Optimising a Fuzzy Logic Real-Time Control System for Sewer Flooding Reduction using a Genetic Algorithm

W. Shepherd[1], S. R. Mounce[1], S. Ostojin[2], M. Abdel-Aal[1], A. Schellart[1], P. Skipworth[2] and S. Tait[1]

[1] Pennine Water Group, Department of Civil and Structural Engineering, University of Sheffield, Sheffield, S1 3JD, UK

[2] Environmental Monitoring Solution, Unit 7, President Buildings, Savile Street East, Sheffield, S4 7UQ, UK

[1]*w.shepherd@sheffield.ac.uk*

ABSTRACT

CENTAUR aims to provide an innovative, cost effective, local autonomous data driven in-sewer flow control system whose operation will reduce urban flood risk. The system comprises of a specially designed flow control device and a wireless local water level monitoring and control system. A data driven algorithm has been developed that is able to analyse the water level data and issue instructions to the flow control device to reduce flood risk at the downstream flooding location. This Fuzzy Logic control algorithm has been linked to a SWMM model to allow virtual testing to take place and provide the basis for a Genetic Algorithm to optimise the Fuzzy Logic membership functions. Methods for generating the initial starting membership functions for input to the Genetic Algorithm have also been investigated. Results confirm that the best Genetic Algorithm optimised Fuzzy Logic controllers reduce flood volume by up to 25% depending on the timestep at which the algorithm is run and the membership function initialisation method.

Keywords: Real time control, sewer flooding, genetic algorithm, fuzzy logic.

1 INTRODUCTION

The effects of climate change, population growth and urbanisation are putting increasing pressure on sewer and drainage networks both in the UK and overseas. The capacity of networks to cope with runoff at the required rate often falls short of requirements leading to localised floods and/or increased CSO spills to receiving waters. Smart Water/ Wastewater Network technologies have the potential to deliver improved service to customers and cost-effective performance improvements for the water industry. Innovative solutions such as Real-Time Control (RTC) to optimise the usage of existing capacity in sewer and drainage networks are both timely and required. CENTAUR [1, 2] is a system designed to take advantage of the local unused storage capacity in the upper parts of many networks, thus attenuating the flow at flood-threatened downstream locations. The system uses a low cost Flow Control Device (FCD) and local sensing system combined with a data driven approach based on Fuzzy Logic (FL), which processes local real-time in-sewer level information and reflects expert and local knowledge of network behaviour, which in this work is further optimised by an evolutionary process. FL is particularly suited to wastewater applications, where phenomena can be understood but where their behaviour are characterised by variability. FL algorithms can capture various information including expert knowledge, the conclusions of laboratory and field experiments, and modelling outputs around a particular phenomenon, and can also cope with their variability.

Developing a robust FL system necessitates the understanding of the control mechanism's impact on water levels in the network. The selection of the input variables was conducted such that the FL algorithm can regulate the FCD in order to reduce downstream flood risk while incorporating fail safe rules in order to prevent any additional upstream flooding.

This paper describes the optimisation of the FL Membership Functions (MF) using a Genetic Algorithm (GA). Unlike in previous work where the vertices of the MFs have been predefined manually and based on expert knowledge [3], this optimisation allows the algorithm to be automatically tuned for sites with different characteristics and also enable it to be re-tuned in case of changes within a sewer network. It can also provide a better understanding of the FL controller sensitivity and provide a useful range of MF parameters.

To investigate the effect of different initialisation on the speed of the GA and objective function value, three methodologies are compared. The GA optimisation has been carried out using a design rainfall event which results in a total flood volume larger than the available storage to give the objective function a suitable target. Performance of these three approaches and different time steps / GA settings are presented for the test network.

2 BACKGROUND

Control algorithms based on Artificial Intelligence (AI) have been widely applied in different engineering fields. FL was selected as the most appropriate AI solution for CENTAUR because of the less onerous data needs (the latest generation of deep learning artificial neural networks require 'big data') and the greater transparency of its logic based reasoning than some other AI solutions and also the ability to incorporate expert knowledge. FL can be used to enable a degree of certainty to be placed on a classification or output.

