A Committee Evolutionary Neural Network for the Prediction of Combined Sewer Overflows

Rosin, T. R.; Romano, M.; Keedwell, E.; Kapelan, Z.

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

A Committee Evolutionary Neural Network for the Prediction of Combined Sewer Overflows

T. R. Rosin[1,2] · **M. Romano[2]** · **E. Keedwell[1]** · **Z. Kapelan[1,3]**

Abstract

Combined Sewer Overflows (CSOs) are a major source of pollution and urban flooding, spilling untreated wastewater directly into water bodies and the surrounding environment. If overflows can be predicted sufficiently in advance, then techniques are available for mitigation. This paper presents a novel bi-model committee evolutionary artificial neural network (CEANN) designed to forecast water level in a CSO chamber from 15 min to 6 h ahead using inputs of past/current CSO level data, radar rainfall data and forecast forecasted rainfall data. The model is composed of two evolutionary artificial neural network (EANN) models. The two models are trained and optimised for wet and dry weather conditions respectively and their results combined into a single response using a non-linear weighted averaging approach. An evolutionary strategy algorithm is employed to automatically select the optimal artificial neural network (ANN) structure and parameter set, allowing the network to be tailored specifically for different CSO locations and forecast horizons without significant human input. The CEANN model was tested and evaluated on real level data from 4 CSOs located in Northern England and the results compared to three other ANN models. The results demonstrate that the CEANN model is superior in terms of accuracy for almost all forecast horizons considered. It is able to accurately forecast the dry weather and wet weather level, predicting the timing and magnitude of upcoming spill events, thus providing information that is of clear use to a wastewater utility.

Keywords Combined sewer overflow prediction · Evolutionary artificial neural network · Radar rainfall nowcasts

✉ T. R. Rosin
  tr319@exeter.ac.uk

1 Centre for Water Systems, University of Exeter, Harrison Building, North Park Road, Exeter EX4 4QF, UK

2 United Utilities Plc, Lingley Mere Business Park, Warrington WA5 3LP, UK

3 Department of Water Management, Delft University of Technology, Stevinweg 1, 2628, CN, Delft, The Netherlands

1 Introduction

Combined Sewer Overflows (CSOs) are designed to prevent properties and sewage treatment works flooding during heavy rainfall. They operate by diverting excess flow to a nearby water body, such as a river, stream or ocean. CSOs are a necessary part of sewage infrastructure. However, overflows contain untreated wastewater from residential, commercial, and industrial sources and can cause significant pollution in the surrounding environment, degrading water quality, threatening public health and causing regulatory failures. Predicted changes to rainfall patterns and intensity due to climate change and urbanisation present a risk of increasing numbers of overflows in the near future. The resilience of wastewater networks against flooding is therefore a growing concern (Salerno et al. 2018).

The Urban Wastewater Treatment (England and Wales) Regulations 1994 places a duty on the Environment Agency (EA) to ensure that pollution due to CSO overflows is limited. The EA operates a consent system that sets out the conditions under which spills are permitted. In recent years there has been growing regulatory pressure to prioritise spill reduction and, in particular, to minimise the number of unconsented spills which occur. The Event Duration Monitoring (EDM) program implemented by the EA, has required utilities to monitor levels at the majority of CSOs in England and Wales since 2020 and report on their performance in terms of the number of discharges. Consequently, there has been an increasing interest in developing methods to accurately forecast CSO performance in near real-time. By predicting overflows in advance preventative measures can be implemented (e.g. maximizing storage, adjusting set points for movable gates and pumping stations) to mitigate negative effects (Joseph-Duran et al. 2014). Even if proactive management cannot be implemented the ability to forecast overflows is still valuable – for example by enabling utilities to post warnings for spills which will affect bathing waters.

