Artificial Intelligence for Wastewater

Professor Pete Skipworth*^, Dr Sonja Ostojin^

**Leeds Sustainability Institute, Leeds Beckett University*

^Environmental Monitoring Solutions Ltd

Many people think of artificial intelligence, or AI, as the *emulation* of human intelligence. This creation of a machine that has common sense (the ability to perceive, understand, and judge things) is known as artificial *general* intelligence, or AGI. Think Blade Runner!

As a science, artificial intelligence – a term coined by John McCarthy in 1955 - is generally recognised as the ability of a computer programme to think and learn; to *imitate* human cognition

The AI that is becoming more prevalent in wastewater - for example fuzzy logic (FL), artificial neural networks (ANN) - allows us to make computers, via carefully designed and enacted mathematical algorithms, capable of specific tasks. This is often referred to as *narrow* AI. The algorithm outputs can be translated into useful operational information to support decisions, or can directly enact decisions.

AI is useful where processes have recurrent patterns that are understood to be complex and variable; ANNs can be trained to learn patterns in these circumstances. Where human knowledge is available but processes are shrouded in variability, FL can be used to rationalise decisions. In summary, wastewater processes – in the network and at the treatment works - are a good fit for the use of AI.

AI has been applied in wastewater in; the detection of CSO failure; the detection of blockages; pump station control and optimisation; the in-line control of pH neutralisation; treatment process control for additives; energy and nitrogen removal optimisation for aeration; control of activated sludge plants; and in investment and operational planning.

This paper discusses the application of three types of AI in wastewater:

- artificial neural networks
    - generally used to learn from "big data" and then to predict or detect events
    - generally used where phenomena are predictable but poorly understood or complex.
- fuzzy logic
    - generally used where expert knowledge is available, and
    - where phenomena are understood in principle but are variable
- genetic algorithms
    - used for optimisation of decisions or systems
    - often used in combination with ANNs and FLs; often for their refinement

The paper considers where each of these is useful - alone or in combination - and where they have drawbacks and advantages in particular applications. It considers the emergence of these types of AI from academia and their practical application for business benefits, citing examples.

The aim of the paper is to promote understanding and uptake such that the very significant business benefits can be realised.

Artificial intelligence it is only termed "artificial" because the intelligence isn't from a biological being. Machine intelligence is arguably a more appropriate but less well used term as it captures this distinction.

Types of AI

The main distinction here is between "weak AI", more usefully termed "*narrow* AI", and "strong AI". This is a distinction between non-sentient machine intelligence used for a narrow task, and sentient machine intelligence with consciousness and mind. The latter remains hypothetical.

People have extended the "strong AI" definition by talking about "artificial general intelligence" (AGI), where a machine has the ability to apply intelligence to any problem. AGI is generally taken as meaning to be at least as smart as a human. Superintelligence is beyond this, referring to AI which surpasses the brightest human intelligence.

However, strong AI isn't on the horizon and isn't currently relevant to the water industry. Narrow AI is coming into use in the water industry, and is likely to come into further use in many more applications in the near future.

Approaches

Narrow AI can be usefully classified by approach, and this tends to be what is done in the water industry. The approaches commonly used in the water industry are all branches of soft computing which can be usefully broken down at a high level as follows:

- Machine Learning
    - the field of computer science that gives computers the ability to learn without being explicitly programmed (Arthur, 1959, Koza et al, 1996)
    - includes **artificial neural networks**
- **Fuzzy Logic**
    - the analysis of analogue input values in terms of logical variables that take on continuous values between 0 and 1, in contrast to classical or digital logic, which operates on discrete values of either 1 or 0 (true or false, respectively) (Pedrycz, 1993, Hájek, 1998)
    - or (more digestibly): logic to deal with concepts that cannot be expressed as the "true" or "false" but rather as "partially true"
- Evolutionary Computation
    - optimisation algorithms inspired by biological evolution; they use candidate solutions in population-based trial-and-error computations
    - these algorithms commonly include **genetic algorithms** (Holland, 1975), ant colony, and particle swarm optimisations
- Probabilistic Methods
    - probabilistic models such as Bayesian networks and Markov chain models have been applied to water industry phenomena such as pipe failure
    - however, this is less recognised as AI and is more traditionally recognised as applied statistics
    - there is no further discussion of this type of AI in this paper

"Chaos Theory" and "Creative Computing" are the other branches of soft computing but are not commented here, where we are focussing on AI in the water industry.