In a FL algorithm, input data are mapped to MFs which relate to a set of predefined rules to create output data. Each input variable can relate to several membership functions and the value of the data is used to assign a degree of membership between 0 and 1 to each relevant MF. FL rules are built from combinations of (or individual) MFs, the input data will have a degree of truth for each rule also between 0 and 1. The outputs of all rules are then aggregated into a single output fuzzy MF. Defuzzification is then applied to the output fuzzy set to obtain a single valued output. FL controllers have been applied in a number of wastewater applications including for treatment [4].

Optimisation of the FL membership functions was investigated through the use of GAs which have been previously applied in similar applications [5]. A GA is a search procedure based on the mechanics of natural selection and genetics [6]. They are highly parallel, mathematical algorithms that transform a set (population) of mathematical objects (typically strings of ones and zeros referred to as genes) into a new population. They function by combining survival of the fittest for individual genes; these are then passed on to the next generation. As the successful (fittest) genes breed over generations they quickly converge to optimal solutions after examining only a small fraction of the search space. Mutation and crossover operations are also included in generations to ensure that a string of genes that may help provide an optimal solution are not lost too early. GAs and other evolutionary algorithms have been successfully applied to many complex engineering optimisation problems and extensively for water resources engineering and management [7]. GAs have the advantage that, in principle, they can be applied to any search space, as long as a point in the space can be represented by a bit string and are particularly useful for optimising problems that are not well formed. However, care must be taken with framing the problem, and constraining the space.

3 METHODS

3.1 FL controller

The FL algorithm uses water level data provided by a local sensing network as input data, the FL rules implement expert knowledge and the output adjusts the setting of the FCD. The selection of the input variables for CENTAUR has to be done such that the FL algorithm can control the FCD to reduce downstream flood risk without causing any additional upstream flooding. Level data is recorded at the downstream flood location and upstream of the FCD. The CENTAUR FL algorithm uses four sets of input data, each has 3 MFs with triangular or trapezoidal shapes defined by 11 vertices in total. The MFs for one of the level data inputs are: Normal (N), High (H) or Very High (VH), as shown in Figure 1, these MF labels give a textual definition of the categorisation, e.g. Normal represents a normally expected depth of flow in dry weather and minor (low return period) rainfall. The output variable, Change Position (CP), has five MF labels, corresponding to changes in the FCD position, e.g. Big Close (BC), Small Close (SC), Zero change (Z). This output is used to adjust the FCD by the given percentage per minute until the next run of the FL.

The FL rules are expressed in the form of IF-THEN fuzzy rules written using expert knowledge. This expert knowledge takes into account the expected response of drainage networks during and after rainfall events and also the concept of when the FCD should activate – e.g. when the water level at the downstream location is close to flooding, close the FCD. The CENTAUR FL control algorithm was developed with the MATLAB Fuzzy Logic toolbox, it uses the Mamdani method [8] and applies the min-max-centroid method. In each rule a minimum value is selected among the input MFs involved in the IF-part. In this way the strength of the rule is formed which defines the intensity of the MF of the output variable involved in the THEN-part. Output MFs of every rule are combined by the maximum value. The output value for CP is calculated by using the centroid method for defuzzification. The FL output is a crisp value but not an integer, post processed as needed.

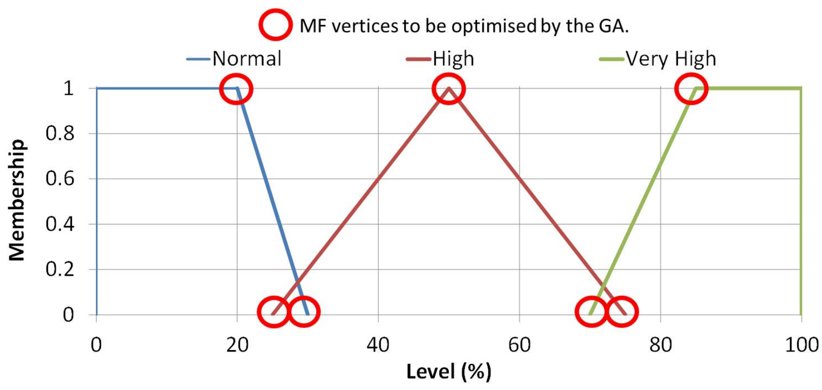


Figure 1 An example of the MFs for level data

3.2 Sewer network modelling

The water levels used as input to the FL during the GA optimisation are generated by a SWMM hydrodynamic model which is controlled from Matlab using MatSWMM [9]. This 'virtual testing' modelling methodology and the sewer network used are as described in [3]. In this paper, a single Flood Estimation Handbook (FEH) design rainfall event of 5 years return period (20% Annual Exceedance Probability) and 120 minutes duration (M5-120) has been used. FEH design events are Gaussian shaped rainfall time series which are scaled to different return periods and duration.