Wastewater utilities have traditionally constructed physical models of the sewer system. However, these models require detailed information of the sewer system and incorporate many parameters, are often difficult to build and calibrate, and are computationally expensive. Data-driven models such as artificial neural networks (ANNs) have become an increasingly popular alternative. Driven by the introduction of the EDM project, and in conjunction with decreasing sewer sensor and data storage costs, and improved computer processing power water wastewater utilities routinely collect large volumes of accurate sewer level data in near real time. This detailed data provides the opportunity for real-time data-driven modelling and management of the wastewater system. Data-driven models generally have low computational costs and fast computational times, which is advantageous when applied to real-time data.

Several ANN models have been successfully applied to CSO level forecasting in recent years. Fernando et al. (2007) designed one of the earliest models, a feed-forward, back-propagation ANN to forecast overflows 75 min ahead using flow rate and rainfall data. The ANN generally performed well, however, the results were heavily dependent on the availability of real-time flow data, which is often not available. Kurth et al. (2008) developed a three hidden-layer feed-forward ANN to predict CSO levels 15 min ahead using level and rain gauge data. Sumer et al. (2007) designed an ANN to detect disruptions in sanitary sewer overflows in real time. Mounce et al. (2014) utilised radar rainfall data to predict CSO depth up to 75 min ahead. Zhang et al. (2018) compared the results of four different types of ANNs (a multilayer perceptron, a wavelet neural

network, a long short-term memory network, and a gated recurrent unit), designed to forecast levels simultaneously in multiple CSOs up to 80 min ahead using a deep-learning approach.

More generally in wastewater drainage many ANN models have been applied to areas such as sewer flow modelling (She and You 2019; Zhang et al. 2019), urban flooding (Berkhahn et al. 2019; Kim and Han 2020) and modelling of sediment transport and accumulation (Al-Ani and Al-Obaidi 2019; Ebtehaj and Bonakdari 2014).

The key deficiencies identified in the above literature models are as follows: (a) Whilst the authors used different ANN based models, the majority are based on fixed ANN structures determined using a trial-and-error approach. As such, these models are rather subjective in nature and time-consuming to build and cannot be easily transferred to different CSO sites. It would therefore not be practical for a wastewater utility to develop ANN models for an entire wastewater network, containing possibly hundreds of CSO chambers; (b) The above modelling approaches do not distinguish between dry and wet weather periods, hence the trained ANNs tend to accurately predict CSO water levels for more frequent/everyday rainfall events but do not perform so well for rarer extreme rainfall events; (c) Prediction accuracy of existing methods is reasonably good but could be still improved, as shown in the case study; (d) Existing models provide accurate CSO level forecasts with rather short leading times (up to 80 min, according to our best knowledge) which does not provide the utility with sufficient advance warning for upcoming overflows.

The methodology proposed here aims to make equivalent predictions up to 6 h ahead, thus providing water utilities with additional time to address the problem. The study consists of a Committee Evolutionary Artificial Neural Network (CEANN) model designed to forecast the level in a CSO chamber from 15 min to 6 h ahead. The CEANN aims to improve on these previous models using two techniques. Firstly, an evolutionary artificial neural network (EANN) is utilised, whereby the architecture and parameterisation of an ANN is evolved using an Evolutionary Algorithm (EA). Compared to ANNs, EANNs significantly reduce the manual effort required to construct the model by a human expert, whilst also achieving, and often outperforming, the quality of the results attained. EANNs have been used successfully since the 1990s (Yao 1993, 1999) and in the last decade there has been significant interest in applying EANNs to hydraulic systems (e.g. Chen and Chang 2009; Moradi and Dariane 2017). However, to our knowledge, EANNs have not yet been applied to CSO level forecasting (and is beneficial for the reason mentioned above).

Secondly, the CEANN model consists of two EANNs, optimised for dry and wet weather periods respectively and combined into a single output, with the aim that the combined result is superior to the constituent models. This is employed as previous work by Rosin et al. (2018) demonstrated that ANN models often perform poorly when predicting CSO levels during heavy rainfall events, especially for long forecast horizons. This is due to data bias caused by imbalanced data - wet weather data is significantly underrepresented compared to dry weather data. As wastewater utilities desire to use CSO level forecasting largely to understand CSO behaviour during heavy rainfall and anticipate spill events in advance this is an important issue. The benefit of combining the responses of two EANNs is the reduced bias of forecasted CSO levels.