Before considering the above identified forms of narrow AI in the water industry, it is worth considering how we encounter them in our everyday lives:

- Artificial Neural Networks
    - Speech recognition commonly uses artificial neural networks (and Markov chain) algorithms in their operation
    - Image recognition for medical diagnostics and smart phone apps that can recognise everyday objects are based on artificial neural networks
- Fuzzy Logic
    - Automatic gear transmission systems use several variables (speed, acceleration, throttle opening, rate of change of throttle opening, engine load) and weight each of these in a fuzzy aggregate to decide on gear changes

    - Antilock brakes, with wheel circumferential speed and vehicle speed as input variables
    - Interpretation of hand writing
    - Washing machines, which sense the load size, detergent amount and then track water turbidity and make decisions on this
    - Television: ambient lighting, time of day and so on to adjust parameters such as screen brightness, colour, contrast and sound
    - Criminal search systems: combining photo and descriptive analysis (tall, young)
    - CGI – huge scale armies created to have random yet orderly movements
- Genetic Algorithms
    - Mostly used "off-line" in design applications (Brainz published a useful list of 15 applications: https://www.brainz.org/15-real-world-applications-genetic-algorithms/)
        - e.g. aircraft wing design
        - e.g. optimising trading strategies
        - e.g. product design to fit the market

AI in Water and Wastewater

Many applications in the water sector, e.g. Skipworth et al, 1999, Huntingdon et al, 2001, remained academic rather than being operationalised. In the last 5 years or so this has changed. AI applications are becoming operationalised and are offering new possibilities and efficiencies to water operators.

Because of their relative prevalence in the water industry, and in particular in wastewater applications, the following are considered:

- Artificial Neural Networks (ANN)
- Fuzzy Logic algorithms (FL)
- Genetic Algorithms (GA)

Application of Artificial Neural Networks in Wastewater

ANNs are generally used to learn from "big data" and then to predict or detect events. They are used where phenomena are poorly understood or very complex but obey patterns.

Data needs to be available of a sufficient extent and quality that an algorithm can be designed to learn from it and to sufficiently generalise within the overall landscape of possibilities. A downfall of ANNs, therefore, can be insufficient data, or a bias towards data which focusses on specific areas of the overall landscape.

A further downfall can be a change in an aspect of a system which is not picked up in the original input data. Take the example where an ANN learns aspects of the behaviour of a wastewater treatment works (WwTW) in order to support operational decision making or detect anomalous behaviour. If the WwTW is significantly changed, then the past learning will be largely irrelevant and the ANN would need to re-learn from newly collected data. This limits the application of ANNs.

ANNs have been used to replicate the behaviour of complex models which are too slow to run for real-time application, for example in urban flood prediction (Duncan, 2011). ANNs have been used in the prediction of CSO discharges (Mounce et al, 2014). ANNs combined with a Multi-objective Genetic Algorithm were used for the integrated management of three interconnected reservoirs previously operated in isolation (Pianosi and Galelli, 2010). These examples show the range of their applicability.

ANNs are being used to tackle emerging climate related problems. Urban flooding is an increasingly prevalent side-effect of climate change and urbanisation. Hence, the focus on modelling has increased. ANNs have been used to model and predict water levels, flow rates and flood volumes in both fluvial environments and in wastewater networks. For example, Campolo (2003) used ANNs to model flow rates of the River Arno at Firenze and successfully predicted floods up to 6 hours in advance, which was an operationally useful warning period.

Application of Fuzzy Logic in Wastewater

In wastewater, FL has been used where expert knowledge is available, and where phenomena are understood in principle but the outputs are variable.

Pedryz (1993) observed that although alternative approaches such as GAs and ANNs can perform just as well as FL in many cases, FL has the advantage that the solution to the problem can be cast in terms that human operators can understand, so that their experience can be used in the design of the controller. This makes it easier to mechanize tasks that are already successfully performed by humans.

For these reasons, FL is finding many applications in wastewater, especially those involving an element of control.

