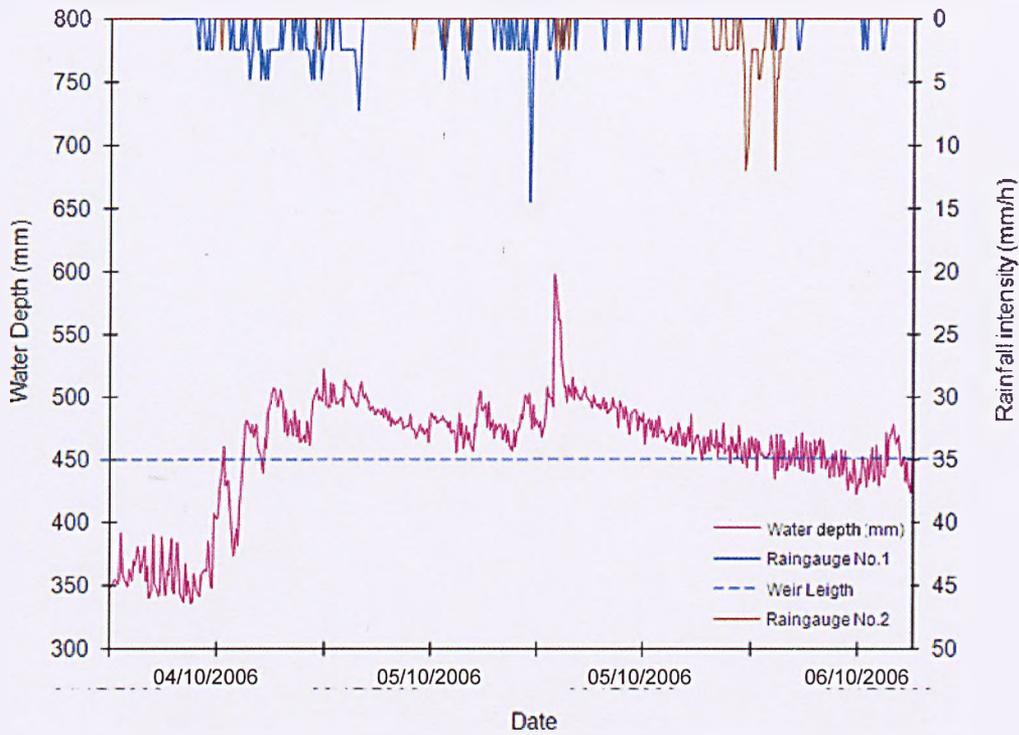


Figure 3.11: Correlation analysis example: two rain gauge values during storm No2

The correlation results for this event are as follows:

$$Corr(X, Y)_{Rain\ gauge\ No.1} = 0.511$$

$$Corr(X, Y)_{Rain\ gauge\ No.2} = 0.243$$

As shown in

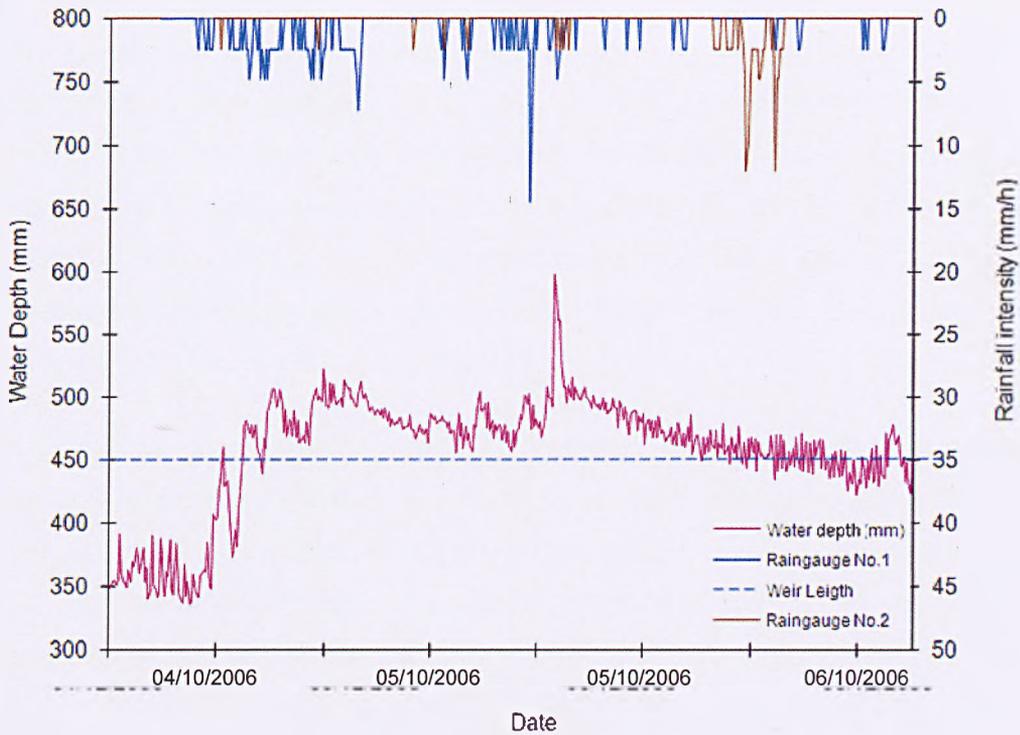


Figure 3.11, the rainfall intensity value, which was recorded by rain gauge No.1, appears to more appropriate to identify the flow depth's performance changing in the CSO chamber. This storm event (No.2) did not pass over the area where rain gauge No.2 was located.

As a conclusion of analysis above, it is clear that the correlation coefficient which relates water depth and rainfall intensity varies as a function of the rain gauge and the spatial and temporal distribution of the rainfall event. As a consequence it was considered necessary for all catchments with multiple rain gauges to establish the correlation coefficient between the pattern of measured depth and the pattern of rainfall measured at each rain gauge. The rain gauge with the highest correlation coefficient was used in the subsequent analysis using the ADALINE prediction model. However, a parallel project that utilises rainfall data measured by radar, ,Shepherd et al. (2010), has shown that radar data provides

useful measurements of rainfall which can be applied to sewer hydraulic models with similar confidence to rain gauge data. The comparison between different rain gauge measurements has highlighted a potential problem with the application of the methodology to take account of the spatial and temporal change in rainfall events over a catchment. The direction of the rainfall has been shown to be particularly important and hence the application of rainfall radar data to detect such spatial and temporal change may be used to improve the modelling capability using rain gauge data. This has been identified as a potential future problem worthy of further research.

As introduced in section 3.1, both CSO chamber water depth performance data and relevant rainfall information were used to develop the CSO ADALINE hydraulic prediction model. All collected CSO chamber water depth performance and original rainfall information had been summarised in the CSO database disc, which is included as Appendix A.

3.5. CSO performance failures

In the period January 2006 to February 2007 a total of 64 Category 1 and 2 pollution incidents were reported for the 34 'problem' CSOs identified in this study. These incidents are summarised in Table 3.3, which also gives an indication of the cause of the problem.

Table 3.3: List of CSO structural information and causes of performance failures

<u>Asset Local Name</u>	<u>Catchment</u>	<u>Date and time to YWS</u>	<u>Overflow Type</u>	<u>Screen</u>	<u>Cause</u>
CHANTRY BRIDGE/CSO	Sheffield Catchment	23/01/2006 13:21	Single Sided High Weir	None	CSO - blocked chamber
CHANTRY BRIDGE/CSO	Sheffield Catchment	09/02/2006 14:39	Single Sided High Weir	None	Sewer - soft blockage
CHANTRY BRIDGE/CSO	Sheffield Catchment	22/02/2006 19:00	Single Sided High Weir	None	CSO - blocked chamber
CHANTRY BRIDGE/CSO	Sheffield Catchment	12/03/2006 12:00	Single Sided High Weir	None	CSO - normal operation
CHANTRY BRIDGE/CSO	Sheffield Catchment	01/05/2006 13:00	Single Sided High Weir	None	Sewer - obstruction
CHANTRY BRIDGE/CSO	Sheffield Catchment	12/05/2006 14:02	Single Sided High Weir	None	Private Problem
CHANTRY BRIDGE/CSO	Sheffield Catchment	27/05/2006 19:00	Single Sided High Weir	None	CSO - blocked control/orifice
CHANTRY BRIDGE/CSO	Sheffield Catchment	31/07/2006 19:30	Single Sided High Weir	None	Sewer - soft blockage
FOULRIDGE/CSO	Bradford Catchment	13/11/2006 00:17	Single Sided High Weir	Mechanical	CSO - normal operation
FOULRIDGE/CSO	Bradford Catchment	16/12/2006 08:05	Single Sided High Weir	Mechanical	CSO - normal operation
FOULRIDGE/CSO	Bradford Catchment	17/12/2006 07:58	Single Sided High Weir	Mechanical	CSO - normal operation
FOULRIDGE/CSO	Bradford Catchment	10/01/2007 08:24	Single Sided High Weir	Mechanical	CSO - normal operation
FOULRIDGE/CSO	Bradford Catchment	13/01/2007 01:15	Single Sided High Weir	Mechanical	CSO - normal operation
FOULRIDGE/CSO	Bradford Catchment	31/01/2007 10:30	Single Sided High Weir	Mechanical	CSO - blocked chamber
BUTCHER TERRACE/CSO	York Catchment	16/01/2006 10:23	Single Sided Low Weir	Mechanical	Not established

BUTCHER TERRACE/CSO	York Catchment	18/01/2006 21:15	Single Sided Low Weir	Mechanical	CSO - normal operation
BUTCHER TERRACE/CSO	York Catchment	25/02/2006 11:00	Single Sided Low Weir	Mechanical	CSO - normal operation
BUTCHER TERRACE/CSO	York Catchment	28/02/2006 10:47	Single Sided Low Weir	Mechanical	Sewer - obstruction
BUTCHER TERRACE/CSO	York Catchment	25/03/2006 17:00	Single Sided Low Weir	Mechanical	Sewer - obstruction
THE MILL/NO 2 CSO	Chesterfield Catchment	16/01/2006 13:20	Stilling Pond	Static	CSO - blocked control/orifice
THE MILL/NO 2 CSO	Chesterfield Catchment	04/02/2006 11:18	Stilling Pond	Static	CSO - blocked hydro-brake
THE MILL/NO 2 CSO	Chesterfield Catchment	13/03/2006 14:45	Stilling Pond	Static	CSO - normal operation
THE MILL/NO 2 CSO	Chesterfield Catchment	08/05/2006 14:30	Stilling Pond	Static	CSO - blocked control/orifice
SHEAF BANK/CSO	Sheffield Catchment	15/03/2006 09:00	Stilling Pond	Mechanical	CSO - blocked control/orifice
SHEAF BANK/CSO	Sheffield Catchment	11/05/2006 12:00	Stilling Pond	Mechanical	CSO - blocked chamber
SHEAF BANK/CSO	Sheffield Catchment	05/07/2006 11:37	Stilling Pond	Mechanical	CSO - blocked control/orifice
MAYFIELD GROVE/CSO	York Catchment	15/04/2006 11:45	Single Sided Low Weir	Mechanical	Sewer - soft blockage
MAYFIELD GROVE/CSO	York Catchment	19/04/2006 09:18	Single Sided Low Weir	Mechanical	Sewer - fat/grease blockage
MAYFIELD GROVE/CSO	York Catchment	16/02/2007 16:00	Single Sided Low Weir	Mechanical	Sewer - fat/grease blockage
TERRY AVENUE/CSO	York Catchment	18/01/2006 18:00	Single Sided Low Weir	Static	CSO - normal operation
TERRY AVENUE/CSO	York Catchment	05/03/2006 12:00	Single Sided Low Weir	Static	SPS - other equipment failure
TERRY AVENUE/CSO	York Catchment	12/11/2006 14:00	Single Sided Low Weir	Static	SPS - other equipment failure
KEARSLEY LANE/CSO	Sheffield Catchment	23/04/2006 15:15	Stilling Pond	Static	

KEARSLEY LANE/CSO	Sheffield Catchment	03/07/2006 11:06	Stilling Pond	Static	
KEARSLEY LANE/CSO	Sheffield Catchment	03/12/2006 11:00	Stilling Pond	Static	CSO - blocked hydro-brake
CARLETON RD SKIPTON/CSO	Bradford Catchment	15/03/2006 04:15	VORTEX	Static	CSO - normal operation
CARLETON RD SKIPTON/CSO	Bradford Catchment	23/11/2006 13:08	VORTEX	Static	CSO - normal operation
CARLETON RD SKIPTON/CSO	Bradford Catchment	16/12/2006 08:45	VORTEX	Static	CSO - normal operation
BROUGH GOLF COURSE/CSO	Hull Catchment	08/02/2006 12:15	VORTEX	None	CSO - blocked chamber
BROUGH GOLF COURSE/CSO	Hull Catchment	22/02/2006 12:50	VORTEX	None	CSO - blocked chamber
DEARNE HALL ROAD/CSO	Barnsley Catchment	16/01/2006 18:23	Double Sided Low Weir	Mechanical	CSO - blocked control/orifice
DEARNE HALL ROAD/CSO	Barnsley Catchment	07/02/2006 21:00	Double Sided Low Weir	Mechanical	CSO - blocked control/orifice
GREEN LANE 125/CSO	Bradford Catchment	06/02/2006 16:30	Stilling Pond	Static	CSO - blocked chamber
GREEN LANE 125/CSO	Bradford Catchment	06/06/2006 14:40	Stilling Pond	Static	CSO - blocked chamber
MYTHOLMES LANE/CSO	Bradford Catchment	13/05/2006 12:55	Single Sided High Weir	Mechanical	CSO - blocked control/orifice
MYTHOLMES LANE/CSO	Bradford Catchment	18/05/2006 18:21	Single Sided High Weir	Mechanical	
SHARLSTON/CSO	Tadcaster Catchment	17/03/2006 12:00	Stilling Pond	Static	CSO - normal operation
SHARLSTON/CSO	Tadcaster Catchment	23/05/2006 16:35	Stilling Pond	Static	
SKELDERGATE BRIDGE/CSO	York Catchment	18/01/2006 18:00	Single Sided Low Weir	None	CSO - normal operation
SKELDERGATE BRIDGE/CSO	York Catchment	11/03/2006 12:00	Single Sided Low Weir	None	SPS - pump failure
WOODBINE COTTAGE/CSO	Barnsley Catchment	07/03/2006 23:55	Double Sided Low Weir	Mechanical	

WOODBINE COTTAGE/CSO	Barnsley Catchment	10/10/2006 13:30	Double Sided Low Weir	Mechanical	CSO - blocked control/orifice
WORTH WAY SUN STREET/CSO	Bradford Catchment	21/08/2006 07:39	Single Sided Low Weir	None	CSO - blocked chamber
WORTH WAY SUN STREET/CSO	Bradford Catchment	28/08/2006 04:10	Single Sided Low Weir	None	CSO - normal operation
WYKE OLD LANE/CSO	Huddersfield Catchment	19/06/2006 17:24	Double Sided Low Weir	None	CSO - 3rd party interference
WYKE OLD LANE/CSO	Huddersfield Catchment	14/08/2006 16:30	Double Sided Low Weir	None	CSO - blocked chamber
BOROUGH BOUNDARY/CSO	Huddersfield Catchment	24/07/2006 14:30	Double Sided Low Weir	None	
CHAPEL LANE/NO 2 CSO	Dewsbury Catchment	26/06/2006 12:30	Single Sided Low Weir	Mechanical	CSO - blocked chamber
BEIGHTON TIP/CSO	Chesterfield Catchment	11/06/2006 20:30	Single Sided Low Weir	None	CSO - blocked control/orifice
CANAL ROAD/CSO	Leeds Catchment	03/10/2006 08:30	Double Sided Low Weir	Mechanical	CSO - blocked control/orifice
DELVES ROAD/CSO	Chesterfield Catchment	26/07/2006 12:00	Single Sided Low Weir	None	CSO - blocked control/orifice
HOLLIN DRIVE/CSO	Leeds Catchment	21/06/2006 14:45	Single Sided High Weir	None	CSO - blocked control/orifice
SKIRLAUGH/CSO	Hull Catchment	16/04/2006 18:30	Stilling Pond	Mechanical	CSO - blocked chamber

From this record the number of incidents for different CSO and screen arrangements can be identified and has been summarised in Table 3.4

Table 3.4: Pollution incidents by CSO chamber and screen type

Chamber type	Numbers of CSO	Numbers of Pollution Incident	Pollution Incident frequency /year
Side Weir	24	46	1.9
Stilling Pond	5	12	2.4
VORTEX	3	6	2

Screen type	Numbers of CSO	Numbers of Pollution Incident	Pollution Incident frequency /year
Mechanical	12	17	1.4
Static	5	15	3
None	13	30	2.3

The side weir chamber type of CSO is the most widely used in YWS but it is clear that all types of chamber have associated pollution incidents, with the average for the problem CSO's used in this study of approximately 2 per annum per CSO. However, considering with the pollution frequency, those CSO assets with Stilling Pond chamber appears to have the highest possibility of a pollution incident. Similarly there are pollution incidents associated with both mechanical and static screens and in chambers that do not incorporate any screen arrangement at all. Clearly however there are far more incidents (on average 3 per CSO per year) for static screens when compared to mechanical screens (1.4 per CSO per annum).

All incident data has been summarised on the database disc, Appendix A.

3.6. Operation and Maintenance records

All maintenance and related information was obtained from the YWS database that holds such information. The recorded depth data for each CSO is used by YWS to trigger alarms and to identify operational and maintenance actions. A typical data set is shown as Figure 3.12.

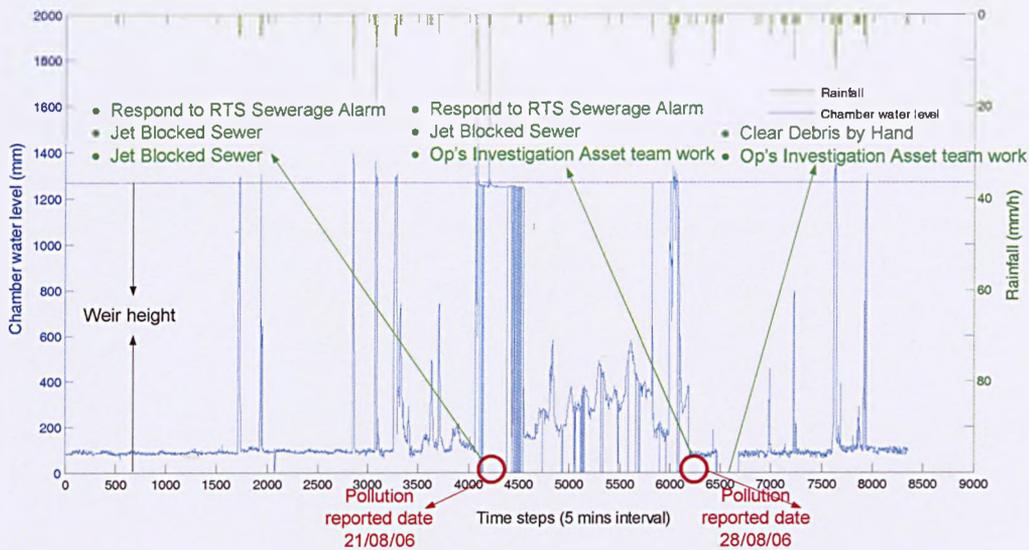
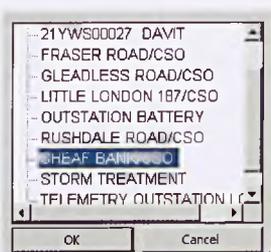


Figure 3.12: CSO performance with maintenance actions

All maintenance actions are located on the time axis at the point where the pollution incident is highlighted. Comparing the implementation date of the maintenance actions with the date of the pollution incidents gives a clear view of the sewerage service provider's response to CSO pollution and the type of maintenance actions carried out with regard to the incidents.

O&M records were obtained for each CSO and a typical record is shown in Figure 3.13:



Description	Actual release	Actual Finish	Actual start	TotSum (actual)	Service product
SHEAF BANK/CSO	31/12/2005		25/01/2006	23.93	Site Inspection CSO (Mechanical) Type 2
SHEAF BANK/CSO	01/04/2006		27/04/2006	21.30	Site Inspection CSO (Mechanical) Type 2
SHEAF BANK/CSO	15/03/2006	15/03/2006	15/03/2006	21.19	Sewage Treatment Inspection
SHEAF BANK/CSO	15/03/2006			0.00	Respond to RTS Sewerage Alarm
SHEAF BANK/CSO	16/03/2006	17/03/2006	17/03/2006	796.01	EMM4 - Mechanical Repair £501 to £1000
SHEAF BANK/CSO	16/03/2006	16/03/2006	16/03/2006	22.79	9248 Assist on Site for Contractor (OPS)
SHEAF BANK/CSO	17/03/2006	17/03/2006	17/03/2006	46.72	9221 Assist on Site for Contractor (CS)
SHEAF BANK/CSO	08/05/2006	08/05/2006	08/05/2006	22.70	Ops Investigation (Field Staff)
SHEAF BANK/CSO	11/05/2006	11/05/2006	11/05/2006	9.13	Sewage Treatment Insp (Pollution)
SHEAF BANK/CSO	01/07/2006		12/07/2006	25.51	Site Inspection CSO (Mechanical) Type 2
SHEAF BANK/CSO	08/09/2006	17/10/2006	11/10/2006	42.59	Op's Investigation Asset Team
SHEAF BANK/CSO	30/09/2006		25/10/2006	12.72	Site Inspection CSO (Mechanical) Type 2
SHEAF BANK/CSO	30/12/2006		12/01/2007	10.06	Site Inspection CSO (Mechanical) Type 2
SHEAF BANK/CSO	09/01/2007	09/01/2007	09/01/2007	27.38	Respond to RTS Sewerage Alarm
SHEAF BANK/CSO	11/01/2007	11/01/2007	11/01/2007	0.94	Respond to RTS Sewerage Alarm
SHEAF BANK/CSO	18/01/2007	18/01/2007	18/01/2007	4.45	Respond to RTS Sewerage Alarm
SHEAF BANK/CSO	10/02/2007	10/02/2007	10/02/2007	32.99	Respond to RTS Sewerage Alarm
SHEAF BANK/CSO	04/03/2007			0.00	Respond to RTS Sewerage Alarm
SHEAF BANK/CSO	31/03/2007		24/04/2007	70.02	Site Inspection CSO (Mechanical) Type 2
SHEAF BANK/CSO	30/06/2007			0.00	Site Inspection CSO (Mechanical) Type 2
SHEAF BANK/CSO	06/07/2007			85.00	Jet Blocked Sewer

Figure 3.13: Maintenance records document

The database information fields that are concerned with maintenance include: Actual Release date (notification of problem), Actual Finish date, Actual Start date, Actual Cost and Service product (actual maintenance action). The data base also contains rules for allocating business costs shared between water and sewerage services and for the General and Support cost subcategories associated with each maintenance action.

This original data was also included in the CSO database disc and was subsequently used in the development of the FL pro-active O&M decision support model.

Summary, all the data presented in this chapter has been stored and archived in such a way that it may easily be retrieved to aid the development of the pro-active O&M decision support model.

4. CSO hydraulic performance prediction model

As detailed in the objectives in section 1.1, the CSO hydraulic prediction model was developed by applying the ADALINE approach. Both CSO hydraulic performance data and rainfall information were used as the model inputs. The prediction process started with the training of the network using the model inputs, i.e. the recorded CSO chamber water depth and rainfall intensity data. The model learned the relationship between CSO water depth and rainfall intensity, and subsequently the model was able to predict the CSO hydraulic performance from rainfall. For example, the prediction model can produce CSO chamber water depth performance which responded to a given set of rainfall information as prediction outputs. All rainfall intensity data, which was collected and used for model development and verification, was evaluated as smaller than the return level of a 1 in 10 year event. Hence the application of the model is limited to those storms that have a return period of less than 10 years.

4.1. ADALINE Methodology

The ADALINE networks were normalised similar to the perception neural network that was introduced in section 2.2.1, but their transfer function was linear rather than hard-limiting. This allowed the ADALINE model's outputs to take on any value, whereas the perceptron output is limited to either 0 or 1. Both the ADALINE and the perceptron could only solve linearly separable problems. However, here the LMS (least mean squares) learning rule, which is much more powerful than the perceptron learning rule, is used. The LMS or Widrow-Hoff learning rule minimises the mean square error and thus moves the decision boundaries as far as it can from the training patterns (Widrow and Sterns, 1985).

This research is to design an adaptive linear system that responds to changes in its environment as it operates. Here, 'the environment is the input data. Linear

networks that are adjusted at each time step based on new input and target vectors. Weights and biases that minimise the network's sum-squared error for recent input and target vectors are then found.

Details of the ADALINE algorithm were introduced in section 2.2.2. The initial methodology of how the ADALINE was applied to develop CSO hydraulic performance model is now explained.

4.1.1. Linear relationship function

The basic linear function of the ADALINE approach can be represented as Equation 4.1:

$$y_p = wx = \sum_i w_i x_i + \varepsilon$$

Equation 4.1

In respect of CSO hydraulic performance, each variable is defined as:

x: Model input (recorded chamber water depth and rainfall intensity value)

w: Weight for each model input

ε : Bias value for each calculation step

i: Index (time step of each input)

y_p : Predicted model output (the predicted chamber water depth) (mm)

Equation 4.1 established the initial linear relationship between model input and output. In the ADALINE CSO model, the desired model output Y is the actual chamber water depth at time step $i + 1$. The model input includes both the previous water depth value and the current rainfall intensity value, and hence Equation 4.1 can be written as:

$$X = [y, u]$$

$$y_p = y_{i+1} = WX = \sum_i w_{y,i}y_i + w_{u,i}u_{i+1} + \varepsilon$$

Equation 4.2

Where,

X : Model input vector (chamber water depth and rainfall intensity value)

W : Weight vector

y : Chamber water depth value (mm)

u : Rainfall intensity value (mm/h)

w_y : Weight for each chamber water depth input

w_u : Weight for each rainfall intensity input

ε : Bias value for each calculation step

i : Index (time step of each input)

y_p : Predicted model output (the predicted chamber water depth) (mm)

An example which explains the linear relationship between model input and output is shown in Figure 4.1. It can be seen that the relationship between the different depth values is assumed linear.

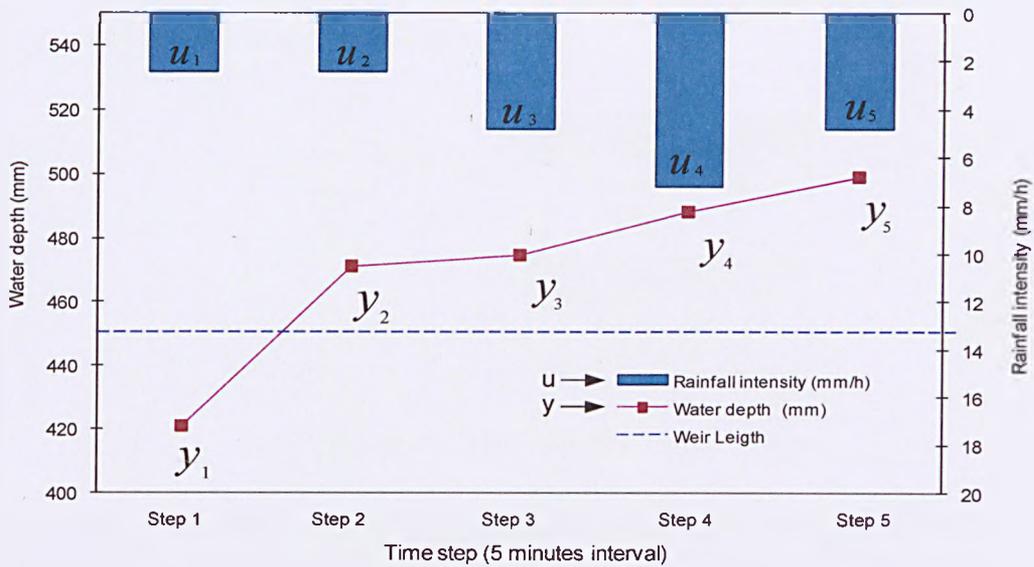


Figure 4.1: Example chart of model linear relationship explanation

Based on the definition of Equation 4.2, the linear function of this example can be written as Equation 4.3:

$$\begin{cases} y_5 = w_{y,4}y_4 + w_{u,4}u_5 + \varepsilon_4 \\ y_4 = w_{y,3}y_3 + w_{u,3}u_4 + \varepsilon_3 \\ y_3 = w_{y,2}y_2 + w_{u,2}u_3 + \varepsilon_2 \\ y_2 = w_{y,1}y_1 + w_{u,1}u_2 + \varepsilon_1 \end{cases}$$

Equation 4.3

4.1.2. Model Input

Both CSO chamber water depth and rainfall intensity values were formatted for use in the model:

- ✓ CSO hydraulic data and the rainfall data in the same period was filled into the same dataset file (txt or xls format which can be recognised in MATLAB)
- ✓ The prepared input dataset was programmed into MATLAB as a file of Mat format

- ✓ As shown in Figure 4.2, CSO chamber water depth was recognized as 'y' and rainfall data recognised as 'u'

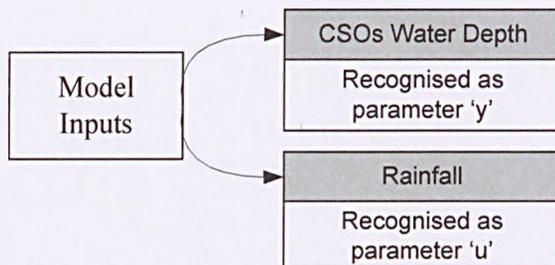


Figure 4.2: Model inputs selection

As detailed in Section 3.4, the distance between the rain gauge station and CSO asset had an impact on the weights that are used in the model. This is due to the time difference between the recorded rainfall and the delay in the response time for the runoff to arrive at the CSO chamber. For example, the longer the distance of the rain gauge from the CSO, the longer the expected delay between the rainfall and the runoff. In addition, the weights used in the model are also a function of many other rainfall factors that include the spatial and temporal distribution of the rainfall over the catchment, including the speed the direction of travel. These factors may also contribute to the delay in runoff response and an example of the delay that based on Terry Avenue CSO is shown in Figure 4.3.

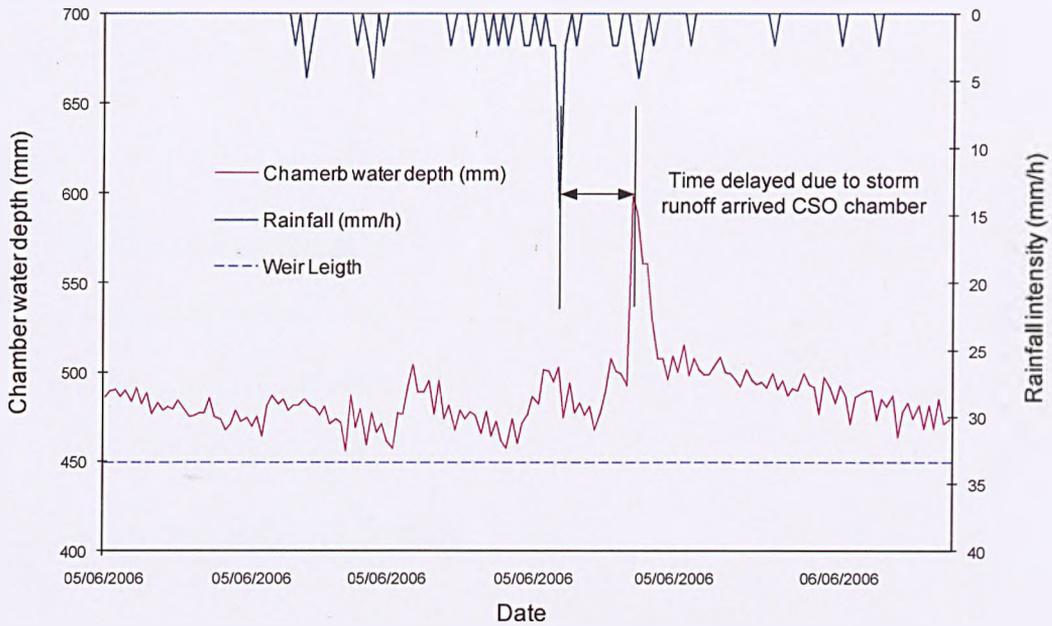


Figure 4.3: Delay response of chamber water depth to rainfall ()

As can be seen from Figure 4.3, the rainfall peak was recorded some time steps earlier than the CSO chamber water depth peak. The delay corresponding to a second storm for the same catchment and CSO was shown in Figure 3.9. This storm has a significantly different time delay and again is a function of the distance between the rain gauge station and CSO asset and the speed and direction of the storm's movement.

The cross and serial correlation between the overflow and rainfall data were investigated. This method, previously successfully used for similar studies (Fernando, 2005) provides useful information to determine the size of the model input in order to capture the underlying function efficiently. Cross and serial correlation functions were represented as Equation 4.4 and Equation 4.5:

Cross correlation analysis (Bracewell, 1965):

$$Xcorr[y, u] = Cov[\bar{y}(-t), u(t)]$$

Equation 4.4

Where,

y : Chamber water depth

u : Rainfall intensity value

Cov : Convolution function

\bar{y} : Complex conjugate of y

Serial correlation analysis (Zwillinger 1995):

$$Autocorr[y, y] = Cov[\bar{y}(-t), y(t)]$$

Equation 4.5

Where,

y : Chamber water depth

Cov : Convolution function

\bar{y} : Complex conjugate of y

The cross-correlation between the overflow rates and rainfall data and serial correlation amongst the overflow rates were determined for each CSO's water depth and rainfall event. Two examples of the analysis output based on the same CSO (Terry Avenue CSO) are shown in Figure 4.4 and Figure 4.5.

From Figure 4.4, it can be concluded that the cross correlation values of rain gauge No 2 increase with increasing lag time, peak around a time lag of approximately 9 units and then decreases with increasing lag time. In general, a high correlation can be observed between approximate lag time units of 5 and 16. Thus, the appropriate rainfall input to forecast $y(t)$ for this example were $u(t-$

9), $u(t-10)$, $u(t-11)$, $u(t-12)$, $u(t-13)$, $u(t-14)$, $u(t-15)$, $u(t-16)$ and so on. The low cross correlation value of rain gauge No.1 was explained in section 3.4.

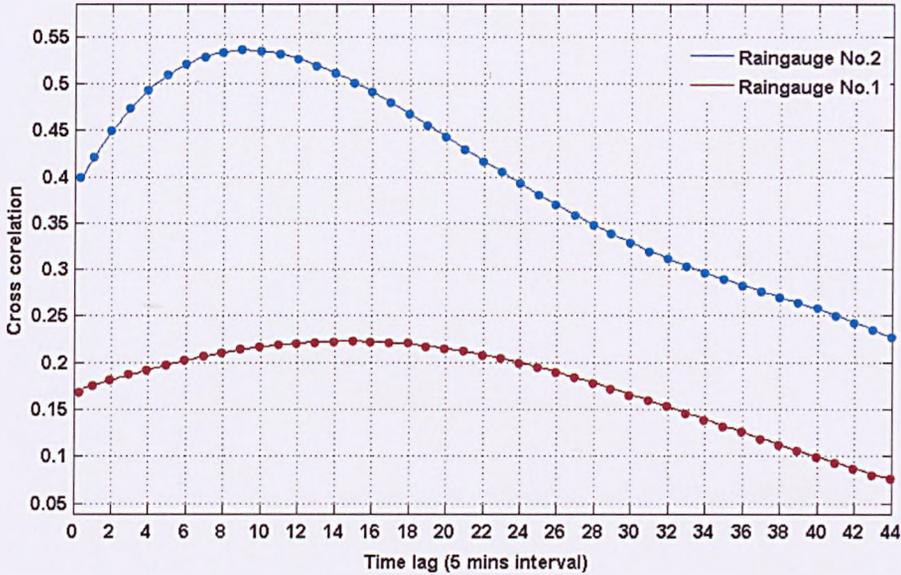


Figure 4.4: Cross correlation analysis of rainfall event example No. 1

From Figure 4.5, the cross correlation values increase with increasing lag time, peak around a time lag of approximately 2 units and then decreases with increasing lag time. Compared with Figure 4.4, the response of chamber water depth performance to the analysis of example rainfall No.2 is faster than No1. In general, a high correlation can be observed between approximate lag time units of 1 and 8. Therefore, for this example, the appropriate rainfall input to forecast $y(t)$ were $u(t-2)$, $u(t-3)$, $u(t-4)$, $u(t-5)$, $u(t-6)$, $u(t-7)$, $u(t-8)$ and so on. Again, the low cross correlation value of rain gauge No.2 was explained in section 3.4.