To analyse the performance of the CEANN model it is compared in this study to three other ANNs:

(i) An ANN model with structure and parameter determined via trial-and-error (representing CSO modelling approaches currently adopted in the literature),
(ii) A single EANN model applied to both wet and dry weather,

(iii)	A model utilising wet and dry EANNs combined using a hard switch between the models rather than a weighted average.

The rest of the paper is structured as follows: after the introduction the methodologies developed for the CEANN and the comparison ANN models are presented. The models are applied to case study CSO sites and the results from the different models are compared and discussed. Finally, conclusions are presented.

2 Methodology

The CSO prediction methodology in this study is composed of a Committee Evolutionary Artificial Neural Network. Individual CEANN models are optimised and trained specifically for different forecast horizons, to ensure that the models are tailored (and thus produce the best results) for each lead time. In this study a total of 24 different models have been constructed with forecast horizons ranging from 15 min to 6 h ahead at 15-min intervals. The CSO levels for the next 6-h period are thus obtained by running all 24 models in parallel.

Different forecast horizons were considered (i) to provide the water utility with as much information as possible and (ii) to analyse the change in model performance when predicting at greater forecast lead times and to understand how far ahead the model is capable of predicting. Note that the model has a forecast range of 6 h as this is the range of the rainfall nowcasts utilised. However, short-range weather forecasts, from 1 to 3 days ahead, are also available, although at a lower temporal resolution. Therefore, a model forecasting further ahead could be investigated in the future.

The following section describe the methods used to construct the CEANN model, and the additional comparison ANN models.

2.1 Artificial Neural Network (ANN) Model

The basic ANN utilised in this study consists of a feed forward ANN with a single hidden-layer trained using the back-propagation method, with a hyperbolic tangent transfer function for the neuron in the hidden layer and a linear transfer function for the neuron in the output layer. This setup was identified by Romano and Kapelan (2014), who developed an ANN model for water demand forecasting, as producing accurate results with fast training times.

The specific inputs and structure of the ANN model are selected individually for each site the model is applied to, in order to ensure good results. Different methods are used to perform this parameter selection process for the various comparison ANN models, described in detail in sections 2.1 to 2.4. The inputs to the model consist of antecedent CSO level data, antecedent rainfall data, and forecast rainfall data. Feed-forward ANNs have no internal memory to store past information and thus cannot process time series data satisfactorily. Therefore, a sliding time window approach is employed, whereby past data is input to the network using a window of lagged data. The size of these sliding windows (i.e. the number of past timesteps of data) are chosen during the selection of the network parameters. This is an important process as selecting insufficient data may mean the model cannot capture the necessary dynamics of the system, however an overlarge window may prolong the learning time and cause unnecessary information to appear as noise.

When using 15-min data the range of window sizes was set as 1–72 timesteps for the level and antecedent rainfall data and 1–24 timesteps for the forecast rainfall data. The maximum window size for the level and antecedent rainfall data was determined based on the times of concentration found during an analysis of CSO sites from Northern England. The maximum size of the forecast data was set at 24 as 6-h forecasts were utilised (i.e. 24 timesteps of 15-min data). However, future models could utilise longer rainfall forecasts if the data were available.

The day of the week and the time of day associated with the forecast horizon are also potentially used as inputs to the model (the inclusion/exclusion of these parameters is determined during the network parameters selection). The time of day and the day of the week are converted into a field representation (i.e. ones and zeros), also called a one-hot vector, to provide a format for categorical data that can be easily input to the ANN (this is not necessary for the other input variables). These two variables are included as flow in sewer pipes display significant hourly and weekday/weekend trends due to varying industrial and residential wastewater production at different times of the day (Butler et al. 1995). These trends are most obvious during dry weather as during rainfall events the CSO level is affected primarily by precipitation in the surrounding catchment.