In wastewater, FL has been widely used in control applications, e.g. pump station control and optimisation of energy use (Ostojin et al., 2011), control of additives in treatment, control of an activated sludge plant, energy saving in the aeration process (Ferrer et al., 1998), in-line control of non-linear pH neutralization, optimisation of nitrogen removal and aeration energy consumption in wastewater treatment plants. FL has also been used in blockage detection, state estimation in anaerobic wastewater treatment (Murnleitner et al., 2002), CSO performance optimisation and management in near-real-time (Mounce et al., 2014).

Application of Genetic Algorithms in Wastewater

GAs are generally used off-line in the optimisation of decisions or systems. This might be in model calibration, or in asset investment optimisation looking over extended time periods.

For network design and rehabilitation planning, much of this began in water distribution networks and progressed then onto wastewater networks and then from infrastructure to non-infrastructure. For example, Savic and Walters (1997) used GAs for the least cost design of a distribution network, and Dandy and Engelhardt (2001) used them for the optimal scheduling of water pipe replacement. Similar studies were published for sewer network design (Weng and Liaw, 2005) and for sewer network rehabilitation (Ward et al, 2014). GAs are now in common use in asset investment planning in the water industry.

Rauch and Harremos (1999) report a number of applications of GAs for rainfall-runoff calibration problems.

Mackle et al (1995) used a simple GA for pump scheduling in a water supply system; their objective was to minimize the costs of pumping by taking advantage of low cost electricity tariffs and additional storage in the system. In a similar vein, Boulos et al (2001) used GAs to develop an optimal operational policy for each pumping station in a water distribution network to minimise energy costs for pumping. The GAs were necessarily used offline but the advantages coming from them were built into operational policies.

In Hajda et al (1998), GAs were used in combination with ANNs for applications in wastewater systems control. As GAs are likely to perform too slowly for online application, ANNs were trained to approximate GA results.

GAs are now commonly used in modelling packages such as Infoworks for optimisation problems, for example in looking at management strategies for large networks.

Case Study: CENTAUR

CENTAUR is an intelligent autonomous system for local urban flood risk reduction. It utilises untapped network capacity. It does this through the operation of a gate to control flow based on an intelligent algorithm which leverages local water level data. CENTAUR is a self-managing, easily deployed system, which can be much less costly than capital and space intensive solutions such as storage tanks, enlarged sewers, or blue-green drainage systems.

In beta format, as shown in Figure 1, it is operational in Coimbra in Portugal, contributing to flood protection of a World Heritage Site. Further implementations are planned in Toulouse, France in parallel to full market launch.

The control intelligence for CENTAUR is a fuzzy logic algorithm. GAs are being trialled off-line in the optimisation of the FL algorithm.

It could be said that the increased network handling capacity and financial benefits (capital avoided) come from the algorithm. However, without the robustness of monitoring and communication technologies as the enablers, the effective application of AI would not be possible.

Much of the design of the system has been focussed on reliability and failsafes. The system uses specially engineered communications to guarantee signal without latency. An online dashboard connects to the system hub and gives operational visibility. Although the dashboard isn't necessary for the operation of the system, it introduces convenience features. The gate technology is purpose-designed for the application and for easy deployment. It has physical fail-safes in the form of overflows which keep upstream risk to an absolute minimum. Sensors are designed to give reliable data at low power, with special installation techniques to avoid in-manhole problems. However, if communication or sensor failures should occur the system remains safe.

Although not necessary for the autonomous operation of the system, an online dashboard connects to the system Hub and gives operational visibility (Figure 2).