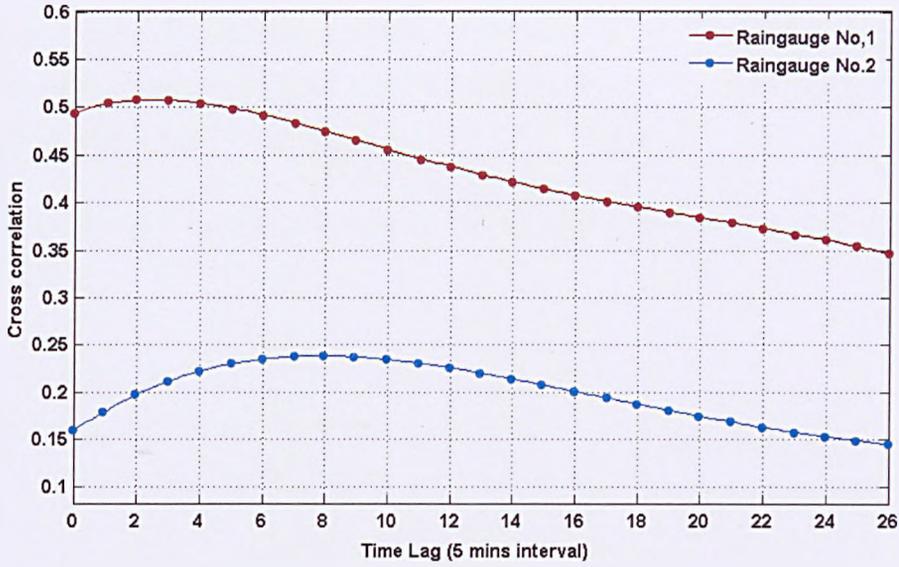


Figure 4.5: Cross correlation analysis of rainfall event example No. 2

Similarly, the serial correlation values for four examples of CSO chamber water depth performance were plotted. The serial correlation is shown in Figure 4.6.

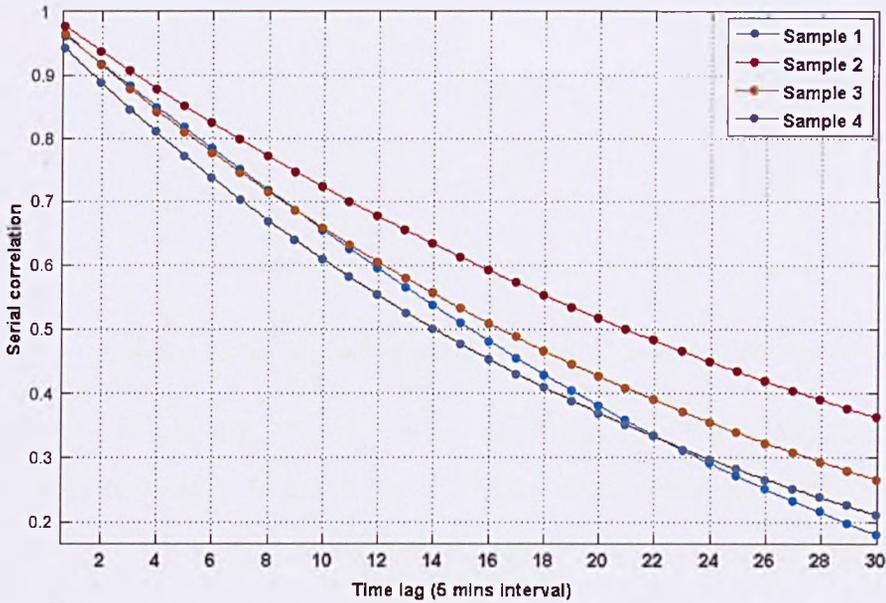


Figure 4.6: water depth serial correlation analysis

For serial-correlation in Figure 4.6, the correlation values decreases gradually, as expected, with increasing lag time. Therefore, it can be concluded that the appropriate water depth input to forecast $y(t)$ were $y(t-1)$, $y(t-2)$, $y(t-3)$, $y(t-4)$, $y(t-5)$, $y(t-6)$, $y(t-7)$, $y(t-8)$ and so on.

The model input selection approach for each training step is demonstrated in Figure 4.7:

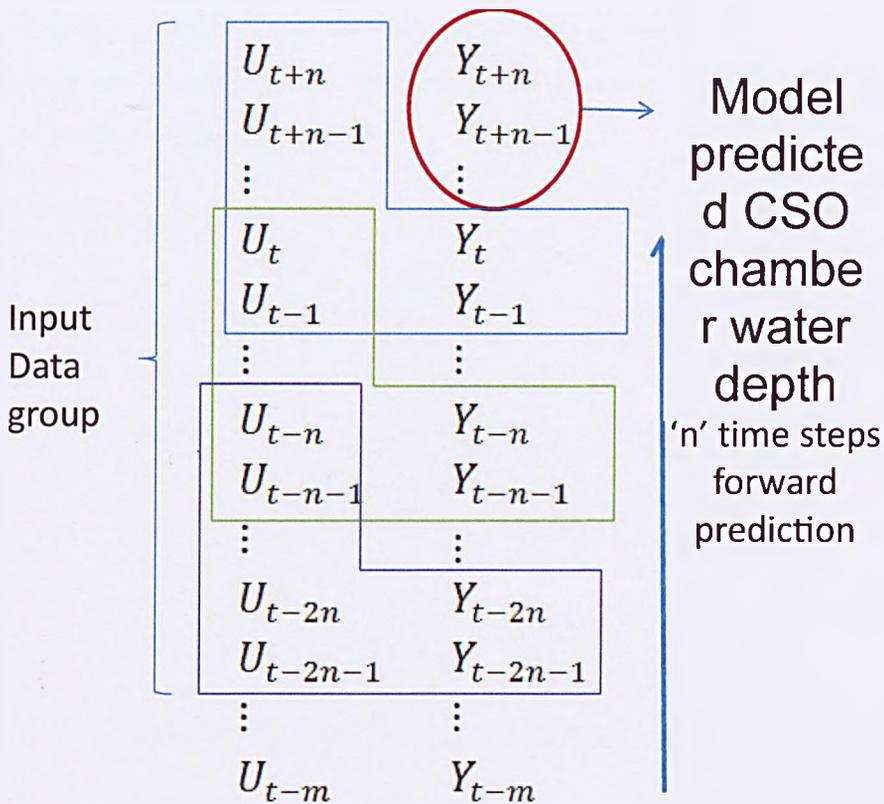


Figure 4.7: Model input selection and model predicted duration

As shown in Figure 4.7, chamber water depth value of $t - m$ step to t step and rainfall intensity value of $t - m$ to $t + n$ steps were used as model input. The chamber water depth performance of $t + 1$ to $t + n$ step were predicted as model output.

The model prediction range is affected by the difference between two groups of input. Theoretically, the longer the range of the advanced rainfall available, the longer the range that the water depth can be predicted into the future. However, the prediction accuracy is discussed in model sensitivity testing section, see Section 4.5.

This section of the research has highlighted a need for further research to fully address this issue and one way forward, as detailed in the section of further work, may be to incorporate rainfall radar data to assist in the prevalent rain gauge selection process as such data may be used to monitor the direction and speed of each individual storm event.

4.1.3. Learning rule

The CSO ADALINE prediction model includes a learning rule that is used to establish the relationship between the input and output variables. This is presented as Equation 4.6:

$$w + \eta(Y - y_p)X \rightarrow w$$

Equation 4.6

Where

X: Model input vector (chamber water depth and rainfall intensity value)

w: Weight vector

y: Chamber water depth value (mm)

y_p : Model predicted chamber water depth (mm)

η : Bias value for each calculation step

Y: Desired model output (actually chamber water depth for the next time step) (mm)

The learning rate of the neural system is defined as η , and Y is the desired output – real value, the identity function $o = y$ is the activation function and the squared error $E = (Y - o)^2$ is the error function. Therefore, the decision boundary is the assumed condition: $E = 0$.

The ADALINE learning rule is justified by determining the gradient descent of the change in the mean square value:

$$\begin{aligned}\frac{\partial E}{\partial w} &= \frac{\partial (y_p - Y)^2}{\partial w} = 2(y_p - Y) \frac{\partial}{\partial w} \\ &= 2(y_p - Y) \frac{\partial y}{\partial w} = 2(y_p - Y) \frac{\partial w \cdot x}{\partial w} \\ &= 2(y_p - Y)x\end{aligned}$$

Equation 4.7

From the equations above, following the calculation of squared error function $(y_p - Y)x$ increases error, conversely, the process $(Y - y_p)x$ is decreasing error.

The process of applying the least mean square learning rule in the ADALINE prediction model is a looped calculation, which is shown in Figure 4.8. During this process, the model is learning and obtaining the most appropriate weight value, which can be used to train the system and provide an efficient prediction.

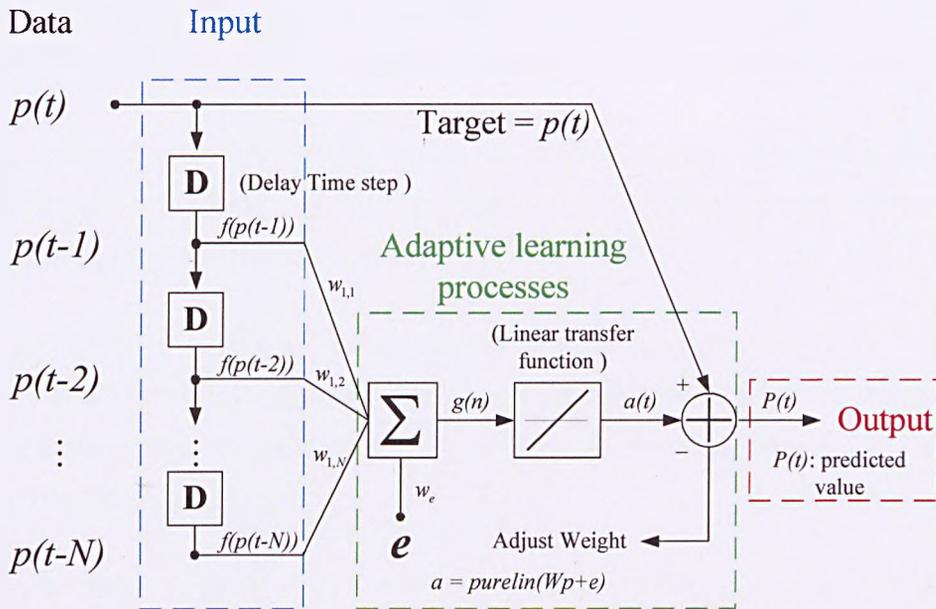


Figure 4.8: ADALINE network system for CSO hydraulic performance prediction model (Developed from Widrow, 1985)

p_1, p_2, \dots, p_R : Elements in input vector – [chamber water depth, rainfall intensity value]

w : Weight of each input element in transfer function (linear)

a_1, a_2, \dots, a_S : Output of linear neural layer

e : summary of bias each transfer function

R : Number of elements in input vector

S : Number of neurons in layer

t : time step

As demonstrated in Figure 4.8, the 'Input' is identified as blue, which is normalised from the original dataset. The newest updated value $p(t)$ is the desired model output and $p(t-1)$, $p(t-2)$, $p(t-N)$ is the model input, which

are the values at the required time steps. The difference between model input time step and output time step were introduced as 'delay' in Figure 4.8.

Weights were added to model inputs to determine the relationship between model inputs and outputs. A system bias of each training calculation was also implemented into the model as e .

The approach of obtaining model outputs was not a single calculation; instead, there was an adaptive learning process with the purpose of adjusting the general weight parameters of model inputs. These weights were an essential component of the prediction process.

As presented in Figure 4.8, the newest updated value $p(t)$ is defined as the 'target output' (actual measured value) of the prediction system; as a consequence therefore, each result of a single calculation was compared with the 'target output' to identify the bias. Bias was used to adjust the weight value following the LMS learning rule. Loop calculations were carried out for each adaptive step and memorised by the model, see section 2.2.3.

4.2. Model framework:

The framework of developing CSO hydraulic performance prediction model is presented in Figure 4.9.

Model Development

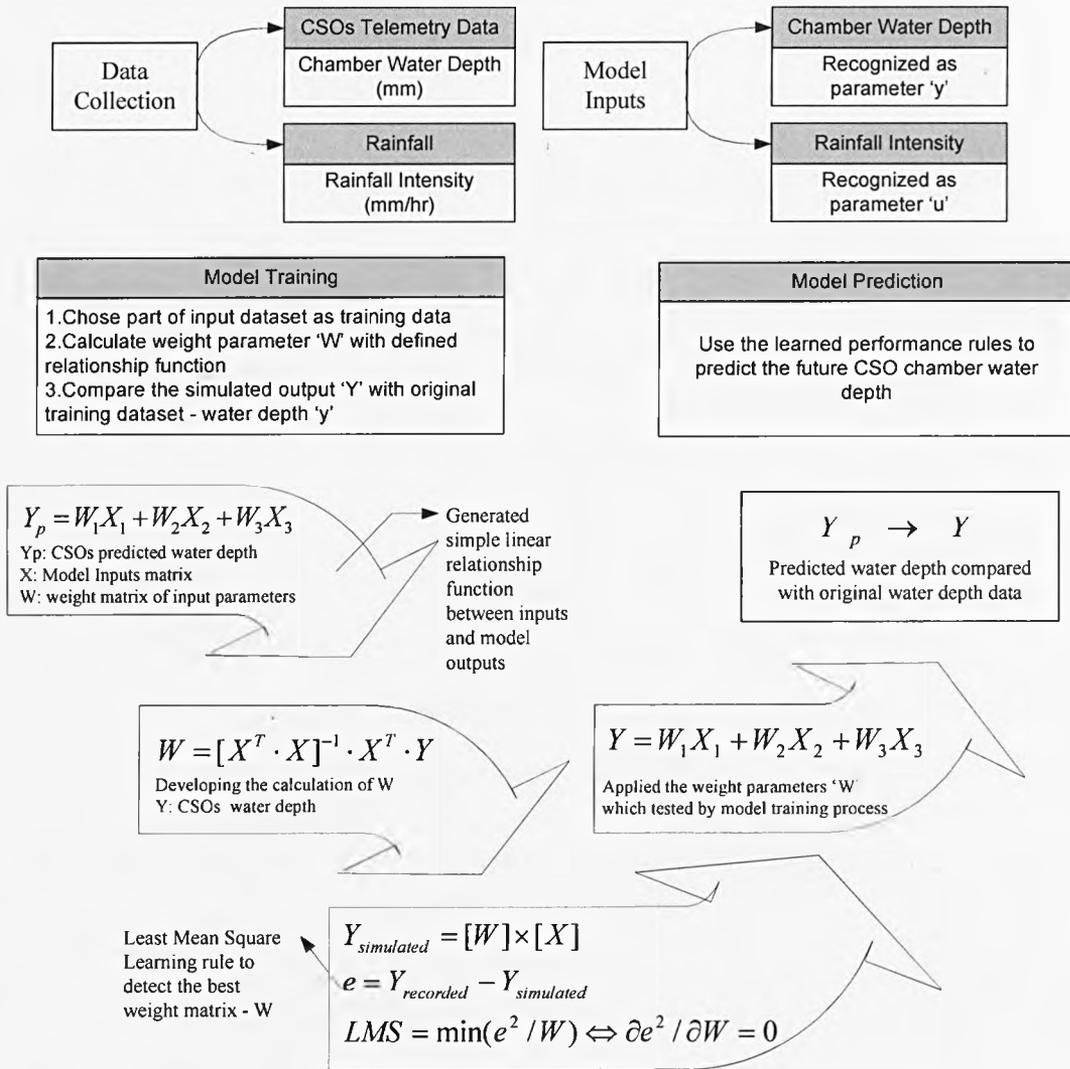


Figure 4.9: CSO hydraulic performance prediction model framework

All data preparation works had been done and introduced in chapter 3. The application of these collected data in developing this CSO hydraulic performance prediction model is introduced from the next section 4.3.

4.3. Model training:

Section 4.1 introduced the linear prediction model training process and was based on a hypothetical linear transfer function which presents the relationship between model inputs and outputs.

The very first step of building this CSO performance prediction model is to develop the linear network according to the ADALINE approach, shown in Equation 4.8:

$$net = newlin (X, Y)$$

$$Y = sim (net, X)$$

Equation 4.8

Where:

X R×Q matrix of Q input vectors

Y S×Q matrix of Q target class vectors

Here the X is presented as a 2×Q dimension matrix that contains:

Rainfall: u_{t+n} to $u_{t+n-Q+1}$

CSO hydraulic performance: y_t to y_{t-Q+1}

Y is the target class vector included a single set of data:

CSO hydraulic performance: y_{t+n} to $U_{t+n-Q+1}$

To use the ADALINE methodology the collected data sets are divided into two or possibly three groups: a set for training, a set for testing and, if required, a set for validating the model. The training data set is a data group used for learning, i.e.

learning the performance rules hidden inside the set of data, to fit the parameters or weights of the classifier. This process is shown as Figure 4.10:

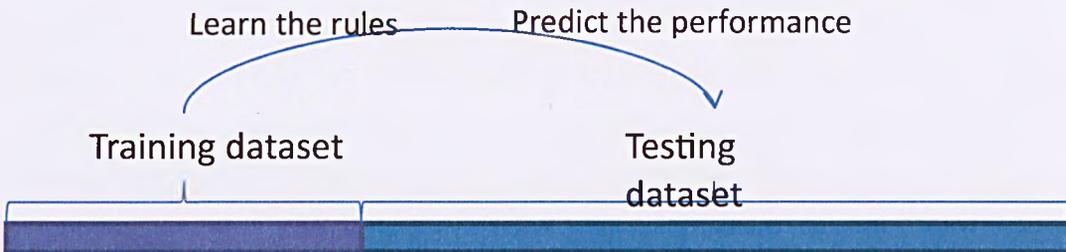


Figure 4.10: Training dataset and testing dataset

Learned rules from the training data set were then used to predict the performance for the testing data set. These predictions were compared with the measured testing dataset to test the model efficiency; details of model testing will be introduced in section 4.5. After model testing, the verified model was applied to predict the CSO chamber water depth performance in advance. The model prediction process will be introduced in next section.

As detailed in section 4.1.2, both CSO hydraulic performance data and rainfall intensity data were arranged for periods of 4 weeks and the interval between each record was 5 minutes. Therefore, normally, each dataset contained more than 8000 values.

All inputs parameters are generally defined as variable X and the output parameter was defined as variable Y . The theoretical transfer function of input and output of this model was developed from Equation 4.1 and presented as Equation 4.9:

$$Y = [W][X]$$

$$Y = W_1X_1 + W_2X_2 + W_3X_3$$

Equation 4.9

$$X = [u_1, u_2, \dots, u_n, y_1, y_2, \dots, y_m, e]$$

$$W = [A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m, C]$$

Variables 'u' and 'y' represented the model input: rainfall data and recorded CSO chamber water depth, 'e' indicated the bias added into this system. A, B, and C represent the weight parameters for each input variables included in the error, n, m were known as the number of each input. The model training process is shown as

Figure 4.11:

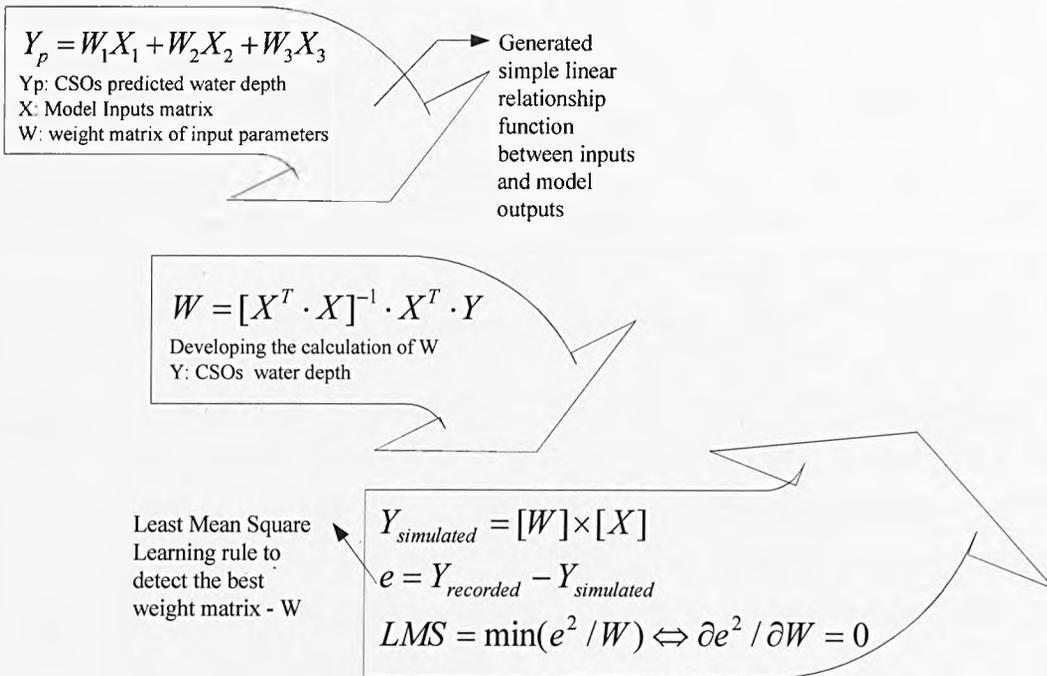
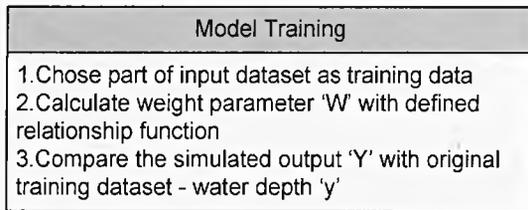


Figure 4.11 Model training process

The learning rule was introduced in the Section 4.1.3.

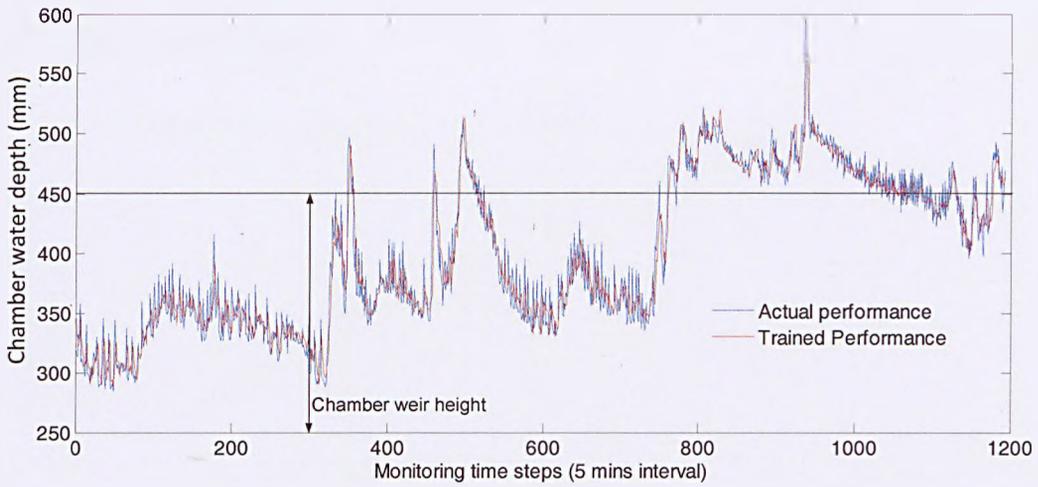


Figure 4.12: Example of model training output

Figure 4.12 shows an example of a 1200 time steps (100 hours duration) of water depth values during rainfall event recorded in Terry Avenue CSO with static screen. Also shown are the water depth values predicted by the ADALINE model. To test the goodness of fit correlation analysis was carried out to indicate the correlation between training output data and original recorded data. The result is shown as Figure 4.13 below:

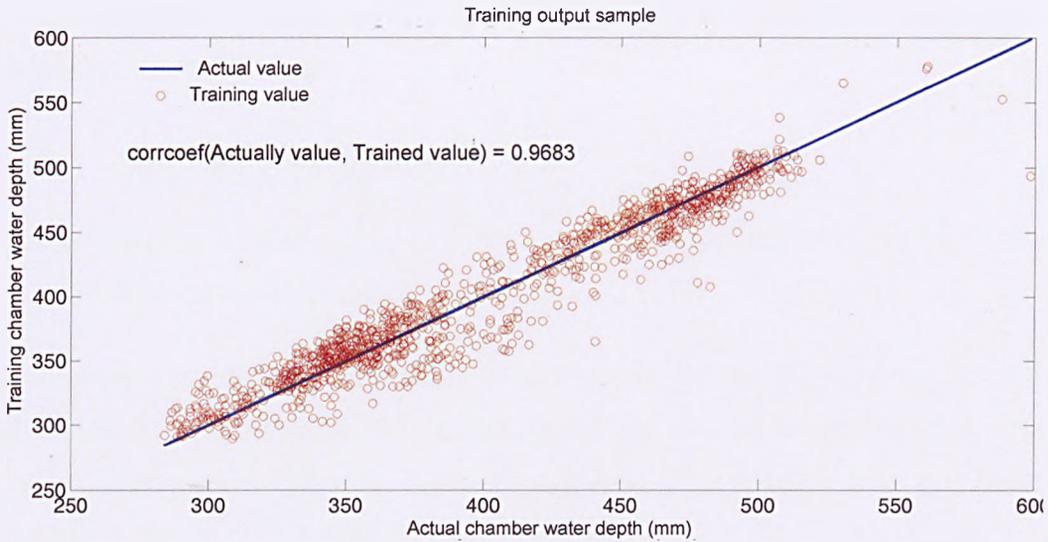


Figure 4.13: Example of correlation analysis between training outputs and original value

A correlation coefficient was established for the data using Equation 4.10

$$Ce(X, Y) = \frac{Cov(X, Y)}{\sqrt{Cov(X, X) \times Cov(Y, Y)}}$$

Equation 4.10

'Cov' is the covariance matrix computed from the matrices described in as Equation 4.11:

$$Cov(X, Y) = E[(X - E[X]), (Y - E[Y])]$$

Equation 4.11

The examples in Figure 4.13 indicate the training output samples based on using 1200 data points as training datasets (Figure 4.12). The model training outputs, which were obtained by applying the trained 'weight' value derived from the training dataset, were used to reproduce the predicted CSO chamber water depth performance. These trained outputs were compared with the original recorded values that were used to evaluate the model training process. Normally, training outputs always appeared highly correlated with the original data.

Subsequently, the trained 'weight' values were applied to the rest of the prepared dataset for model testing.

4.4. Model predicting:

The relationship between input and output was 'learned' from the model training step that was introduced in section 4.3.

The model prediction process consisted of applying the relationship or rules, which was obtained from model training, in order to predict further chamber water depth performance. As the inversed process of Equation 4.9, the model prediction can be represented as Equation 4.12:

$$[W][X] = Y \quad \leftrightarrow \quad W_1X_1 + W_2X_2 + W_3X_3 = Y$$

Equation 4.12

$$X = [u_1, u_2, \dots, u_n, y_1, y_2, \dots, y_m, e]$$

$$W = [A_1, A_2, \dots, A_n, B_1, B_2, \dots, B_m, C]$$

As introduced in section 4.3, the model training approach was intended to train the dataset to obtain performance rules. For this project specifically, it was to indicate the relationship between rainfall and recorded CSO chamber water depth performance. The learned performance rules – presented as a matrix populated with weight parameters linked with all inputs (rainfall, recorded CSO chamber water depth and a constant added as system error) were applied to another section of dataset selected from prepared data as testing.

Tested weight parameters were then implemented to predict the CSO performance. The framework of model prediction is shown in Figure 4.14:

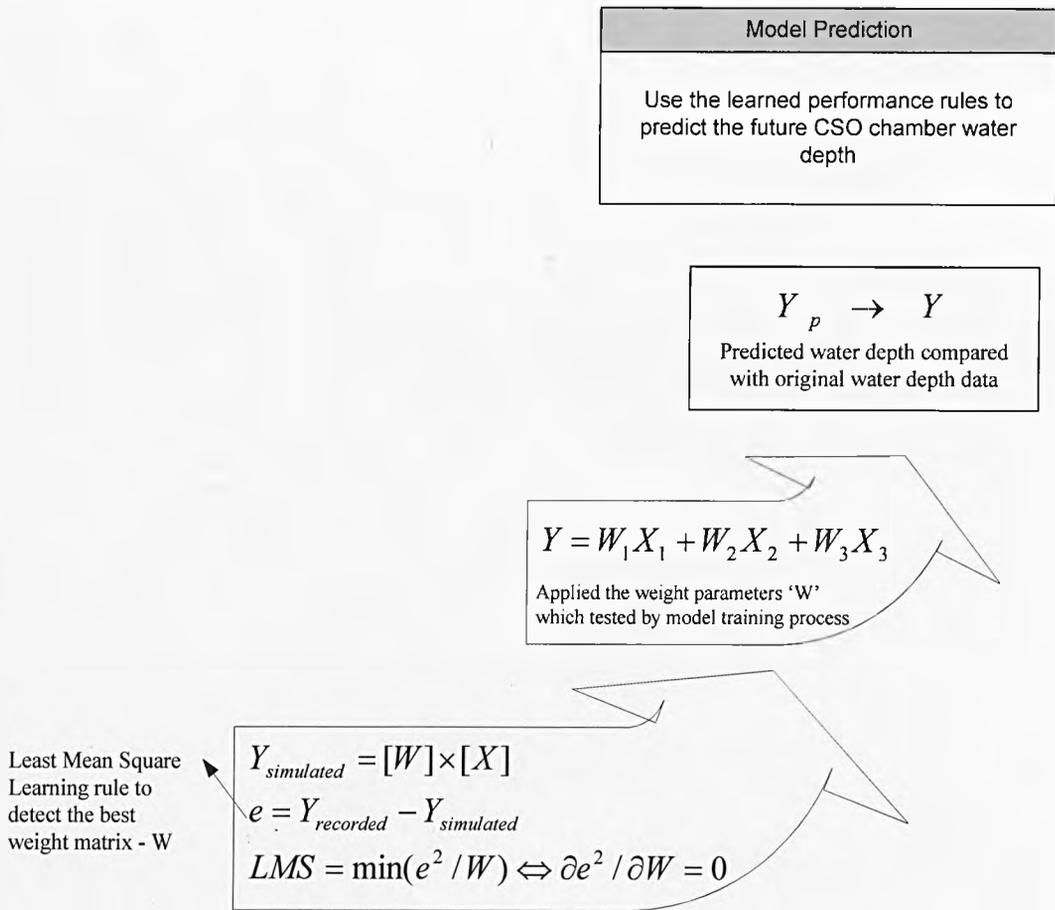


Figure 4.14: Model prediction process

As seen from Figure 4.14, this working flow diagram for the model prediction process, shown in Figure 4.11 was, essentially a reverse application of the applied learned relationship between CSO chamber water depth and rainfall intensity to predict further chamber water depth with new given rainfall information. An example of the model predicted output is shown in Figure 4.15:

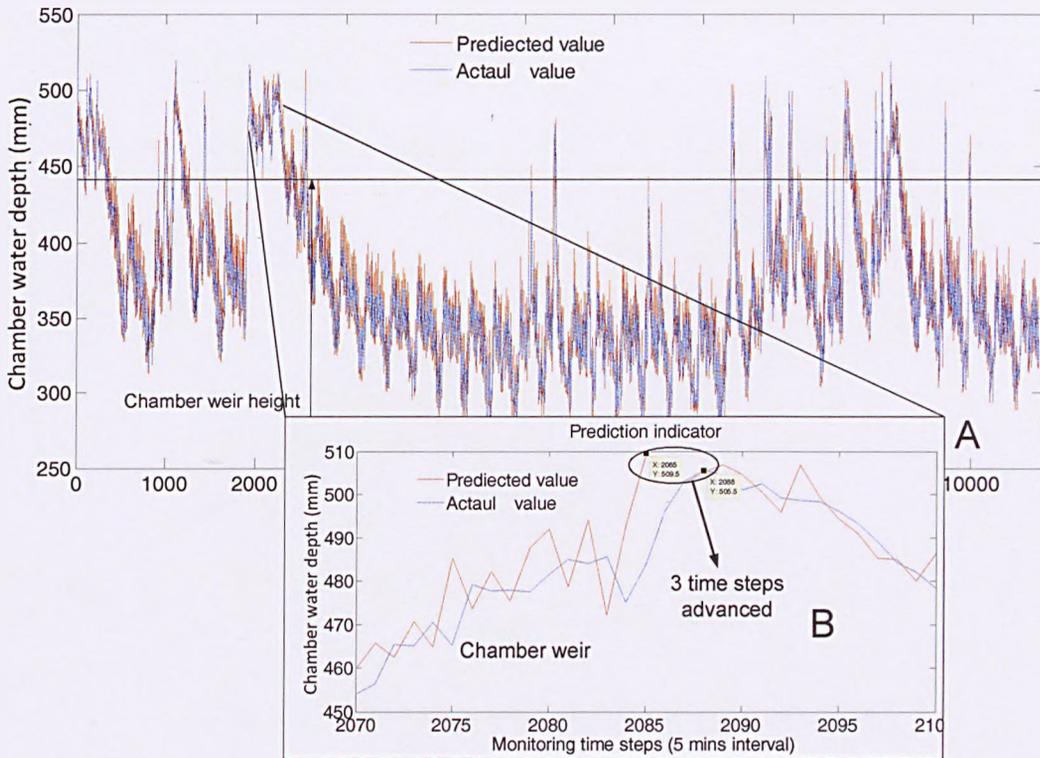


Figure 4.15: Model prediction output samples.

Figure 4.15, also includes a zoomed in section of the water depth over the chamber weir, (shown as figure B). The prediction outputs were shown 3 time steps in advance, as reported in Section 4.3. The expected changes in water depth in response to rainfall were accurately predicted. Figure B also shows the relationship between the predicted and actual measured depth was reasonably well predicted 3 time steps in advance. To test the goodness of fit of the example was shown in Figure 4.15, correlation analysis was used. Figure 4.16 indicates the correlation between predicted outputs and original CSO chamber water depth values. In this example, 1200 data points were used for model training and 10000 data points were used for the model testing.