The data for each input to the model is normalised between −1 and 1 via Min-Max normalisation, which has been demonstrated to improve the accuracy and efficiency of ANNs (Sola and Sevilla 1997).

The number of hidden neurons, the number of training cycles and the coefficient of weight decay regularisation of the ANN are also selected during the selection of model parameters. The range of values for these variables was set as 1–100, 50–500 and 10^{-5}– 10^3 respectively, these values were identified by Romano and Kapelan (2014) as likely to enable finding an optimal solution for the problem considered. The selection of the number of hidden neurons is important as a network with too few hidden neurons may have poor accuracy, however an excessive number may decrease the generalisation ability of the model due to overfitting. The number of training cycles and the coefficient of weight decay regulation are designed to prevent overfitting, by controlling the number of training cycles and applying a penalisation coefficient to the weights of the ANN model respectively.

2.2 Evolutionary Algorithm Optimisation

As explained above when developing ANNs the selection of the model architecture and inputs is of great importance and can have a significant impact on the model accuracy. Different CSOs often exhibit very different behaviours and responses to rainfall events. Indeed, even for the same CSO chamber different forecast horizons require different inputs and structures to produce optimal results. Consequently, using pre-defined parameters can result in poor predictions. However, when forecasting CSO levels on a large scale (e.g., a sewer network containing hundreds of CSOs) the selection of parameters and input structure using manual trial-and-error is not feasible.

Therefore, an EANN model has been employed here. EANNs are a class of ANN where evolutionary algorithms (EAs) are used in the model designing and/or training. EAs are a class of stochastic search and optimisation techniques inspired by the process of natural selection and are designed to perform searches over complex spaces without any substantial human involvement required. The EANN used here employs an evolutionary strategy algorithm (Schwefel 1998) to automatically select the optimal (i.e. that yields the best forecasting performance) ANN input structure and parameter set for a specific set training data

and forecast horizon. The EA methodology is based on the design presented by Romano and Kapelan (2014).

During the EA process the network structure and input parameters, described in Section 2.1, are initially randomly selected. For each cycle of the algorithm the model prediction error (or 'fitness') is computed on the test dataset using the Nash-Sutcliffe efficiency coefficient (NSE). The fittest networks are allowed to survive while the weak networks are replaced, i.e. 'survival of the fittest', thus increasing the overall fitness of the population. The surviving networks reproduce by generating copies of their genotypes with the addition of changes (or mutations) introduced. These new networks then compete with their parents for a place in the next generation. An isotropic Gaussian mutation operator is employed here to introduce the mutations. This is a well-known operator which perturbs each component independently using a random number from a Gaussian distribution with zero mean and constant standard deviation. This operator is advantageous as it is flexible and supports both fine tuning of solutions and searching the domain. The process is repeated for a number of generations until a pre-defined termination criterion is satisfied, resulting in the selection of the parameter combination which produce the lowest error during testing (Benbassat and Sipper 2013).

The parameters used in the EA are as follows: the number of parents per generation and the number of children per generation are set as 10 and 20 respectively. The termination criterion is defined as the number of fitness function evaluations run, i.e. the number of cycles of the EA, and is set as 210. The probability of a parameter being perturbed is set as 0.6. The standard deviation (σ) is set as 0.75, this parameter determines the strength of the introduced mutation - a large value increases the exploration of the search space whilst a small value promotes exploitation of the parent information. Lastly the selection operator is set as '+', meaning both the parent and child networks are included in the selection pool evaluated to pick the fittest networks, rather than only including the newly generated child networks. The range of value for these parameters were identified as producing good results by Romano and Kapelan (2014).

