Artificial Neural Networks Application to Support Plant Operation in the Wastewater Industry

Ivan Pisa, Ramon Vilanova, Ignacio Santín, Jose Lopez Vicario, Antoni Morell

HAL Id: hal-02295245
https://inria.hal.science/hal-02295245

Submitted on 24 Sep 2019

Artificial Neural Networks Application to Support Plant Operation in the Wastewater Industry

Ivan Pisa[1,2], Ramon Vilanova[1], Ignacio Santín[1], Jose Lopez Vicario[2], Antoni Morell[2],

[1] Advanced Systems for Automation and Control (ASAC) Group,
[2] Wireless Information Networking (WIN) Group,
Escola d'Enginyeria, Universitat Autònoma de Barcelona,
08193 Bellaterra, Spain
{Ivan.Pisa, Ramon.Vilanova, Ignacio.Santin, Jose.Vicario, Antoni.Morell}@uab.cat

Abstract. This communication presents the main aim, contextual and development framework of the PhD that is being conducted by the first author. In this PhD, main aim is the application of data driven methods to industrial processes in order to improve and support industrial operations. In this case, Wastewater Treatment Plants (WWTPs) are adopted as the industry where data driven methods will be applied. WWTPs are industries devoted to managing and process residual water coming from urban and industrial areas. Those type of industries apply highly-complex and nonlinear processes to reduce the contamination of water. Therefore, among the different data driven methods, in this PhD we will focus on the application of Artificial Neural Networks (ANNs) in order to improve and support the operations performed in this type of industries. ANNs are considered due to their ability in the modeling of highly-complex and nonlinear processes such as the WWTPs processes.

Keywords: Wastewater Industries, Data Driven methods, Artificial Neural Networks.

1   Introduction

Nowadays, the incursion of Industry 4.0 has arisen as a paradigm becoming true, industrial systems are becoming able to communicate and cooperate with other systems and also with humans by means of Wireless Sensor Networks (WSN) in order to take decentralized decisions to manage the industrial plants [1]. Consequently, large amount of measurements and data from industrial processes will be stored and processed [2]. Thus, the aim of this research is to support and improve the performance of certain industrial processes by means of data driven methods and especially Artificial Neural Networks (ANNs). In addition, these methods will be applied over data coming from Wastewater Treatment Plants (WWTPs) which are industrial plants devoted to reducing the pollutant levels of residual waters.

As introduced, investigated prediction schemes will be based on ANNs, whose most popular tasks when dealing with industrial data correspond to predictive tasks

such as predictive maintenance or anomaly detection [3]. Thus, predictions can support industrial processes operations by providing them with information about future industrial events or measurements. In this work, predictions will support the WWTP plant operations by feeding control strategies providing information to them to anticipate actuations and improve the overall WWTP performance.

2   Relationship to Industrial and Service Systems

As it has been stated, Industry 4.0 is becoming a reality. Industries have been provided with systems able to communicate and cooperate with others. Thus, the connectivity and communication among industrial systems will be performed by means of WSNs which will generate thousands of bytes of information per day. Consequently, data can be processed in order to extract and generate value from it.

In terms of WWTPs industries, the ones in which this work is based, WSN are adopted in order to measure incoming water ratios, pollutant concentrations, oxygen demand and other parameters which characterize the WWTPs' behavior [4]. Therefore, large amount of concentration measurements is stored during a day and consequently, the application of data driven methods to these measurements is widely encouraged in order to extract certain characteristics of the plant. For instance, in [5] the WWTP's influent ratio profile is processed to determine the expected weather and therefore prepare the WWTP to deal with all the incoming flow rate.

In that sense, ANNs are widely adopted as service systems which feed industrial systems or simply obtain certain measurements which gives information about the industry where they are deployed. In addition, their adoption is motivated by the fact that they only require input and output information of the industrial processes to generate a mathematical model able to yield information of the industrial process. This information will be used to complement other systems and consequently, generate value from the stored data. Thus, ANNs can be seen as soft-sensors adopting data-driven methods which work as service systems. Their purpose is to model high-complex and nonlinear systems in order to yield either prediction or classification of process outputs which will fulfill other industrial processes or drive monitoring and decision modules. In that manner, the overall complexity of the industrial processes can be simplified just requiring input and output data of these processes.