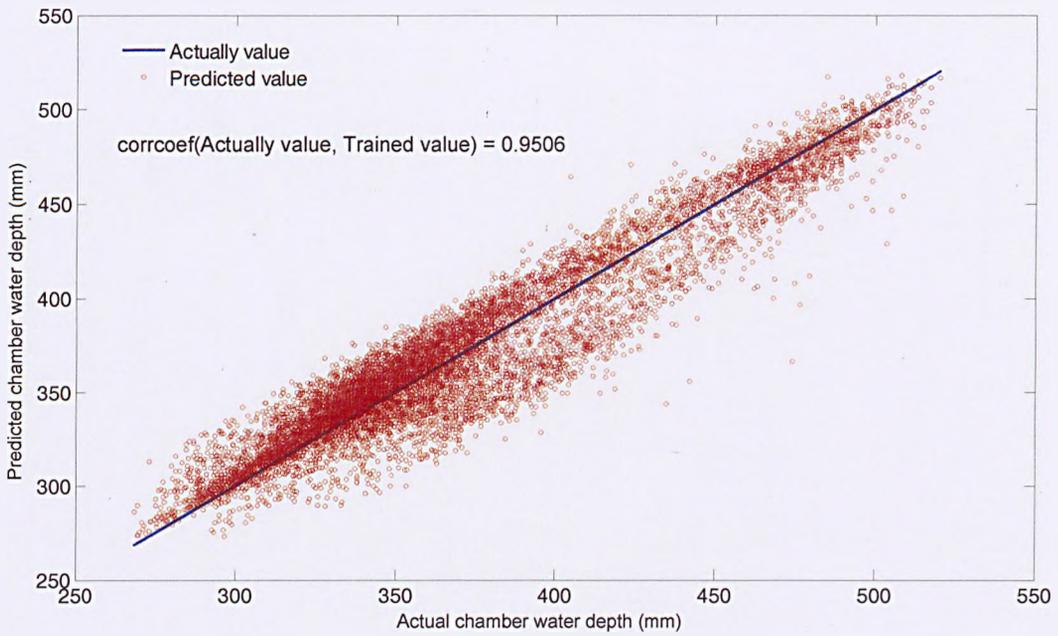


Figure 4.16: Correlation analysis of sample in Figure 4.15

During the development of this CSO prediction model, some chamber water depth values appeared to be corrupted, for example, as shown in the Figure 4.17 below:

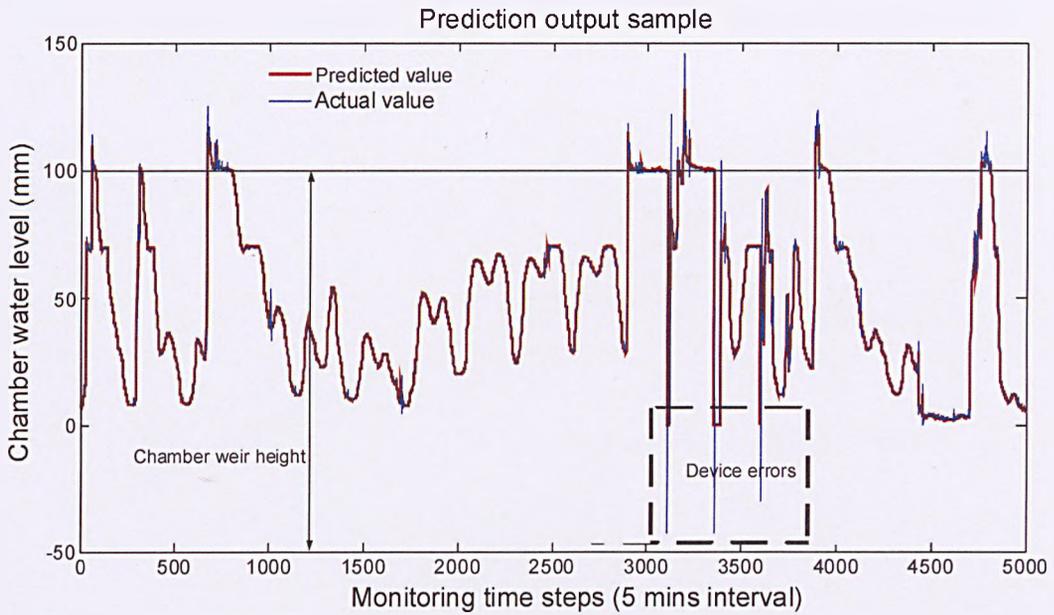


Figure 4.17: comparison between predicted and actual value of model testing. Chamber water depth from CSO asset with telemetry device errors

Several water depth data point values were recorded in the data base as '0' or 'negative' between 3000th to 4000th time steps. These were caused by telemetry system errors. If the model inputs contained a significant number of telemetry errors, the model will 'learn these errors and make incorrect predictions. These telemetry errors can be identified, from a correlation analysis of model predicted data and recorded data as shown in Figure 4.18:

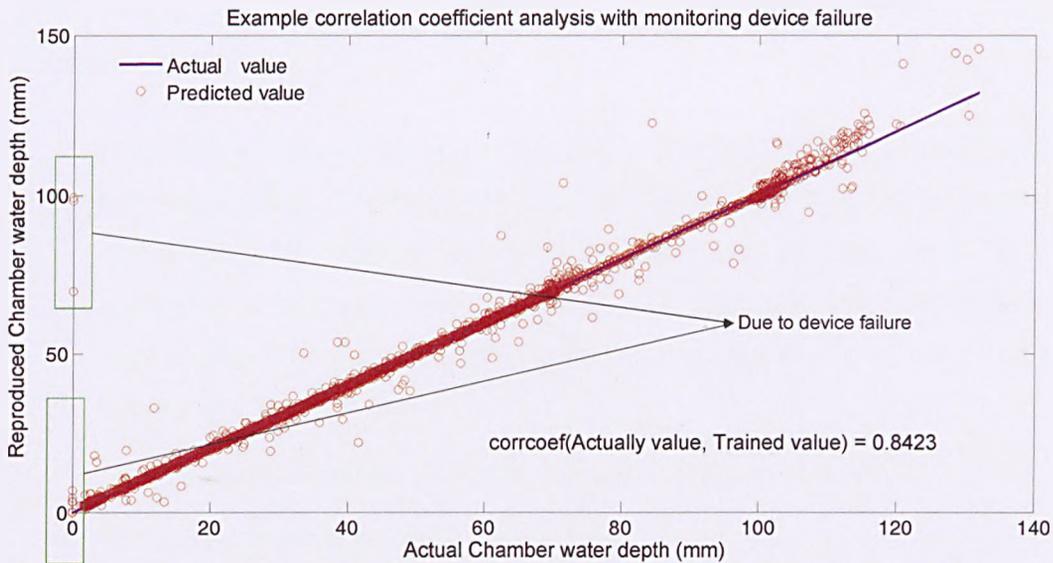


Figure 4.18: Error example of correlation coefficient analysis in prediction testing.

Compared with the previous correlation analysis outputs, the 'corrcoef' result corresponding to the Actual chamber water depth value in Figure 4.18 is 0.8423, which appeared a lower prediction coefficient. Hence, according to the learning and prediction mechanism of ADALINE, such 'error' values had a negative influence on model predicting process. However these results may be usefully used in the analysis as they may be used to detect a malfunction in the telemetry system.

4.5. Model sensitivity testing:

After the development of this CSO prediction model, as series of sensitivity tests were completed using the model in an attempt to improve the accuracy in the application of the model and to optimise the data inputs.

Model sensitivity tests were also completed to determine the factors that could have an impact on the model prediction results. During the progress of building the model and experiences referred to by previous researchers (Fernando et al.

2006), three factors were considered in the sensitivity testing. Four are summarised below:

- ✓ Over fitting problem. Length of dataset, which was selected during model training process, was known as the potential cause of over*fitting problem. Noise within the input dataset could mislead the prediction model in the training and learning process. Similarly for example, the CSO chamber water depth data appeared to contain error values which were caused by the monitor function failure.
- ✓ Test the trained system with using different length of dataset.
- ✓ The model prediction range.

Each of the three tests will be discussed individually in the following sections.

4.5.1. Avoidance of over fitting

Normally in neural network analysis, 40% of input data was used as network training data to learn the relationships between inputs and outputs (Wasserman, 1993). In this project, a correlation coefficient analysis loop calculation was developed to ensure that a sufficient length of input data was used in the model to prevent problems associated with "less fitted", which means the trained value had a low fitness to original measured value due to narrow range of input data and also to prevent problem with "over fitted" (a low fitness between trained value and measured value due to too many data points being used for network training. Therefore, a wide range of input data points spanning from 10% to 80% of the total length of dataset (4 weeks length of both CSO chamber water depth and rainfall, 8046) were used as network training. In line with most other researchers, for example (Fernando et. al., 1993), the last 10% of CSO water depth data was used for network simulation testing. The result of the predictions were compared, by correlation analysis, to identify the best solution in terms of the number of data points to be used to optimise the training outputs i.e. when the trained output gave the highest correlation coefficient.

In Figure 4.19, the X coordinate identifies the number of inputs from 100 to 4000. The Y coordinate indicates the 'total error square'; the blue line indicates the total squared trained error (TTE) and the red line the total squared predicted error (TPE). According to the basic mechanism of the ADALINE prediction model, the longer the dataset was selected as the model training the more system characteristics were learned and an example is shown in Figure 4.19:

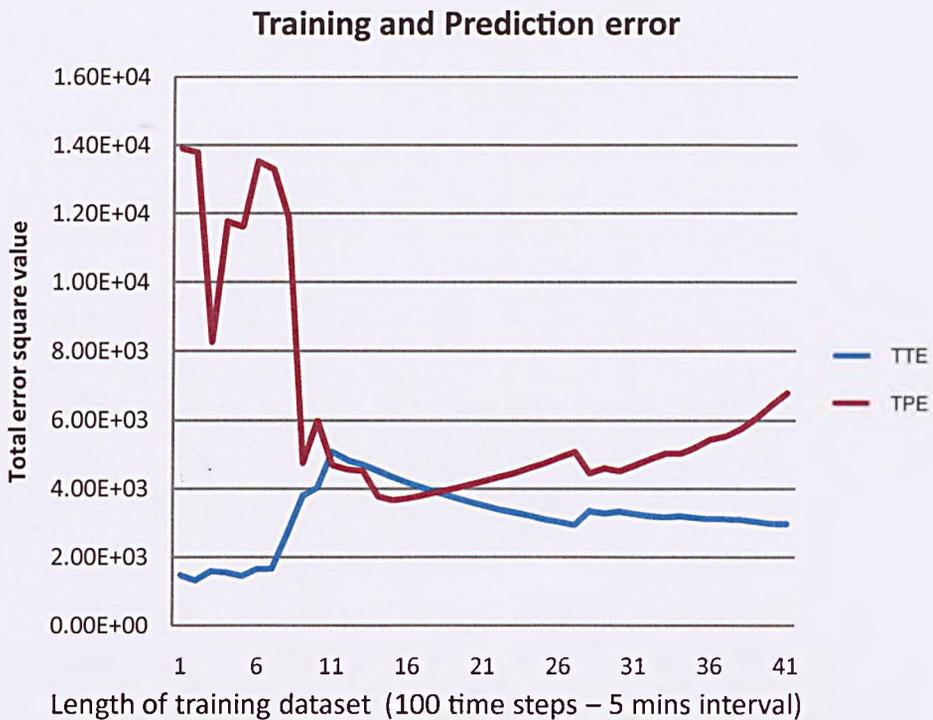


Figure 4.19: Example 1 of model over fitting

This highlights both under and over fitting. The total squared trained error (TTE) increased as up to the time when 1200 data values were selected, whilst the TPE was shown to reduce rapidly over this interval. This indicates that the model prediction accuracy was improving with an increased length of the dataset that was used for model training.

A further increase in the number of data points showed that the TTE peaked at approximately 1200 data points and then gradually reduced to the end of the data set. However, the TPE increased after the training set was longer than 1400 time steps. The TTE and TPE had similar values up to a data set of 1700 values after which the TPE was observed to gradually increase. This meant that the model was now becoming over fitted with an increase in the number of data points. A second example of over fitting is shown in Figure 4.20:

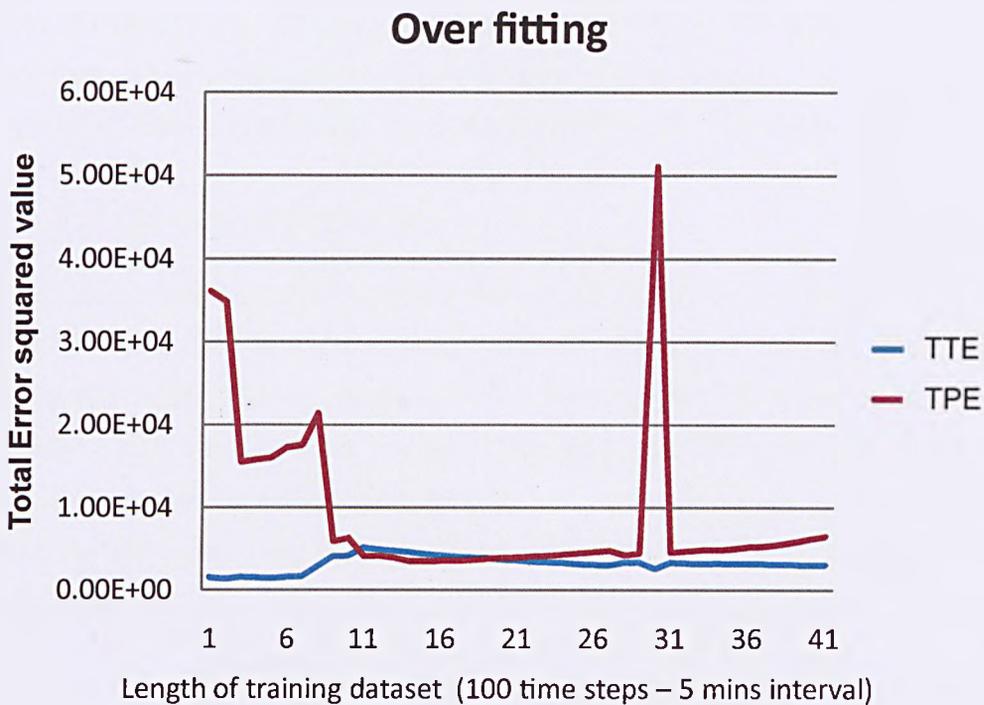


Figure 4.20: Example2 of model over fitting

The data set shown in Figure 4.17, which contained some zero and negative water depth values due to instrument malfunction was also used to demonstrate that over fitting problems maybe caused by unexpected performance features. In Figure 4.20, it is again shown that the TTE and TPE converge to a minimum value at a dataset length of approximately 1200. Hence it may be argued that this

length of data set is appropriate for the ADALINE methodology as this gives the most efficient solution for model prediction.

Figure 4.20 also shows that there was a significant peak in the value of the TPE corresponding to the training dataset length of about 3000. This corresponded to the time that the instrument malfunctioned.

As a conclusion, the selection of model training dataset length has a direct impact on the model's prediction accuracy. For the data presented here a training data set length of 1200 was found to be appropriate for this particular CSO chamber. As a consequence a model over fitting analysis was completed for each CSO assets to establish the most efficient solution for model prediction.

4.5.2. Test learned system

Other tests were also completed to test the sensitivity of the model to a number of different input parameters. The first of these was the sensitivity of the model to the length of the training data set, with a view to optimising the accuracy (best correlated) of model output. Initially 1200 steps of CSO chamber water depth values were used for the model training and to test the sensitivity of the model three further tests were carried out by using 3000, 6000, 9000 steps of data respectively.

Comparisons between actual values and predicted values and the resultant correlation coefficient analysis are shown in Figure 4.21 to Figure 4.26:

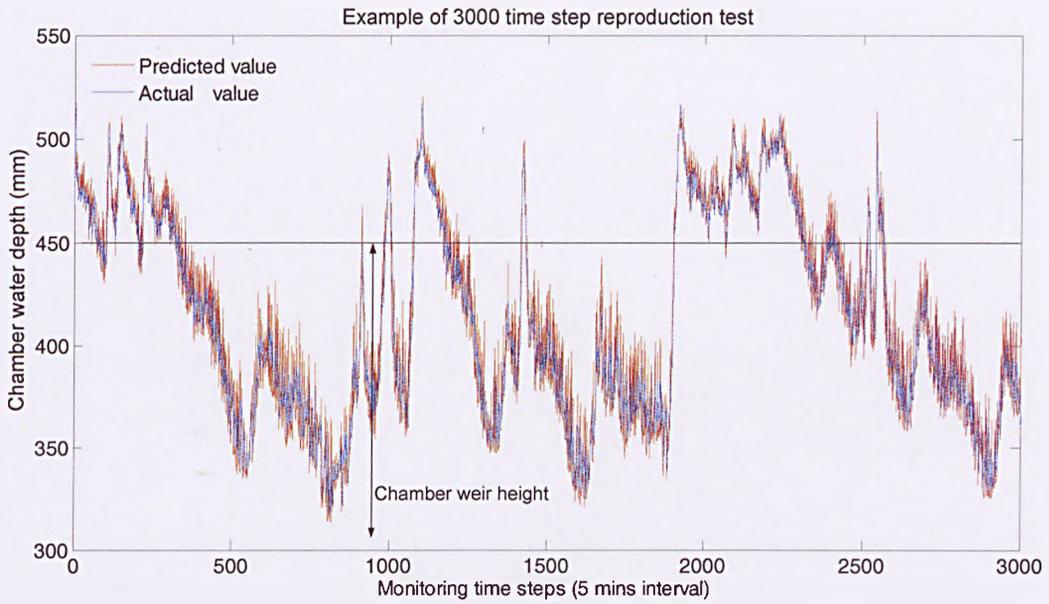


Figure 4.21: Example of output by using 1200 values for training and 3000 values for testing

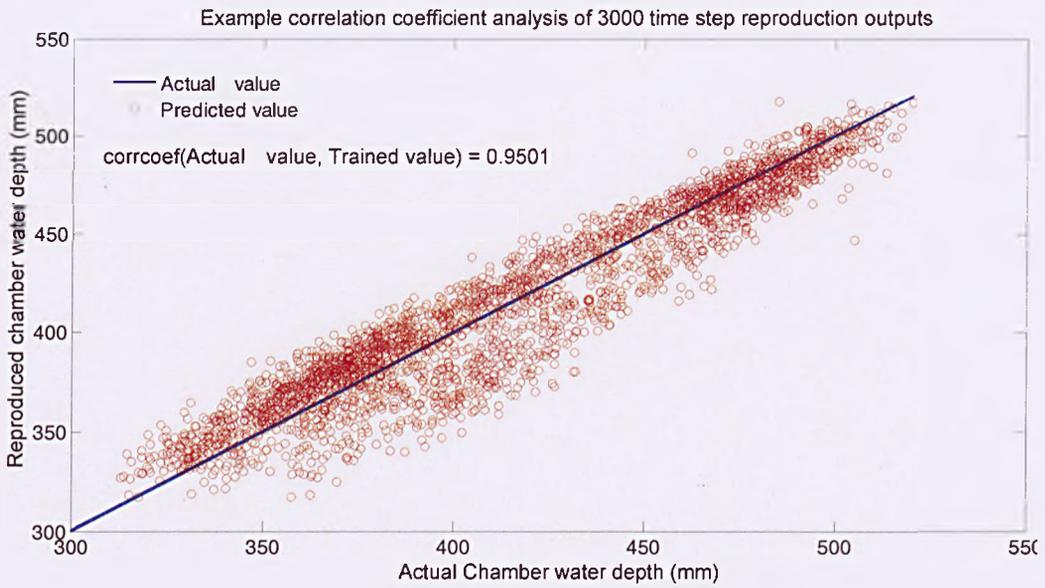


Figure 4.22: 'corrcoef' analysis of model by using 1200 values for training and 3000 values for testing

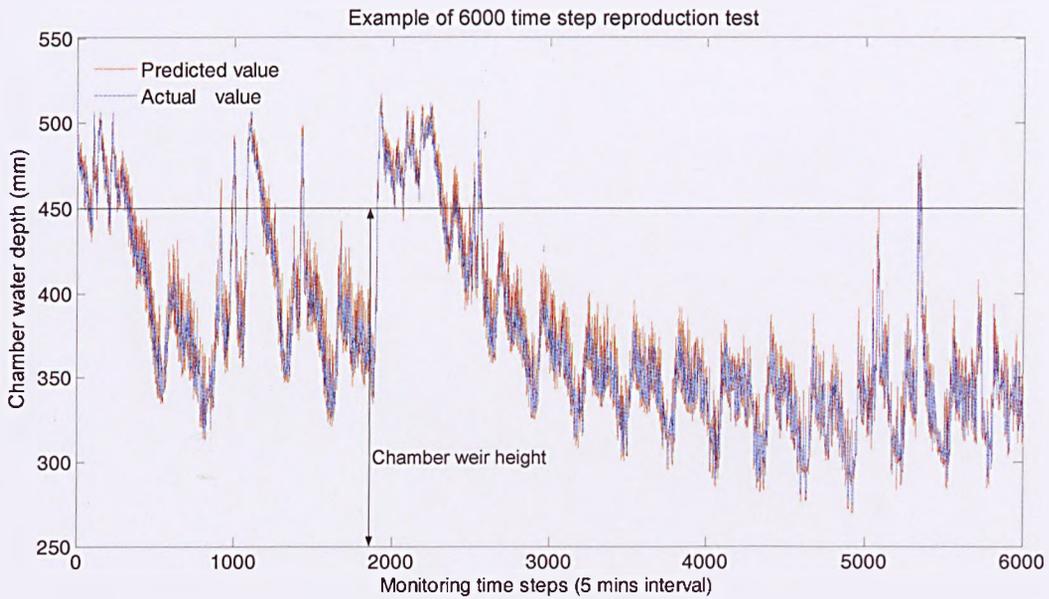


Figure 4.23: Example of output by using 1200 values for training and 6000 values for testing

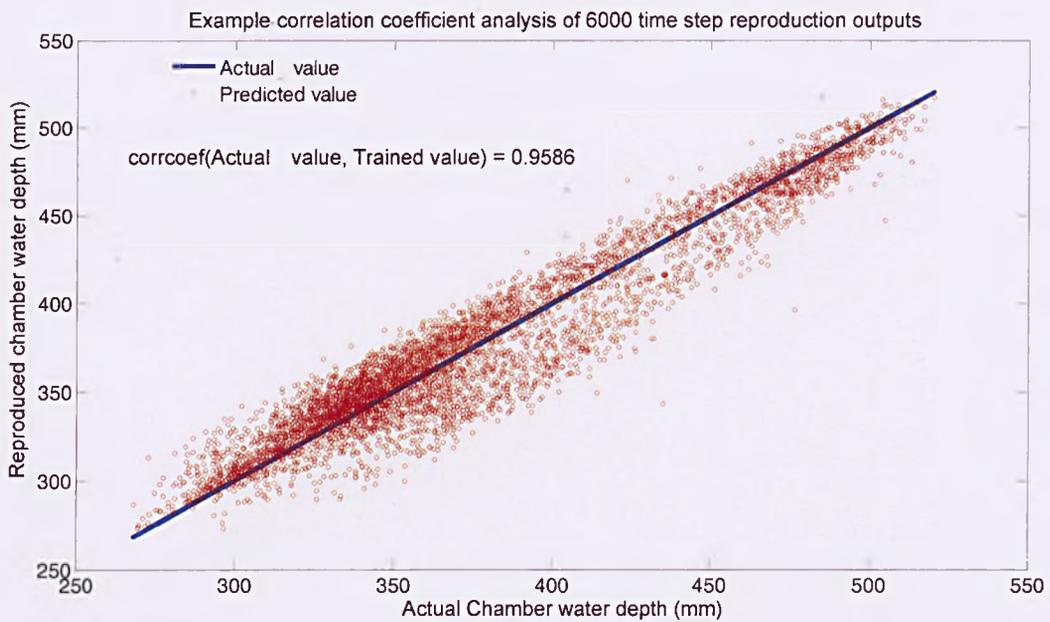


Figure 4.24: 'corrcoef' analysis of model by using 1200 values for training and 6000 values for testing

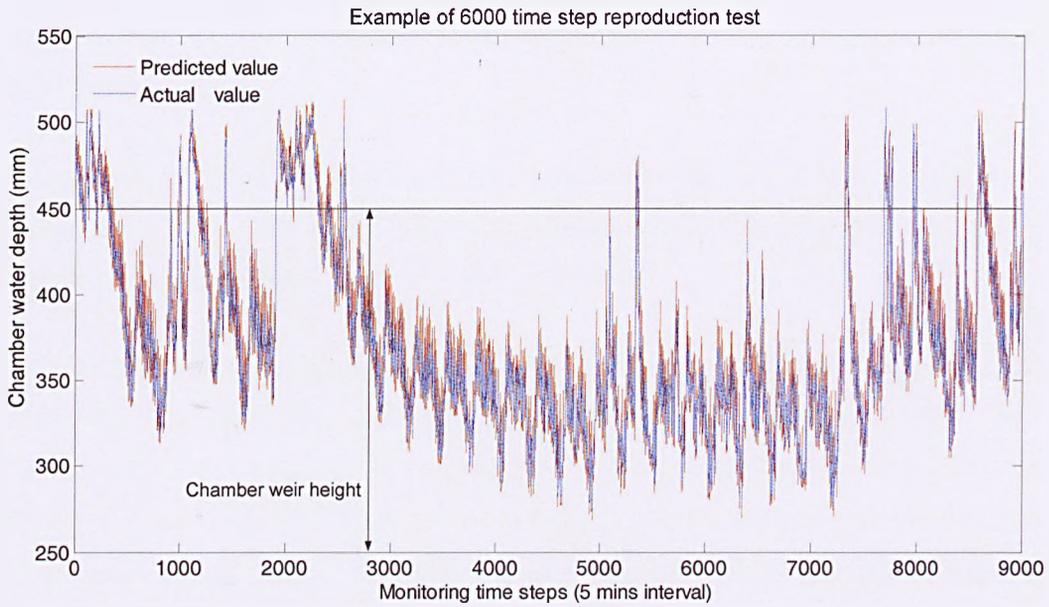


Figure 4.25: Example of output by using 1200 values for training and 9000 values for testing

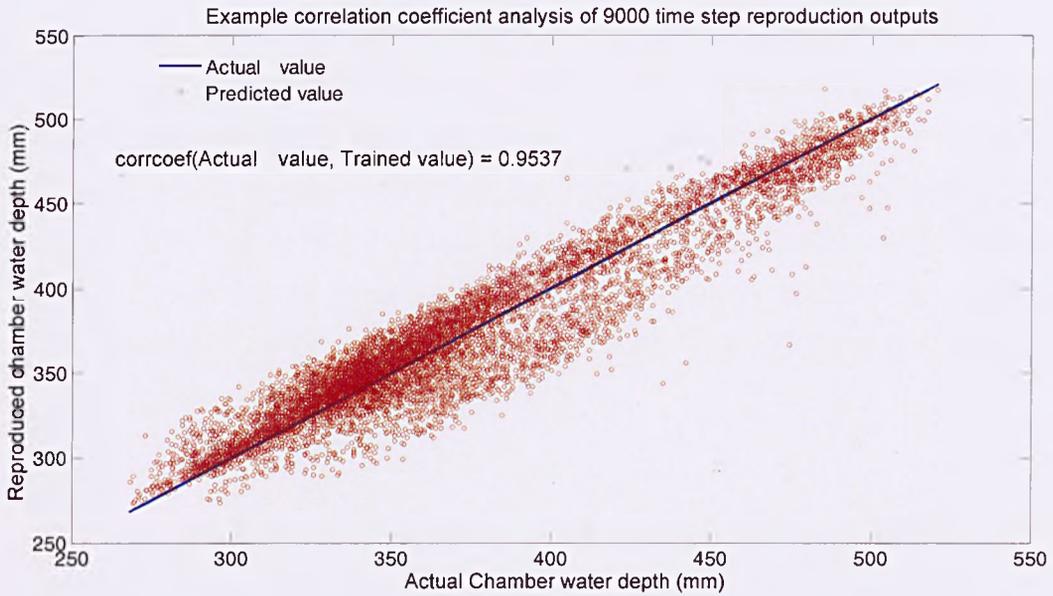


Figure 4.26: 'corrcoef' analysis of model by using 1200 values for training and 9000 values for testing

The prediction outputs of each model were shown in Figure 4.21, Figure 4.23, and Figure 4.25. Each figure demonstrated that the prediction output compared with original recorded chamber water depth with different length of data steps from 3000 to 9000.

Correlation coefficient analysis of model tests were shown in Figure 4.22, Figure 4.24, and Figure 4.26 indicated the correlation coefficient analysis results. Here 'corrcoef' indicated the calculation as Equation 4.10,

The correlation analysis of all three tests showed that the model prediction accuracy was high and was very similar for each test. Actual vales were:, $Ce_{3000} = 0.9501$, $Ce_{6000} = 0.9586$, $Ce_{9000} = 0.9537$. These may be compared and contrasted with the correlation coefficient $Ce_{10000} = 0.9506$ that was obtained with 1200 data points. As a conclusion, it was argued that 1200 data points were sufficient to train the ADALINE system.

4.5.3. Prediction range

The approach of selecting model input was introduced in section 4.1.2. As explained in Figure 4.7, the difference of input data steps indicated the model prediction range. Which highlights that the data steps of input parameter u (rainfall intensity) was always more than the data steps of input parameter y (chamber water depth), such that the model outputs may be predicted in advance. The relationship between model prediction accuracy and prediction range was established as part of the sensitivity analysis and an explanation of model prediction range is shown in Figure 4.27. This Figure also includes an example of the model predicted output:

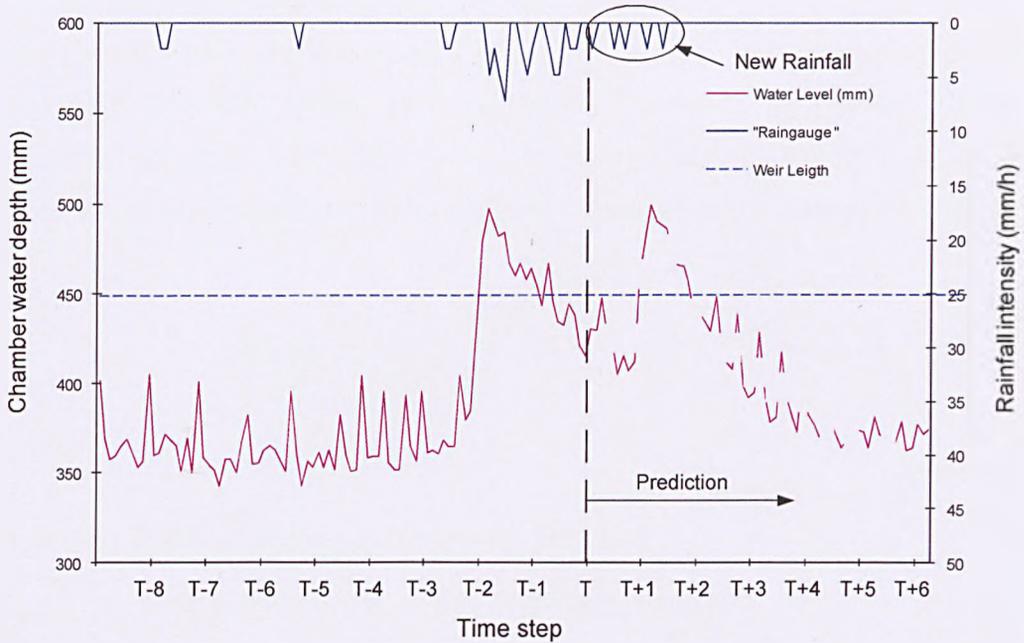


Figure 4.27: Demonstration of model prediction (from T+1 time step) Wednesday

In Figure 4.27, 'T' time step is considered as the 'current' chamber water depth, the measured rainfall intensity value after 'T' time is defined as new rainfall information and the chamber water depth values after 'T' time describe the model predicted values. Because all CSO chamber hydraulic performance data used in developing this ADALINE prediction model was collected historically from a CSO telemetry system database, the prediction shown in Figure 4.27 was actually completed as an off line process. Predicted CSO chamber water depth values were compared with recorded values in order to evaluate the model's efficiency.

Generally, the time interval between measurements for both rainfall intensity and water depth was 5 minutes. As explained in Figure 4.7, the model was sensitive to the number of time steps in advance that the model was used to predict i.e. the length of the model prediction range. This is indicated by the increased number of steps of rainfall intensity values that were used as input compared with the number of CSO chamber water depth values.

To test the model prediction accuracy due to different lengths of the prediction range, the mean error between predicted outputs and actual value were calculated and the values were compared according to different length of prediction range for each CSO asset. The mean error of prediction outputs was calculated using Equation 4.13 to indicate the prediction accuracy of model.

$$\text{Mean error} = \sqrt{\frac{\sum_{i=1}^n (Y_i - Y_{i_{\text{predicted}}})^2}{n}}$$

Equation 4.13

Y_i : Actual chamber water depth value for time step i

$Y_{i_{\text{predicted}}}$: Predicted chamber water depth value for step i

n : predicted time steps

To evaluate the prediction accuracy through different CSO assets, the value of error in percentage, which can be calculated by Equation 4.14, was used for each CSO on each step of prediction.

$$\frac{\text{Mean error}}{\text{Predicted water depth value}} \times 100\%$$

Equation 4.14

Applying Equation 4.14, three examples of prediction analysed results based on three different CSO assets are summarised in Figure 4.28:

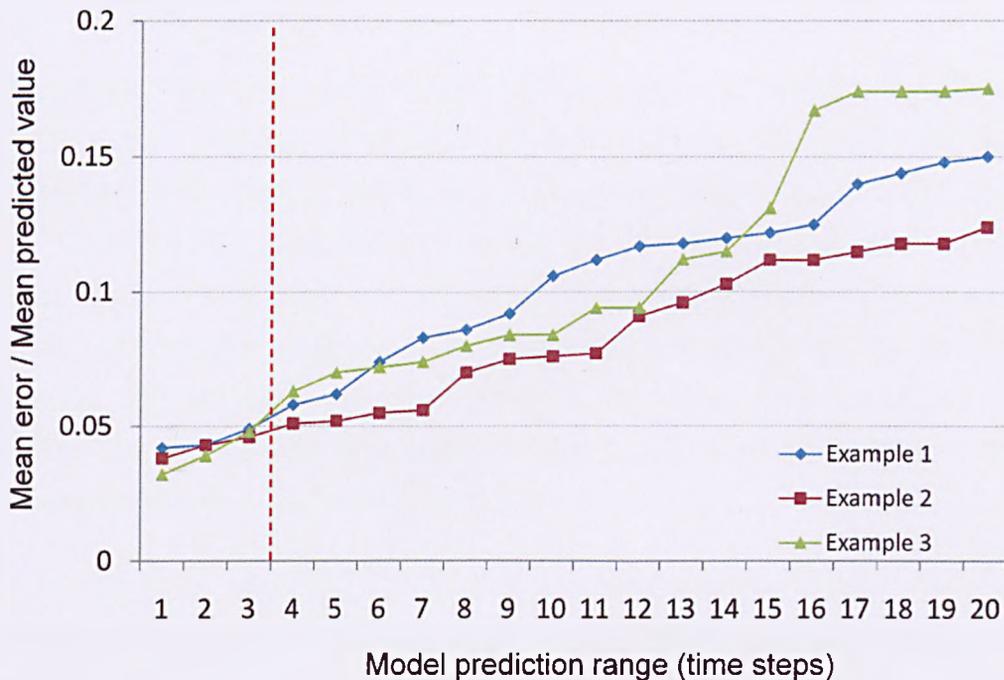


Figure 4.28: Model total mean error square compared with length of prediction period

The methodology was applied to predict the outputs at 1, 2 and 3 time steps in advance (corresponding to example 1, 2 and 3 respectively) in Figure 4.28 for time steps between 5 minutes to 100 minutes as shown in Figure 4.28. These tests were completed using 3000 data points of which 1200 values were used for model training. As can be seen from Figure 4.28, the mean error of the prediction output for example 3 (3 time steps in advance) were lower than 5% (0.05 on the plot) compared with value of mean predicted chamber water depth. The testing output, which locates on the right side of the red dash line, indicates that the model prediction accuracy is reduced as the prediction range is increased.

As a conclusion, from testing the model prediction range and accuracy for each CSO asset is 3 time steps (15 minutes in advance). For this range the prediction outputs of the CSO ADALINE model were recognised as confident with an error less than 5% comparing with the mean predicted value.