2.3 Committee EANN Model

A committee approach is utilised here to overcome bias caused by imbalanced data. A committee neural network, also known as a committee machine, employs the principle of 'divide and conquer'; rather than using only one ANN the results from multiple networks are combined into a single output, with the aim that the overall result is superior and more robust than any single network acting alone (Tadeusiewicz 1995). Using this approach committee machines can produce significantly improved results with little extra computational effort.

Class imbalance, where the number of training samples of one class are significantly fewer than other classes, is a common but serious problem in machine learning. As ANNs, and most other machine-learning algorithms, aim to optimize overall classification accuracy, the learning algorithm tends to be biased towards the majority class and so the minority class is misclassified. In the case of CSO modelling dry weather dominates and rainfall events are relatively infrequent. A study of radar rainfall data from Northern England found that approximately 98% consisted of rainfall under 1 mm/h. EANN level forecasting models generally produce very accurate results during dry and light rainfall periods, but poor results during heavy rainfall events, especially at long forecast horizons.

The committee machine used here is a bi-model CEANN, consisting of two EANN models constructed for dry and wet weather conditions respectively. The model input data is normalised between -1 and 1 separately for the dry weather and wet weather training sets. This ensures the data is appropriately normalised, despite the large dynamic range that occurs across wet and dry weather periods. Committee ANNs generally combine multiple different models. Here only two networks are used, as the main purpose of the committee is to effectively forecast both dry and wet data. However, in the future additional networks could be incorporated if this were demonstrated to improve the performance.

There are a number of other techniques available to overcome imbalanced data, such as over or under sampling, penalised models and synthetic data generation (He and Ma 2013). However, these approaches become more complex when working with time-series data as the time dependency among the observed values must be considered. The CEANN provides a straightforward, easy-to-use method.

There is no universal definition to categorise dry and wet weather. A binary threshold is used here to classify the data based on cumulative rainfall over a past number of timesteps. The threshold is defined as

$$Z_i = \sum_{t=0}^{n} R_{i-t} \tag{1}$$

$$\begin{cases} \text{Wet if } Z_i > \theta \\ \text{Dry otherwise} \end{cases} \tag{2}$$

where R is rainfall intensity, θ is the wet weather threshold and n is the number of past timesteps considered. θ and n are here set as 0.5 mm and 10 respectively. These values were determined by analysing historical CSO level and rainfall data and identifying for which values rainfall during dry weather had a negligible effect on CSO level.

There are several different methods available for combining the individual EANN model outputs. The most common approach is simple averaging which assigns equal weights to all the component models. Other methods include weighted averaging where the contribution of each model is weighted according to its estimated performance, and majority voting where the utilised result is the one chosen by the most ANNs. For this study, the aim of the model blending is to use the most appropriate model for the current rainfall conditions. Therefore, a weighted averaging approach has been selected, dependant on the rainfall intensity of the current timestep. The models are combined using a non-linear weighted average based on the sigmoid function:

$$S_t = \frac{e^{x_t}}{1+e^{x_t}} \tag{3}$$

$$L_t = S_t W_t + (1 - S_t) D_t \tag{4}$$

where x_t is the cumulative rainfall over a past number of timesteps, L_t is the overall ensemble output, W_t is the output from the wet-weather model and D_t is the output from the dry-weather model, at time t.

A sigmoid function was selected due to its characteristic 'S'-shaped curve which exists continuously between 0 and 1, providing a continuous, or 'soft', transition between the two EANNs. During dry weather (i.e. when $S=0$) only the dry-weather model is utilised, during heavy rainfall (i.e. when $S=1$) only the wet-weather model is utilised, and for all other times a combination of the two models is used, weighted to the rainfall intensity.