3.3 GA optimisation

A software module to optimise the FL input MFs has been developed using the Global Optimisation toolbox. As detailed in 3.1, the rules and the shapes of MFs were predetermined based on expert knowledge. MFs were chosen as the decision variable for the GA as during manual tuning of the FL algorithm it was found to be sensitive to changes in the locations of the input MFs vertices. Conversely the rule base design was based on expert knowledge of the problem domain and the rules followed robust logic with little flexibility for any reasonable further adjustment.

The GA optimisation module is therefore used to determine the optimum locations of the vertices of the MFs' relevant edges for the four sets of input data, the vertices which are part of the optimisation are highlighted for one of the level MFs in Figure 1, each input dataset has 7 MF vertices which can be optimised, hence a total of 28 values are optimised. The objective of the GA is to minimise the total flood volume.

To investigate the influence of the starting point on the rate of convergence and hence reaching an optimal solution in the quickest possible time, three methodologies were compared: the default (expert) set of MFs, a randomised set and a pseudo-randomised set (perturbation of the expert set). Figure 2 shows a flow diagram of the GA optimisation module for the randomised and pseudo-randomised sets, the default set has a single starting point which is the result of manual tuning of the input MFs and hence skips from the first step straight to 'Run main GA …'. The starting point for the randomised and pseudo-randomised sets is selected from the results of 10 mini-runs as described in Figure 2. Prior to running the FL, the randomised / pseudo-randomised values are first sorted to maintain the shape and cross-overs (unless otherwise specified). Appropriate lower / upper bounds, linear inequality constraints and tolerance checks were implemented at this stage. The M5-120 rainfall event was selected for the optimisation because it results in a total flood volume larger than the available storage, thus giving the objective function a suitable target. The GA stop criteria is based on having no improvement in the objective function for a number of generations.

The results presented in this paper are from 25 generations with a population size of 200, a stall limit of 5 generations with a function tolerance of 1×10^{-6}, there has been some experimentation with these GA parameters and these values were found to be a useful compromise between run times and improvement in the objective function.