4.6. Summary

This chapter has introduced the whole process associated with the development of the CSO hydraulic performance prediction model. The initial purpose of developing this prediction model was to study the relationship between rainfall and CSO hydraulic performance and to subsequently predict CSO chamber water depth at time steps into the future. The overall intention of the research was to provide water companies with a predictive tool to provide early warning of potential spill events from individual CSO chambers, and subsequently to improve the management and operational strategies for sewer systems in order to reduce flooding and pollution incidents.

As a conclusion, this ADANLNE CSO performance prediction model was shown to be effective providing efficient chamber water depth prediction with a 15min lead time. The prediction of CSO performance created the opportunity to develop an advanced proactive alarm mechanism to predict potential asset performance failures. The next section of this thesis was to introduce the development of a pro-active asset operation and maintenance decision support model based on the outputs provided by the hydraulic prediction model.

5. CSO pro-active operation and maintenance decision support model

The goal of this pro-active O&M decision support model is to provide an efficient pro-active maintenance decision support tool that may be used to respond to CSO hydraulic performance failures that were predicted by the ADALINE model.

In terms of the proposed model of CSO asset O&M prediction, there are several challenges to predicting an appropriate O&M schedule:

- ✓ Multiple parameters, a number of parameters were recognised to have a potential impact on the requirement for O&M actions. Each of these parameters needs to be considered.
- ✓ Complicated relationships exist between each input parameter and the O&M requirement, for example it is not feasible to use simple linear or non-linear equations to represent the relationship
- ✓ Linguistic classification of input parameters, such as the CSO chamber types and screen types, cannot be applied using a conventional numerical definition.

Compare with the conventional reactive O&M actions that occur in response to reported performance failures. The pro-active O&M approach is described in Figure 5.1:

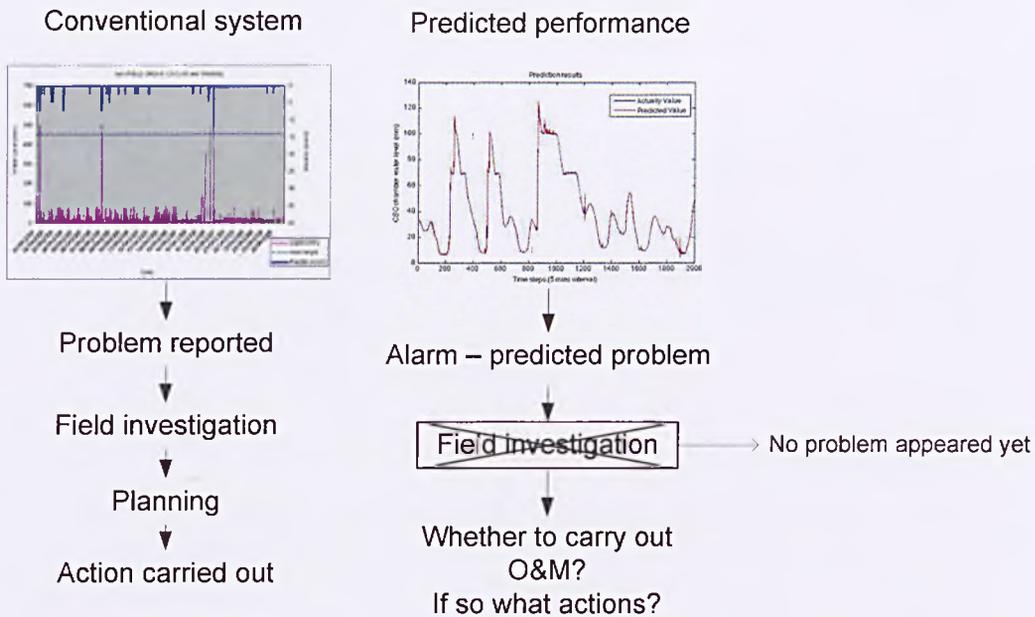


Figure 5.1: Predicted performance of system for proactive maintenance

Figure 5.1 leads to the question as to whether actions are indeed needed to remedy the potential problem, and as to what kind of action would be useful to solve the problem. To answer these questions, a decision support model utilising FL was developed and is introduced in this chapter.

5.1. Methodology

In this project, the potential applicability of a FL approach to predict the operation and maintenance requirement of a CSO asset was based on twenty CSOs which were selected from the YWS CSO pollution incident list. Seven hydraulic performance characteristics were considered as key influencing factors to the CSO O&M requirements, which are also to be used as input sets of this FL prediction model, see Table 5.1:

Table 5.1: Concerning chamber feature and hydraulic parameters of FL model

✓ Asset Structural Characteristic

CSO chamber type Installed screen type Chamber weir height
✓ CSO Designed Hydraulic Characteristics Flow rate at 1st spill event Spill rate during 1 in 5 year return rainfall incident
✓ Recorded Asset Hydraulic Performance Total spill duration Total spill volume

Three steps comprise the process of the theoretical FL application:

Step 1: Fuzzification

Fuzzification is introduced in Section 5.4. All the data sets collected for the seven input parameters were defined as a crisp set, which means the data set containing the actual value each parameter were defined by a sensible unit. The FL model required all crisp data sets needed to be fuzzified into fuzzy sets which can be used for model fuzzy inputs.

The process of fuzzification was to define the input fuzzy set using linguistic descriptions like “Low, medium, high” by developing membership functions. Each membership function was defined as a curve or mathematic expression that describes how each point in the input space is mapped to a membership value or degree of membership between 0 to 1.

Step 2: Define Rules

Fuzzy input acted as the subject of the FL model which was the “fuzzy” part of FL model. The “logic” part is defined as a fuzzy operator with the function of defining rules of the whole FL model. FL rules were defined according to historical records, experience and logic relationships. In this CSO O&M FL model, fuzzy operational rules were defined according to the actual relationship between influence factors and the requirement for a CSO O&M action.

Step 2: Defuzzification

Defuzzification was the process of producing a quantifiable result in the FL model. Normally, a fuzzy system has a number of rules that transform a number of variables into a “fuzzy” result which were described in terms of membership in fuzzy sets. There were many different defuzzification methods available.

Details of this model’s development are described in section 5.2.

5.2. CSO O&M Fuzzy System

As introduced in section 5.1, the component requirements of the CSOs’ O&M model are demonstrated in Figure 5.2.

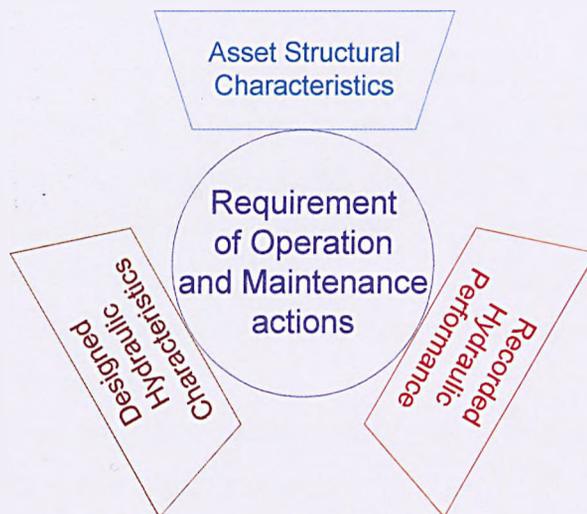


Figure 5.2: Proposed consideration of CSO O&M system

As indicated in Figure 5.2, there were three general factors that were considered in this research as potential influences on the requirements of CSO O&M actions:

- ✓ Asset characteristics: CSO chamber type, screen type and weir height
- ✓ Designed asset hydraulic features: flow rate as 1st spill event, spill rate during 1 in 5 year rainfall event

- ✓ Recorded asset hydraulic performance: the pro-active O&M decision support model was based on a performance failure alarm that was provided by the ADALINE prediction model. Therefore, the hydraulic performance features, which were recorded before the performance alarm, were considered as important influences on the requirement for CSO O&M actions. For example recorded CSO spill duration and total spill volume were considered as important variables to describe the potential for screen blinding to occur. (discussed in section 5.4.4)

The initial calculation engine used for this modelling was the FL Toolbox that was combined in MATLAB. The FL toolbox in MATLAB provided a computing environment with functions for designing a system based on a FL approach. Steps for building a FL model were identified as follows:

1. Normalise the fuzzy inference system: indicate all parameters which should be included so as to clarify system input and output
2. Develop the fuzzy inference system by identifying a membership function for each individual parameter (both input and output)
3. Define rules for fuzzy expert system: train the fuzzy inference system as an expert system by defining model simulation rules which were drawn from practical records
4. Graphical presentation of system output: visual display of the model result and used for further decision support.

Figure 5.3 shows the brief outline of the expressions used in the model, which was built with the MATLAB FL toolbox.

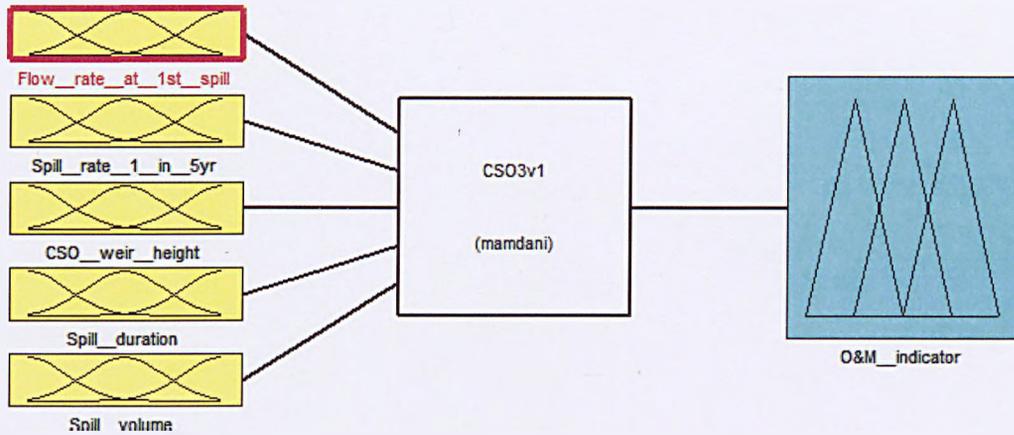


Figure 5.3: Fuzzy system in FL tool box

Five initial parameters (as shown in Table 5.1, except the chamber and screen type), which had an impact on the requirement for CSO pro-active operation and maintenance actions, were considered as model inputs as these were recognised as important factors, in consultation with the water company, as these parameters may be used as design criteria. The CSO chamber type and screen type introduced in section 5.1 were not considered as model input directly. However, in this research, the FL model was developed based on each individual chamber and screen combination, such as a CSO with side weir chamber and mechanical screen installed.

Details of the model development that are based on these three steps are introduced from Section 5.3 to Section 5.8.

5.3. Fuzzy Sets

In this research, both model inputs and output sets were developed as fuzzy sets, which included the five model inputs and the CSO O&M action information (see Figure 5.3). Fuzzy sets were simply qualitative descriptions of the chosen domains of the inputs, each of which was thought to have a specific effect on the output. Figure 5.4 shows the fuzzy sets used in the study.

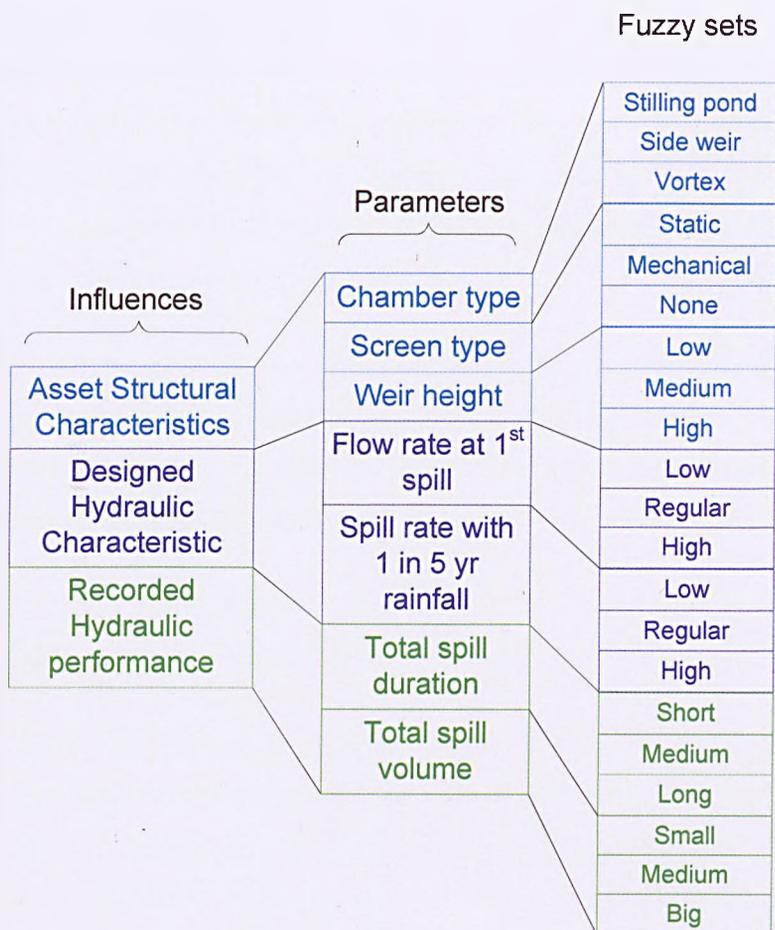


Figure 5.4: Fuzzy input linguistic description

In this research, the model is built by considering a certain combination of chamber and screen type – the side-weir chamber with a mechanical screen was the most commonly used system and hence the research initially focussed on this group of assets. CSO assets with static screen installed and side-weir chambers with no screens installed were less common. By testing the side weir chambers with and without screens, the outputs could then be compared to reach conclusions on the relationship between O&M requirements for CSO chambers both with and without screens.

The relationship between the actual values of each of these inputs was categorised as a linguistic description and was defined by membership functions.

Terms like 'low' and 'short' were defined as the smallest value of each parameter with a degree of truth equal to 1, similarly, the largest values were defined as 'high' or 'long' with the truth degree equal to 1. The average values of each input parameter were defined as terms 'Medium' and 'regular' with the truth degree equal to 1. Clearly, these definitions could be changed as more CSO asset and performance information is made available.

The CSO maintenance records including all O&M operations that were carried out on each individual CSO and screen were collected from the Water Company over a one year period. As seen in section 3.6, these records of O&M actions were grouped into two categories:

- Routine asset maintenance
- Emergency incident maintenance

Routine asset maintenance actions were based on the general asset management schedule of the water companies and were designed to guarantee and maintain an acceptable level of service and performance. The Table 5.2 indicates the O&M actions used in routine practice as recorded O&M information:

Table 5.2: Routine maintenance action List

9202-W/C Pollution Inspection
Routine Inspection CSO weekly
9257 - Op's Investigation Asset Team
CSO Inspection by Constructor
Jet/Flush Sewer From Routine Jet Point
Pilot R&M powered screen CSO inspection

Emergency incident maintenance actions, in contrast, were based on responding to CSO pollution incidents and were not planned in the water company asset management schedule. In this research, recorded responsive O&M actions were used as valuable experimental knowledge to develop the CSO O&M requirement

fuzzy expert system performance rules. Reactive O&M action types in the collected O&M information are shown in Table 5.3:

Table 5.3: Emergency incident maintenance action List

9255-Respond to RTS Sewerage Alarm
Jet Blocked Sewer
Flush and/or Jet Sewer/MH as Instructed
Clear Sewage Trash from Watercourse
EMM4 - Mechanical Repair
CLEAN/INSP/CHECK OPP OF ANCILLARY
EME1 - Electrical Repair
Rod Blocked Sewer

In the process of developing this FL decision support model, only those emergency maintenance actions which were carried out as responses to CSO performance failures were considered and used to build the expert system.

5.4. Model Input Parameters

The five parameters that were considered as model input were defined in Figure 5.3. Based on FL theory, model input parameters were represented as linguistic descriptions, which were defined by certain membership functions. Shown in Figure 5.3, the membership functions are the graphs that define how each point in the input space is mapped to a membership of degree of truth between 0 and 1. In terms of a graphical expression, there were generally four commonly used membership functions:

- ✓ Straight line
- ✓ Trapezoidal
- ✓ Gaussian
- ✓ Triangular

Selection of different membership functions depended on the actual graphic expression of real value points mapped in the space of inputs.

In order to organise a logical expression of CSO structural and operational information, a data normalisation process was applied to the developed database. During the normalisation process, all values in each dataset were normalised into a range from 0 to 1. See Equation 5.1:

$$[A_N] = \sum_1^i \left[\frac{A_1 - A_{min}}{A_{max} - A_{min}}, \dots, \frac{A_i - A_{min}}{A_{max} - A_{min}} \right]$$

Equation 5.1

A_N : Subjective dataset

A_i : Sample of data

A_{min} : Smallest value in the set

A_{max} : Biggest value in the set

Linguistic descriptions were defined based on the normalised value of each of the input factors. Note that, in this step, a normalised value between 0 and 1 represented a different characteristic to the value of 'degree of truth' which was also ranged from 0 to 1.

In this model, terms defined by the membership function as "Low weir" and "High weir" which indicated that a modified Gaussian function was used as the Gaussian combination membership function. The 'lowest' or 'shortest' value was considered as the truth degree of 'low' or 'high' as 1, similarly, the 'highest' or 'longest' value was considered as the truth degree of 'high' or 'long' as 1. Further membership functions follow the same rule in being categorised as "low" and "high" or "short" and "long" linguistic labels. All 'Regular' and 'Medium' labels are defined with considering the mean value as truth degree as 1 to develop the membership function by using a Gaussian function. From Section 5.4.2 to

Section 5.4.5 and model output membership function in Section 5.5 are following these definitions.

Membership functions were developed in this project by applying a mathematical curve fitting approach. Types of the membership function, which were introduced in section 2.3.2, were used to test the curve fitting. Examples of the fitting outputs' root-mean-square error (RMSE) value of are shown in Table 5.4:

Table 5.4: Examples of curve fitting RMSE analysis

		Gauss2mf	Gaussmf	Trapmf	Psigmf	Pimf
Weir height	Low	0.01415	0.0490	0.1002	0.0537	0.0590
	Medium	0.03991	0.0674	0.1082	0.0573	0.0778
	High	0.02975	0.0788	0.1130	0.0615	0.0778
Flow rate at 1st spill	low	0.0285	0.0588	0.1133	0.0542	0.0645
	Regular	0.05493	0.0631	0.1175	0.0558	0.0738
	High	0.03652	0.0667	0.1230	0.0714	0.0628
Spill rate at 1 in 5 yr rainfall	Low	0.01013	0.0568	0.1096	0.0636	0.0600
	Regular	0.05464	0.0719	0.1144	0.0670	0.0641
	High	0.04078	0.0725	0.1246	0.0681	0.0703
Total spill duration	Short	0.01964	0.0534	0.1016	0.0593	0.0701
	Medium	0.0295	0.0587	0.1019	0.0619	0.0783
	Long	0.03403	0.0638	0.1089	0.0636	0.0829
Total spill volume	Small	0.02544	0.0588	0.0998	0.0644	0.0778
	Medium	0.0301	0.0619	0.1007	0.0760	0.0789
	Big	0.05185	0.0622	0.1116	0.0783	0.0818

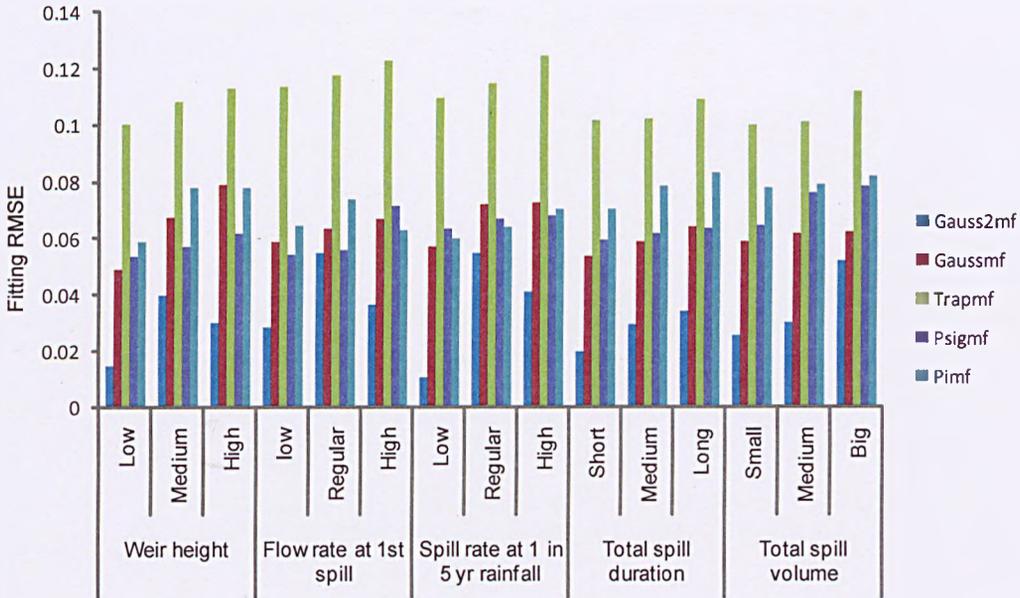
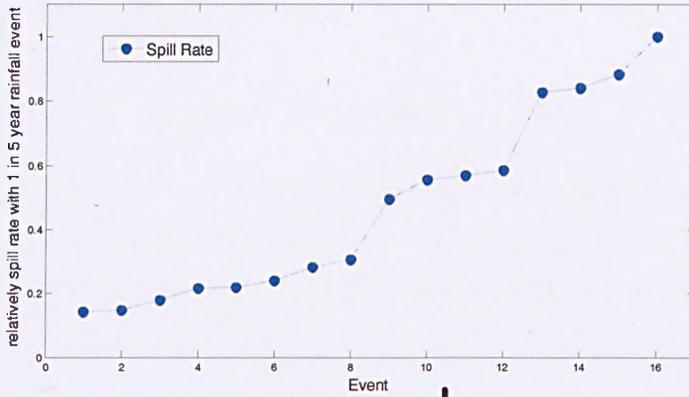


Figure 5.5: Fitting RMSE comparison between different type of membership functions

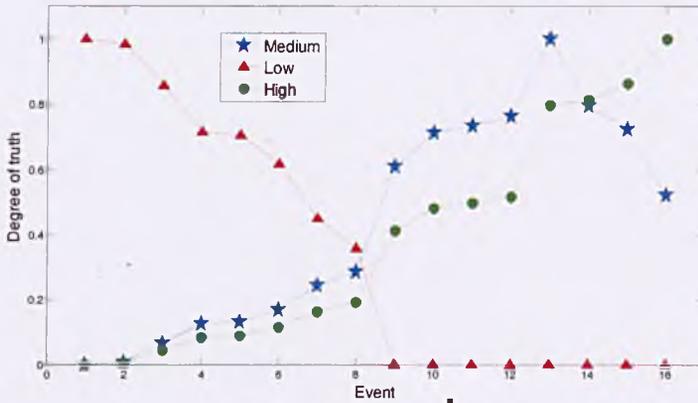
As seen from Figure 5.5, the application of Gaussian combination function provided the best fitting in developing. Therefore, Gauss2mf function, which was built in 'Matlab tool box', was used to develop the membership function by fitting the summarised values of five inputs.

The development process of the model input membership functions are illustrated in Figure 5.6. There are three descriptive categories for each input and output fuzzy set. Step one was to analyse the crisp data set to sort them into "low", "medium" and "high".

Normalised input from original collected data
(sample as spill rate at 1 in 5 year rainfall, original unit is litre per second)



Three classes of inputs were defined based on linguistic labels as : Low, regular and high. Normalised input were calculated in the sense of degree of truth



Functional expression were developed with the application of curve fitting approach. In this example, Gaussian equation was identified as initial expression of membership function

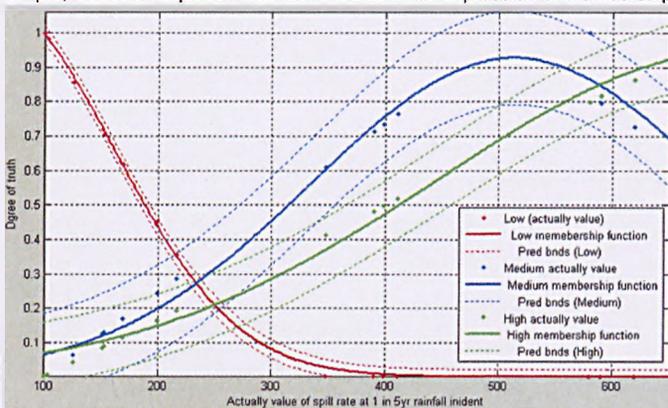


Figure 5.6: Demonstration of model input development

Process of developing five model inputs' membership functions are introduced as following sections.

5.4.1. Input 1: Chamber Weir Height

In general the CSO chambers with screens had CSO chamber dimensions and overall design criteria established using the *CSO design guide* published by WaPUG (WaPUG, 2006). The chamber weir height values of the assets used in the study are shown in Table 5.5:

Table 5.5: Summary of chamber weir height (mm)

ID	Asset Local Name	Weir Height
1	CHANTRY BRIDGE/CSO	370
2	FOULRIDGE/CSO	710
3	BUTCHER TERRACE/CSO	650
4	THE MILL/NO 2 CSO	660
5	SHEAF BANK/CSO	970
6	MAYFIELD GROVE/CSO	460
7	TERRY AVENUE/CSO	950
8	TERRY AVENUE/ NO 2 CSO	780
9	KEARSLEY LANE/CSO	700
10	CARLETON RD SKIPTON/CSO	450
11	BROUGH GOLF COURSE/CSO	400
12	DEARNE HALL ROAD/CSO	890
13	GREEN LANE 125/CSO	345
14	MYTHOLMES LANE/CSO	800
15	SHARLSTON/CSO	650
16	SKELDERGATE BRIDGE/CSO	1020
17	WOODBINE COTTAGE/CSO	600
18	WORTH WAY SUN STREET/CSO	1000
19	WYKE OLD LANE/CSO	317
20	BOROUGH BOUNDARY/CSO	470
21	CHAPEL LANE/NO 2 CSO	430
22	BEIGHTON TIP/CSO	350
23	CANAL ROAD/CSO	750
24	DELVES ROAD/CSO	890

25	HOLLIN DRIVE/CSO	375
26	SKIRLAUGH/CSO	650

Fuzzified into linguistic labels are shown as:

- ✓ $\leq 317(\text{mm})$ Defined as “Low” degree of truth = 1
- ✓ = 630(mm) (Mean value of all records) defined as “Regular”, degree of truth = 1
- ✓ $\geq 1020(\text{mm})$ Defined as “High”, degree of truth = 1

From applying a curve fitting approach with typical Gaussian functions, the three membership functions that categorized the description of chamber weir height were established. The Goodness of fit to the relationship was also established and shown in Table 5.6. Goodness of fit was defined by determining three parameters: sum of square error (SSE), coefficient of determination (R-squared) and root mean square error (RMSE).

Table 5.6: Membership function definition of weir height

Low weir (red curve in Figure 5.7)	
Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.001801
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.999
For $x \geq 317 \text{ mm}$	Adjusted R-square: 0.9988
Coefficients (with 95% confidence bounds):	RMSE: 0.01415
a1 = 1.216 (1.17, 1.263)	F(x) = 1
b1 = 273.1 (269.1, 277.2)	For $x < 317 \text{ mm}$ were considered as
c1 = 164.8 (155.2, 174.3)	“Low” as weir height with truth degree as 1.
Medium weir (blue curve in Figure 5.7)	
Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.01434

$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9885
Coefficients (with 95% confidence bounds):	Adjusted R-square: 0.986
a1 = 0.9246 (0.8721, 0.9772)	RMSE: 0.03991
b1 = 578 (565.6, 590.4)	
c1 = 255 (235.5, 274.5)	
High weir(brown curve in Figure 5.7)	
Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.007968
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9926
For $x \geq 970mm$	Adjusted R-square: 0.9909
Coefficients (with 95% confidence bounds):	RMSE: 0.02975
a1 = 1.05 (1.006, 1.095)	$F(x) = 1$
b1 = 1117 (1073, 1161)	For $x > 1020mm$ were considered
c1 = 617.4 (565.4, 669.3)	as "Low" as weir height with truth
	degree as 1.

Curve fitting for low, medium and high membership functions:

The Figure 5.7 summarizes the curve fitting results which also highlights the 95% confidence band.

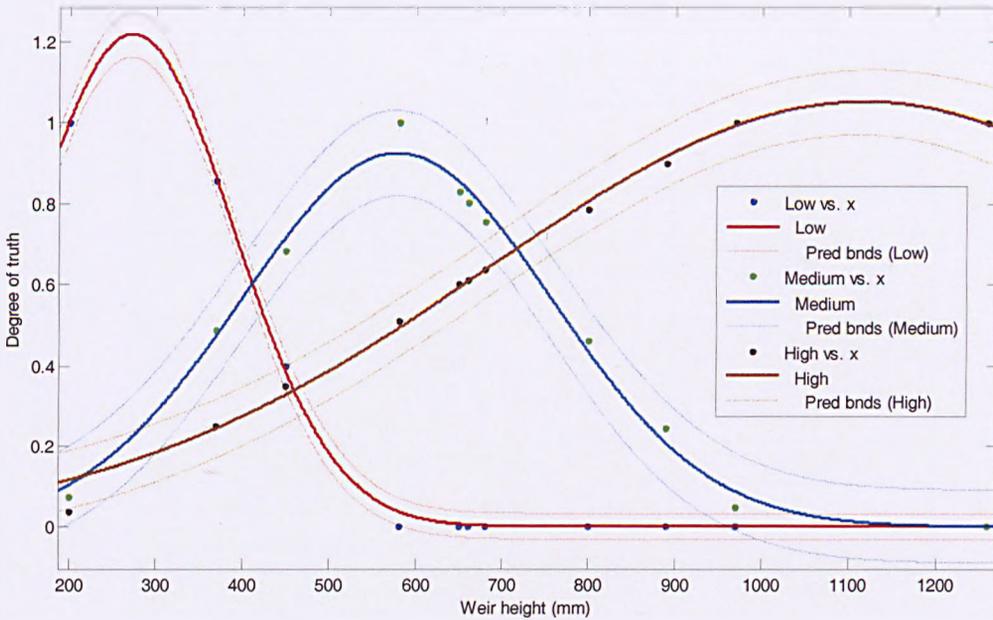


Figure 5.7: Curve fitting result of chamber weir height

In Figure 5.7 the X axis presents the actual value of chamber weir height for each CSO asset, while the Y axis presents the defined degree of truth for each categorised description. The Gaussian distribution curve was the fitted membership function of each description, the “point” with the same colour indicates actual value and the related degree of their linguistic definitions, for example, red curve indicated the membership function of the definition of ‘low weir height’ of CSO asset, which considered the weir height $\leq 317(mm)$ as truth degree of ‘low’ as 1. $317(l/s)$ is the smallest value of weir height in those CSO assets which were used to develop this model. Similar curves are shown for the ‘medium’ (blue) and ‘high’ (brown) weir heights.

5.4.2. Input 2: Flow rate at 1st spill

The flow rate at 1st spill was considered as one of the initial parameters in CSO chamber design, and represents a measure of the CSO setting i.e flow rate to treatment when a spill event first occurs. Due The detail value of each asset's flow rate at 1st spill is shown in Table 5.7:

Table 5.7: Summary of asset flow rate at 1st spill event (l/s)

ID	Asset Local Name	Flow rate at 1 st spill
1	CHANTRY BRIDGE/CSO	82
2	FOULRIDGE/CSO	164
3	BUTCHER TERRACE/CSO	98
4	THE MILL/NO 2 CSO	125
5	SHEAF BANK/CSO	80
6	MAYFIELD GROVE/CSO	138
7	TERRY AVENUE/CSO	131
8	KEARSLEY LANE/CSO	121
9	CARLETON RD SKIPTON/CSO	35
10	BROUGH GOLF COURSE/CSO	75
11	DEARNE HALL ROAD/CSO	142
12	MYTHOLMES LANE/CSO	16
13	SHARLSTON/CSO	125
14	SKELDERGATE BRIDGE/CSO	72
15	WOODBINE COTTAGE/CSO	122
16	WORTH WAY SUN STREET/CSO	156
17	WYKE OLD LANE/CSO	129
18	BOROUGH BOUNDARY/CSO	28
19	CHAPEL LANE/NO 2 CSO	160
20	BEIGHTON TIP/CSO	95
21	CANAL ROAD/CSO	141
22	DELVES ROAD/CSO	105
23	HOLLIN DRIVE/CSO	120

Due to the data availability, 23 out of the 26 CSO assets' flow rate at 1st spill events were collected in Table 5.7. The linguistic definitions are:

- ✓ $\leq 16(l/s)$ Defined as "Low", degree of truth = 1 (in the dataset collected for this model development, value of 16(l/s) is the lowest flow rate at 1st spill event)
- ✓ = 107(l/s) (Mean value of all records) defined as "Medium", degree of truth = 1

✓ $\geq 160(l/s)$ Defined as “High”, degree of truth = 1 (in the dataset collected for this model development, value of 160(l/s) is the highest flow rate at 1st spill event)

Again, applying a curve fitting approach with typical Gaussian functions, the three membership functions categorised descriptions for flow rate at 1st spill were created and these are detailed in Table 5.8:

Table 5.8: Membership function definition of flow rate at 1st spill

Low flow (red curve in Figure 5.8)	
Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.01137
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9948
For $x \geq 16(l/s)$;	Adjusted R-square: 0.994
Coefficients (with 95% confidence bounds):	RMSE: 0.0285
a1 = 0.9855 (0.9022, 1.069)	F(x) = 1
b1 = 6.7 (1.245, 12.15)	For $x < 16(l/s)$ were considered as
c1 = 32.51 (27.43, 37.58)	“Low” of flow rate at 1 st spill with truth degree as 1.
Medium flow (blue curve in Figure 5.8)	
Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.04224
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9688
Coefficients (with 95% confidence bounds):	Adjusted R-square: 0.9644
a1 = 0.9479 (0.864, 1.032)	RMSE: 0.05493
b1 = 98.77 (96.01, 101.5)	
c1 = 54.4 (49.45, 59.35)	
High flow (brown curve in Figure 5.8)	
Curve fitting	Goodness of fit:

General model Gauss1:	SSE: 0.01867
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9893
For $x \leq 160(l/s)$;	Adjusted R-square: 0.9878
Coefficients (with 95% confidence bounds):	RMSE: 0.03652
$a1 = 0.9938$ (0.8926, 1.095)	$F(x) = 1$
$b1 = 195.9$ (174.4, 217.3)	For $x > 160(l/s)$ were considered
$c1 = 116.5$ (100.2, 132.8)	as "High" of flow rate at 1 st spill with
	truth degree as 1.

Curve fitting for CSO first spill data:

Figure 5.8 summarises the curve fitting result for flow rate at the 1st spill of CSOs, with 95% confidence.