2.4 Comparison of ANN Models

To assess the improvements obtained by using a Committee EANN model the results were compared against the following three ANN models:

1. Trial-and-error ANN

ANN models were constructed which used manual trial-and-error to select the model inputs and parameters, in contrast to the EANN models which applied an EA. This type of model is used as representative of existing literature models and hence is a reference point for comparison. Selection of the network structure and parameters was accomplished by constructing models with different configurations of parameters and input structures. The ANN with the best generalisation capacity was selected by identifying, for each parameter, which value produced the best performance across all 24 forecast horizons (evaluated using the NSE). Overall, 15 different ANN configurations were trained and tested for each of the 24 forecast horizons, i.e. a total of 360 models (the results are not shown here due to space restrictions). This was a time-consuming, labour-intensive process, however, it ultimately considered only a small number of all possible parameter configurations. This process was performed separately for each CSO case study site considered, rather than constructing a single set of general parameters for all sites, to ensure good results.

2. All-weather EANN

Single EANN models were trained using the whole dataset, i.e. not optimised for wet and dry weather conditions by using separate wet and dry weather datasets. The models were constructed separately for each forecast horizon and case study site.

3. Combined Wet/Dry Weather EANN

As with the committee EANN, separate wet and dry EANN models were constructed. However, rather than combining the results using a weighted average (as utilised by the CEANN) a discontinuous switching method was used, whereby values for wet timesteps are obtained from the wet model only and dry timesteps from the dry model.

The threshold for designating timesteps as wet or dry is the same as that given above. Unlike the CEANN, where the sigmoid function provides a continuous transition between the two EANNs, this model uses a discontinuous transition, switching from one model to the other. It is possible, therefore, that a small change in output from the two input models could produce a large change in the overall output.

The performance of the different ANN models are evaluated using the following performance metrics: the Nash-Sutcliffe Efficiency Index (NSE), the Mean Absolute Percentage Error (MAPE), the Mean Square Error (MSE) and the Structural Similarity Index (SSIM). The SSIM, also known as the Wang-Bovik index, is a metric originally developed to compare the quality of digital images and videos to quantify image degradation caused by data processing (Wang et al. 2004). The index measures how similar the processed image is from a reference (i.e. perfect) image with respect to the structure within a convolution window. Unlike traditional metrics, the SSIM is designed to capture the perceived structural variation rather than the simple elementwise error.

For images X and Y (computed as matrices of pixels) the SSIM between two windows x and y of common size $N \times N$ is calculated as

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{\left(\mu_x^2 + \mu_y^2 + c_1\right)\left(\sigma_x^2 + \sigma_y^2 + c_2\right)} \tag{5}$$

$$c_1 = \left(k_1 L\right)^2 c_2 = \left(k_2 L\right)^2 \tag{6}$$

where L is the dynamic range of the image and k_1 and k_2 are arbitrarily taken as 0.01 and 0.03 respectively and are used to ensure that near-zero denominators do not cause instability.

The mean SSIM for the global image similarity is then given as

$$\text{mean SSIM}(X, Y) = \frac{1}{n} \sum_{j=1}^{n} \text{SSIM}(x_j, y_j) \tag{7}$$

where the mean SSIM = 1 if the images are identical.

Although the SSIM was developed for image analysis, it can be applied to time-series data by considering the signal as an $N \times 1$-pixel image and using a one-dimensional convolution window over the timeseries. Using this approach SSIM has been used successfully in disciplines such as biological neurogram signals (Hines and Harte 2012), speech signals (Hines et al. 2012) and aeroacoustics (Breakey and Meskell 2013).

The metric has not, to our knowledge, been applied to hydraulic modelling. However, Mo et al. (2013) stated that the SSIM could have a novel application potential in hydrology. Evaluating the applicability of four non-traditional similarity metrics for hydrometeorology data, they concluded that the SSIM has an advantage as it considers the pattern correlation between two compared objects. In this study the SSIM is calculated by taking x and y as the observed and forecast level and L as the range of the CSO level.

3 Case Study

3.1 Description

To evaluate the performance of the methodology all four ANN models were applied to real data from four case study sites, located in Northern England. The sites were selected to represent the characteristics of different types of CSOs; the CSO chambers are of different sizes and situated in both rural and urban regions.

