Science of the Total Environment 645 (2018) 449–459

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Performance prediction of an aerobic granular SBR using modular multilayer artificial neural networks

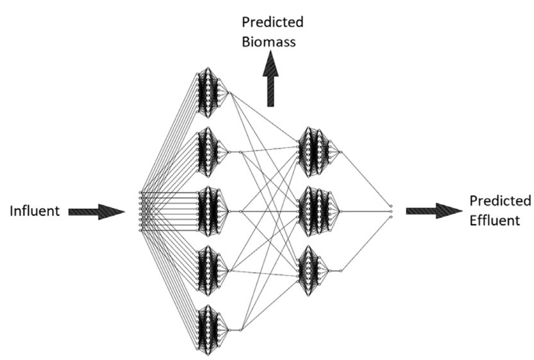
Mohamed Sherif Zaghloul *, Rania Ahmed Hamza, Oliver Terna Iorhemen, Joo Hwa Tay

Department of Civil Engineering, University of Calgary, 2500 University Drive NW, Calgary, AB., Canada. T2N 1N4

HIGHLIGHTS

- A modular neural networks model was developed to simulate aerobic granulation.
- The first sub-model receives 8 influent parameters and predicts biomass characteristics.
- The second sub-model predicts the effluent quality using biomass characteristics.
- A prediction accuracy (R^2 > 99%; RMSE < 5%) was achieved in all predicted parameters.
- This model is the core for adaptive aerobic granular reactors.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 18 May 2018
Received in revised form 10 July 2018
Accepted 11 July 2018
Available online xxxx

Editor: Paola Verlicchi

Keywords:
Aerobic granulation
Modelling
Neural networks
Wastewater treatment
Data-driven modelling

ABSTRACT

Aerobic granulation is a complex process that, while proven to be more effective than conventional treatment methods, has been a challenge to control and maintain stable operation. This work presents a static data-driven model to predict the key performance indicators of the aerobic granulation process. The first sub-model receives influent characteristics and granular sludge properties. These predicted parameters then become the input for the second sub-model, predicting the effluent characteristics. The model was developed with a dataset of 2600 observations and evaluated with an unseen dataset of 286 observations. The prediction R^2 and RMSE were >99% and <5% respectively for all predicted parameters. The results of this paper show the effectiveness of data-driven models for simulating the complex aerobic granulation process, providing a great tool to help in predicting the behaviour, and anticipating failures in aerobic granular reactors.

1. Introduction

1.1. Background

Aerobic granulation has proven to be a more efficient biological wastewater treatment method than conventional aerobic methods (Tay et al., 2001). This is due to the outstanding settleability, high biomass retention, strong microbial structure, high resilience to toxic chemicals, and good ability to handle high organic and shock loading

* Corresponding author.
E-mail addresses: mohamed.abdelsamie@ucalgary.ca (M.S. Zaghloul), rania.sayedeid@ucalgary.ca (R.A. Hamza), oliver.iorhemen@ucalgary.ca (O.T. Iorhemen), jhatay@ucalgary.ca (J.H. Tay).

https://doi.org/10.1016/j.scitotenv.2018.07.140

rates compared to flocculent sludge (Tay et al., 2009; Franca et al., 2018; Nancharaiah and Reddy, 2017). Granulation of biomass is achieved by applying shear stress on the bacteria and enforcing denser particle selection by the means of controlling the settling time of the sequencing batch reactor (SBR) (Zheng et al., 2006). Detailed studies on the start-up and operation of aerobic granulation reactors are discussed in the literature (Su and Yu, 2005; Ni and Yu, 2010; Zhao et al., 2014; Long et al., 2014), often with contradicting results (Ni and Yu, 2010; Khan et al., 2013).

Several factors affect the granular reactor including: influent composition (Beun et al., 2002; Adler et al., 2017), organic loading rate (OLR) (Szabó et al., 2017; Corsino et al., 2018), hydrodynamic shear (Jungles et al., 2017; Sarma et al., 2017), dissolved oxygen (DO) concentration (Morales et al., 2015; Rathnayake et al., 2015), food to microorganism (F/M) ratio (Wu et al., 2018; Wilén et al., 2018), hydraulic retention time (HRT) (Liu et al., 2016), volumetric exchange ratio (Liu and Tay, 2015; Liu et al., 2016), pH (Rezasoltani et al., 2015; Sarma et al., 2017), and temperature (Ab Halim et al., 2015). These factors are not only hard to control, but also inter-connected and affect each other, posing challenges in controlling the process and predicting the performance of the aerobic granular reactor (Ni and Yu, 2010). The complexity of the process is the main cause for the difficulty of scaling up of this new technology (Khan et al., 2013; Sarma and Tay, 2018).

Modelling is a great tool to expand our understanding of the process by providing simulations of the process and predicting performance (Ni and Yu, 2010). This tool can be used in the optimization of the reactor design and operation, and in decision making for operational parameters adjustment based on changes in the feed characteristics or the environmental conditions.