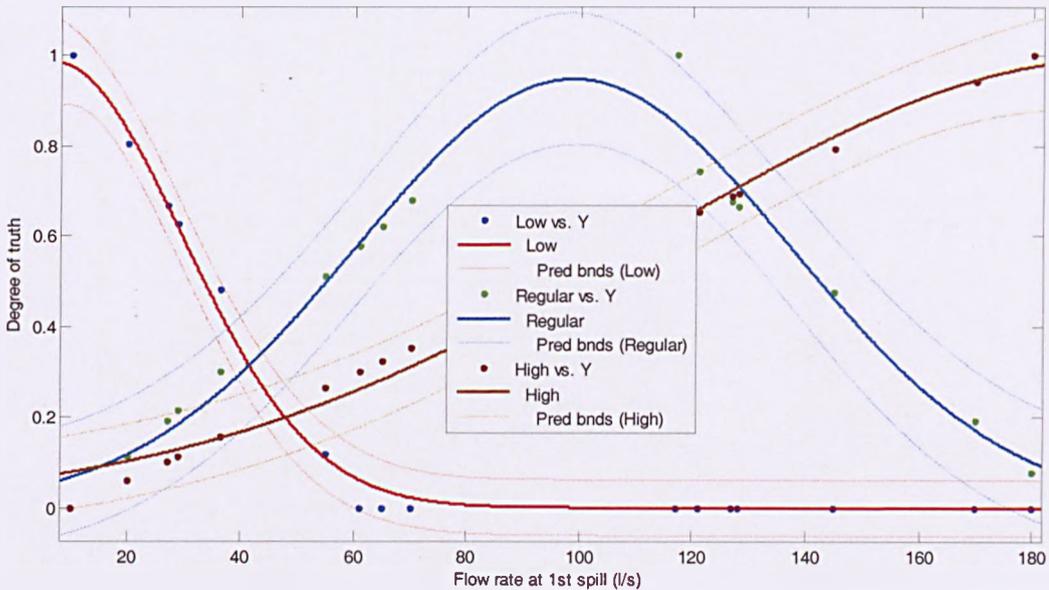


Figure 5.8: Curve fitting result of flow rate at 1st spill

In Figure 5.8 the X axis presents the value of actual flow rate at the 1st spill incident for each CSO asset, and the Y axis presents the defined degree of truth for each categorised description. The blue curve indicates the membership

function of the definition of 'Regular flow rate' at the 1st spill of CSO asset, which considered the flow rate = 107(l/s) as truth degree of 'Regular' as 1. 107(l/s) is the average value of flow rate at 1st spill of all collected CSO assets for this model.

5.4.3. Input 3: spill rate at 1 in 5yr rainfall

Input 3 was concerned with a measure of the design spill rate of CSO assets corresponding to a specific rainfall event falling on the upstream catchment. The spill rate at a 1 in 5 year return period rainfall was used as the second parameter representing the asset design hydraulic parameters. Same as Table 5.7, 23 CSOs' Spill rate values are shown in Table 5.9:

Table 5.9: Summary of asset spill rate at 1 in 5 year rainfall event (l/s)

ID	Asset Local Name	Spill rate at 1 in 5 yr rainfall
1	CHANTRY BRIDGE/CSO	361
2	FOULRIDGE/CSO	722
3	BUTCHER TERRACE/CSO	431
4	THE MILL/NO 2 CSO	550
5	SHEAF BANK/CSO	352
6	MAYFIELD GROVE/CSO	707
7	TERRY AVENUE/CSO	576
8	KEARSLEY LANE/CSO	632
9	CARLETON RD SKIPTON/CSO	254
10	BROUGH GOLF COURSE/CSO	330
11	DEARNE HALL ROAD/CSO	625
12	MYTHOLMES LANE/CSO	100
13	SHARLSTON/CSO	550
14	SKELDERGATE BRIDGE/CSO	317
15	WOODBINE COTTAGE/CSO	537
16	WORTH WAY SUN STREET/CSO	686
17	WYKE OLD LANE/CSO	568
18	BOROUGH BOUNDARY/CSO	619
19	CHAPEL LANE/NO 2 CSO	678
20	BEIGHTON TIP/CSO	518
21	CANAL ROAD/CSO	620

22	DELVES ROAD/CSO	462
23	HOLLIN DRIVE/CSO	528

Fuzzy linguistic classifications were defined:

- ✓ $\leq 100(l/s)$ Defined as “Low”, degree of truth = 1 (in the dataset collected for this model development, value of 100(l/s) is the lowest spill rate under 1 in 5 year rainfall event)
- ✓ $= 510(l/s)$ (Mean value of all records) defined as “Medium”, degree of truth = 1
- ✓ $\geq 710(l/s)$ Defined as “High”, degree of truth = 1 (in the dataset collected for this model development, value of 710(l/s) is the highest spill rate under 1 in 5 year rainfall event)

By using a curve fitting approach with typical Gaussian functions, the three membership functions categorised descriptions of spill rate at 1 in 5yr rainfall were generated together with a measure of the goodness of fit as shown in Table 5.10:

Table 5.10: Membership function definition of spill rate

Low flow rate (red curve in Figure 5.9)	
Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.001333
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9994
For $x \geq 100$;	Adjusted R-square: 0.9994
Coefficients (with 95% confidence bounds):	RMSE: 0.01013
a1 = 1.097 (1.023, 1.172)	$F(x) = 1$
b1 = 51.2 (32.5, 69.91)	For $x < 100(l/s)$ were considered
c1 = 154.5 (139.4, 169.5)	as “Low” of spill rate at 1 in 5 yr
	rainfall with truth degree as 1.
Medium flow rate (blue curve in Figure 5.9)	

Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.03881
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9768
Coefficients (with 95% confidence bounds):	Adjusted R-square: 0.9733
a1 = 0.9269 (0.8602, 0.9936)	RMSE: 0.05464
b1 = 513.2 (497.2, 529.2)	
c1 = 252.8 (227.6, 278)	
High flow rate (brown curve in Figure 5.9)	
Curve fitting	Goodness of fit:
General model Gauss1:	SSE: 0.02162
$F(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$	R-square: 0.9876
For $x \leq 710(l/s)$	Adjusted R-square: 0.9857
Coefficients (with 95% confidence bounds):	RMSE: 0.04078
a1 = 0.9506 (0.8614, 1.04)	F(x) = 1
b1 = 715.3 (639.5, 791)	For $x > 710(l/s)$ were considered
c1 = 378.6 (317.5, 439.8)	as "High" of spill rate at 1 in 5 yr
	rainfall with truth degree as 1.

Curve fitting for spill flow corresponding to 1 in 5 year rainfall event:

The Figure 5.9 summarises the results, again with 95% confidence bands.

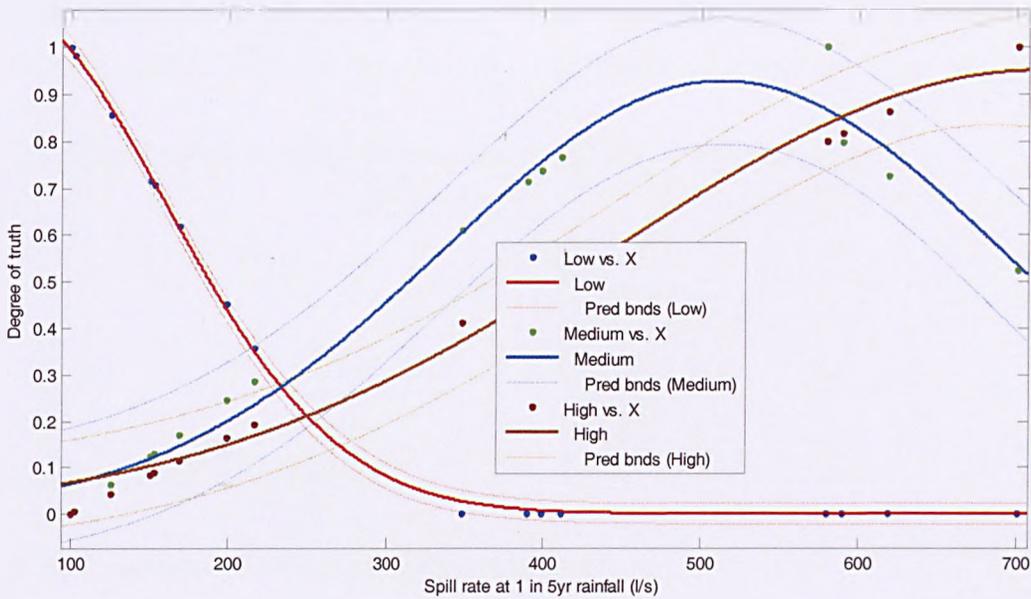


Figure 5.9: Curve fitting result of spill rate at 1 in 5yr rainfall

In Figure 5.9, Brown curve indicated the membership function of the definition of 'High spill rate' at 1 in 5 year rainfall event of CSO asset, which considered the flow rate $\geq 710(l/s)$ as truth degree of 'High' as 1. Flow rate of $710(l/s)$ is the highest value of spill rate at 1 in 5 year rainfall event of those CSO assets which were used to develop this model.

5.4.4. Input 4: Total spill duration over inspection period

Unlike previous fuzzy model inputs, the recorded CSO hydraulic performance parameters proposed in this model were developed based on telemetric performance data (chamber water depth). In order to develop decision support for creating pro-active CSO asset O&M action plans, it was considered sensible to use the period of time (termed the inspection period) immediately prior to the hydraulic performance failure at the time at which incidents were reported. To this end, hydraulic performance failures were summarised using a general aggregation approach over the selected inspection period and applied as model inputs in order to represent asset hydraulic performance statuses in the O&M prediction model.

Wherever water depth is higher than weir height a spill flow was considered to have occurred. Total spill duration was calculated for the period of spill.

The total duration of spill flow was calculated with the equation:

$$T_{totalspill} = \sum_1^k T_{t1spill}$$

Equation 5.2

$T_{totalspill}$: Total spill duration

k: Spill duration of 3 weeks performance record

$T_{t(1-k) spill}$: Duration of each spill event

Total spill duration was calculated by including all individual spill events over the inspection period. In this study the investigation duration of each sample CSO asset was four weeks based on the introduction of Section 3.3. According to data collection mechanism that was adopted, the record of chamber water depth in the three weeks before the recorded pollution incident and for one week following the failure were collected and used in the study. Here in this model development, the total spill duration and subsequently the total spill volume were calculated over the three weeks duration prior to a maintenance incident

A typical 3 week length of recorded water depth is shown in Figure 5.10. Spill flow occurred when the telemetry chamber water depth was greater than the height of the weir. In Figure 5.10, the line in the middle of figure indicated the height of the chamber weir whilst the spill flow was represented by the dashed area.

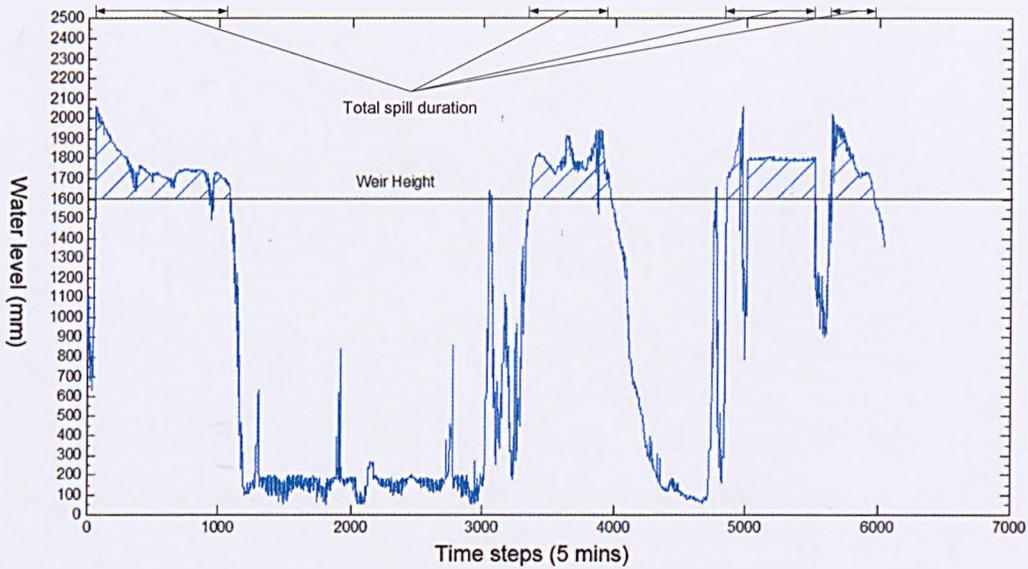


Figure 5.10: Example of spill duration of a CSO chamber water depth performance

Values of each CSO asset's spill duration are shown in Table 5.11:

Table 5.11: Summary of asset spill duration (mins)

ID	Asset Local Name	Spill duration
1	CHANTRY BRIDGE/CSO	2880
2	FOULRIDGE/CSO	2130
3	BUTCHER TERRACE/CSO	4000
4	THE MILL/NO 2 CSO	1760
5	SHEAF BANK/CSO	1275
6	MAYFIELD GROVE/CSO	1370
7	TERRY AVENUE/CSO	1040
8	TERRY AVENUE/ NO 2 CSO	1080
9	KEARSLEY LANE/CSO	630
10	CARLETON RD SKIPTON/CSO	990
11	BROUGH GOLF COURSE/CSO	510
12	DEARNE HALL ROAD/CSO	820
13	GREEN LANE 125/CSO	770
14	MYTHOLMES LANE/CSO	670
15	SHARLSTON/CSO	700

16	SKELDERGATE BRIDGE/CSO	740
17	WOODBINE COTTAGE/CSO	640
18	WORTH WAY SUN STREET/CSO	350
19	WYKE OLD LANE/CSO	510
20	BOROUGH BOUNDARY/CSO	440
21	CHAPEL LANE/NO 2 CSO	330
22	BEIGHTON TIP/CSO	250
23	CANAL ROAD/CSO	220
24	DELVES ROAD/CSO	120
25	HOLLIN DRIVE/CSO	180
26	SKIRLAUGH/CSO	60

In terms of fuzzified linguistic labels, the classes were defined as:

- ✓ $\leq 60 \text{ min}$ Defined as “short” spill duration, degree of truth = 1 (60 mins is the shortest of all collected data)
- ✓ = 1800 *min* (Mean value of all records) defined as “Medium” spill duration, degree of truth = 1
- ✓ $\geq 4000 \text{ min}$ Defined as “long” spill duration, degree of truth = 1 (4000 mins is the longest of all collected data)

Using the familiar Gaussian approaches the three membership functions’ for short, medium and long spill durations are shown in Table 5.12:

Table 5.12: Membership function definition of spill duration

General model for short spill duration (red curve in Figure 5.11):

$$f(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$$

for $x \geq 12$ which presented the time steps with 5 minute intervals

Coefficients (with 95% confidence bounds):

$$a1 = 1.067 \quad (0.957, 1.177)$$

$$b1 = -188.9 \quad (-301.1, -76.6)$$

$$c1 = 661.5 \quad (585.7, 737.4)$$

Goodness of fit:

SSE: 0.006946

R-square: 0.9975

Adjusted R-square: 0.9973

RMSE: 0.01964

$f(x) = 1$

for $x < 12$

Spill duration shorter than 60 minutes (12 time steps) was considered as "short" spill duration with truth degree as 1.

General model for Medium spill duration (blue curve in Figure 5.11):

$f(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$

Coefficients (with 95% confidence bounds):

$a1 = 0.924$ (0.8808, 0.9672)

$b1 = 377.1$ (370.6, 383.5)

$c1 = 239.3$ (230.2, 248.4)

Goodness of fit:

SSE: 0.01566

R-square: 0.9901

Adjusted R-square: 0.989

RMSE: 0.0295

General model Long spill duration (brown curve in Figure 5.11):

$f(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$

for $x \leq 800$ which presented the time steps with 5 minute intervals

Coefficients (with 95% confidence bounds):

$a1 = 1.032$ (0.9143, 1.149)

$b1 = 925.5$ (812.1, 1039)

$c1 = 586.5$ (504.1, 669)

Goodness of fit:

SSE: 0.01737

R-square: 0.9917

Adjusted R-square: 0.9906

RMSE: 0.03403

$$f(x) = 1 \quad \text{for } x > 800$$

Spill duration shorter than 4000 minutes (12 time steps) was considered as "Long" spill duration with truth degree as 1.

Curve fitting results for spill duration:

Figure 5.11 shows the summarised curve fitting results for the short, medium and long spill durations.

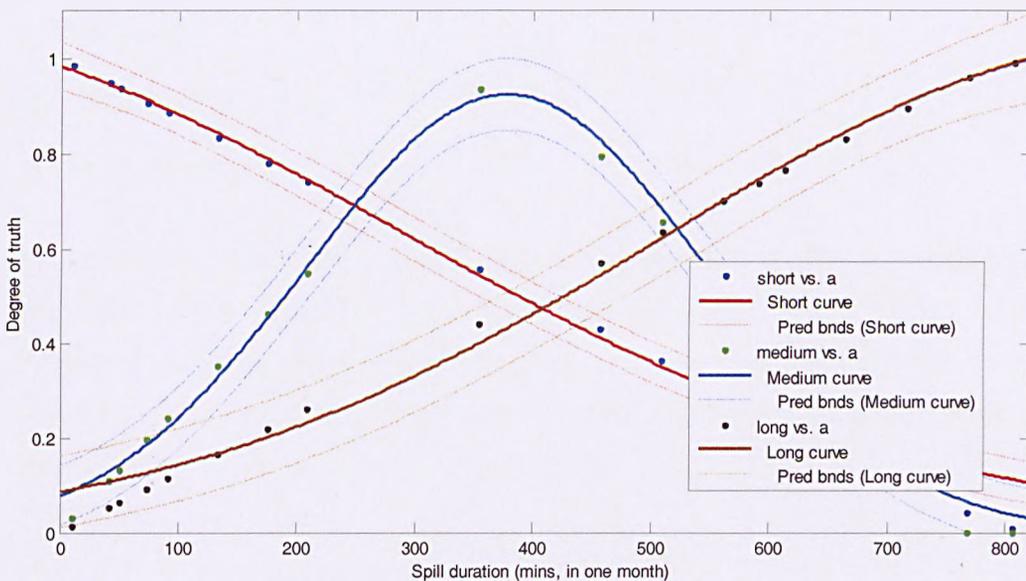


Figure 5.11: Membership function representation of spill duration

These results were used to develop the FL rules as will be detailed in Section 5.6 of the thesis.

5.4.5. Input 5: Total spill volume over the duration of the inspection period

The duration of a spill event indicated both when a spill event occurred and for

how long the overflow continued. Therefore, combined with the information on CSO chamber water depth over the weir at the time of the spill events, the general volume of spill flow can be calculated by using the spill flow equation 5.3:

$$V_{spill} = Q_w \times t = \int C_D^{\frac{2}{3}} \sqrt{2g} L H_w^{\frac{3}{2}} \times t$$

Equation 5.3

V_{spill} : Total spill volume

Q_w : Spill flow rate

C_D : Discharge coefficient for weir

L : Weir length

H_w : Head over the weir

t : Total spill duration

In line with the CSO design guide published by WaPUG (2006), a C_D value of 0.6 was used in this study. The spill volume was calculated as the integral value of individual spill magnitude per unit duration and summing this over the complete duration of the spill event. A trapezoid integral approach was used to calculate the spill flow volume as shown in Figure 5.12:

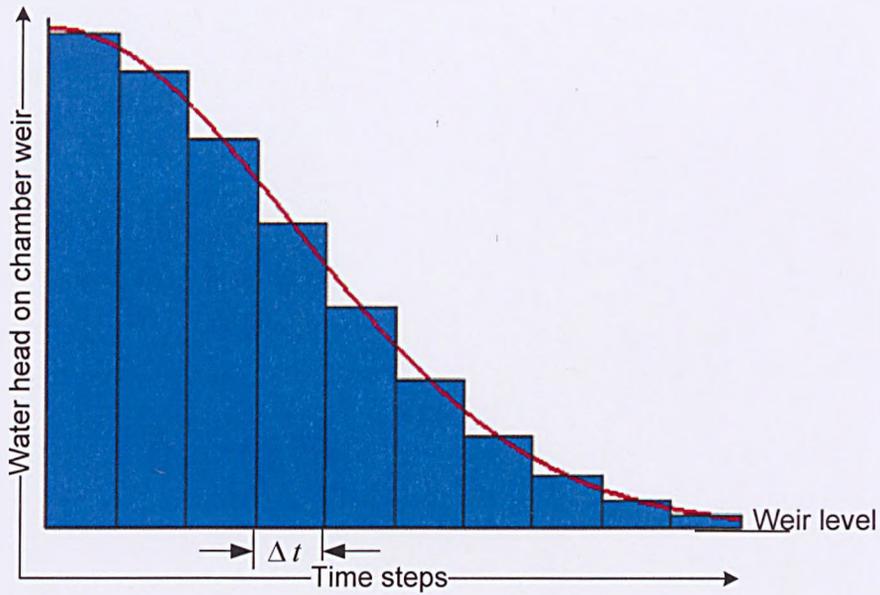


Figure 5.12: Demonstration graph of Trapezoid rules

Using trapezoid rules, the area under the curve was estimated to give the spill flow volume over the duration of the spill event, where:

$$V_{spill} = a \times S_{area}$$

Equation 5.4

$$a = C_D^{\frac{2}{3}} \sqrt{2gL}$$

Equation 5.5

$$S_{area} = H_w^{\frac{3}{2}} \times \Delta t$$

Equation 5.6

V_{spill} : Total spill volume

Q_w : Spill flow rate

C_D : Discharge coefficient for weir

L : Weir length

H_W : Head on the weir

a : Constant

S_{area} : Area under the curve

Δt : Interval width – spill duration

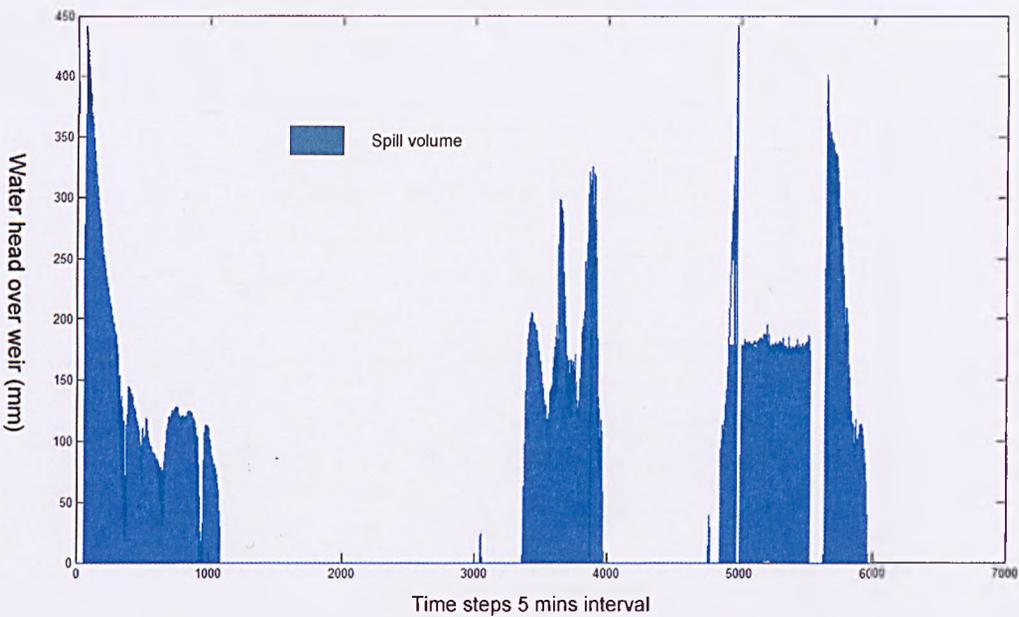


Figure 5.13: Water head on chamber weir

A typical flow distribution is shown in Figure 5.13 and to establish the fuzzified linguistic labels, the classes of small, medium and large spill flow volumes were extracted from the records of spill flow volume over the total record of data. Calculated asset spill volume values are shown in Table 5.13:

Table 5.13: Summary of asset spill volume (m^3)

ID	Asset Local Name	Spill volume
1	CHANTRY BRIDGE/CSO	2057
2	FOULRIDGE/CSO	1521

3	BUTCHER TERRACE/CSO	2299
4	THE MILL/NO 2 CSO	1354
5	SHEAF BANK/CSO	981
6	MAYFIELD GROVE/CSO	1054
7	TERRY AVENUE/CSO	800
8	TERRY AVENUE/ NO 2 CSO	982
9	KEARSLEY LANE/CSO	573
10	CARLETON RD SKIPTON/CSO	900
11	BROUGH GOLF COURSE/CSO	464
12	DEARNE HALL ROAD/CSO	745
13	GREEN LANE 125/CSO	700
14	MYTHOLMES LANE/CSO	609
15	SHARLSTON/CSO	636
16	SKELDERGATE BRIDGE/CSO	673
17	WOODBINE COTTAGE/CSO	582
18	WORTH WAY SUN STREET/CSO	318
19	WYKE OLD LANE/CSO	464
20	BOROUGH BOUNDARY/CSO	400
21	CHAPEL LANE/NO 2 CSO	300
22	BEIGHTON TIP/CSO	208
23	CANAL ROAD/CSO	147
24	DELVES ROAD/CSO	80
25	HOLLIN DRIVE/CSO	120
26	SKIRLAUGH/CSO	5

These were defined as:

- ✓ $\leq 5 \text{ m}^3$ Defined as “Low” spill volume, degree of truth = 1 (5 m^3 is the lowest volume of spill in collected performance)
- ✓ $= 754 \text{ m}^3$ (Mean value of all records) defined as “Medium” spill volume, degree of truth = 1
- ✓ $\geq 2300 \text{ m}^3$ Defined as “High” spill volume, degree of truth = 1 (2300 m^3 is the highest volume of spill in collected performance)

The following membership functions were used to categorise the total spill volume as detailed in Table 5.14:

Table 5.14: Membership function definition of spill volume

Low spill volume (red curve in Figure 5.14)

General model Gauss1:

$$f(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$$

for $x \geq 5$

Coefficients (with 95% confidence bounds):

$$a1 = 1.014 \quad (0.9116, 1.116)$$

$$b1 = -357.3 \quad (-689.1, -25.41)$$

$$c1 = 1757 \quad (1493, 2021)$$

Goodness of fit:

SSE: 0.007119

R-square: 0.9945

Adjusted R-square: 0.9935

RMSE: 0.02544

$$f(x) = 1$$

for $x < 5$

Spill volume smaller than 5 m³ was considered as "Low" total spill volume with truth degree as 1.

Medium spill volume (blue curve in Figure 5.14)

General model Gauss1:

$$f(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$$

Coefficients (with 95% confidence bounds):

$$a1 = 0.9271 \quad (0.7697, 1.085)$$

$$b1 = 989.3 \quad (863.6, 1115)$$

$$c1 = 722.6 \quad (580.4, 864.9)$$

Goodness of fit:

SSE: 0.07139

R-square: 0.933

Adjusted R-square: 0.9139

RMSE: 0.101

High spill volume (brown curve in Figure 5.14)

General model Gauss1:

$$f(x) = a1 * \exp(-((x-b1)/c1)^2)$$

for $x \leq 2300$

Coefficients (with 95% confidence bounds):

$$a1 = 1.002 \quad (0.8095, 1.195)$$

$$b1 = 2546 \quad (1951, 3142)$$

$$c1 = 1611 \quad (1149, 2074)$$

Goodness of fit:

SSE: 0.0242

R-square: 0.9728

Adjusted R-square: 0.9668

RMSE: 0.05185

$$f(x) = 1$$

for $x > 2300$

Spill volume bigger than 2300 m³ was considered as "High" total spill volume with truth degree as 1.

Curve fitting for total spill volume:

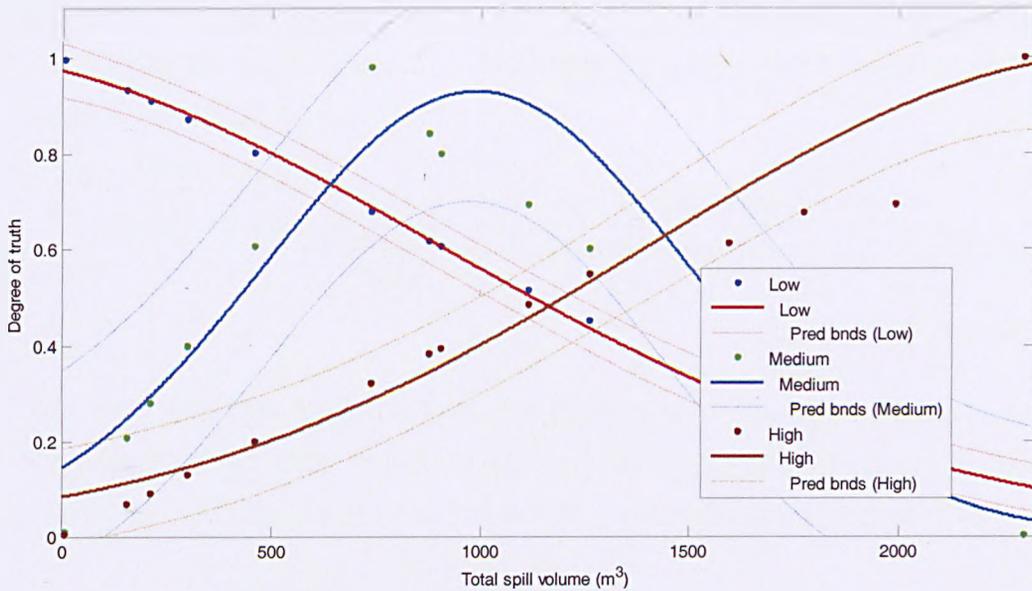


Figure 5.14: Membership function figure of CSO total spill volume

In Figure 5.14, values of 5 m^3 and 2300 m^3 were described the boundary values of the 'lowest' and 'highest' value of spill volume for the CSO assets used in the study

5.5. Output Parameter: Determine O&M action

The recorded number of asset operation and maintenance actions were collated from the data records and summarised with the purpose of developing the outputs to the FL model. As has been the case in previous chapters, all CSO O&M actions were categorised either as routine or responsive actions. However, in terms of responsive actions, three further sub-categories were defined:

- ✓ Site visit in response to a performance alarm
- ✓ Asset inspection to determine the reason of performance alarm (not routine asset inspection)
- ✓ Cleaning actions (such as jet, rod, hand clean)

The model output was designed to provide decision support for O&M action

planning. The proportion of actual cleaning actions in responsive site visits was considered as the parameter to determine the actual requirement of operating actions, shown as Equation 5.7.

$$O\&M\ indicator = \frac{\text{Number of actual clean actions}}{\text{Number of field inspections responded to alarm}}$$

Equation 5.7

The linguistic labels to define the fuzzified classes in Equation 5.7 were based on the data from the asset database and are shown below with the corresponding membership functions shown in Table 5.15 and distributed as Figure 5.15:

- ✓ ≤ 0.1 Defined as “No action needed”, degree of truth = 1 (the value of O&M indicator 0.1 is the smallest in collected information)
- ✓ $= 0.42$ (Mean value of all records) defined as “Inspection needed”, degree of truth = 1
- ✓ ≥ 0.8 Defined as “Action needed”, degree of truth = 1 (the value of O&M indicator 0.8 is the highest in collected information)

Table 5.15: Membership function definition of model output

<p>No action needed (red curve in Figure 5.15)</p> <p>General model Gauss1:</p> $f(x) = a1 \cdot \exp(-((x-b1)/c1)^2)$ <p>for $x \geq 0.1$</p> <p>Coefficients (with 95% confidence bounds):</p> <p>a1 = 18.18 (-178.7, 215.1)</p> <p>b1 = -30.94 (-132.8, 70.92)</p> <p>c1 = 18.28 (-7.79, 44.36)</p> <p>Goodness of fit:</p>

SSE: 0.03046

R-square: 0.9731

Adjusted R-square: 0.9693

RMSE: 0.04665

$$f(x) = 1$$

for $x < 0.1$

Proportion of actual actions in response to performance alarm smaller than 0.1 was considered as "No action needed" decision truth degree is 1.

Inspection needed (blue curve in Figure 5.15)

General model Gauss1:

$$f(x) = a1 * \exp(-((x-b1)/c1)^2)$$

Coefficients (with 95% confidence bounds):

$$a1 = 0.876 (0.824, 0.9281)$$

$$b1 = 8.062 (7.772, 8.353)$$

$$c1 = 5.951 (5.504, 6.398)$$

Goodness of fit:

SSE: 0.03906

R-square: 0.9681

Adjusted R-square: 0.9636

RMSE: 0.05282

Clean action needed (brown curve in Figure 5.15)

General model Gauss1:

$$f(x) = a1 * \exp(-((x-b1)/c1)^2)$$

for $x \leq 0.8$

Coefficients (with 95% confidence bounds):

$$a1 = 1.013 (0.9487, 1.078)$$

$$b1 = 18.85 (17.16, 20.53)$$

$$c1 = 13.31 (11.91, 14.72)$$

Goodness of fit:

SSE: 0.007344

R-square: 0.9947

Adjusted R-square: 0.994

RMSE: 0.0229

$f(x) = 1$

for $x > 0.8$

Proportion of actually actions in responses to performance alarm bigger than 0.8 was considered as "Action needed" decision truth degree is 1.

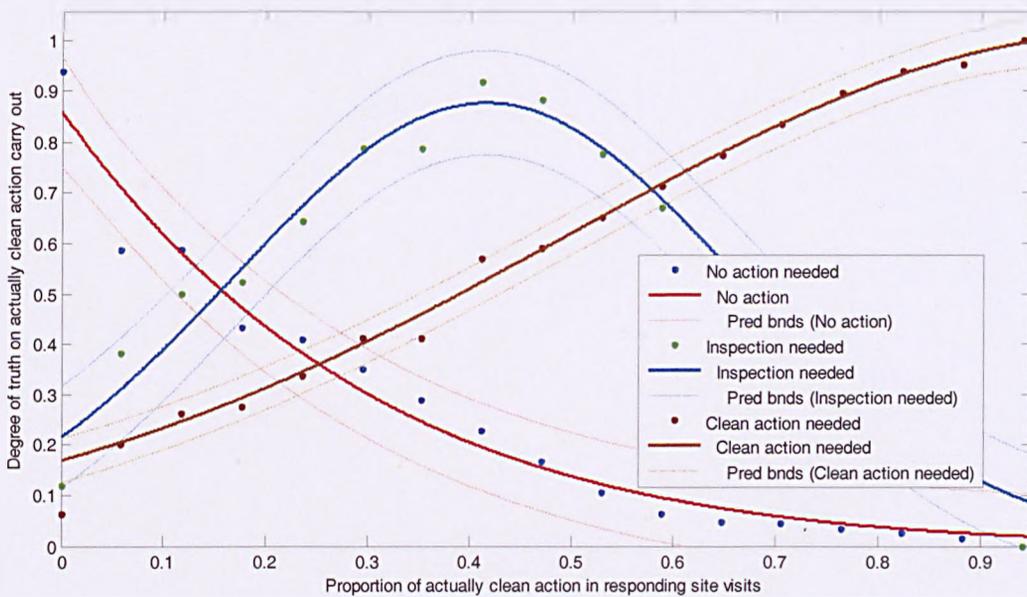


Figure 5.15: Curve fitting result of model output – requirement of CSO O&M action

Membership functions of 'No action needed', 'Inspection needed' and 'clean action needed' are demonstrated. In Figure 5.15 the X axis indicates the proportion of actual cleaning actions that were completed in response to reactive site visits, while the Y axis presents the defined degree of truth for each categorised description.

