Contents

List of Figures xiii

List of Tables xv

Glossary xvii

I Introduction and proposal of hydraulic sectors 1

1 Introduction 3

1.1 The aims of the thesis . . . 6

1.2 Outline of the thesis . . . 6

1.3 Contributions of the thesis . . . 9

2 Urban water management by hydraulic sectors 11

2.1 Preliminaries . . . 12

2.2 Essentials of water supply network management . . . 14

2.2.1 Leak detection . . . 16

2.2.2 Demand management . . . 17

2.2.3 Screening of system vulnerabilities . . . 18

2.3 Implementation of hydraulic sectors in water supply networks . . . 20

2.3.1 Viability studies . . . 20

2.3.2 Design and enforcement of sectors . . . 21

2.4 Classical and recent trends in hydraulic management by sectorised networks . . . 22

2.5 Case-study . . . 23

2.5.1 Main case-study . . . 24

2.5.2 A simple real case . . . 26
2.6 Summary and comments . . . 26

II Establishment of supply clusters 29

3 Water supply clusters using semi-supervised learning 31
3.1 Clustering processes . . . 32
3.1.1 K-means algorithm . . . 34
3.1.2 Cluster evaluation . . . 35
3.1.2.1 Analysis of the silhouette . . . 36
3.2 Graph clustering and water supply network data . . . 37
3.2.1 Review of some graph theory topics . . . 37
3.2.2 Graph clustering process and spectral methods . . . 40
3.2.3 Graphs and water supply network data . . . 41
3.2.4 Spectral clustering . . . 43
3.2.4.1 Spectral clustering algorithm . . . 44
3.3 Graph clustering on kernel spaces . . . 45
3.3.1 Introduction to kernel methods . . . 45
3.3.1.1 Other properties of kernels . . . 46
3.3.1.2 Most common kernels . . . 46
3.4 The proposed semi-supervised clustering algorithm . . . 47
3.4.1 Semi-supervised clustering . . . 47
3.4.2 Kernel-based semi-supervised clustering . . . 49
3.4.3 The proposed algorithm . . . 50
3.5 Experimental process . . . 52
3.5.1 Specifying the data matrices to kernelise . . . 52
3.5.2 Results . . . 54
3.6 Summary and comments . . . 55

4 Agent-division of water distribution systems into supply clusters 57
4.1 Intelligent agents . . . 58
4.1.1 A first approach to intelligent agents . . . 58
4.1.2 Formal definitions of agents and their environment . . . 59

4.2 Multi-agent systems . . . 60
4.3 Water network abstraction in a multi-agent environment . . . 62
4.4 Graph multi-agent clustering . . . 63
4.4.1 Negotiating the boundaries . . . 65
4.5 Agent algorithms and the experimental processes . . . 65
4.5.1 Proposing a MAS clustering algorithm . . . 66
4.5.2 MAS clustering algorithm implementation . . . 67
4.5.3 Results . . . 69
4.6 Summary and comments . . . 71

5 Multi-agent adaptive boosting on semi-supervised water supply clusters 73

5.1 Clustering large graphs: the case of real WSN division into supply clusters 74
5.1.1 Improving number of operations . . . 75
5.1.2 Sampling graphs techniques . . . 76
5.1.3 Multi-agent support to sampling subgraphs . . . 77
5.1.3.1 Sampling by simulation of virus propagation . . . 77
5.2 Boost-clustering for semi-supervised subgraphs . . . 78
5.2.1 Semi-supervised boosting and the pre-clustering phase . . . 79
5.2.2 Semi-supervised boost-clustering in WSN . . . 80
5.3 The proposed MAS-boost clustering algorithm . . . 81
5.3.1 Pre-clustering phase . . . 81
5.3.2 Sampling subgraph by multi-agent graph exploration . . . 82
5.3.3 Assigning pseudo-source nodes . . . 83
5.3.4 Semi-supervised clustering . . . 84
5.3.5 Cluster evaluation: silhouette for graphs . . . 84
5.3.6 Re-assigning weights to sampling subgraphs . . . 85
5.3.7 Multi-agent voting system . . . 86
5.4 Experimental process . . . 87
5.4.1 Results . . . 88
5.5 Summary and comments . . . 89

III Results and applications **91**

6 Results of dividing a real water network into supply clusters **93**

6.1 Results and first conclusions 93
6.1.1 Application of semi-supervised clustering 96
6.1.2 Application of multi-agent clustering 98
6.1.3 Application of MAS-boost clustering 99
6.1.4 Comparison of the methodologies 102
6.2 Summary and comments 106

7 Water network management based on supply clusters: working proposals **107**

7.1 Predictive models of water demand 108
7.1.1 Exploratory analysis of the data 109
7.1.1.1 The prediction task 113
7.1.2 Experimental study 113
7.1.2.1 Monte Carlo estimates 114
7.1.2.2 The evaluation statistics 115
7.1.2.3 Model building 116
7.1.2.4 Results of the Monte Carlo comparisons 117
7.1.2.5 Comparing the model building strategies 117
7.1.2.6 The best model variants for each evaluation metric . . . 118
7.1.2.7 Using the best model 120
7.2 Anomaly causes in a water supply system 121
7.2.1 Methodology 125
7.2.1.1 EPANET simulation 125
7.2.1.2 Detecting anomalies 126
7.2.1.3 Kernel-based causal algorithm 127
7.2.1.4 Action taking phase 128
7.2.2 Results 129
7.3 Ranking nodes in Water Supply Networks 129
7.3.1 Brief introduction to Google's PageRank algorithm 131
7.3.1.1 Computation of PageRank vector 132
7.3.2 Adaptation of PageRank algorithm to a WSN 132