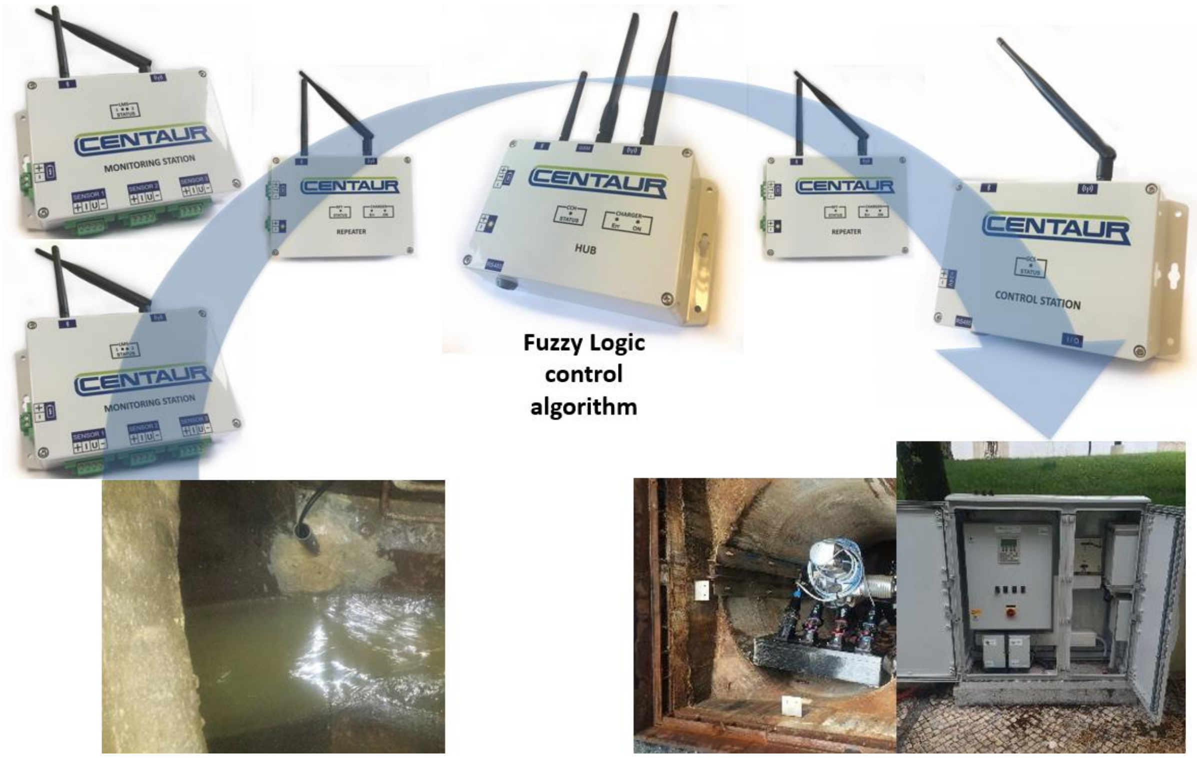

Figure 1: Modular Arrangement of CENTAUR (beta version)