5.6. Rules of FL system

The model calculation mechanism used to develop the FL expert system was to create FL rules following an "if-and-then" approach. These rules were derived from the collected historical asset performance data. For all five model inputs eighteen model performance rules were defined based on thirty-four CSO chambers which had recorded serious performance failures. The process of system rule development is shown in Figure 5.16:

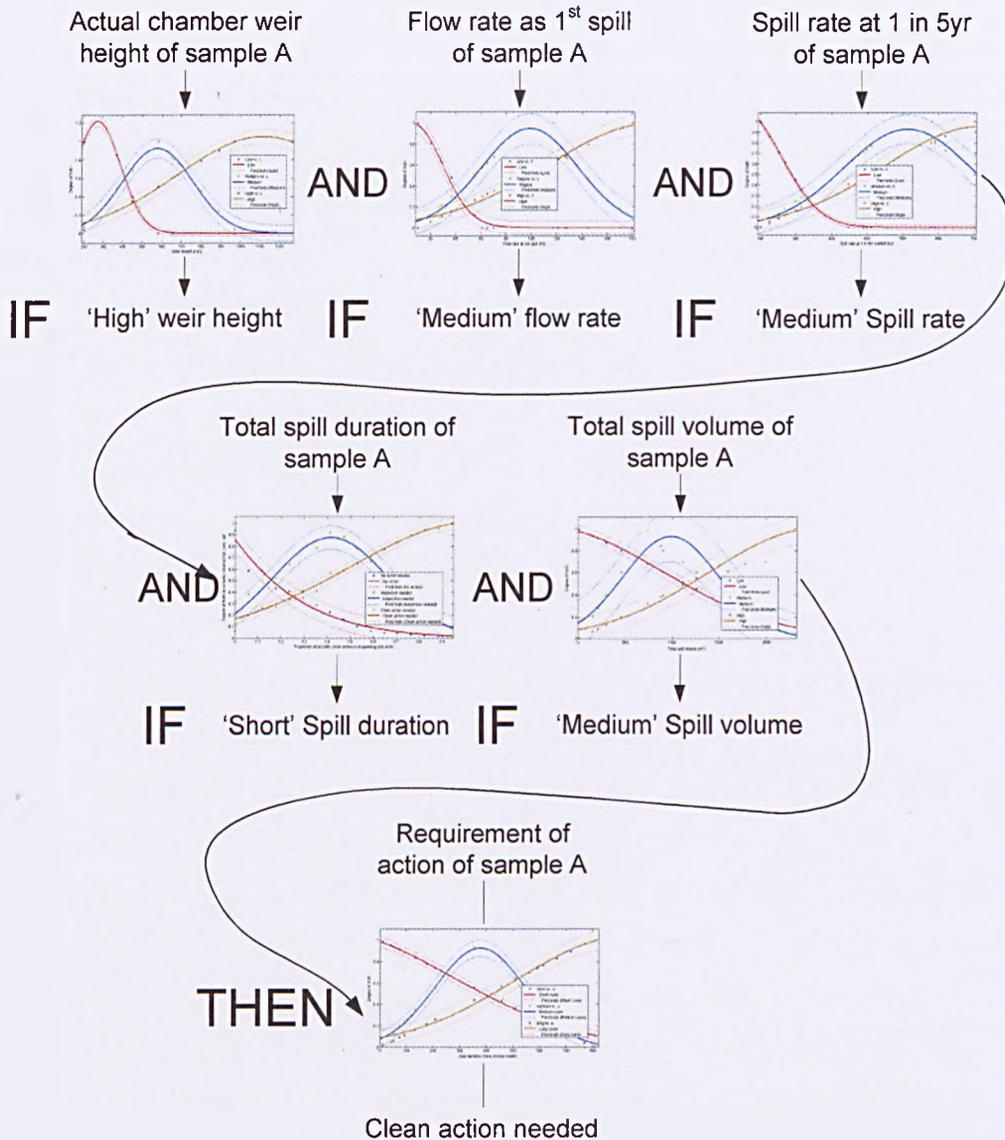


Figure 5.16: Process of system rules development

Based on the process introduced in Figure 5.16, all 'IF, THEN' rules were summarised from a review of all CSO asset information and O&M records. Eighteen rules were developed as the basis of this FL expert system and these are shown in Table 5.16:

Table 5.16: FL system rules

1. if (flow rate at 1 st spill is 'medium') and (spill rate at 1 in 5yr is 'high') and (CSO weir height is 'high') and (spill duration is 'short') and (spill volume is 'medium') then (O&M indicator is 'Clean action needed')
2. if (CSO weir height is 'low') and (spill duration is 'short') and (spill volume is 'low') then (O&M indicator is 'inspection needed')
3. if (CSO weir height is 'high') and (spill duration is 'long') and (spill volume is 'high') then (O&M indicator is 'No action needed')
4. if (CSO weir height is 'medium') and (spill duration is 'short') and (spill volume is 'medium') then (O&M indicator is 'Clean action needed')
5. (CSO weir height is 'medium') and (spill duration is 'short') and (spill volume is 'low') then (O&M indicator is 'Clean action needed')
6. if (CSO weir height is 'low') and (spill duration is 'short') and (spill volume is 'low') then (O&M indicator is 'No action needed')
7. if (CSO weir height is 'medium') and (spill duration is 'short') and (spill volume is 'low') then (O&M indicator is 'No action needed')
8. if (flow rate at 1 st spill is 'medium') and (spill rate at 1 in 5yr is 'medium') and (CSO weir height is 'high') and (spill duration is 'short') and (spill volume is 'medium') then (O&M indicator is 'Clean action needed')
9. if (flow rate at 1 st spill is 'low') and (spill rate at 1 in 5yr is 'medium') and (CSO weir height is 'low') and (spill duration is 'long') and (spill volume is 'high') then (O&M indicator is 'Clean action needed')
10 if (CSO weir height is 'medium') and (spill duration is 'medium') and (spill volume is 'high') then (O&M indicator is 'inspection needed')
11 if (CSO weir height is 'high') and (spill duration is 'long') and (spill volume is 'high') then (O&M indicator is 'inspection needed')
12 if (flow rate at 1 st spill is 'high') and (spill rate at 1 in 5yr is 'low') and (CSO weir height is 'low') then (O&M indicator is 'No actionneeded')
13 if (flow rate at 1 st spill is 'medium') and (spill rate at 1 in 5yr is 'low') and (CSO weir height is 'low') then (O&M indicator is 'inspection needed')

14 if (flow rate at 1 st spill is 'low') and (spill rate at 1 in 5yr is 'low') and (CSO weir height is 'low') then (O&M indicator is 'Clean action needed')
15 if (flow rate at 1 st spill is 'low') and (spill rate at 1 in 5yr is 'medium') and (CSO weir height is low) then (O&M indicator is not 'Clean action needed')
16 if (flow rate at 1 st spill is 'medium') and (spill rate at 1 in 5yr is 'medium') and (CSO weir height is 'high') then (O&M indicator is 'inspection needed')
17 if (flow rate at 1 st spill is 'medium') and (spill rate at 1 in 5yr is 'medium') and (CSO weir height is 'medium') then (O&M indicator is not 'Clean action needed')
18 if (flow rate at 1 st spill is 'medium') and (spill rate at 1 in 5yr is 'high') and (CSO weir height is 'high') then (O&M indicator is 'No action needed')

All fuzzified linguistic labels were highlighted with different colours. Rule viewers of the 18 'If-and-then' rules were defined in the FL model - Figure 5.17:

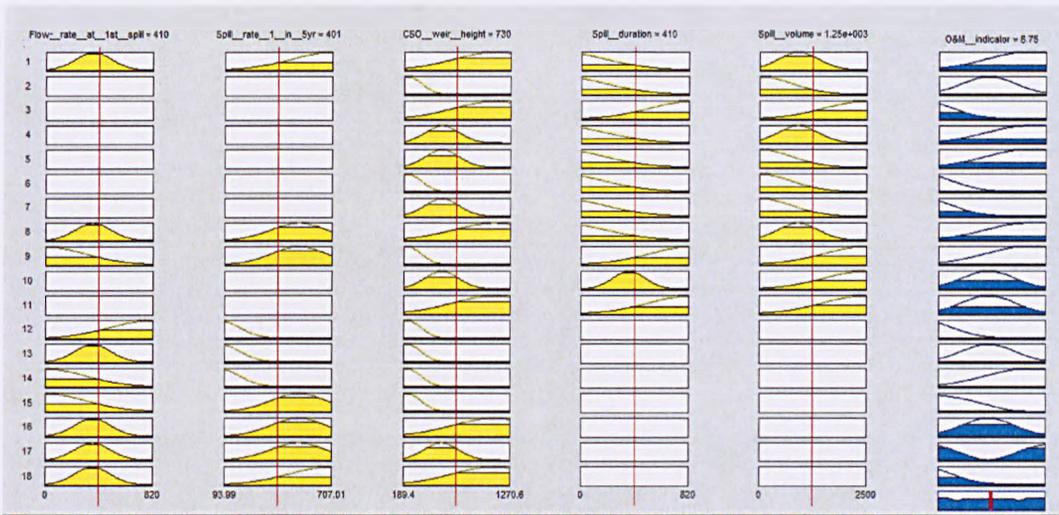


Figure 5.17: Graphic expression of defined rules

Figure 5.17 presents a graphical view of FL model rules which were used by the FL expert system as an initial calculation mechanism. According to the methodology of the FL approach, more expert system rules could be defined if the sample base was expanded, as discussed in section 2.3.

5.7. Model Surface View

Based on the model input and output parameters described in Sections 5.4 and 5.5 and the 18 FL expert system rules in Section 5.6, it has been feasible to develop 3D model surface views of the outcomes of the model. These were generated by applying the MATLAB FL tool box. The methodology has been applied to the data on side weir chamber CSO assets with mechanical screens. Five input parameters (introduced in section 5.4) were demonstrated through 3D model surface viewer:

One parameter was studied as the output during the model building – requirement of asset operation and maintenance. The requirement of O&M was identified by the value of O&M indicator - proportion of actual clean actions in responsive field inspection in Equation 5.7.

Model surface views are the expression of the result that produced by FL model based on defined inputs, output and system rules, which indicated a completed FL expert system was built. Testing of this FL system will be introduced in Section 5.8 – ‘model testing’. Description of five versions of model surface view samples is demonstrated from Figure 5.18 to Figure 5.22. Because of the dimensional limitation, only two input parameters can be indicated in each figure:

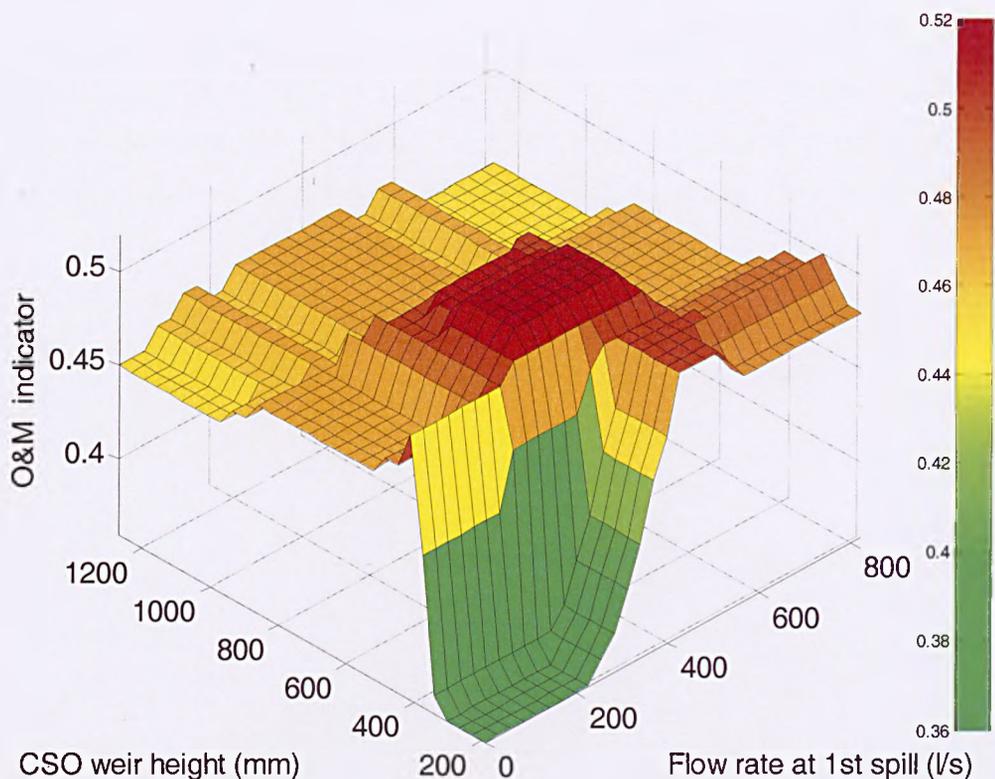


Figure 5.18: Model viewer - Weir height Vs. Flow rate at 1st spill

Figure 5.18 indicated the relationship between O&M indicator (requirement of O&M which considered as the output parameter) and two input parameters: Weir height and Flow rate at 1st spill. Colour green indicated safe – ‘no action needed’, colour yellow gradually to red indicated more actions were needed. The level of asset O&M requirement was defined from model output parameter’s membership function introduced in Figure 5.15. Read from Figure 5.18, the FL system considered that, an asset with low weir height (less than 300mm, here the low is different from the linguistic definition ‘Low weir height’ in Section 5.4.1) appeared to need more actual actions (inspection or clean action) with the flow rate increasing, especially, those asset with these input parameters with the value located in the red surface of Figure 5.18. According to this FL system, the decision of the O&M requirement (action needed or not, what action was needed)

of a CSO asset cannot be made based on one single model surface view, all other three input parameters which were not included in Figure 5.18 should also be evaluated. Similarly, from Figure 5.19 to Figure 5.22 presented other four samples of model surfaces which included different groups of input parameters, and the relationship to the only concerning output – asset O&M indicator.

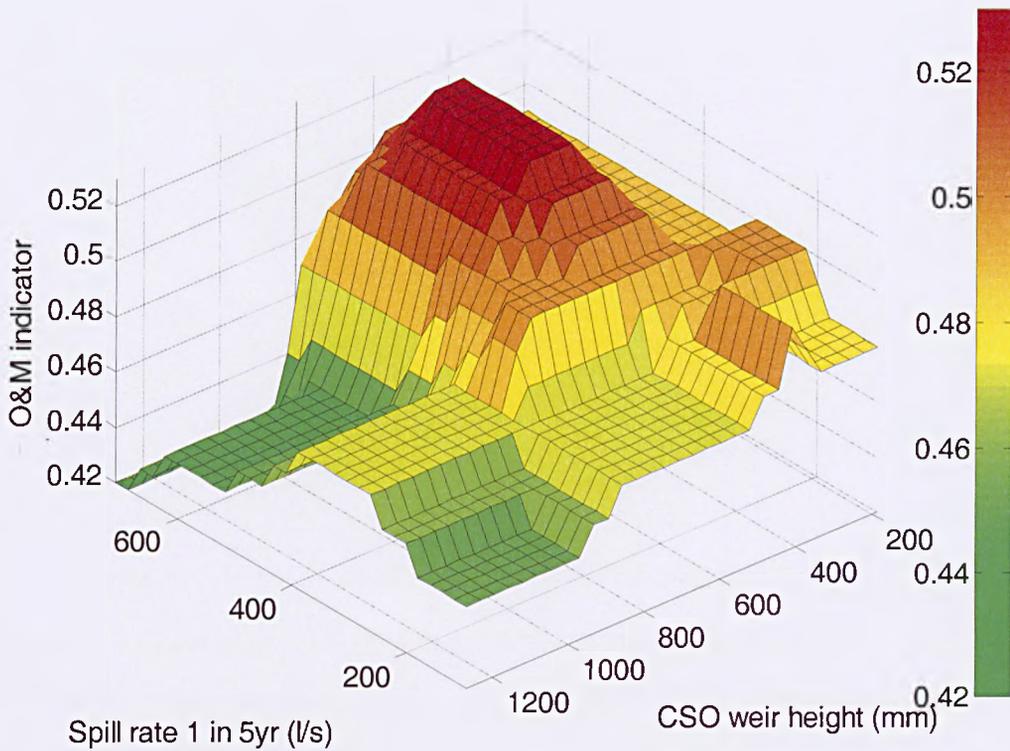


Figure 5.19: Model viewer - Weir height Vs. Spill rate at 1 in 5 year rainfall

Figure 5.19 indicated that, if a performance alarm was reported on a CSO asset with weir height lower than 800mm and spill rate at 1 in 5 year rainfall higher than 400 l/s, the chance of 'clean action needed' is very high. Accurately decision can be made with considering other three model input parameters together.

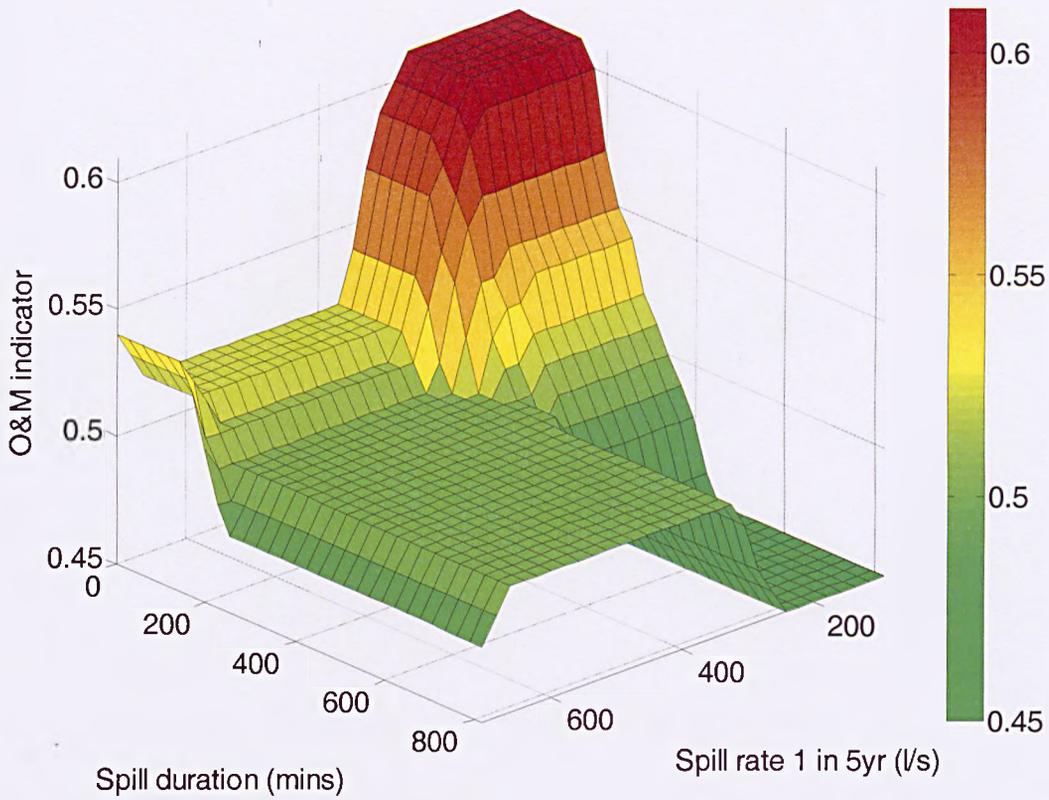


Figure 5.20: Model viewer – Spill rate at 1 in 5 year rainfall Vs. Spill duration

Figure 5.20 indicates the O&M requirement when considering the impact from both overflow spill duration and spill flow rate using a 1 in 5 year rainfall event.

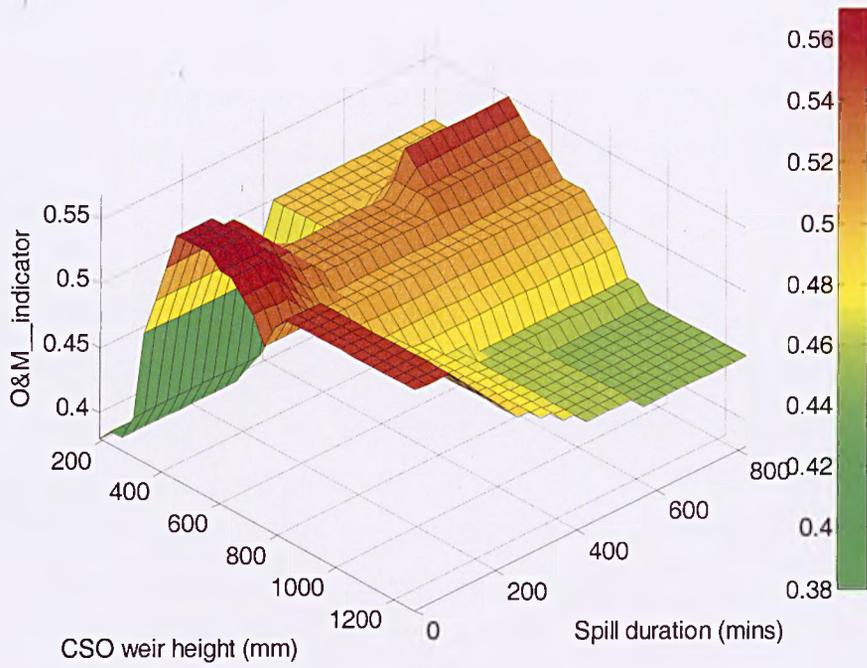


Figure 5.21: Model viewer – CSO weir height Vs. Spill duration

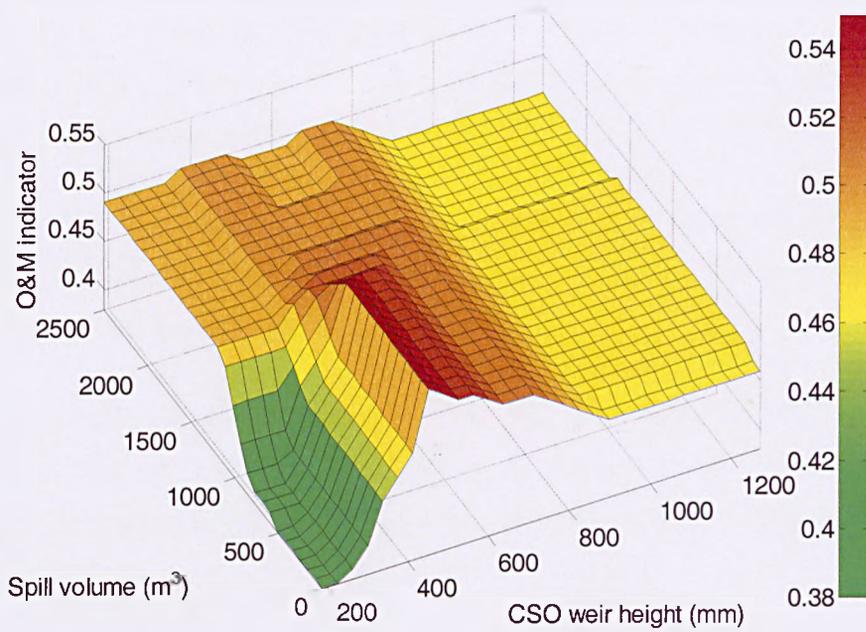


Figure 5.22: Model viewer - Spill volume Vs. weir height

Decision of asset O&M requirement cannot be made according to any single model surface figure, as only two inputs parameters were presented in 3D views. Each model surface view can only presented the relationship between two input parameters and the requirement of O&M actions which was delivered by this FL export system.

5.8. Model Testing

A sensitivity analysis was performed in this FL model for the FL operator AND, and for the methods of implication, aggregation and defuzzification. The results of changing a single operator or method while the rest of the model was held constant were compared with the results from the baseline model. 18 new selected CSO asset together with all required information were used in this sensitivity test. The results were evaluated on the basis of correct linguistic matches. As mentioned above, all data used were based on side- weir chamber with mechanical screen assets.

The testing process follow the introduction of FL expert system development, each of 18 new selected CSO assets were evaluated by applying every input parameters into the membership functions defined in Section 5.4 and follow the process demonstrated in Figure 5.16:

The 'IF-THEN' rules output – predicted requirement of O&M will be compared with actual asset O&M records with the purpose of testing the efficiency of this FL export system on decision making support.

Based on this sensitivity analysis, the AND operator 'minimum' and the implication method 'minimum' were found to perform better than the product method. This supports the use of these operators for independent data input.

Table 5.17: Test outputs of different operators

<u>Minimum Operator</u>			
	Linguistic Matches		
	No action needed	Inspection needed	Cleaning action needed
Actual Number of action	7	3	8
FL model output indication	6	4	8

<u>Product Operator</u>			
	Linguistic Matches		
	No action needed	Inspection needed	Cleaning action needed
Actual Number of action	7	3	8
FL model output indication	4	8	6

Other than the logical calculation mechanical modification, the model's defuzzification approach which introduced in section 5.1 was also considered as an initial parameter which delivered significant influence on model output prediction. In Table 5.17, the comparison of the outputs generated by two defuzzification methods (Bisector and Centroid) is presented:

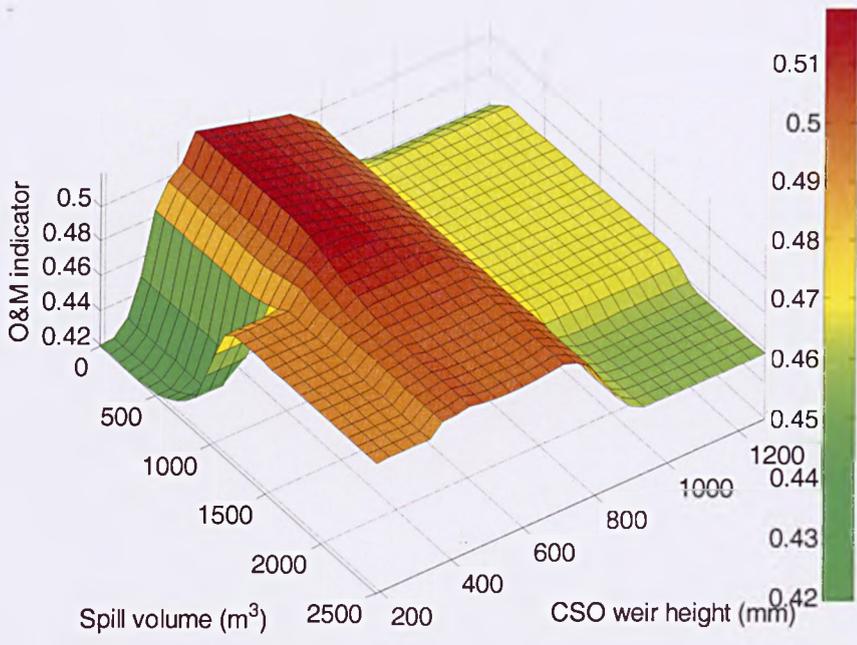
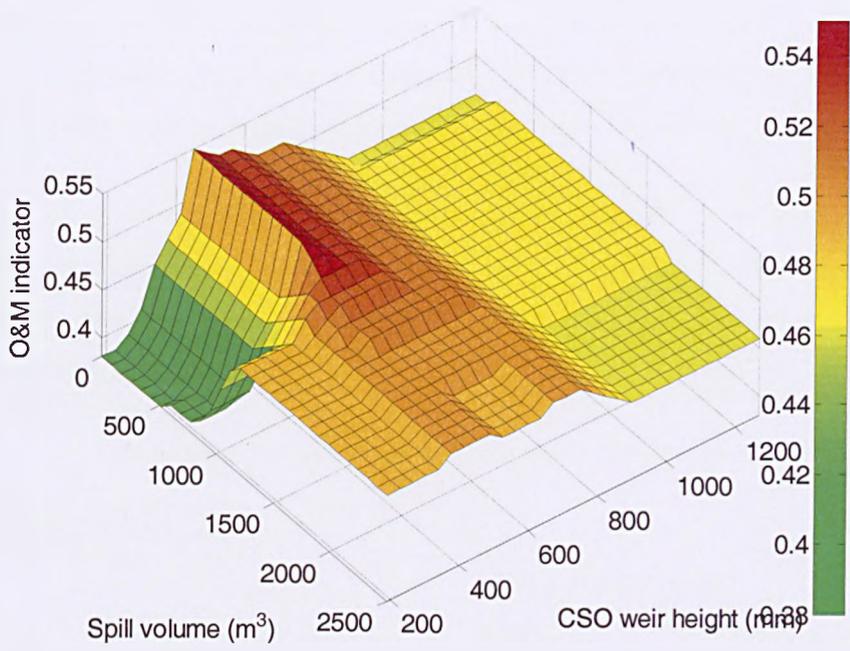


Figure 5.23: Different defuzzification outputs comparison

In Figure 5.23, the figure A presented the output surface by using bisector defuzzification method and figure B is using Centroid approach, yellow-red area can be noticed that the predicted accuracy difference between two defuzzifying algorithms. Bisector approach appeared to be more veracious

Figure 5.23, displayed the comparison between outputs developed by applying different defuzzification methods. The bisector approach appeared to be more sensitive in the graphical view than the output given by Centroid defuzzification. The comparison results are shown in Table 5.18below:

Table 5.18: Test outputs of different Defuzzification approach

<u>Centroid Defuzzification</u>			
	Linguistic Matches		
	No action needed	Inspection needed	Cleaning action needed
Actual Number of action	7	3	8
FL model output indication	7	5	6

<u>Bisector Defuzzification</u>			
	Linguistic Matches		
	No action needed	Inspection needed	Cleaning action needed
Actual Number of action	7	3	8
FL model output indication	6	4	8

Based on this, a prototype model configuration was developed: using minimum for the AND operator; product for the implication; and maximum for the aggregation method. The model results were most sensitive to the method of defuzzification. The Centroid and Bisector methods produced better results than

the smallest, median, and largest of maxima methods. The summary of all 18 CSO asset tested result is indicated in Section 5.6.

5.8.1. Comparison of different screen devices

The previous FL decision supporting model was developed based on the CSO asset designed as a side weir chamber. In order to use the CSO structural features as a model input parameter, the model surfaces were compared between CSOs with different screen devices installed and similar structures without screens.

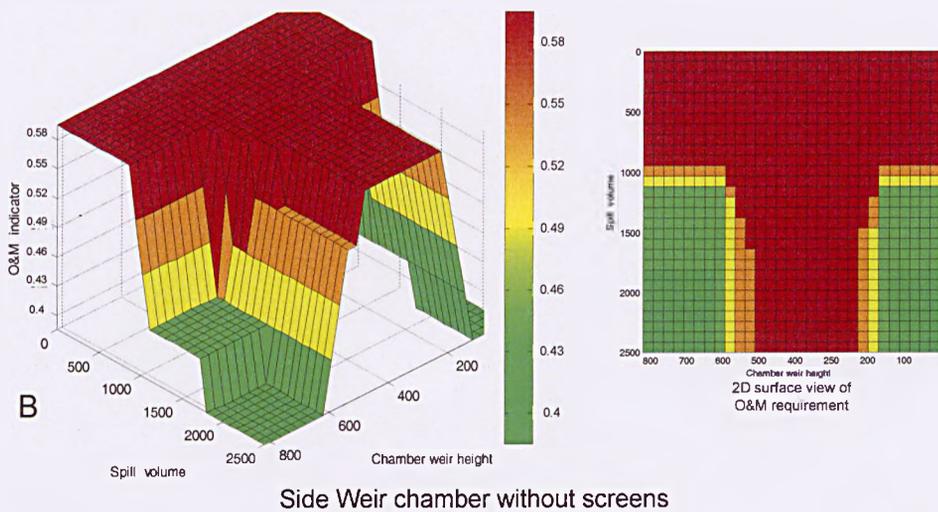
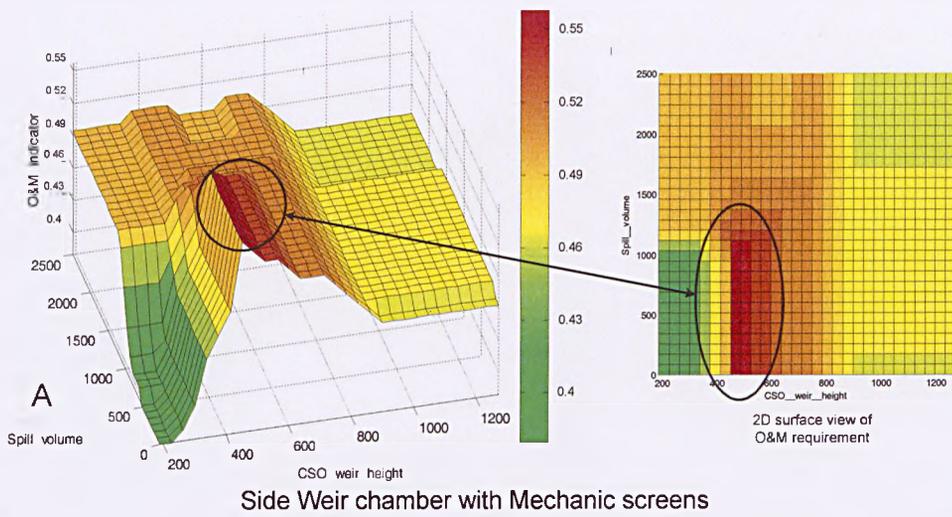
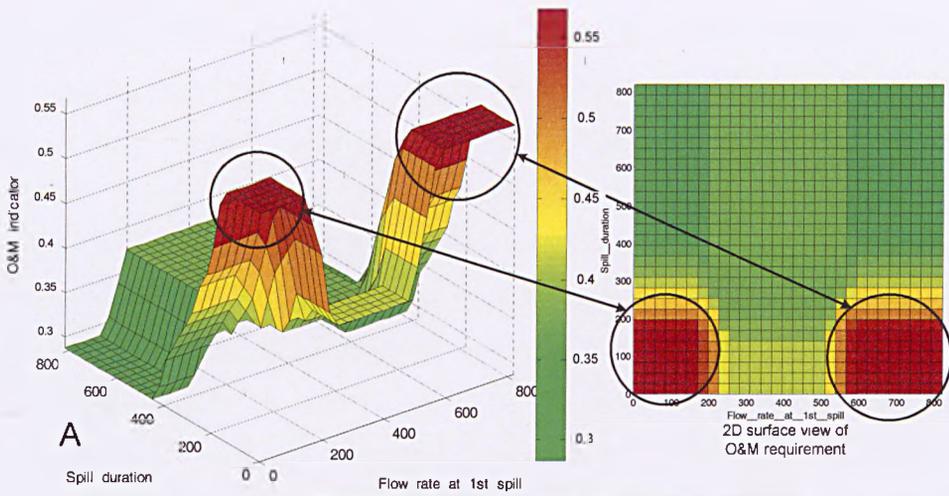
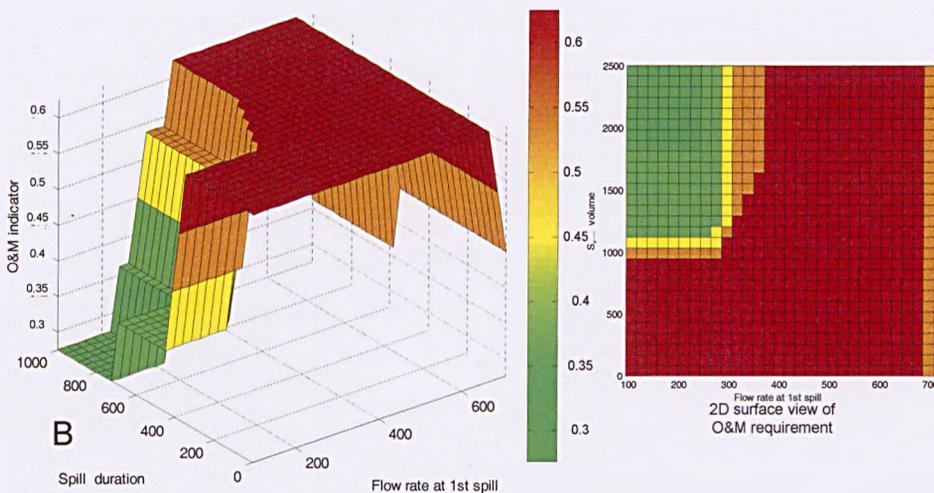


Figure 5.24: Output viewer of CSO with mechanic screen and non-screen example 1

In Figure 5.24, figure a – on the left is the 3D surface of model output for CSO asset with mechanic screens (spill volume and CSO weir height as inputs), on the right is the 2D view indicated the requirement of actual maintenance actions. Figure B on the left is the 3D surface of model output for CSO asset with mechanic screens (spill volume and CSO weir height as inputs), on the right is the 2D view indicated the requirement of actual maintenance actions.



Side Weir chamber with Mechanic screens



Side Weir chamber without screens

Figure 5.25: Output viewer of CSO with mechanic screen and non-screen example 2

In Figure 5.25, in figure A, on the left is the 3D surface of model output for CSO asset with mechanic screens (spill duration and flow rate as 1st spill as inputs), on the right is the 2D view indicated the requirement of actual maintenance actions. In figure B, on the left is the 3D surface of model output for CSO asset with mechanic screens (spill volume and CSO weir height as inputs), on the right is the 2D view indicated the requirement of actual maintenance actions.