There are two types of models for wastewater treatment: mathematical deterministic models and data-driven models. Mathematical modelling is used in systems that are, or assumed to be, fully understood. Examples of well-developed mathematical models are the Activated Sludge Models (ASM1), (ASM2), (ASM2d), and (ASM3) (Henze et al., 2000). In more complex systems, mathematical models can be challenging in the development and calibration of the differential equations and parameters. They are also usually limited to certain ranges of operational conditions (Henze et al., 2000; Ni and Yu, 2010). Kinetic models are more suitable for understanding the process on a micro level, i.e., the rates of change of bacterial groups, substrate fractions, and by-products. However, they are not very useful if the objective is daily monitoring and control to predict and prevent failures. Data-driven models provide a viable substitute to simulate these complicated processes for that purpose (El-Din et al., 2004).

The majority of aerobic granulation are mechanistic mathematical models based on process kinetics and mass transfer (Xavier et al., 2007; Su et al., 2013). Kinetics are important for explaining the details of the process (Ni and Yu, 2010; Su and Yu, 2006); however, they lack generalization. A kinetic model needs to be re-calibrated for different reactors, wastewater types, or environmental conditions. Kinetic models can also become mathematically complicated and computationally intensive if all aspects of the process were to be included in one model for the simulation. Therefore, the need arises for a different modelling approach.

There is a need to develop a comprehensive and adaptive model that can be used for daily operation of aerobic granulation. Machine learning data-driven models such as neural networks provide a good alternative for mathematical models in this area. Neural networks learn the behaviour of the reactors from historical data and use that to predict future scenarios. They are highly adaptive and generalizable as they can continue the learning process as the database expands with continuous operation.

Neural networks are starting to gain attention as a strong tool to simulate aerobic granulation (Mahmod and Wahab, 2017; Gong et al., 2018). The main advantage of neural networks is that they overcome all the disadvantages of mechanistic models mentioned earlier. Thus, they are more suited for monitoring, control, and failure prevention. Neural networks have also proven to be effective for other wastewater treatment modelling applications, including complicated anaerobic processes (Tay and Zhang, 1999; Elnekave et al., 2012), and plant wide modelling (Hamed et al., 2004; Mjalli et al., 2007; Nadiri et al., 2018).

This work aims at providing a comprehensive, neural network–based, model that predicts the performance of an aerobic granular SBR. The parameters used are commonly measured in any wastewater treatment plant, which makes it easier to collect data. The model predicts the performance in two stages: the first stage is for the prediction of biomass characteristics, while the second is for the prediction of effluent quality. The data used for training the model is collected from three up-flow granular SBRs and the collective data-set contains 2886 cycles of operation, providing a wide coverage of operating conditions and influent characteristics of the aerobic granulation process, with various scenarios of successful operation and failures.

1.2. Modelling techniques

There are other modelling techniques that are gaining interest in the field of water and wastewater treatment: Response Surface Methodology (RSM), and Support Vector Regression (SVR) (Rastegar et al., 2011; Shi and Xu, 2018). RSM is a statistical and mathematical modelling tool that is used for optimization. It requires a set of designed experiments to study the relationships between the input parameters and the response (output) parameters, to determine the best operation/input conditions (Yousefzadeh et al., 2018). Although RSM is widely used in different fields, it is only an approximation where it is required to find the closest functional representation between the inputs and outputs.

Support Vector Machines (SVM) is also another popular modelling technique (Shirzad et al., 2014). When used for regression problems, as predictive models, it is called Support Vector Regression (SVR). SVR is a statistical learning method that can be trained with data like the Artificial Neural Networks (ANN), and can provide comparable accuracy. A study by Gupta (2010), compared the performance of RSM, SVR and ANN, and showed the superiority of SVR and ANN prediction over RSM and multi-variate regression.

In aerobic granulation, these methods are yet to be explored. The objective of each model will be a key factor in the choice of the modelling technique. Aerobic granulation still needs process optimization and the development of predictive models, hence, the use of RSM, SVR and ANN will be important for the future of this area of research.

In this work, ANN were chosen for the simulation over SVR because ANN requires less work for model development. The main work required for ANN is data preparation for the training and testing of the model. However, a comparison of the performance of SVR and ANN is essential and will be presented in future work.

1.3. Artificial neural networks

Artificial neural networks are quite popular for simulating complex processes (McCulloch and Pitts, 1990). They were first developed to simulate how the human brain works, but quickly evolved to a much wider range of applications in all fields of science (Sammut and Webb, 2011). The neural network (perceptron) structure consists of layers: input layer, one or more hidden layers, and an output layer. Each layer contains nodes (neurons) that perform a simple mathematical operation on its inputs. The neurons of each layer are connected to the following neurons via connections (synapses) that carry weight factors which represent the relative significance of the node output to the next operation. The output layer compiles all the data inside the network into the desired output parameters, where every output has its own node (Lippmann, 1987).

Inside the neuron, there are two types of processing functions: propagation (integration) functions, and activation (transfer) function