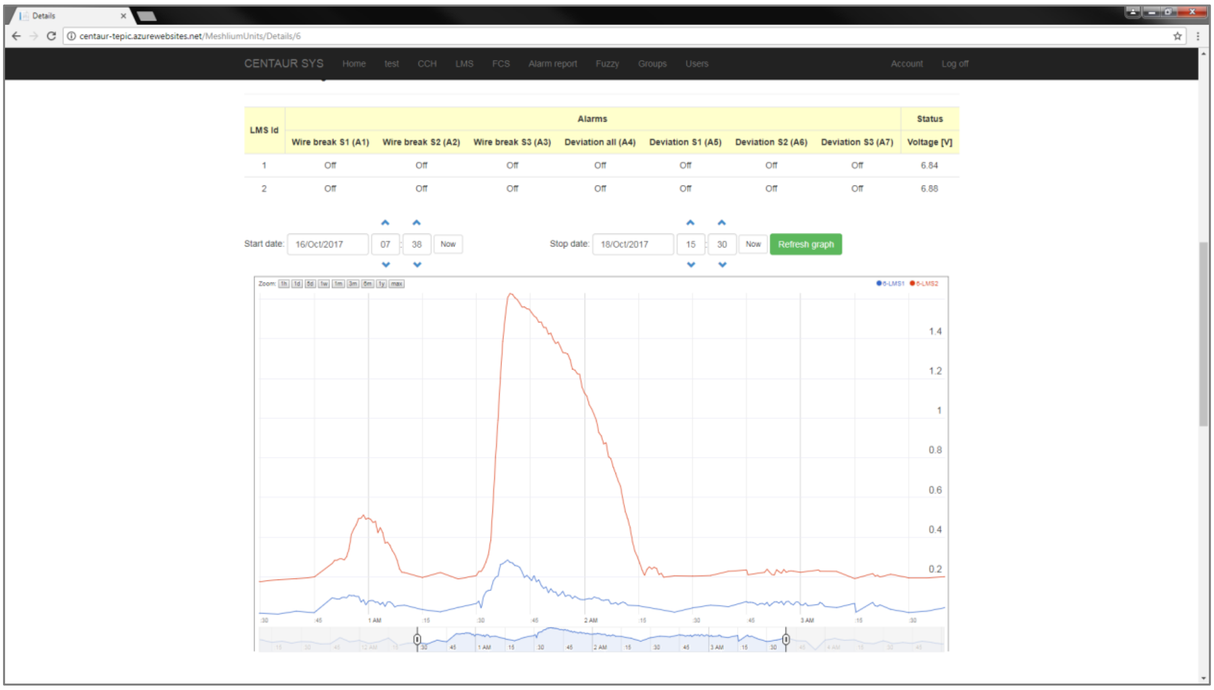

Figure 2: CENTAUR Dashboard for Remote Visibility

The lessons from this case study are in the way the technology has been enacted to leverage the advantages of AI whilst guarding against any downsides. This design and application diligence is as important as the AI itself. With this, it is important to note that the technology has been through a rigorous and expensive development process, pooling expertise in modelling, hardware, communication and software design from across the EU in a Horizon 2020 project.

TRL 1	Basic principles observed and reported.	concept and modelling	2015
TRL 2	Technology concept and/or application formulated.		
TRL 3	Analytical and experimental critical function and/or characteristic proof-of-concept.		
TRL 4	Technology basic validation in a laboratory environment.	prototype development and deployment in lab	2016
TRL 5	Technology basic validation in a relevant environment.		
TRL 6	Technology model or prototype demonstration in a relevant environment.		
TRL 7	Technology prototype demonstration in an operational environment.	prototype in field	Q1/Q2 2017
TRL 8	Actual Technology completed and qualified through test and demonstration.	beta version in field	Q2/Q3 2017
TRL 9	Actual Technology qualified through successful mission operations.	market	late 2017 onwards

Figure 3: CENTAUR Development Path*

* Technology Readiness Levels (TRL) are a method of estimating technology maturity of technologies. The use of TRLs enables consistent, uniform discussions of technical maturity across different types of technology

The Dangers of AI

There are some obvious dangers in the use of AI in that bad decisions can be made unchecked if left to operate autonomously. However, this can only be possible with poor design and poorly considered application of AI. We can see from the case study that good design (hardware/software) can avert such dangers.

Operators may be put off by the newness of AI and be reluctant to take on technologies they don't understand. For this reason, it may be a mistake to put the algorithm "front-of-house". PCs and PLCs in everyday use in the water industry use many clever algorithms, but it isn't necessary for the user to understand these and they aren't put front-of-house when their virtues are under review. Having said this, the human readability of FL may be part of the reason why this type of AI is finding more traction in the water industry.

There have been many misuses of such arguments against automation to satisfy the human tendency to steer away from change, most notably the Luddite fallacy (see below). Where there are significant economic advantages, it may be down to those in power – company leaders, economic regulators – to make sure the potential of AI in the water industry isn't ignored.

> The Luddite fallacy is the simple observation that new technology does not lead to higher overall unemployment in the economy. New technology doesn't destroy jobs – it only changes the composition of jobs in the economy. The introduction of the loom, the printing press, the mechanisation of agriculture, the introduction of the personal computer have all lead to economic growth and the change in the composition of jobs rather than unemployment.
>
> There are those arguing that it is different this time; that AI will actually exhaust what we can do as alternative employment. Just as we went through "peak horse" early in the twentieth century, Chase (2016) argues that we may be now going through "peak human".

Discussion and Conclusions

We are in a notoriously conservative sector. Additionally, the economic regulatory framework for the water industry in the UK struggles to reward technical innovation. Although Ofwat manages a framework of incentives, part of the privatisation framework is to allow water companies to make sufficient profit to raise money on the markets giving them a significant safety net and little reason to properly innovate.

However, the power of these technologies if designed and applied properly can have such large operational and economic advantages that this apathy may be overcome.

To realise the benefits of AI, we need to internalise the power in better products and more robust platforms which have been designed to guard against the downsides of AI. These downsides include the inability to communicate them effectively.

Acknowledgement

The CENTAUR project in the Case Study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 641931.