In Figure 5.24 and Figure 5.25, 3D model outputs surfaces indicate the requirement of actual maintenance actions, shown as a red colour. To the right, a 2D view is provided to demonstrate the chance of need actual maintenance actions in general for predicted potential asset performance failures. In Figure 5.24 A, using total spill volume and CSO chamber weir height as inputs parameters, the chance of need actual clean action for the CSO with mechanical screen installed were 2.89% in general predicted potential performance failures. The chance of need actually clean action for CSO asset without screen installed, indicated in figure B, was 55.56%. Similarly, in Figure 5.24, using total spill duration and flow rate at 1st spill event as inputs parameters, the chance of need actual clean actions required for CSOs with mechanical screens installed is 11.7% in general predicted potential performance failures, whilst the chance of CSO asset without screens installed, indicated in figure B, was 76.2%.

Direct observation of both 2D output views in Figure 5.24 and Figure 5.25, showed that a greater 'red' colour area indicated that, under the same upstream hydraulic conditions and similar rainfall inflows, CSO assets with mechanical screens installed appeared to be more reliable in performance and fewer actual maintenance actions were required compared with those CSO assets without screen devices.

Static screens were not considered in developing this FL O&M prediction model, as very few CSO assets were shown to have one installed in the preparatory work for this research.

5.9. FL model Summary

In order to try and both predict CSO hydraulic performance failures and to provide pro-active O&M action decision support, a FL approach was adopted and the model building process was detailed in this section.

The FL CSO O&M prediction model was created with the intention of providing support for decisions concerning pro-active O&M actions based on predicted

CSO hydraulic performance failure alarms. Three general types of parameter with seven sub-features were considered as initial influences to the requirement of asset operation and maintenance actions.

The model calculation mechanism of FL expert system used FL rules following an "if-and-then" approach, which were summarized from collected asset practical operation performances. With all seven model inputs, 18 general model performance rules were defined based on 34 CSO samples which had recorded serious performance failures.

In terms of CSO asset structural design this research focused on side-weir chambers both with mechanical screens and without for its comparison. This highlighted that the chance of the need for actual CSO clean operations for CSO asset without screen installed when there is a performance alarm reported were higher than those CSO assets with mechanical screen installed.

To verify the efficiency of this decision support model, another eighteen CSO assets (of the same asset design) were used for individual prediction tests in this project. It can be seen that there was a high correlation coefficient between the actual and the predicted maintenance requirements. Such approaches were used to confirm the validity of the FL approach. A summary of the model prediction outputs are shown in Figure 5.26:

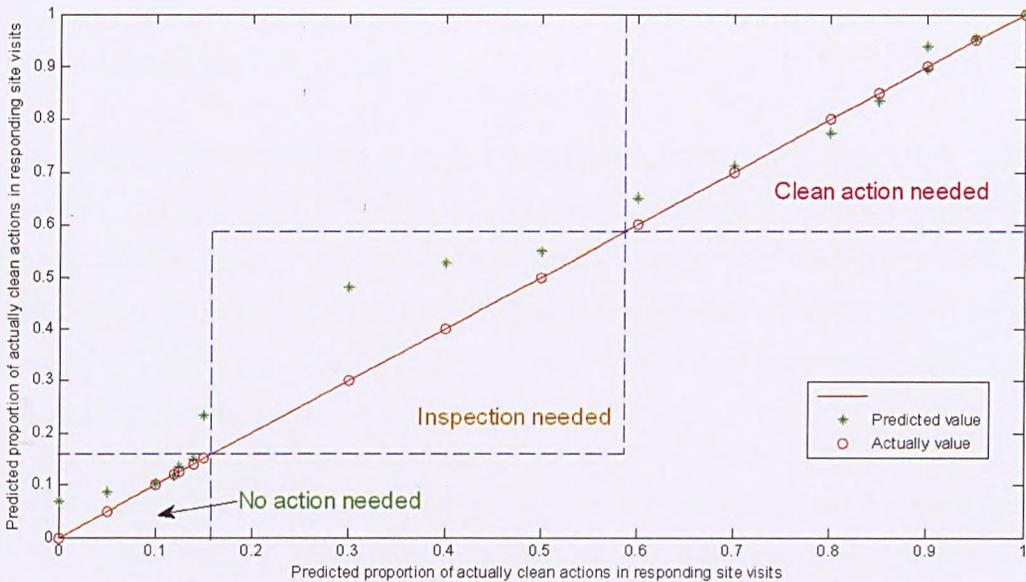


Figure 5.26: Prediction accuracy of FL CSO O&M model

In Figure 5.26, both decision provided by FL model and recorded actions are presented. Actions are divided into three phases: 'No action needed', 'Inspection needed' and 'Clean action needed'. Figure 5.26 is a summary of the outputs presented in Table 5.18, only one 'No needed' decision is miss-provided as 'Inspection needed'.

According to the membership function of the model output definitions in Figure 5.15, the categories of 'no action needed', 'Inspection needed' and 'Clean action needed' were identified in the figure above. As indicated above, there was a "No action needed" range requirement inaccurately forecasted as "Inspection needed", but all "Clean action needed" ranges were captured successfully and with a high correlation between actual cleaning events and predicted need for clean action accurately.

5.10. Conclusion:

After having developed the off line neural network hydraulic performance prediction model, the predicted performance failures (potential spill risks) were

used in a forecasting alarm, which was the initial goal of creating the CSO O&M action decision support model.

The ability to predict performance failure combined with a real-time alarm system provides the opportunity to prepare 'pro-active' operation actions rather than simply conventional 'reactive' maintenances, which are realised to be less efficient for practical services. Using the FL prediction model, better decision making for pro-active operation actions can help to avoid potential performance failures and further negative influences.

The FL prediction model had the further benefit of being useful in evaluating current O&M routine efficiency. This can be achieved by using the model to determine which CSO chamber types performed best with which screen devices to provide better services and lower maintenance requirements.

Combined with the previously hydraulic performance prediction model, the pro-active operation strategies can be developed to improve the serviceability of CSO assets and reduce operational expenditures.

A case study detailing how these two CSO asset management models are merged together will be introduced in the next chapter.

6. Practical application

6.1. Case Study Introduction

With the purpose of testing the CSO hydraulic performance off line prediction model and pro-active O&M decision support model, 20 further CSO assets were drawn from the water company's database to be assessed in terms of their structure, hydraulic performance and recorded O&M actions. This section details a case study intended to develop a more comprehensive approach to CSO asset management through the application of the previously introduced models.

The hydraulic performance records (chamber water depth value) and rainfall data were collected following the same process described in Chapter 3. The chamber and screen features of all 20 CSO assets were specified in previously introduced CSO asset database (see attached DVD). The asset names and their design information are shown in Table 6.1

Table 6.1: 20 CSO asset list for case study

ID	Asset Local Name	Chamber Type	Screen Type
1	BONDGATE CSO	SHARPE & KIRKBRIDE	Static
2	BROADFIELD ROAD CSO	STILLING POND	Mechanical
3	BRUNSMERE SCHOOL CSO	SHARPE & KIRKBRIDE	Mechanical
4	BULLS HEAD CSO	SINGLE SIDED HIGH WEIR	Static
5	BURLEY LODGE CSO	DOUBLE SIDED HIGH WEIR	Static
6	CARR FORGE ROAD CSO	SINGLE SIDED LOW WEIR	Static
7	COWBAR CSO	SINGLE SIDED LOW WEIR	Mechanical
8	CRESCENT TERRACE CSO	SINGLE SIDED LOW WEIR	Mechanical
9	DAYLANDS AVENUE NO.2 CSO	SINGLE SIDED LOW WEIR	Mechanical
10	EBOR WAY CSO	STILLING POND	Static
11	GILSTEAD LANE 130 2 CSO	DOUBLE SIDED LOW WEIR	Mechanical
12	HANGINGWATER ROAD CSO	STILLING POND	Static
13	HILDERTHORPE ROAD CSO	DOUBLE SIDED LOW WEIR	Mechanical
14	HOUGH SIDE WORKS CSO	SINGLE SIDED LOW WEIR	Static
15	LIMEKILN LANE CSO	DOUBLE SIDED LOW WEIR	Mechanical
16	SHIBDEN PARK CSO	STILLING POND	Static
17	SPA MILLS BRIDGE ST CSO	VORTEX (CENTRAL)	Static

18	TADCASTER EAST CSO	DOUBLE SIDED LOW WEIR	Mechanical
19	THIRSK FINKLE STREET NO 2 CSO	SINGLE SIDED LOW WEIR	Mechanical
20	WATH DONCASTER ROAD CSO	SINGLE SIDED HIGH WEIR	Static

This case study will make use of two models previously discussed:

1. The CSO hydraulic performance model (detailed in Chapter 4) was used to predict chamber water depth value based on rainfall data. The aim is to predict the CSO spill events and identify the potential performance failure for each CSO asset.
2. The CSO pro-active operation and maintenance decision support model (developed in Chapter 5) was also applied. The objective of applying this model is to provide local knowledge specific to an individual CSO and show how its application will lead to immediate benefits with improved proactive maintenance strategies and reduced costs.

This case study will demonstrate a new comprehensive asset management approach, which includes hydraulic performance prediction of spill event and pro-active operation and maintenance decision support.

6.2. Asset Management Approach

As introduced in section 6.1, all information concerning the 20 CSO assets which were used in this case study was collected following the process introduced in chapter 3. The framework of asset management approach is shown in Figure 6.1:

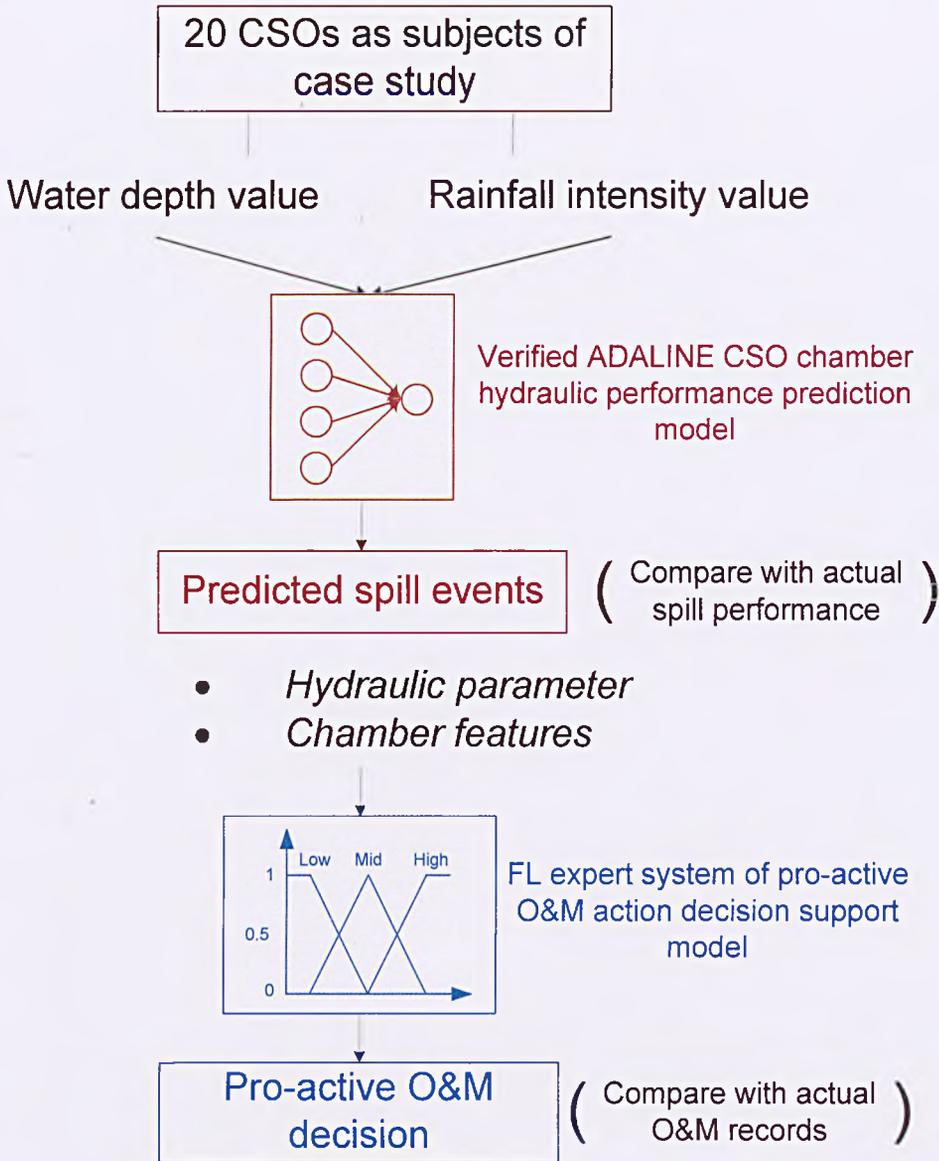


Figure 6.1: Asset management case study frameworks

As the model building process introduced in chapter 4, four weeks' CSO chamber water depth performance were used as one parameter of model prediction. The rainfall intensity information was also collected following the rain gauge selection method explained in section 3.4. Details of return level and duration of each rainfall are shown in Table 6.2:

Table 6.2: 20 Rainfall return level and duration

ID	Asset Local Name	Rainfall date	Level& duration
1	BONDGATE CSO	13-May-06	M2-75min
		22-May-06	M2-60min
2	BROADFIELD ROAD CSO	12-Apri-06	M5-40min
3	BRUNSMERE SCHOOL CSO	01-Jun-06	M3-100min
4	BULLS HEAD CSO	22-Jun-06	M5-30min
		27-Jun-06	M4-40min
5	BURLEY LODGE CSO	13-Mar-06	M2-70min
		16-Mar-06	M2-60min
6	CARR FORGE ROAD CSO	06-Apr-06	M2-30min
7	COWBAR CSO	15-Aug-06	M2-60min
8	CRESCENT TERRACE CSO	08-May-06	M4-100min
		15-May-06	M2-40min
		17-May-06	M2-30min
9	DAYLANDS AVENUE NO.2 CSO	07-Jul-06	M3-100min
10	EBOR WAY CSO	03-Apr-06	M3-20min
11	GILSTEAD LANE 130 2 CSO	03-jun-06	M2-60min
12	HANGINGWATER ROAD CSO	03-Mar-06	M3-40min
		15-Mar-06	M3-60min
13	HILDERTHORPE ROAD CSO	20-Oct-06	M5-20min
14	HOUGH SIDE WORKS CSO	16-Aug-06	M2-25min
		20-Aug-06	M5-40min
15	LIMEKILN LANE CSO	07-Apr-06	M5-75min
16	SHIBDEN PARK CSO	19-Jun-06	M2-60min
		14-Jun-06	M2-100min
17	SPA MILLS BRIDGE ST CSO	15-May-06	M2-40min
18	TADCASTER EAST CSO	02-Mar-06	M2-30min
19	THIRSK FINKLE STREET NO 2 CSO	27-Jun-06	M3-40min
20	WATH DONCASTER ROAD CSO	21-May-06	M5-60min

As discussed in chapter4, the model is capable reliably predicting CSO chamber water depth performance 15 minutes in advance. In this case study, predicted spill events (where the chamber water depth value is higher than chamber weir height) were considered as a potential performance failure. Predicted spill events are compared with actual CSO chamber water depth performance to determine the model prediction accuracy. Therefore, the ADALINE prediction model was

acted as a performance alarm setting mechanism in the asset management approach.

With the performance failure alarm, predicted potential risks can be identified before they actually occur. Subsequently, all concerned hydraulic parameters and CSO chamber information are considered as inputs of the FL asset O&M decision support model and evaluated by the verified expert system. The outputs are then applied to identify the pro-active actions for each spill predicted spill event to prevent performance failures.

The outputs of applying two models to a case study are presented in the following sections of this chapter.

6.3. Outputs

The outputs of applying two CSO management models will be presented in sequence in order to systematically and clearly indicate the achievements of the case study.

6.3.1. Hydraulic performance prediction

Off line prediction of chamber water depth performance was successfully completed by using the ADALINE technique. A typical prediction output example (Crescent Terrace CSO) is shown in Figure 6.2:

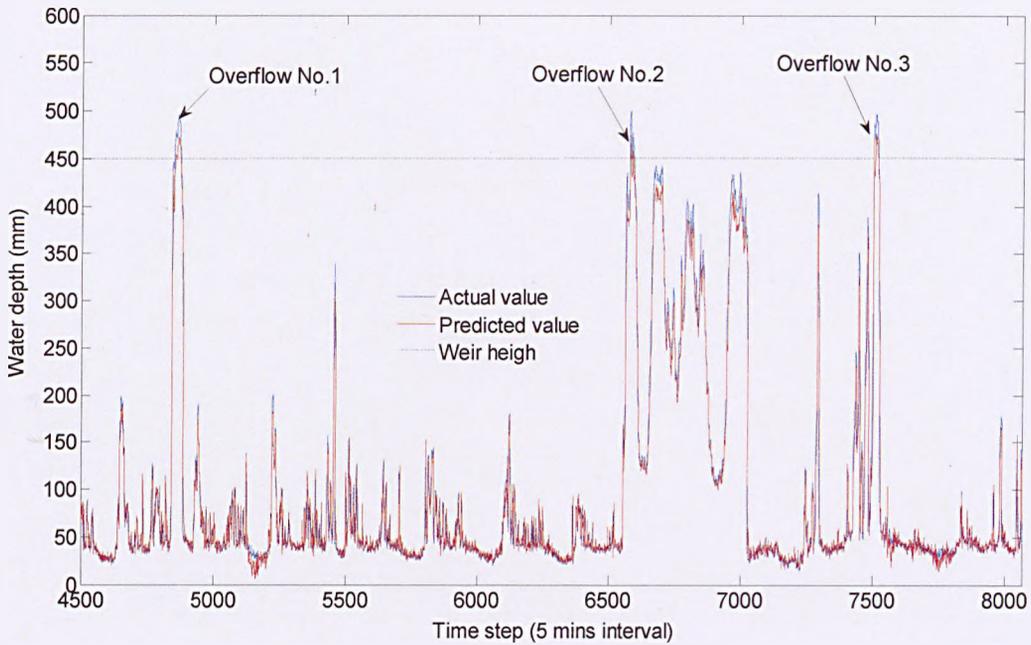


Figure 6.2: Example of NN model output of CRESCENT TERRACE CSO

For this prediction, the model inputs were selected as:

Rainfall intensity value: $u(t + 3)$, $u(t + 2)$, $u(t + 1)$, $u(t)$, $u(t - 1)$, $u(t - 2)$, $u(t - 3)$, $u(t - 4)$ and $u(t - 5)$;

Water depth value: $y(t)$, $y(t - 1)$, $y(t - 2)$, $y(t - 3)$, $y(t - 4)$ and $y(t - 5)$;

The prediction outputs: $y(t + 1)$, $y(t + 2)$ and $y(t + 3)$

For this example, 1200 steps data was used for model training. The prediction correlation coefficient were calculated and summarised in Table 6.3:

Table 6.3: Prediction correlation coefficient summary

ID	Asset Local Name	Prediction output corr
1	BONDGATE CSO	0.986
2	BROADFIELD ROAD CSO	0.974
3	BRUNSMERE SCHOOL CSO	0.957
4	BULLS HEAD CSO	0.963

5	BURLEY LODGE CSO	0.983
6	CARR FORGE ROAD CSO	0.975
7	COWBAR CSO	0.962
8	CRESCENT TERRACE CSO	0.98
9	DAYLANDS AVENUE NO.2 CSO	0.959
10	EBOR WAY CSO	0.985
11	GILSTEAD LANE 130 2 CSO	0.96
12	HANGINGWATER ROAD CSO	0.965
13	HILDERTHORPE ROAD CSO	0.972
14	HOUGH SIDE WORKS CSO	0.984
15	LIMEKILN LANE CSO	0.956
16	SHIBDEN PARK CSO	0.977
17	SPA MILLS BRIDGE ST CSO	0.968
18	TADCASTER EAST CSO	0.988
19	THIRSK FINKLE STREET NO 2 CSO	0.97
20	WATH DONCASTER ROAD CSO	0.969

As seen in Table 6.3, prediction outputs for the selected 20 CSO assets are of high accuracy, the correlation coefficients are better than 0.96.

Seen from Figure 6.2, three spill events were predicted for this CSO, from the overview of 20 CSO assets, all predicted spill events are compared with actual recorded spill performance and summarised in Table 6.4:

Table 6.4: Predicted spill event numbers

ID	Asset Local Name	Actual spill event No	Predicted spill event No.
1	BONDGATE CSO	2	2
2	BROADFIELD ROAD CSO	1	1
3	BRUNSMERE SCHOOL CSO	1	1
4	BULLS HEAD CSO	2	2
5	BURLEY LODGE CSO	2	2
6	CARR FORGE ROAD CSO	1	1
7	COWBAR CSO	1	1
8	CRESCENT TERRACE CSO	3	3
9	DAYLANDS AVENUE NO.2 CSO	1	1
10	EBOR WAY CSO	1	1
11	GILSTEAD LANE 130 2 CSO	1	1
12	HANGINGWATER ROAD CSO	2	2
13	HILDERTHORPE ROAD CSO	1	1

14	HOUGH SIDE WORKS CSO	2	2
15	LIMEKILN LANE CSO	1	1
16	SHIBDEN PARK CSO	2	2
17	SPA MILLS BRIDGE ST CSO	1	1
18	TADCASTER EAST CSO	1	1
19	THIRSK FINKLE STREET NO 2 CSO	1	1
20	WATH DONCASTER ROAD CSO	1	1

All 28 recorded spill events were predicted by using ADALINE prediction model with an advanced period of at least 3 time steps (15 minutes). The prediction analysis of each predicted spill performances are also discussed. Examples of the three predicted spill events are shown in Figure 6.3 to Figure 6.8.

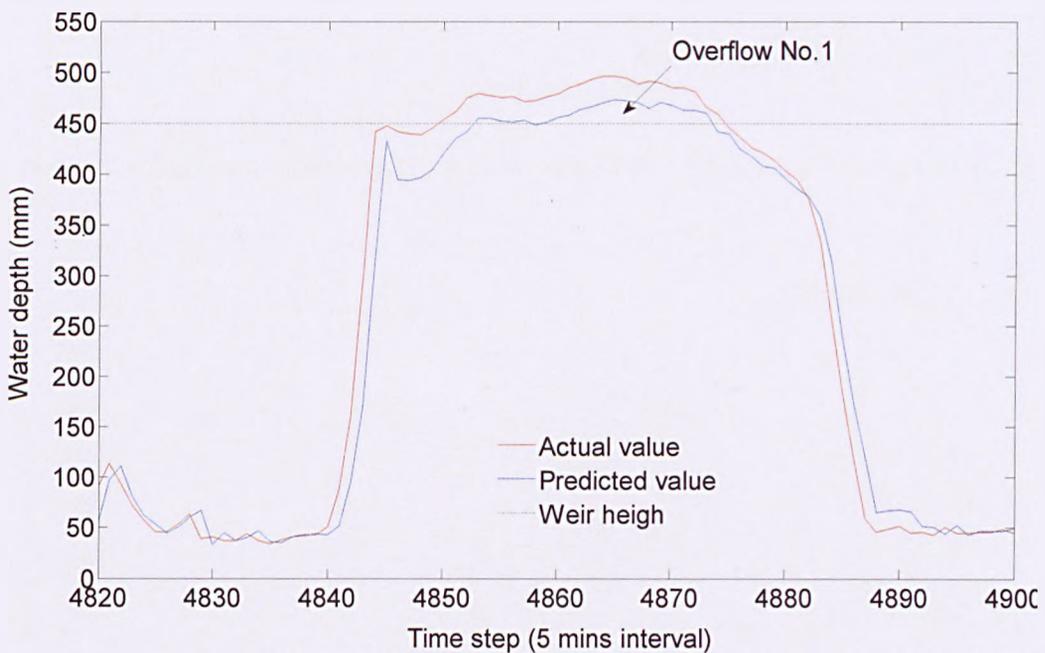


Figure 6.3: Predicted spill event No.1 for CRESCENT TERRACE CSO

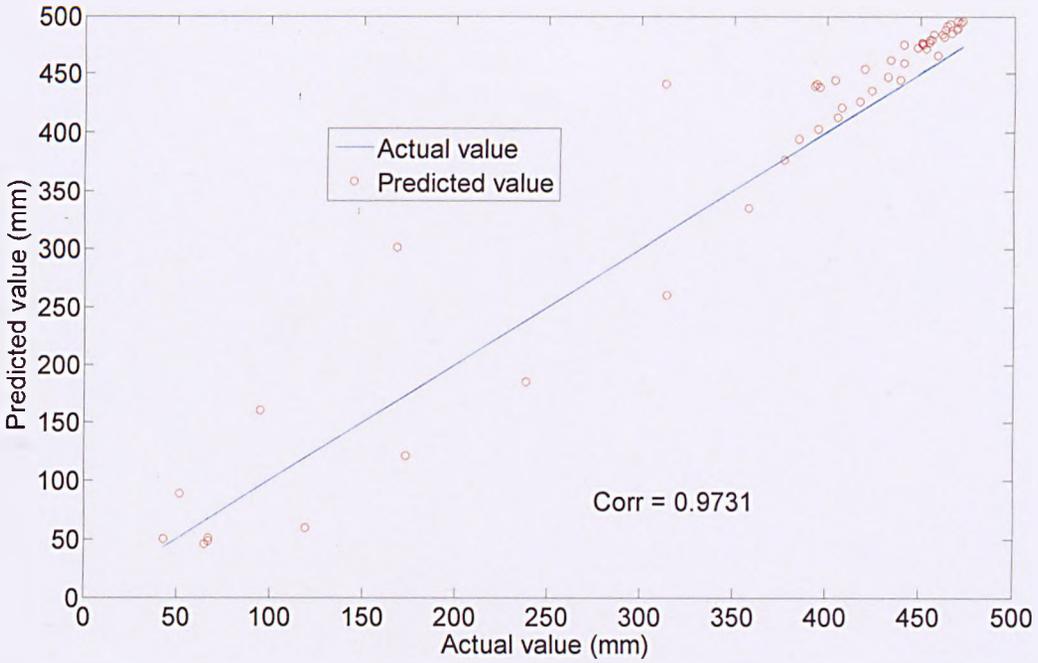


Figure 6.4: Correlation analysis of predicted of spill No 1 (CRESCENT TERRACE CSO)

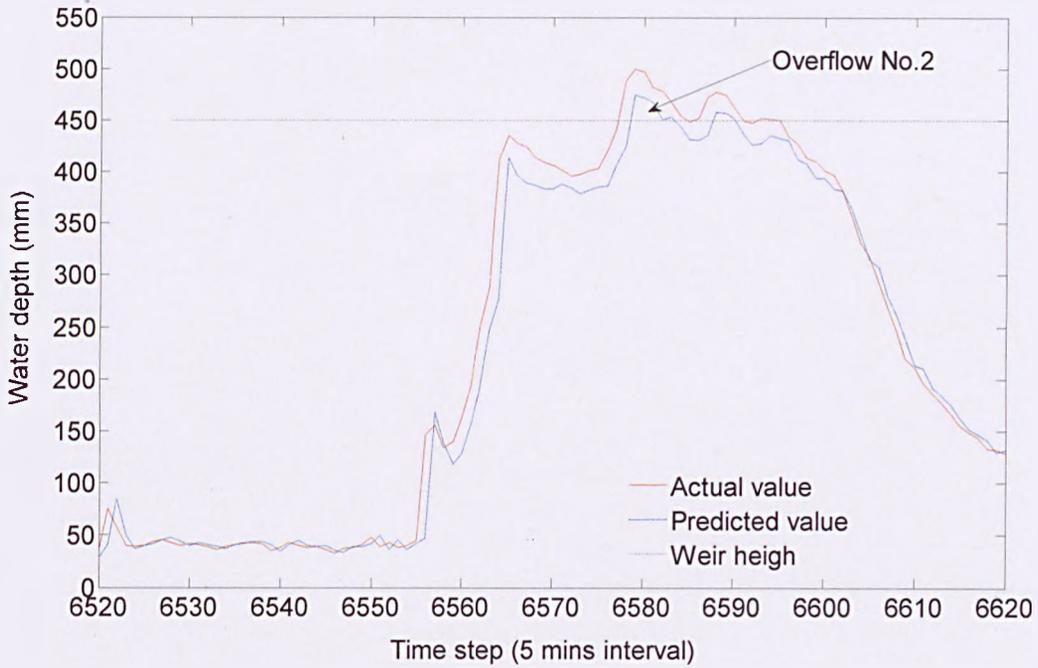


Figure 6.5: Predicted spill event No.2 for CRESCENT TERRACE CSO

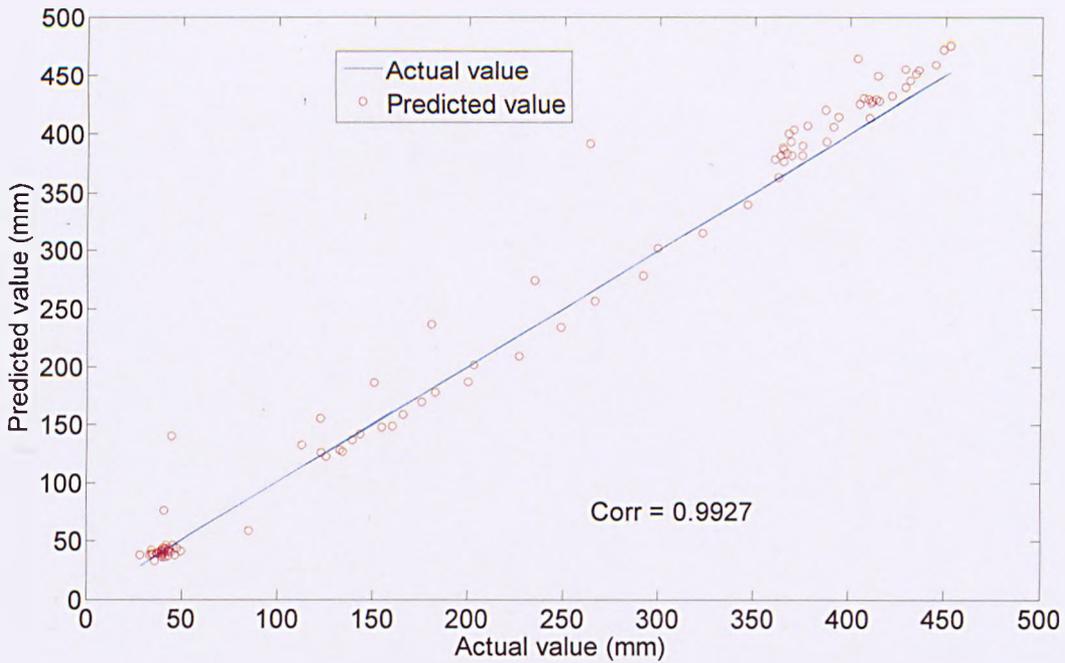


Figure 6.6: Correlation analysis of predicted of spill No 2 (CRESCENT TERRACE CSO)

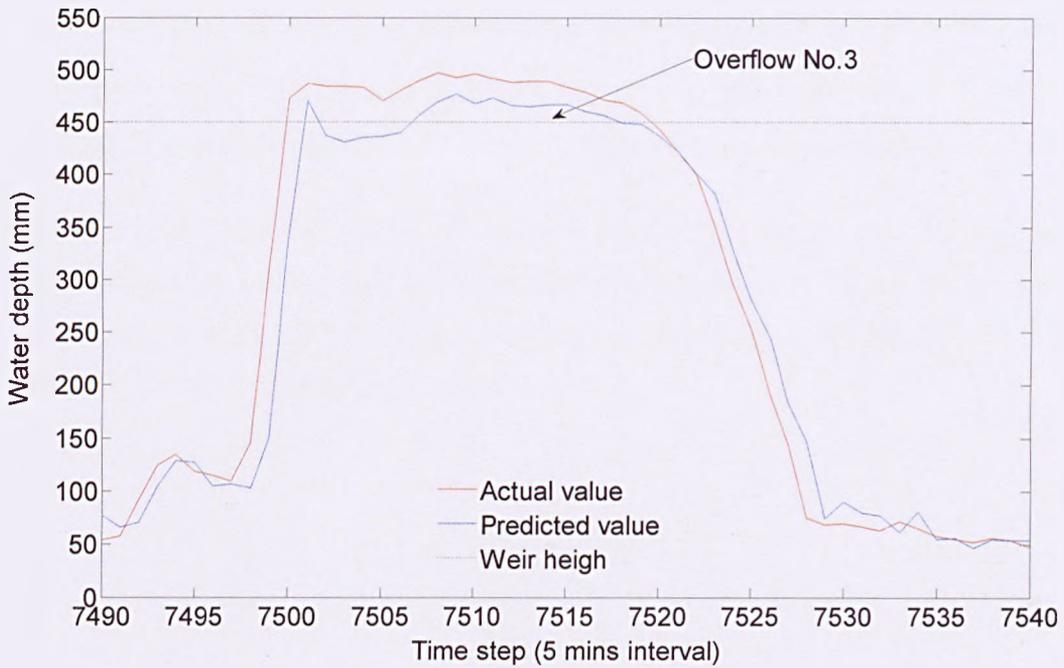


Figure 6.7: Predicted spill event No.3 for CRESCENT TERRACE CSO

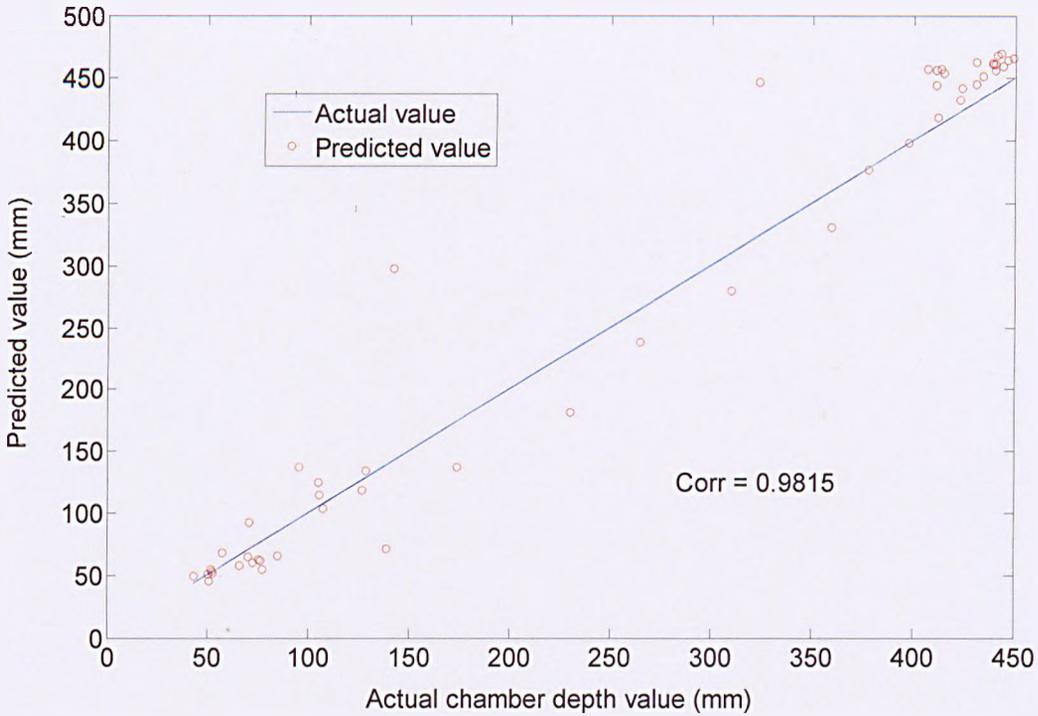


Figure 6.8: Correlation analysis of predicted of spill No 3 (CRESCENT TERRACE CSO)

Seen from Figure 6.4, Figure 6.6 and Figure 6.8, the water depth prediction during spill event also appeared high correlation with actual performance.