3   State-of-the Art

The increasing interest in the Internet of Things (IoT) and Industry 4.0 [6] has motivated the adoption of ANNs for forecasting purposes in industrial systems. ANNs have been considered in different works where predictions are adopted to complement different industrial processes and systems. For instance, the adoption of neural networks as service systems dealing with sewer systems' data is proposed in [7], [8].

In WWTPs context, ANNs have been considered to perform different tasks. For instance, ANNs have been considered as soft-sensors predicting hard-to-measure WWTPs' values. Predictions of measurements such as Chemical Oxygen on Demand

(COD) and aeration tank's dissolved oxygen concentration (SO) are performed in [9] showing a Percentage Error (MAPE) in the prediction around 4.48%.

Other approaches are based on the adoption of ANN in the reduction of pollution effluent concentrations. This process is performed by means of highly-complex and nonlinear biochemical and biological processes which are defined by the Activated Sludge Model No.1 (ASM1) [10]. Those are devoted to reducing the pollutants derived from nitrogen and phosphorus in order to reduce the pollution of the environment where treated water flows. Besides, local administrations have established limits to pollutants such as ammonium ($S_{NH,e}$) and total nitrogen ($S_{Ntot,e}$) present at the WWTP effluent. Violations of them implies economical punishments to the WWTP [11].

In order to avoid violation, predictions can be adopted by different control strategies to let them actuate before violations are really produced. For instance, in [12] a hierarchical control strategy based on Model Predictive Control (MPC) and Fuzzy Logic control are proposed to reduce the concentrations of ammonium ($S_{NH,e}$) and total nitrogen ($S_{Ntot,e}$) concentrations. Both strategies are fed with effluent concentration predictions performed by Multilayer Perceptron (MLP) ANN. They predict the maximum effluent concentration that will be observed during a day. Moreover, MLPs do not consider the time correlation which characterize industrial signals. In [13], the adoption of ANNs to forecast WWTPs effluent concentrations and feed control strategies is presented. However, the considered ANN does not take into account the time correlation between influent and effluent measurements. Instead, it adopts the previously observed effluent concentration as an input.

4 Research Problem

Aforementioned, ANNs predictions feeding some control strategies have been applied in previous works. However, there exist pollutant violations as shown in [12], where the hierarchical control strategy is not able to totally remove the violations of total nitrogen. They are produced during a total of 8.72 days during a year. Some of these violations are produced due to the fact that the control strategy does not have enough time to process the pollutant.

In order to solve this, data driven methods and specially ANNs considering the time correlation between influent and effluent data are proposed in the realization of this PhD. In such a context, those ANNs are adopted at WWTPs industries to complement the control strategies and, therefore improve the performance of the overall WWTP operability exploiting the time-correlation of data. In the case of this work, we will focus on the nitrogen removal process. Thus, $S_{NH,e}$ and $S_{Ntot,e}$ concentrations in the effluent are the process outputs to be forecasted by the considered ANNs. A violation is prone to occur whenever the prediction exceeds the limits shown in Table 1.

Table 1. Effluent limits. Exceeding these limits is punished with high economical fines.

Concentration	Value
Ammonium ($S_{NH,e}$)	*4 mg/L*
Total Nitrogen ($S_{Ntot,e}$)	*18 mg/L*

5   Materials and Methods

In accordance to what has been seen in the previous section, the purpose of this work is to adopt ANNs to perform predictions of the WWTP effluent pollutant levels to complement the control strategies. Thus, the required materials and methods are described in this section: (i) Benchmark Simulation No.2 will be considered in the generation of input and output data, whereas (ii) ANNs and especially Long-Short Term Memory cells will be adopted to perform predictions.