References

Boulos, P.F., Orr, C.H., De Schaetzen, W., Chatila, J. G., Moore, M., Hsiung, P. & Thomas, D. (2001). Optimal pump operation of water distribution systems using genetic algorithms. In Proc. of AWWA Distribution System Symposium. American Water Works Association, Denver.

Campolo, M., Andreussi, P., and Soldati, A. (2003). Artificial neural network approach to flood forecasting in the river Arno, Hydrol. Sci. J., 48(3), 381-398, 2003.

Chase, C., (2016). The Economic Singularity: Artificial intelligence and the death of capitalism. Three Cs.

Duncan, A., Chen, A.S., Keedwell, E., Djordjevic, S., Savic, D.A. (2011). Urban flood prediction in real-time from weather radar and rainfall data using artificial neural networks, in: IAHS Red Book series no. 351, 58. Presented at the Weather Radar and Hydrology International symposium, International Association of Hydrological Sciences, Exeter, UK.

Dandy, G.C., Engelhardt, M.O. (2001). Optimal Scheduling of Water Pipe Replacement Using Genetic Algorithms. Journal of Water Resources Planning and Management, 127, 214-223.

Ferrer J., Rodrigi M.A., Seco A., Penya-roja, J.M. (1998). Energy savings in the aeration process by fuzzy logic control, Water Science and Tech., 38, (3), pp. 209-217.

Hajda, P., Novotny, V., Feng, X. & Yang, R. (1998). Simple feedback logic, genetic algorithms, and artificial neural networks for real-time control of a collection system. Water Science and Technology, 38(3), 187-195.

Hájek, P. (1998). Metamathematics of fuzzy logic (4 ed.). Springer Science & Business Media.

Holland, J.H. (1975). Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor.

Huntingdon, J.E., Skipworth, P.J., Saul, A.J. (2001). Modelling humic substance derived colouration of surface waters in peatland catchments using artificial neural networks. Earth System Processes, Edinburgh.

Koza, J.R., Bennett, F.H., Andre, D., Keane, M.A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. Artificial Intelligence in Design ’96. Springer, Dordrecht. pp. 151–170.

Mackle, G., Savic, D.A. & Walters, G.A. (1995). Application of genetic algorithms to pump scheduling for water supply. In Proc. of Genetic Algorithms in Engineering Systems: Innovations and Applications, GALESIA'95. IEE Conference Publication 414, 400–405.

Mounce, S. R., Shepherd, W., Sailor, G., Shucksmith, J. and Saul, A. J. (2014). Predicting CSO chamber depth using Artificial Neural Networks with rainfall radar data. IWA Water Science and Technology. 69(6), pp. 1326-1333.

Murnleitner, E., Becker, T.M. and Delgado, A. (2002). State detection of overloads in the anaerobic wastewater treatment using fuzzy logic. Wat. Res., 36, 201–211.

Ostojin, S., Boxall J,B., Mounce S.R. (2011). An artificial intelligence approach for optimising pumping in sewer systems, J. Hydroinf., vol. 13 no. 3, pp. 295-306, 2011.

Pedrycz, W. (1993). Fuzzy Control and Fuzzy Systems (2 ed.). Research Studies Press Ltd.

Pianosi F., Galelli S. (2010). Neural Network and evolutionary algorithm for water resources management, 9th International Conference on Hydroinformatics, HIC 2010, Tianjin, China.

Rauch W., Harremos P. (1999). On the potential of genetic algorithms in urban drainage modelling. Urban Water, 1, 79–89.

Samuel, A. (1959). Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development.

Savic, D., Walters, G. (1997). Genetic Algorithms for Least Cost Design of Water Distribution Networks. Journal of Water Resources Planning and Management, 123, 67-77.

Skipworth, P.J., Saul, A.J., Machell, J. (1999). Predicting water quality in distribution using artificial neural networks. ICE J. Water, Maritime and Energy Vol 136 Issue 1, pp1-8.

Ward, B., Kawalec, M., Savic, D. (2014). An optimised total expenditure approach to sewerage management. ICE J. Municipal Engineer, Vol 167, Issue ME4, 191-199.

Weng, H.T, Liaw, S.L. (2005). Establishing an Optimization Model for Sewer System Layout with Applied Genetic Algorithm. Journal of Environmental Informatics, 5(1), 26-35.