Based on the ADALINE CSO performance prediction model's methodology and model sensitivity testing that were discussed in chapter 4.5, the predicted water depth values from $y(t + 1)$ to $y(t + 20)$ were tested. Example of CRESCENT TERRACE CSO is shown in Figure 6.9:

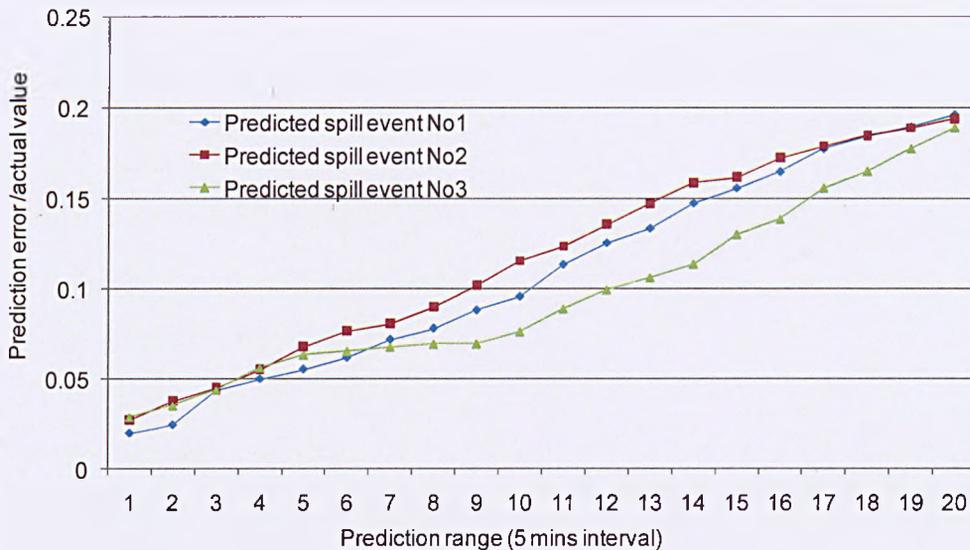


Figure 6.9: Prediction error comparison

Seen from Figure 6.9, the predicted output for time step 1, 2 and 3 (5 minutes, 10 minutes and 15 minutes) appeared that the biases are less than 5%. Three time steps allowed that 15 minutes advanced chamber water depth performance could be predicted with a satisfactory degree of confidence. The prediction range analysis of all 20 CSOs are summarised in Table 6.5:

Table 6.5: Summary of predicted overflow events

ID	Asset Local Name	spill event No	The 3 rd step Prediction error/actual value
1	BONDGATE CSO	2	0.039
			0.042
2	BROADFIELD ROAD CSO	1	0.041
3	BRUNSMERE SCHOOL CSO	1	0.043
4	BULLS HEAD CSO	2	0.046
			0.042
5	BURLEY LODGE CSO	2	0.047
			0.043
6	CARR FORGE ROAD CSO	1	0.036
7	COWBAR CSO	1	0.039
8	CRESCENT TERRACE CSO	3	0.043

			0.045
			0.044
9	DAYLANDS AVENUE NO.2 CSO	1	0.047
10	EBOR WAY CSO	1	0.044
11	GILSTEAD LANE 130 2 CSO	1	0.043
12	HANGINGWATER ROAD CSO	2	0.039
			0.038
13	HILDERTHORPE ROAD CSO	1	0.038
14	HOUGH SIDE WORKS CSO	2	0.042
			0.04
15	LIMEKILN LANE CSO	1	0.041
16	SHIBDEN PARK CSO	2	0.047
			0.044
17	SPA MILLS BRIDGE ST CSO	1	0.037
18	TADCASTER EAST CSO	1	0.046
19	THIRSK FINKLE STREET NO 2 CSO	1	0.042
20	WATH DONCASTER ROAD CSO	1	0.039

In the range of 15 minutes, all predicted 28 water depth performance during spill events were approved that the errors were lower than 5%.

As a conclusion, the off-line CSO water depth performance was successfully predicted by applying ADALINE model. With the high prediction accuracy, predicted spill events were then considered as highlighted subjects for subsequently pro-active action evaluation.

6.3.2. Proactive O&M strategy

The combination of CSO chamber and screen of the 20 CSO assets were shown in Table 6.1. Following the introduction of FL model in chapter 5, other five hydraulic parameters were collected for the 28 predicted spill events:

- ✓ Input 1: CSO chamber weir height
- ✓ Input 2: Flow rate as 1st spill event
- ✓ Input 3: Spill rate as 1 in 5 years return period rainfall
- ✓ Input 4: Total spill duration of inspection period
- ✓ Input 5: Total spill volume during inspection

Each of these five parameters of the 28 predicted spill events were evaluated follow the process demonstrated in Figure 6.10:

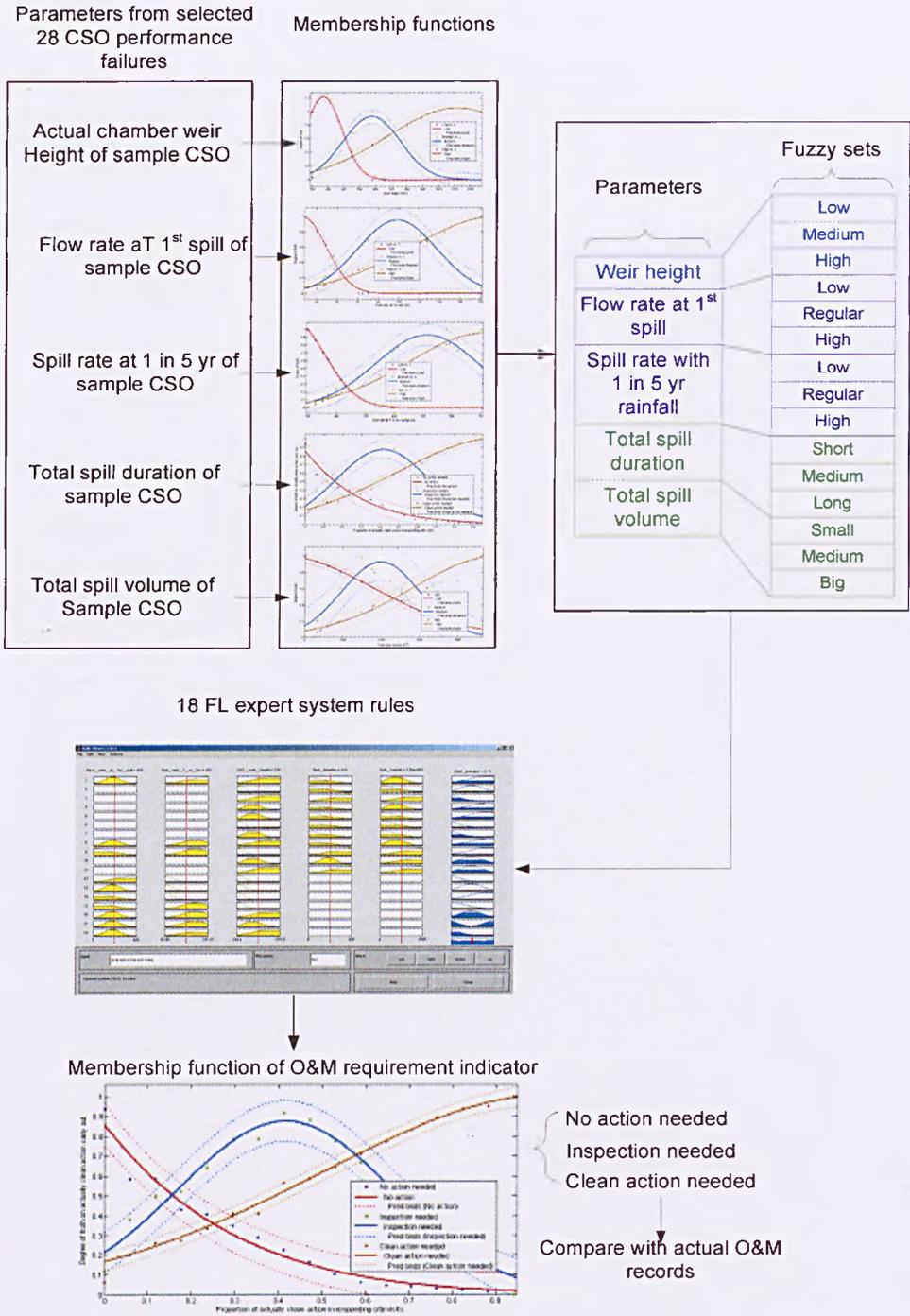


Figure 6.10: FL expert system application process

The predicted pro-active actions were compared with O&M records for each asset, and the list is shown in Table 6.6.

Table 6.6: List of predicted O&M actions

ID	Asset Local Name	Recorded O&M Action	Predicted Action
1	BONDGATE CSO	Respond to RTS Sewerage Alarm	No action
		Respond to RTS Sewerage Alarm	No action
2	BROADFIELD ROAD CSO	Jet Blocked Sewer	Clean
3	BRUNSMERE SCHOOL CSO	INSP/CHECK OPP OF ASSET	Inspection
4	BULLS HEAD CSO	Respond to RTS Sewerage Alarm	No action
		Jet Blocked Sewer	Inspection
5	BURLEY LODGE CSO	Respond to RTS Sewerage Alarm	No action
		Rod Blocked Sewer	Clean
6	CARR FORGE ROAD CSO	Respond to RTS Sewerage Alarm	No action
7	COWBAR CSO	Respond to RTS Sewerage Alarm	No action
8	CRESCENT TERRACE CSO	Respond to RTS Sewerage Alarm	No action
		Flush and/or Jet Sewer	Clean
		Jet Blocked Sewer	Clean
9	DAYLANDS AVENUE NO.2 CSO	Respond to RTS Sewerage Alarm	No action
10	EBOR WAY CSO	Flush and/or Jet Sewer	Clean
11	GILSTEAD LANE 130 2 CSO	Respond to RTS Sewerage Alarm	No action
12	HANGINGWATER ROAD CSO	INSP/CHECK OPP OF ASSET	Inspection
		Jet Blocked Sewer	Clean
13	HILDERTHORPE ROAD CSO	Respond to RTS Sewerage Alarm	No action
14	HOUGH SIDE WORKS CSO	Jet Blocked Sewer	Clean
		Respond to RTS Sewerage Alarm	No action
15	LIMEKILN LANE CSO	INSP/CHECK OPP OF ASSET	Inspection
16	SHIBDEN PARK CSO	Respond to RTS Sewerage Alarm	No action
		Respond to RTS Sewerage Alarm	No action
17	SPA MILLS BRIDGE ST CSO	Flush and/or Jet Sewer	Clean
18	TADCASTER EAST CSO	Respond to RTS Sewerage Alarm	No action

19	THIRSK FINKLE STREET NO 2 CSO	Jet Blocked Sewer	Clean
20	WATH DONCASTER ROAD CSO	INSP/CHECK OPP OF ASSET	Inspection

The pro-active decisions were summarised according to the different types of screen installed and are shown in Table 6.7:

Table 6.7: Evaluation results of applying FL model

Static

	Linguistic Matches		
	No action needed	Inspection needed	Cleaning action needed
Actual Number of actions	6	2	8
FL model output	5	3	8

Mechanical

	Linguistic Matches		
	No action needed	Inspection needed	Cleaning action needed
Actual Number of action	4	2	6
FL model output	4	2	6

As shown in Table 6.7, in terms of CSO with mechanical screen, the decisions regarding pro-active O&M actions for the 28 predicted spill events were proved to be completely correct when compared with O&M records. For the CSOs with static screens, there was a mismatch between the observed and predicted number of performance alarms on one out of 16 predicted performance alarms. Hence, it is concluded that, in general, the results are satisfactory and describe the first step in an enhanced O&M methodology to better understand system performance and to reduce the potential for failures.

When compared with current reactive maintenance strategies, a significant number of performance failures and subsequent pollution accidents can thus be avoided, and unnecessary site visits can be reduced. These advantages can both improve CSO serviceability and reduce asset operational expenditure for the water company.

6.4. Project Achievement

To summarise the contributions of this CSO management approach the project achievements have been listed as:

Data collection

- ✓ Data collection process and the subsequent arrangement into a series of initial analysis of database provided comprehensive information for the development of two models to enhance the understanding of CSO performance and the predictive requirements for pro-active maintenance.

CSO neural network hydraulic performance prediction model:

- ✓ This component of the research examined telemetry system error and presented a procedure to accommodate missing data caused by remote device functional failures.
- ✓ Simulated the influence of rainfall on CSO chamber hydraulic performance.
- ✓ The ADALINECSO model results clearly gave reliable results on CSO performance and hence this model was used as the initial basis to provide alarms for subsequent pro-active decision support.

FL CSO unexpected O&M prediction model:

- ✓ Based on predicted CSO hydraulic performance caused by rainfall incidents which could potentially set off a performance alarm, the CSO FL

model was developed to predict whether proactive O&M action was needed thus providing advanced decision support to enhance CSO hydraulic performance management strategies.

- ✓ Compared with current reactive maintenance mechanisms, the methodology developed procedures to avoid a significant number of performance failures and the subsequent pollution incidents.
- ✓ From reducing unnecessary site visits to improve CSO serviceability and reduce CSO asset operational expenditure.

7. Conclusion

7.1. Project summary

The aim of this work is to use advanced modelling to gain a better understanding of the performance of CSO assets, and subsequently use this understanding to develop more efficient CSO management strategies. This research selected a total of forty CSOs (31 CSOs combined with screens) from the water company's records. Of these, twenty operated satisfactorily, the other twenty however, were known to have a high risk of failure or required a high frequency of maintenance visits. Related information was collected into a database onto DVD disc which is appended to the thesis.

An ADALINE approach was used to develop a CSO hydraulic performance prediction model which proved to be efficient in establishing the relationship between rainfall intensity and corresponding CSO asset chamber water depth. The model was successfully used to produce a 15 minutes off line prediction of chamber water depth. It also helped to establish the feasibility of using predicted abnormal CSO performances to create a warning system for potential asset failures. Together with the CSO FL O&M model, this generated a powerful decision support tool for effective both routine and pro-active O&M strategies, thereby minimising the need for reactive maintenance.

After having developed and verified the two models, a general CSO asset management case study applied the methodology to a practical water industry asset management project. The case study proved the feasibility and advantages of applying both NN and FL models in CSO asset operation. The achievements of this research are summarised as follows:

CSO hydraulic performance prediction model:

- ✓ Developed and proved the ADALINE methodology in predicting CSO hydraulic performance
- ✓ Demonstrated the reliability of using the ADALINE approach to predict CSO performance, especially in regard of the performance at times of dry weather and due to rainfall level under 1 in 10 year return event
- ✓ Reliably achieved an accurate prediction of CSO chamber depth with a lead time of 15 minutes. This indicates the feasibility of developing an advance alarm mechanism for the CSO asset control system

FL CSO pro-active operation and maintenance decision support model:

- ✓ Developed a methodology of using FL approach in providing decision support on CSO asset pro-active O&M actions
- ✓ Indicated the advantage of installing a mechanic screen when compared with the CSO chambers without a screen by reviewing the maintenance requirement of each reported performance failure. The method of using FL prediction to determine service efficiency of CSO assets with different chamber and screen types was shown to be reliable
- ✓ The CSO FL model reliably predicted whether pro-active O&M actions were required, thus providing a useful decision support tool to improve CSO hydraulic performance management strategies
- ✓ The FL model showed that assets with a static screen installed appeared to need more operations than those with mechanical screens. Comparison between recorded O&M and FL predicted results indicated the high reliability of this decision support model

This new CSO asset management approach provided a far better understanding of hydraulic performance for current CSO assets and helped to develop more efficient operation and maintenance strategies. The methodology was applied to both off line hydraulic performance prediction and pro-active requirements to improve asset service and reduce performance failures and, subsequently, pollution incidents.

7.2. Future works

7.2.1. Rainfall radar data application

As mentioned in section 3.4, rainfall intensity data that provided by rainfall radar devices can be potentially used into ADALINE CSO performance prediction model.

Currently, 85% of the UK has a resolution of 2km and better resolution data for urban catchments (Met Office, 2007). Radar reflectivity is measured by a series of radar sweeps, this data is processed by the Met Office in order to convert the reflectivity measurement into rainfall intensity and to correct potential errors such as attenuation by intervening rainfall and ground clutter (Met Office, 2007). Rainfall radar data is supplied at a time resolution of 5 minutes at near real time.

Conventionally, sewer hydraulic models were validated from flow surveys which included rainfall information provided by rain gauges. The limitation of rain gauge data was explained in section 3.4, however, rainfall radar offers a data solution for long term records and without spatial limitation, which can be applied to CSO hydraulic models and which can be used with near real time operation strategies.

The application of rainfall radar to urban drainage has been discussed for over two decades. However, the resistances of applying rainfall radar data can be summarised as:

- ✓ A limited understanding of the data
- ✓ Concerns over accuracy
- ✓ Data availability
- ✓ Cost of the data

Previous studies have investigated various topics including the accuracy of radar data compared to rain gauges (Jessen et al. 2005), suggested methodologies for the application of radar data to urban drainage systems (Einfalt et al.

2004). Some studies have investigated the use of rainfall radar data in sewer hydraulic models (Kramer et al. 2005).

Research on comparing the application of rain gauge and rainfall radar data was reported by Shepherd et al., (2010). In their research, rainfall radar data was purchased from the Met Office, this data was produced by a network of C-band radars which covered the UK. Data was supplied at spatial resolutions of 1, 2 or 5km, dependant on the distance from the radar station. The work compared predicted flows from InfoWorks with both rainfall inputs from rain gauge and radar methods together with actual measured flow in the sewer. The analysis was carried out by using a verified InfoWorks CS sewer hydraulic model. An example of the analysis is shown in Figure 7.1:

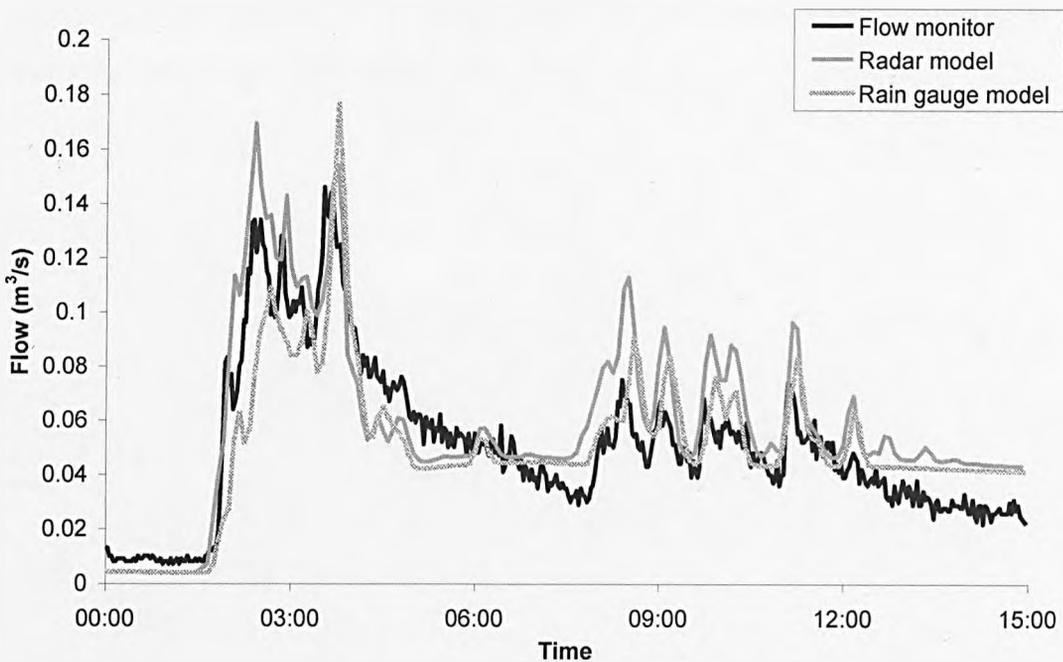


Figure 7.1: Comparison of measured and modelled flows (Shepherd et al. 2010)

Seen from Figure 7.1, flows from the two model runs generally compare well to the measured data Shepherd et al., (2010).

Compared with rain gauge data, the advantage of applying rainfall radar data can be summarized as:

- ✓ Instead of the measurement of rainfall at a single point, radar rainfall intensity data is averaged over the whole given spatial resolution, no more rainfall and water depth correlation analysis is needed.
- ✓ Rainfall radar data represents the real time rainfall condition; there is no 'time lag' which caused by the distance between rainfall measurement and CSO asset.
- ✓ By using predicted rainfall radar data, the ADALINE model's prediction range can be extend and act as real time prediction.

With this further work, potential CSO performance failure (spill event) can be predicted much earlier. As a result, there will be longer responding time for maintenance team to deliver pro-active actions.

8. References

- Ainger, C M., Armstrong, R J. and Butler, D. (1998) 'Dry weather flow in sewers' CIRIA report 177, 1998.
- Andoh, R Y., Smith, B P., and Saul, A J. (1999) 'The screen efficiency of a novel self-cleaning CSO', the 8th International Conference on Urban Storm Drainage, Sydney, Australia, 1999, pp205-212
- Aqil, M., Kita, I., Yano, A. and Nishiyama, S. (2007) 'Neural networks for real time catchment flow modelling and prediction' *Water Resources Management*, Volume 21, Issue 10 pp1781–1796.
- Ashley, R M., and Bertrand-Krajewski, J L. (2004) *Solids in sewers: characteristics, effects and control of sewer solids and associated pollutants*. IWA Publishing.
- Bairaktaris, D., Delis, V., Emmanouilidis, C., Frondistou-Yannas, S., Gratsias, K., Kallidromitis, V. and Rerras, N. (2007), 'Decision-Support System for the Rehabilitation of Deteriorating Sewers' *Journal of Performance of Constructed Facilities*, Volume 21, Issue 3, pp. 240-248.
- Balmforth, D J., Saul, A J. and Clifford, I T. (1994) 'Guide to the Design of Combined Sewer Overflow Structures' FWR Report No. FR0488, November, 1994.
- Balmforth, D.J., Digman, C.J., Kellagher, R. and Butler, D. (2006) 'DEFRA Integrated Urban Drainage Scoping Study Final Report. DEFRA, London. UK.
- Berardi, L., Kapelan, Z., Giustolisi, O. and Savic, D.A. (2008) 'Development of Pipe Deterioration Models for Water Distribution System using EPR' *Journal of Hydroinformatics (IWA)*, Volume 10, Issue 2, pp 113-126.

- Burt, D., Corton, M., Hetherington, D. and Balmforth D J, (2002) 'Multiphase Modelling and the Prediction of Retention Efficiency in a Side Weir', *9th International Conference on Urban Drainage*, Portland, Oregon, September 2002.
- Butler, D. and Graham, J D, (1995) 'Modelling dry weather wastewater flow in sewer network' *Journal of Environmental Engineering*, Volume 121, Number 2 February. P 116-173.
- Butler, David, and Davies, John, W. (2008) *Urban Drainage*, second edition chapter 17, pp 380-400. Taylor & Francis, 2008
- Cashman A., Saul A J., Savic D., Ashley R M. (2002). Whole life costing approach to sewer asset management. Proc. Conf on Sewer Operation and Management, Bradford. Nov. 26-28th
- Cashman, A., Shepherd, W. and Saul, A J (2005) 'COST-S: a new methodology and tools for sewerage asset management based on whole life costs' IWA Publishing.
- Christodoulides, J. and Saul, A.J. (2002) *Monitoring and Control of CSOs Guideline Document*,
<http://ciwem.co.uk/groups/wapug/S2001paper5.pdf>
- Dawson, C W., Abrahart, R J., Shamseldin, A Y. and Wilby, R L. (2006) 'Flood estimation at ungauged sites using artificial neural networks' *Journal of Hydrology*, volume 319, pp 391-409.
- Draper, N R. and Smith, H., (1998) 'Applied Regression Analysis, Wiley-Interscience; Third edition, April 23, 1998, p 89-92.
- Einfalt, T., Arnbjerg-Nielsen, K., Golza, C., Jensen, N., Quirnbach, M., Vaes, G., & Vieux, B. 2004. Towards a roadmap for use of radar rainfall data in urban drainage. *Jnl of Hydrology*, 299, 186-202.

Environment Agency: 'Pollution incidence categories

(2007):<http://www.environmentagency.gov.uk/yourenv/eff/1190084/pollution/296030/296054/297197>. (Accessed Dec. 2010)

Fach, S. Sitzenfrei, R. and Rauch, W., (2008): 'Assessing the relationship between water level and combined sewer overflow with computational fluid dynamics', 11th International Conference on Urban Drainage, Edinburgh, Scotland, UK, 2008

Fernando, A.K. and Kerr, T., "Runoff forecasting with Artificial Neural Network Model", The 3rd Pacific Conference on stormwater and aquatic resource protection, Auckland, New Zealand Water and Wastes Association publication of conference proceedings in CD-ROM, 2003

Fernando, A K., Zhang, Xiujuan., and Peter, F K. (2006) 'Combined sewer overflow forecasting with feed-forward back-propagation artificial neural network'. Proceedings of world academy of science, volume 12, March 2006.

FWR, (2005): *CSO SCREENS – A DESIGN AND INSTALLATION REVIEW*, Online, available <http://www.fwr.org/wapug/csoscrns.pdf>. (Accessed Dec. 2010)

Goguen, Joseph A. (1967) 'L-fuzzy sets'. Journal of Mathematical Analysis and Applications Volume 18, pp 145-174.

Hagan, M T., O. De Jesus, and Schultz, R. (1999): 'Training Recurrent Networks for Filtering and Control,' Chapter 12 in Medsker, L. and Jain, LC. Eds., *Recurrent Neural Networks: Design and Applications*, CRC Press, pp. 311–340.

Hajek, P. (1998) *Meta mathematics of fuzzy logic*, Springer, 31 Aug 1998, p 197-202.

Hessami, M., Anctil, F. and Viau, AA. (2004) 'Selection of an Artificial Neural Network Model for the Post-calibration of Weather Radar Rainfall Estimation', *Journal of Data Science* 2, pp. 107-124.

Hopfield, J. (1982): Neural networks and physical systems with emergent collective computational abilities. Conference: Proceedings of the National Academy of Sciences of the USA, 9(2554).

Hundecha, Y., Bardossy A. and Theisen, H W. (2001) 'Development of a fuzzy logic-based rainfall-runoff model' *Hydrological Sciences-*, Volume 46(3), June 2001.

Hydro International,

<http://www.hydro-international.biz/cso/heliscreen.php>. (Accessed Dec. 2010)

HydrokUK

<http://www.hydrok.co.uk/wms-products-hydrok-mecmex.htm>,
(Accessed Dec. 2010)

Jacquin, A P. and Shamseldin, A Y. (2008), 'Development of Rainfall-Runoff Models Using Mamdani-Type Fuzzy Inference System.' *Water science and technology library*, Volume 68, part III, pp 189-200.

Jessen, M., Einfalt, T., Stoffer, A. and Mehlig, B. 2005. Analysis of heavy rainfall events in North Rhine–Westphalia with radar and rain gauge data. *Atmospheric Research*, 77, 337-346.

Keener, V W., Feyereisen, G W., Lall, U., Jones, J W., Bosch, DD. and Lowrance, R. (2010) 'El Niño/Southern Oscillation(ENSO) impudence on monthly NO₃ load and concentration, stream flow and precipitation in the Little River Watershed, Tifton, Georgia(GA)' *Journal of Hydrology*. Journal 15, v. 381, no. 3-4, February, pp. 352-363.

Kramer, S., Verworn, H.-R., & Ziegler, J., 2005. Radar rainfall time series for the performance assessment of sewer systems. In: Proceedings of 10th International Conference on Urban Drainage, Copenhagen, Denmark, 21-26 August 2005.

Kumar, N V., Mathew, S. and Swaminathan, G., (2010), 'Multi-factorial Fuzzy Approach for the assessment of groundwater quality' *Journal of Water Resource and Protection*, 2010, Volume 2, pp 597-608.

Kurth, A., Saul, A J., Mounce, S., Shepherd, W. and Hanson, D. (2008)'Application of Artificial Neural Networks (ANNs) for the prediction of CSO discharges,'*11th International Conference on Urban Drainage*, Edinburgh, Scotland, UK, 2008.

Li He, Guo-He Huang, Guang-Ming Zeng and Hong-Wei Lu (2008)'Wavelet-based multiresolution analysis for data cleaning and its application to water quality management systems. '*Expert Systems with Applications*Volume 35, Issue 3, October 2008, pp1301-1310.

Mahabir, C., Hicks, F E. and Fayek A R. (2003) 'Application of fuzzy logic to forecast seasonal runoff' *Hydrological Process*. Volume 17, pp3749–3762.

Mamdani, E H. and Assilian, S. (1974) 'Application of fuzzy algorithms for control of simple dynamic plant, Proceedings of.' IEEE volume 121, pp.1585-1588.

Met Office 2007. Fact sheet No. 15 – Weather Radar.

<http://www.metoffice.gov.uk/corporate/library/factsheets/factsheet15.pdf>

f

McCulloch and Pitts (1943) 'A logical calculus of the ideas immanent in nervous activity', *Bulletin of Mathematical Biology* Volume 5, pp.115-133.

- Medsker, L. and Liebowith, J. (1999) 'Design and development of expert systems and neural networks', ISBN 0-02-380131-X
- Minsky, M. and Papert, S. A. (1969): *Perceptron*, MIT Press, Cambridge, MA.
- Mounce, S R., Machell, J. and Boxall, J B. (2006): *Development of artificial intelligence system for analysis of water supply system data*, WDSA, Cincinnati.
- Ofwat, (2004): *Capital Maintenance Review - Independent assessment of Ofwat's PR04 process -initial review.*
- Prechelt Lutz (1998) 'Automatic early stopping using cross validation: quantifying the criteria' *NeuralNetworks*, Volume 11, number 4, pp 761-767.
- Rosenblatt, R. (1961)*Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*, Spartan, New York.
- Rumelhard, D E., Hinton, G E. and Williams, R J. (1986) 'Learning internal representations by error propagation', in Rumelhard, D E.and McClelland, J L. (ed.) *Parallel Distributed Processing 1 (8)*, MIT Press, Cambridge, Massachusetts.
- SadiqRehan, Kleiner Yehuda, and RajaniBalvant (2007) *Water Quality Failures in Distribution Networks—Risk Analysis Using Fuzzy Logic and Evidential Reasoning*, *Risk Analysis*, Volume 27 Issue 5, pp 1381 – 1394.
- SamehAli, Tarek Zayed and Mohamed Hegab, (2007) 'Modelling the Effect of Subjective Factors on Productivity of Trenchless Technology Application to Buried Infrastructure Systems' *Journal of Construction Engineering and Management*, Volume 133, Issue 10, pp. 743-748.
- Saul, A J., et al. (1997) *Sewers – Rehabilitation and new construction, repair and renovation*, Arnold, London.

- Saul, A J. (1998) 'CSO state of the art review: a UK perspective' *UDM'98, Fourth Int. Conf. on developments in urban drainage modelling*. September, London, UK.
- Saul, A J. (2000), 'Screen Efficiency Proprietary Designs', UKWIR Report 99/WW/08/5.
- Savic, D., Giustolisi, O., Beradi, L., Shepherd, W., Djordjevic, S. and Saul A J, (2006)'Modelling Sewer Failure by Evolutionary Computing' *WaterManagement, Volume159*, June 2006, Issue WM2, pp 111-118.
- Schulze FH, Wolf H, Jansen HW, van der Veer P, 2005, Applications of Artificial Neural Networks in integrated water management: fiction or future? *Water Sci Technol.* 2005;52(9):21-31
- Seo, F. and Sakawa, M. (1985)' Fuzzy Multi-attribute Utility Analysis for Collective Choice'*IEEE Transactions on Systems, Man, and Cybernetics*, Volume 15, Number 1.pp. 45-53.
- Shamseldin, A Y. (1997) 'Application of a neural network technique to rainfall-runoff modelling', *Journal of Hydrology*Volume199, pp. 272-294.
- Shepherd, W., Cashman, A., Djordjevic, S., Dorini, G., Saul, A., Savic, D., and Lewis, L, (2005): 'Investigation of blockage relationships and the cost implications for sewerage network management'. *10th International Conference on Urban Drainage*, Copenhagen/Denmark, 21-26 August 2005
- Shepherd, W., Schellart, A.; Saul, A., and Hanson, D., (2010), 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 20092010 pp. 763-766

- Siskos, Y., (1982) 'A way to deal with fuzzy preferences in multi-criteria decision problems.' *European Journal of Operational Research*, Volume 10, number 3, pp:314--324, 1982.
- Solomatine, D P. andDulal, K N. (2003): 'Model trees as an alternative to neural networks in rainfall-runoff modelling', *Hydrological Sciences–Journal–des Sciences Hydrologiques*, Volume 48(3), pp. 399-411.
- Sugeno, M. and Kang, G T., (1988) 'Structure identification of fuzzy model', *Fuzzy Sets and Systems*, Volume. 28, pp. 15-33.
- Sumer, D., Gonzalez, J. and Lansey, K. (2007) 'Real-Time Detection of Sanitary Sewer Overflows Using Neural Networks and Time Series Analysis' *Journal of Environmental Engineering*, Volume 133, Number 4, April 1, 2007.
- Takagi, T. and Sugeno, M. (1985): 'Fuzzy identification of systems and its application to modelling and control' *IEEE Transaction on Systems, Man and Cybernetics*, Volume 15, pp. 116-132.
- Timothy, J., Ross. (2007) *Fuzzy logic with engineering applications*, Wiley-Blackwell, second edition, chapter 1, page 11, and chapter 4, page 90-119.
- Tokar, A. S. and Markus, M. (2000): 'Precipitation-runoff modelling using artificial neural networks and conceptual models', *Journal of Hydrologic Engineering*, pp. 156-161 (April 2000).
- Tsukamoto, Y. (1979) 'An approach to fuzzy reasoning method' *Advances in fuzzy set theory and applications*, North-Holland, Amsterdam, pp. 137-149.

- UKWIR, (1997), *CSO performance evaluation results of a field programme to assess the total solids retention efficiency performance of side weir and stilling pond chambers*. Report ref. 97/WWI/0811
- UKWIR, 2002, *Capital Maintenance Planning: A Common Framework*, UKWIR 2002, ISBN: 1-84057 265 5.
- WaPUG Guide (2001): 'The Design of CSO Chambers to Incorporate Screens', Online found at <http://www.wapug.org.uk/> Accessed Dec. 2010
- WaPUG (2005): 'Sewer Asset Management Through Long Term CSO Monitoring' *WaPUG Spring conference*, Coventry 2005.
- Wasserman, P D. (1993), *Advanced methods in neural computing*, Van Nostrand Reinhold, New York, 1993.
- Weiss, G., Brombach, H. and Wohrle, Ch. (2006)'Monitoring of Combined Sewer Overflow Tanks', *Water Practice & Technology*, Volume 1, Number 1.
- Widrow, B., (1962) *Generalization and information storage in networks of ADALINE Neurons in Self Organizing system*, (M.C. Yovitz, G.T. Jacobi, and G.D. Goldstein, eds) pp.435-461, Washington, DC: Spartan Books.
- Widrow, B. and Sterns, S D.(1985) *Adaptive Signal Processing*, New York, Prentice-Hall.
- Widrow, B. (1995) *Handbook of Brain Theory and Neural Networks*, edited by Michael A. Arbib, pp. 719-724. The MIT press, 1995.
- WRc, (1994) *Sewerage rehabilitation manual* volume1, volume 2, Rational Approach.Planning.3rd edition.

Yorkshire Water (2006) *Research and Development Strategic Partnership
ROAME statement*

Yorkshire Water (2006): *YWS Annual CSO pollution incidents list 2006-2007*

Yorkshire Water (2007): *New CSO grouping report 2007*

Zadeh, L. A. (1965) 'Fuzzy sets', *Information and Control*, Volume 8pp 338-353.

Zadeh, L. A. (1968) 'Fuzzy algorithms'. *Information and Control* Volume 12
number 2, pp. 94–102.