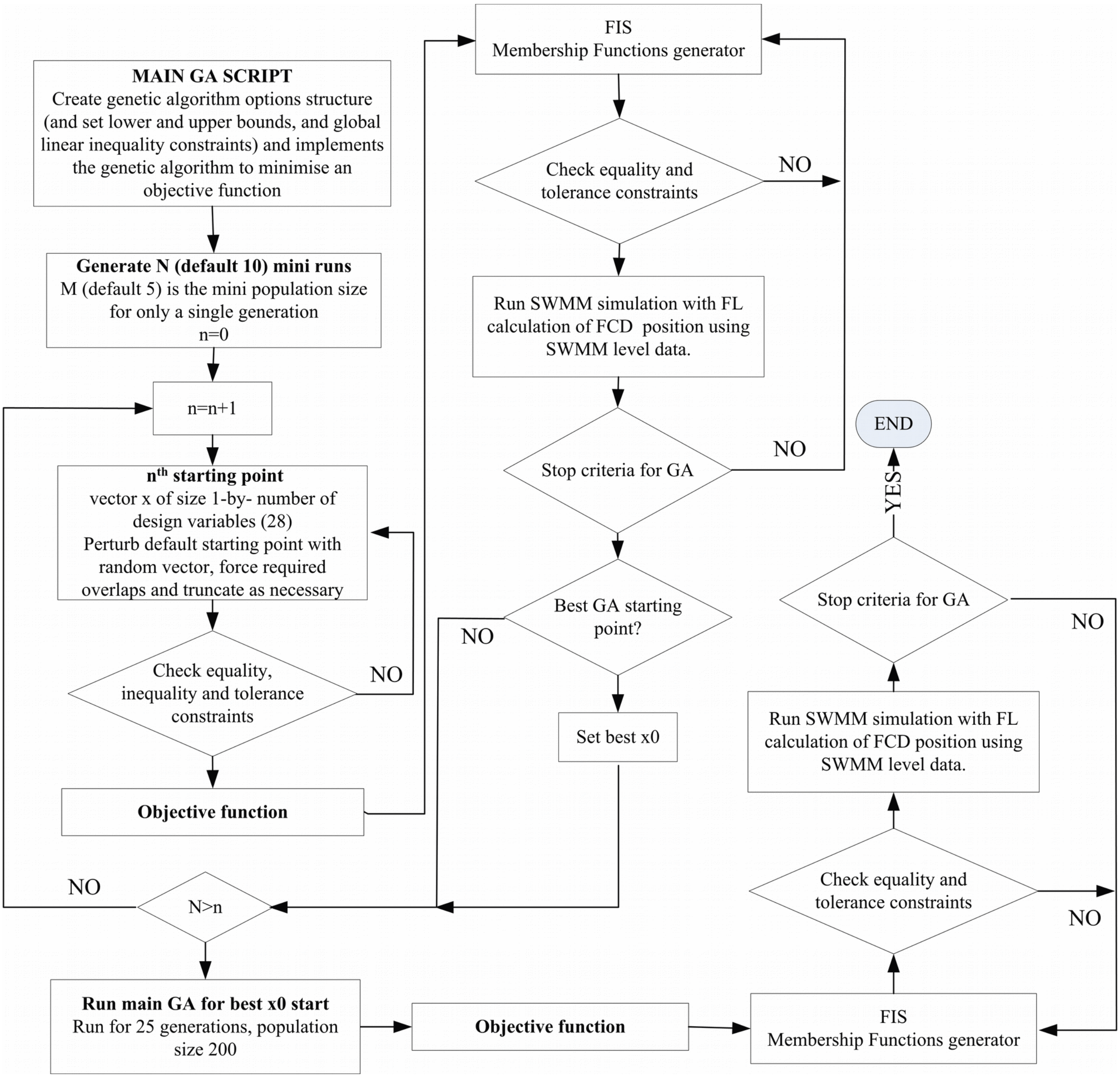
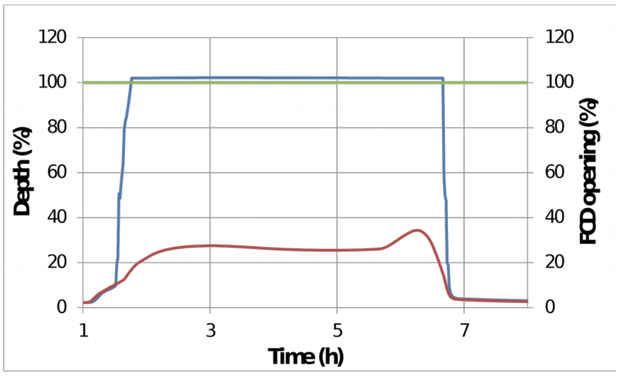