7.3.3 Experimental study: PageRank on a real WSN 133
7.3.3.1 Results 133
7.4 Summary and comments 135

IV Conclusions **139**

8 Conclusions **141**
8.1 Contributions of this thesis 141
8.1.1 Water supply clusters made using semi-supervised learning . . . 142
8.1.2 Agent-division of water distribution systems into supply clusters 142
8.1.3 MAS boosting semi-supervised supply clusters 143
8.1.4 Water network management by supply clusters 143
8.1.4.1 Predictive models 143
8.1.4.2 Screening anomalies 144
8.1.4.3 Ranking nodes 144
8.1.5 Developed and employed software 144
8.2 Publications in relation to this thesis 145
8.2.1 Journal papers 145
8.2.2 Chapters of books 146
8.2.3 Conference papers 147
8.3 Future work 150
8.3.1 Water supply clusters using semi-supervised learning 150
8.3.2 Agent-division of water distribution systems into supply clusters 150
8.3.3 MAS boosting semi-supervised supply clusters 151
8.3.4 Water network management by supply clusters 151

V Appendices **153**

A Developed and employed software **155**
A.1 NetLogo and multi-agent systems 156
A.1.1 Practical implementation issues 157
A.2 R Language: the flexible programming environment 161
A.2.1 igraph library and page.rank function 163

A.2.2 kernlab library and specc function . . . 164
A.3 Exchange framework between EPANET, R Language and NetLogo . . . 166
A.3.1 NetLogo - R - Extension . . . 166
A.3.2 REPANET exchange . . . 167
A.3.2.1 RimpEpa.c . . . 167
A.3.2.2 RexpEpa.c . . . 169

B Overview of predictive models developed in water demand 171
B.1 The used predictive models . . . 171
B.1.1 Artificial Neural Networks (ANN) . . . 171
B.1.1.1 Tuning ANN . . . 172
B.1.2 Projection Pursuit Regression (PPR) . . . 172
B.1.2.1 Tuning PPR . . . 173
B.1.3 Multivariate Adaptive Regression Splines (MARS) . . . 174
B.1.3.1 Tuning MARS . . . 175
B.1.4 Support Vector Regression (SVR) . . . 175
B.1.4.1 Tuning SVR . . . 177
B.1.5 Random Forests . . . 177
B.1.5.1 Tuning Random Forests . . . 177
B.1.6 Weighted pattern-based model for water demand forecasting . . . 178

C Conclusiones 181
C.1 Conclusiones . . . 181
C.1.1 Clusters de abastecimiento de agua mediante aprendizaje semi-supervisado . . . 182
C.1.2 División-agente de los sistemas de distribución de agua en clusters de abastecimiento . . . 183
C.1.3 Remuestreo multi-agente de clusters semi-supervisados . . . 183
C.1.4 Gestión de la red dividida en clusters de abastecimiento . . . 184
C.1.4.1 Modelos predictivos . . . 184
C.1.4.2 Detección de anomalías . . . 184
C.1.4.3 Importancia relativa de los nodos . . . 185
C.1.5 Software empleado y desarrollado . . . 185
C.2 Trabajo futuro . . . 185

C.2.1 Clusters de abastecimiento de agua mediante aprendizaje semi-supervisado . . . 186
C.2.2 División-agente de los sistemas de distribución de agua en clusters de abastecimiento . . . 186
C.2.3 Remuestreo multi-agente de clusters semi-supervisados . . . 187
C.2.4 Gestión de la red dividida en clusters de abastecimiento . . . 187

D Conclusions **189**
D.1 Conclusions . . . 189
D.1.1 Clusters de proveïment d'aigua mitjançant aprenentatge semi-supervisat . . . 190
D.1.2 Divisió-agent dels sistemes de distribució d'aigua en clusters de proveïment . . . 191
D.1.3 Remostreig multi-agent de clusters semi-supervisats . . . 191
D.1.4 Gestió de la xarxa dividida en clusters de proveïment . . . 191
D.1.4.1 Models predictius . . . 192
D.1.4.2 Detecció d'anomalies . . . 192
D.1.4.3 Importància relativa dels nodes . . . 193
D.1.5 Software emprat i desenvolupat . . . 193
D.2 Treball futur . . . 193
D.2.1 Clusters de proveïment d'aigua mitjançant aprenentatge semi-supervisat . . . 194
D.2.2 Divisió-agent dels sistemes de distribució d'aigua en clusters de proveïment . . . 194
D.2.3 Remostreig multi-agent de clusters semi-supervisats . . . 194
D.2.4 Gestió de la xarxa dividida en clusters de proveïment . . . 195

VI References **197**

References **199**