5.1   Benchmark Simulation Model No.2

BSM No.2 corresponds to a simulation scenario which models the default structure of a WWTP [14]. This default structure can be split into two differentiated parts: (i) the water line and (ii) the sludge treatment. The water line corresponds to five biological reactor tanks (Activated Sludge Reactors) where pollutant concentrations reduction process is performed (See Fig.1).
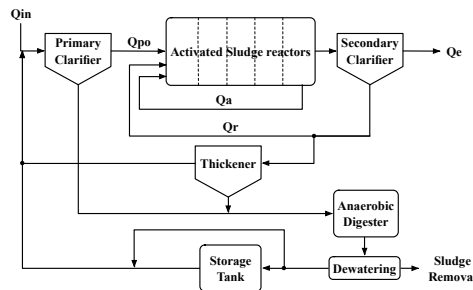
Fig. 1. BSM No.2 model where the different WWTP elements are observable.

First two tanks are anoxic (working without oxygen) whereas the remaining three correspond to aerated ones. An internal recycle flow transporting sludge from the last aerobic tank to the first anoxic one is considered. In addition, two clarifiers where the sludge sedimentation process is performed are considered. Concerning the Sludge treatment part, it corresponds to those elements devoted to treat the WWTPs' sludge. This part encompasses elements such as the Thickener, the Storage tank, the Dewatering and the Anaerobic Digester elements. Different flows are considered in BSM No.2 scenario: Qin corresponds to the influent, Qpo is the primary clarifier overflow, Qa is the internal recycle flow rate, Qr is the sludge internal recycle flow rate and finally Qe is the effluent flow rate.

5.2 Long-Short Term Memory Cells

ANNs consist in a set of layers where each layer considers a certain number of hidden neurons. These neurons are characterized by an activation function which relates the input of the neuron with the output. Each layer is connected to its neighbors by means of weights and biases which are obtained in the ANNs' training process.

Moreover, among the different ANNs, we will consider the adoption of Recurrent Neural Networks (RNN) and specially the Long-Short Term Memories (LSTM) due to their performance dealing with time series and highly-time correlated signals [15]. Thus, if influent and effluent profiles are considered as highly-time correlated parameters, the application of LSTM is widely suitable. LSTMs, as other gated RNNs, are designed to alleviate the well-known RNN's problem: the vanishing and exploding gradients [16]. They replace the RNN's hidden neurons by memory cells with four gates: the forget, the input the state update and the output gate. Forget gate is in charge of resetting the LSTM's memory cell which will be filled with the information coming from the input gate (2) and the state update (3).

$$\mathbf{f_t} = \sigma(\mathbf{W_f}^T\mathbf{x_t} + \mathbf{U_f}^T\mathbf{h_{t-1}} + \mathbf{b_f}) \ . \tag{1}$$
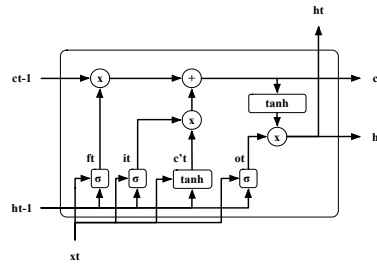
$$\mathbf{i_t} = \sigma(\mathbf{W_i}^T\mathbf{x_t} + \mathbf{U_i}^T\mathbf{h_{t-1}} + \mathbf{b_i}) \ . \tag{2}$$

$$\mathbf{c'_t} = \tanh(\mathbf{W_c}^T\mathbf{x_t} + \mathbf{U_c}^T\mathbf{h_{t-1}} + \mathbf{b_c}) \ . \tag{3}$$

Output gate determines if the output of the memory cell will influence on another LSTM cell or not [18, Chapter 10].

$$\mathbf{o_t} = \sigma(\mathbf{W_o}^T\mathbf{x_t} + \mathbf{U_o}^T\mathbf{h_t} + \mathbf{b_o}) \ . \tag{4}$$

In all cases **W, U** correspond to the weights of each gate, **b** to their biases and **x** and **h** to the input and cell state values. LSTM structure is observable in Fig. 2.

Fig. 2. LSTM Structure. f_t, i_t and o_t correspond to the forget, input and output gate. σ and tanh corresponds to the activation function adopted at each gate. Each gate is formed by different ANNs with the same number of hidden neurons. c_t corresponds to the hidden state of the memory cell. h_t is the cell's output.