Figure 2. Flow diagram for GA optimisation module

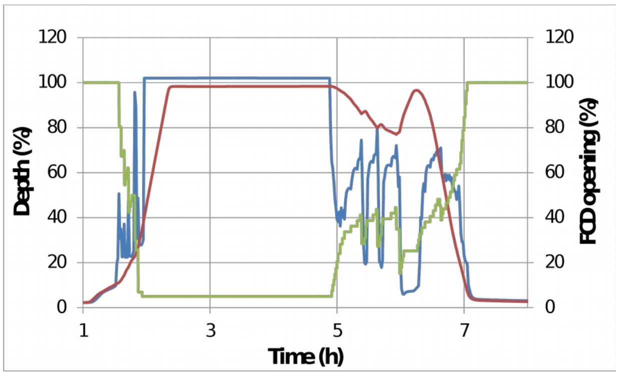
4 RESULTS AND DISCUSSION

4.1 Results

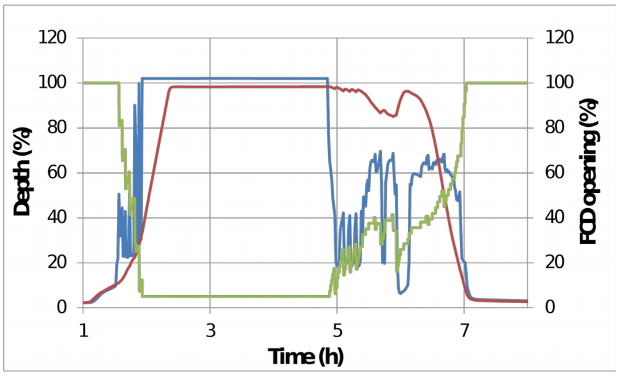
Figure 3 presents time series results for four control cases using the network model and rainfall event described in 3.2, these results all use a 1 minute FL time step (i.e. the FL algorithm runs and can adjust the FCD once per minute). The water depth in the sewer is presented as a percentage, where 100% is the ground level and hence values greater than 100% indicate flooding, an FCD opening of 100% indicates the FCD being fully open. For the control case when the FCD remains fully open (Figure 3a), it can be seen that flooding occurs for 4.9 hours and results in a total flood volume of 247.14 m^3. Using the manually tuned FL (Figure 3b), the flooding duration drops to 2.9 hours and the volume to 111 m^3. The optimisation using the manually tuned FL as the starting point (Figure 3c) reached a stall point after 15 generations, the duration of flooding is still 2.9 hours, but there is a small improvement in the volume of flooding to 108.4 m^3. Finally, the randomised starting point (Figure 3d) reduces the flooding duration to 2.8 hours and the volume to 103.0 m^3.



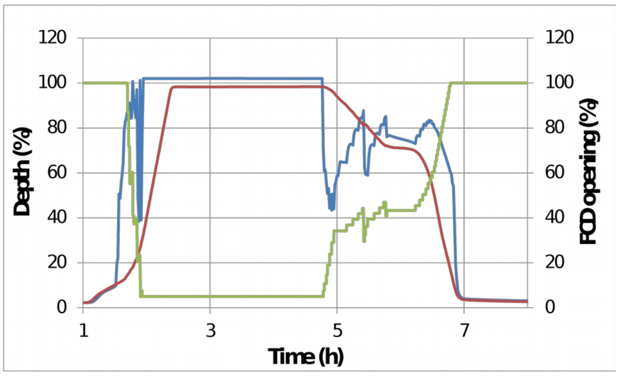
a) FCD inactive (control).



b) Default expert, un-optimised.



c) Default expert, single start optimised.



d) Random, sorted, forced cross-over.

—Flood Location Depth —Depth Upstream of FCD —FCD Opening

Figure 3. Time series results for control, un-optimised and optimised cases, FL run at 1 minute step

Table 1 presents the total flood volume for a greater range of starting points and includes results for the three different FL time steps investigated. Also included is a completely randomised starting point (i.e. not sorted or forced cross-overs), which results in a relatively early stall and the minimum flood volume is the highest of all optimised versions, however it does beat the manually tuned MFs which resulted in a flood volume of 141.2 m^3 for the 5 minute FL step. For reference, the manually tuned MFs give a flood volume of 113.4 m^3 for the 2 minute FL step.

Table 1. GA results of total flood volume for varying time steps and starting points

Starting point	FL time step (min)	Generations run	Flood volume (m^3)
Completely randomised	5	9 (Stall)	122.1
Default expert	5	25	114.9
Default expert	2	25	109.3
Default expert	1	15 (Stall)	108.4
Randomised	5	13 (Stall)	107.0
Randomised	2	21 (Stall)	103.7
Randomised	1	11 (Stall)	103.0
Pseudo-random	5	22 (Stall)	106.4
Pseudo-random	2	23 (Stall)	104.1
Pseudo-random	1	9 (Stall)	103.1

4.2 Discussion

The optimisation results in a flooding reduction of 2-25% when compared to the flood volumes resulting from the un-optimised expert membership functions. The larger benefits (14-25%) are however seen with the 5 minute FL time step which has the largest un-optimised flood volume. For the 1 and 2 minute FL time steps, the improvement is between 2 and 9%. While the un-optimised results show that the control algorithm is sensitive to the time step at which it is run (this could be expected as the FL time steps investigated are within the same order of magnitude as the time over which significant changes in flow could occur), the optimisation can reduce the impact of running the FL algorithm less frequently, which may be desirable to optimise battery life.

Investigating the different staring points indicates that the random and pseudo-random approaches converge more quickly to a better solution than the default expert starting point, the resulting flood volumes being 5-7% better than the default expert starting position. The sorting with forced cross over appears to benefit the final solution, the completely randomised starting point gives a 6% higher flood volume. The number of generations to reach a stalling point varies inconclusively between the randomised and pseudo-randomised starting points, however both do appear to reach the final solution more quickly than the default expert starting point.