Time-series level data (mm) was obtained using ultrasonic depth monitors installed in the CSO chambers, measured at a uniform resolution of 2 min. Observed radar rainfall intensity data (mm/h, 5-min temporal resolution and 1×1 km spatial resolution) and forecast rainfall intensity data (mm/h, 15-min temporal resolution and 2×2 km spatial resolution) were obtained from the UK Met Office, with forecast rainfall lead times from 15 min to 6 h. During data pre-processing all the datasets were interpolated to a uniform resolution of 15 min using linear interpolation. Linear interpolation was selected as it is fast and easy to use, and has been demonstrated to produce good results when imputing time-series data (Norazian et al. 2008).

Data was available from April 2016 to October 2018. 50% of data (April 2016 – January 2017) was used for model training and 25% (January 2017 – June 2017) to test the trained

models. The remaining 25% of data (June 2018 to October 2018) was then used to validate the models on unseen data in a simulated online fashion, i.e. as the models would operate in real-time. It was ensured that all the data sets contained a representative amount of wet and dry periods. All the results presented in the following section are given for the validation datasets only.

3.2 Results

Figure 1(a) presents the NSE, MSE, MAPE and mean SSIM performance index for CSO site 1. Each point on the graph represents a model with a different forecast horizon. The NSE for the other three sites are presented in Fig. 1(b). The full results are shown for one site only due to space limitations, however, each case study showed similar outcomes – factors such as catchment size and type did not appear to affect the models' performances.

As can be seen, all the metrics demonstrate that the CEANN model produces the best results for almost all horizons for all sites. In addition, the Combined Wet/Dry EANN consistently outperforms the other two models, and the all-weather EANN somewhat outperforms the trial-and-error model. The improvements gained by using the CEAAN model are most significant when predicting further ahead.

Figure 2 displays an example of the level forecasts for each model 15 min and 3 h ahead during the validation period for CSO site 1 during: (i) dry weather and (ii) a rainfall event which causes an overflow. All four models accurately forecast the level 15 min ahead during both dry and wet weather. When forecasting 3 h ahead the improvements attained by optimising the separate wet and dry EANN models are clearly demonstrated. During dry weather the trial-and-error ANN and all-weather EANN are unable to predict the diurnal level pattern, forecasting an almost straight line. During the rainfall event the two models forecast the increase in level, however both significantly underpredict the extent of the increase and so do not predict the spill (i.e. the forecast level is under the spill level of the chamber). Analysing the entire validation dataset for all the case study sites this is a frequent occurrence. In contrast, the Combined Wet/Dry EANN model and the CEANN both accurately forecast the level 3 h ahead during dry and wet weather, predicting both the timing and magnitude of the spill - information that would be of use to a utility in real time. The 3 h ahead CEANN prediction is smooth during dry periods and noisier during wet periods – this is due to the noisy nature of the rainfall data (the effect is also present in forecasts with a higher lead time (i.e. 3.25 to 6 h horizons)).

It is evident therefore that the use of separate EANN models to overcome the issues caused by the imbalance between the minority wet and majority dry weather data is effective. Furthermore, combining the models using a weighted average (CEANN) is shown to be superior to using a discontinuous switch approach, although this method still produces good results. Although the accuracy of the CEANN model decreases when predicting further ahead, the model still produces reasonably good results up to 6 h ahead, based on the case study results. Thus, if utilised by a utility in real time the model would have the capability to forecast CSO levels during intense precipitation and provide information on possible spills, provided the rainfall events are similar in nature to those used for the model training. This can be ensured by the training approach and the selection of suitable training and validation time periods.