Its training process is performed by means of the backpropagation through time algorithm, which is based on the iterative update of the network weights towards the

opposite of the gradient of the loss function w.r.t the weights. Training process is performed considering the well-known split of data (70% for training the ANN and 30% for testing purposes).

6   Expected Outcomes

6.1   Prediction Approach

This PhD is based on the application of ANNs to improve and support WWTPs. Therefore, the reduction of violations and the overall operational cost is sought. A first approach adopting WWTP's effluent predictions corresponds to an ANN-based system, which is able to predict the WWTP's effluent concentrations. It will generate an alarm whenever a violation is predicted. Further details on the ANN-based system can be found in [17]. The steps followed by the system are:

1.   Gather available measurements of the WWTP plant (in real time)
2.   Normalize the data to reach ANN converge towards an optimal solution
3.   The system will order measurements considering the time correlation
4.   Effluent concentrations will be predicted by means of LSTM-based ANNs
5.   Predictions will be contrasted with the effluent limits to determine is violations will be performed or not.
6.   The effluent limits can be modified by WWTP's operator to set the accuracy of alarm detection process by means of ROC Curves.
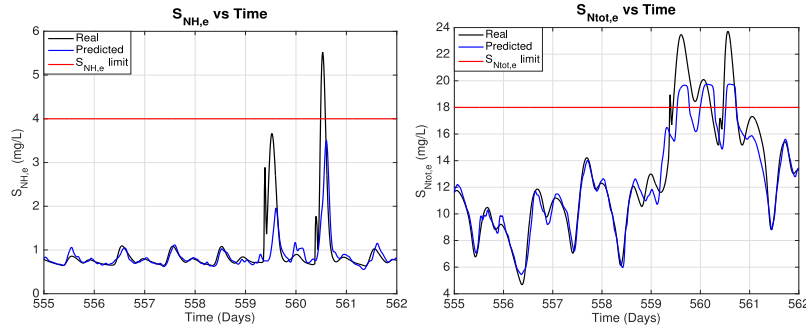
6.2   Prediction Results

Predictions of the proposed ANN-based system [17] are performed by means of two LSTM-based prediction structures: (i) Ammonium prediction structure which considers two stacked LSTM cells with 50 hidden neurons per gate; (ii) and the Total nitrogen prediction structure which considers two stacked LSTM cells with 10 hidden neurons per gate. Table 2 shows the performance of both prediction structures where Mean Absolute Percentage Error (MAPE) and the false positive (P_{fa}) and miss-detection (P_{miss}) probabilities are computed [17]. Predictions are observable in Fig.3

Table 2. Prediction structures' performance.

Prediction Structure	MAPE	P_{fa}	P_{miss}
Ammonium	7.93%	0.03	29.85
Total Nitrogen	4.69%	0.44	31.75

Although performance shows P_{fa} and P_{miss} around 29.85% and 31.75%, they can be enhanced lowering the limits at which effluent predictions are considered a violation. Receiver Operating Characteristic (ROC) curves, where P_{fa} and P_{miss} are faced, showing an Area under the Curve (AuC) around 0.99 (1 means perfect performance).

Fig. 3. Performed predictions. First figure corresponds to Ammonium predictions whereas second figure corresponds to the Total Nitrogen predictions.

Results show that LSTM cells predict accurate effluent concentrations in real time considering the time correlation between values. Thus, predictions performed in previously commented work [12] are improved with the deployment of the proposed system by considering the time correlation between input and output data.

6.3   Future Work

The future research will be oriented along the following main streams:

- On one side, predictions improvement by adopting K-Fold as the training algorithm in order to deal with the data unbalanced problem (the largest part of values is below the effluent limits). Moreover, predictions will be adopted to feed control strategies and consequently reduce the number of violations.
- Direct use of predicted output data for computing of predictive control actions. In this case, the use of data driven predictive control algorithms is sought
- The third step will be to explore the possibility to directly generate control action on the basis of ANN. In this case, ANN could be trained to do the work of the predictive controller or, instead, explore the use of the actor-critic scenario where all the players are implemented as ANN.