This paper presents the initial work on the GA development, there is clearly significant further work to be undertaken, including use of non-Gaussian rainfall events, using multiple events in the GA and applying the GA to different drainage networks. It is also intended to investigate the inclusion of output MFs in the optimisation (total 38 variables), to test different GA options, such as the algorithmic cross-over functions, and to investigate multi-objective optimisation (e.g. include number of FCD movements, total time storage is used, total flood duration, etc.). The CENTAUR system has been installed at a pilot site in Coimbra, PT, hence this will provide an alternative network and an opportunity to test the optimised MFs in a live situation.

5 CONCLUSIONS

The CENTAUR FL control algorithm has been constructed to autonomously adjust the FCD opening in order to reduce downstream flood risk without increasing upstream flood risk. A genetic algorithm optimisation tool which uses the outputs from a SWMM hydrodynamic model has been written to optimise the input MFs of the control algorithm. Results confirm that the best GA optimised input MFs result in up to a 25% decrease in flood volume compared to the MFs selected by experts, although the benefit is significantly higher for the longer FL time step of 5 minutes than 1 and 2 minute time steps. The randomised and pseudorandomised approaches for the initial starting point give 5-7% lower flood volumes than using the expert default MF starting position.

Alongside computer model studies, CENTAUR is being developed and tested in a full scale laboratory facility and is currently being installed into a live sewer network in Coimbra, PT. Additional information on the CENTAUR project is available at www.shef.ac.uk/centaur.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 641931. The contents of this paper reflect the view of the authors, the Executive Agency for Small and Medium-sized Enterprises (EASME) of the European Commission is not responsible for any use that may be made of the information it contains. The authors gratefully acknowledge the support of Innovyze® for the supply of InfoWorks® CS research licences.

References

[1] M. Abdel-Aal, W. Shepherd, S. Mounce, S. Ostojin, A. Schellart, J. Shucksmith, P. Skipworth and S. Tait, "Alleviating The Risk of Sewer Flooding using Fuzzy Logic in a Real Time Control System – An Experimental Study" in *SPN8: 8th International Conference on Sewer Processes and Networks*, Rotterdam, The Netherlands, 2016, DOI: 10.5281/zenodo.399119

[2] M. Abdel-Aal, W. Shepherd, S. Ostojin, S. Mounce, J. Shucksmith, A. Schellart, P. Skipworth, J. Steinhardt and S. Tait, "Developing and testing a fuzzy logic algorithm to alleviate the risk of flooding by controlling a flow control device in a laboratory setting", submitted to 14th International Conference on Urban Drainage, Prague, Czech Republic, 2017.

[3] W. Shepherd, S. Ostojin, S. Mounce, P. Skipworth and S. Tait, "CENTAUR: Real time flow control system for flood risk reduction", in *CIWEM Urban Drainage Group Autumn Conference*, Blackpool, UK, 2016, DOI: 10.5281/zenodo.400969.

[4] G. Vijayaraghavan, and M. Jayalakshmi, "A Quick Review on Applications of Fuzzy Logic in Waste Water Treatment," *IJRASET*, vol. 3 no. 5, pp. 421-425, 2015.

[5] S. Ostojin, S.R. Mounce, and J.B. Boxall, "An artificial intelligence approach for optimising pumping in sewer systems," *J. Hydroinf.,* vol. 13 no. 3, pp. 295-306, 2011.

[6] D.E. Goldberg, *Genetic algorithms in search optimization and machine learning*, Addison-Wesley, Boston, 1989.

[7] J. Nicklow, et al., "State of the art for genetic algorithms and beyond in water resources planning and management," *J. Water Resour. Plann. Manag.*, vol. 136 no. 4, pp. 412–432, 2010.

[8] E.H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *Int. J. Man-machine Studies* vol. 7, pp. 1–13, 1975.

[9] G. Riaño-Briceño, J. Barreiro-Gomez, A. Ramirez-Jaime, N. Quijano and C. Ocampo-Martinez, "MatSWMM – An open-source toolbox for designing real-time control of urban drainage systems", *Environ. Modell. Software*, vol. 83, 2016, DOI: 10.1016/j.envsoft.2016.05.009.