7   Conclusions

The adoption of ANNs and especially LSTM cells as service systems applied at industrial environments has been studied in this work. Here, two LSTM-based structures to predict effluent concentrations at WWTPs have been considered. They show a good performance yielding low MAPE values around 4.69% for $S_{Ntot,e}$ and 7.93% for $S_{NH,e}$ concentrations. Moreover, P_{fa} and P_{miss} probabilities and ROC curves have been also computed showing an AuC around 0.99 (1 is perfect performance).

References

1. Sarvari, P.A., Ustundag, A., Cevikcan, E., Kaya, I., Cebi, S.: Technology Roadmap for Industry 4.0. In: Industry 4.0: Managing The Digital Transformation. Springer Series in Advanced Manufacturing, pp. 95--103. Springer, Cham (2018)
2. Rashid, B., Rehmani, M.H.: Applications of wireless sensor networks for urban areas: A survey. J. Net. And Com. Appl. 60, pp. 192--219 (2016)
3. Martí, L., Sanchez-Pi, N., Molina, J.M., Garcia, A.C.B.: Anomaly detection base on sensor data in petroleum industry applications. Sensors 15, pp. 2774--2797 (2015)
4. Fernandez de Canete, J., Del Saz-Orozco, P., Baratti, R., Mulas, M., Ruano, A., Garcia-Cerezo, A.: Soft-sensing estimation of plant effluent concentrations in a biological wastewater treatment plant using an optimal neural network. Exp. Syst. Appl. 60, pp. 8--19 (2016)
5. Qiao, J.F., Hou, Y., Zhang, L., Han, H.G.: Adaptive fuzzy neural network control of wastewater treatment process with multiobjective operation. Neurocomputing 275, pp. 383--393 (2018)
6. Wollschlaeger, M., Sauter, T., Jasperneite, J.: The future of industrial communication: Automation networks in the era of the internet of things and industry 4.0. IEEE Ind. Electron. Mag 11, pp. 17--27 (2017)
7. Zhang, D., Martinez, N., Lindholm, G., Ratnaweera, H.: Manage Sewer In-Line Storage Control Using Hydraulic Model and Recurrent Neural Network. Wat. Resour. Mang. 32, pp. 2079--2098 (2018)
8. Zhang, D., Hølland, E.S, Lindholm, G., Ratnaweera, H.: Hydraulic modeling and deep learning based flow forecasting for optimizing inter catchment wastewater transfer. J. Hydrol., (2017)
9. Güçlü, D., Dursun, Ş.: Artificial neural network modelling of a large-scale wastewater treatment plant operation. Biopro. Biosyst. Eng. 33, pp. 1051--1058 (2010).
10. Henze, M.: Activated sludge model No.1. IAWPRC Sci. Tech. Reports 1 (1987)
11. Vilanova, R., Santín, I., Pedret, C.: Control y Operación de Estaciones Depuradoras de Aguas Residuales: Modelado y Simulación. RIAI – Rev. Iberoam. Autom. e Inform. Ind. 14, pp. 217--233 (2017)
12. Santín, I., Pedret, C., Vilanova, R., Meneses, M.: Advanced decision control system for effluent violations removal in wastewater treatment plants. Cont. Eng. Prac. 49, pp. 60--75 (2015).
13. Foscoliano, C., Del Vigo, S., Mulas, M., Tronci, S.: Predictive control of an activated sludge process for long term operation. Chem. Eng. J. 304, pp. 1031--1044 (2016).
14. Jeppsson, U., et.al.: Benchmark simulation model no 2: general protocol and exploratory case studies. Water Sci. Technol. 56, (2007)
15. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge, Massachusetts. (2016)
16. Bengio, Y., Simard, P., Frasconi, P.: Learning Long-Term Dependencies with Gradient Descent is Difficult. IEEE Trans. Neural Networks 5, pp. 157--166 (1994)
17. Pisa, I., Santín, I., Vicario, J.L., Morell, A., Vilanova, R.: A Recurrent Neural Network for Wastewater Treatment Plant effluents' prediction. In: XXXIX Jornadas de Automática, pp. 621--268 (2018)