Table 1 presents the parameters selected for CSO case study site 1 by the EA for the all-weather EANN model, and the dry and wet EANNs utilised by the Combined Wet/Dry EANN model and the CEANN. Selected forecast horizons are presented only due to space
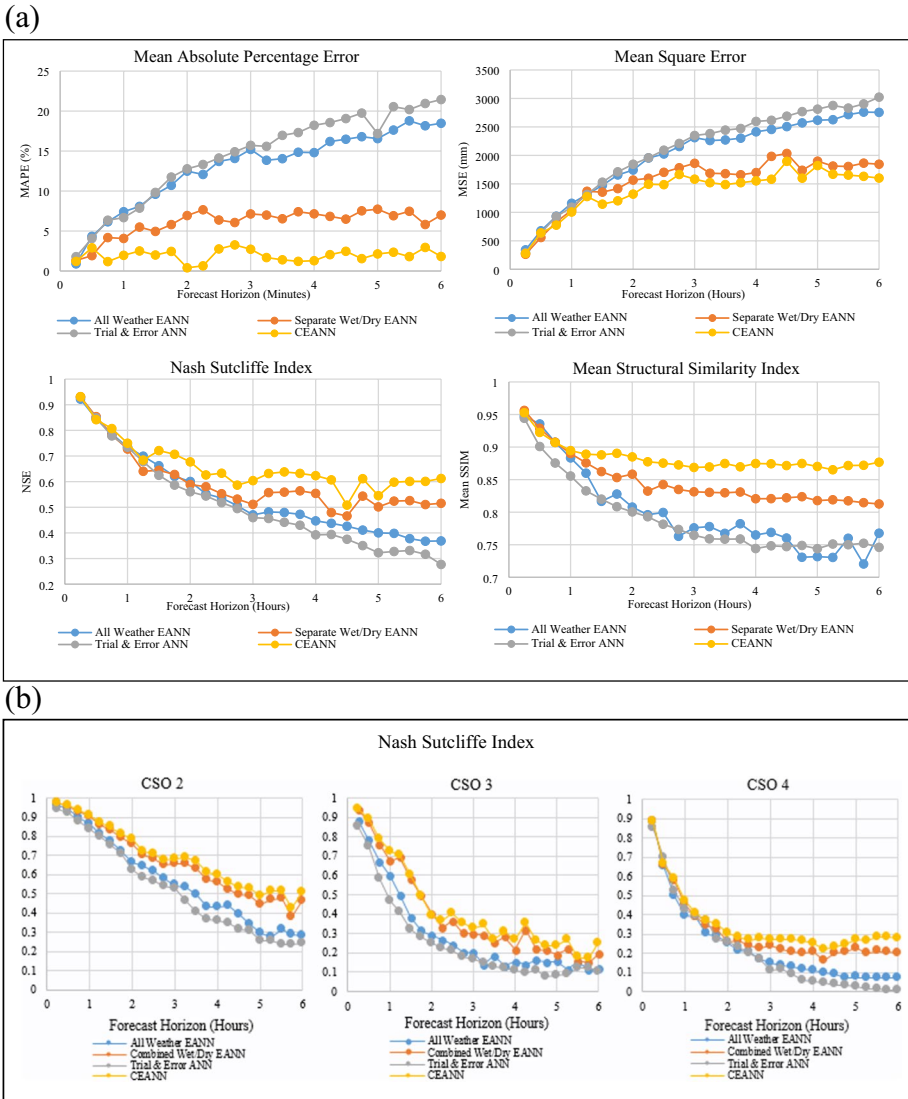
(a)

(b)

Fig. 1 **a** Nash-Sutcliffe Index for Case Study Sites 2, 3 & 4, and **b** Comparison of model performance for different ANN models for CSO site 1

limitations. The selected parameters demonstrate that the forecast horizon and model type considerably influence the selection process. The window of antecedent rainfall and forecast rainfall data generally increases for higher forecast horizons for all models. It is assumed that these horizons are more difficult to predict and so additional data is beneficial. The time of day is utilised only by the all-weather and dry-weather model for higher forecast horizons. Under normal dry weather conditions water levels in the sewer are strongly correlated to the time of day, whereas during wet weather the level is influenced primarily by rainfall. The day of the week, in contrast, is utilised only by the wet-weather model during high forecast horizons
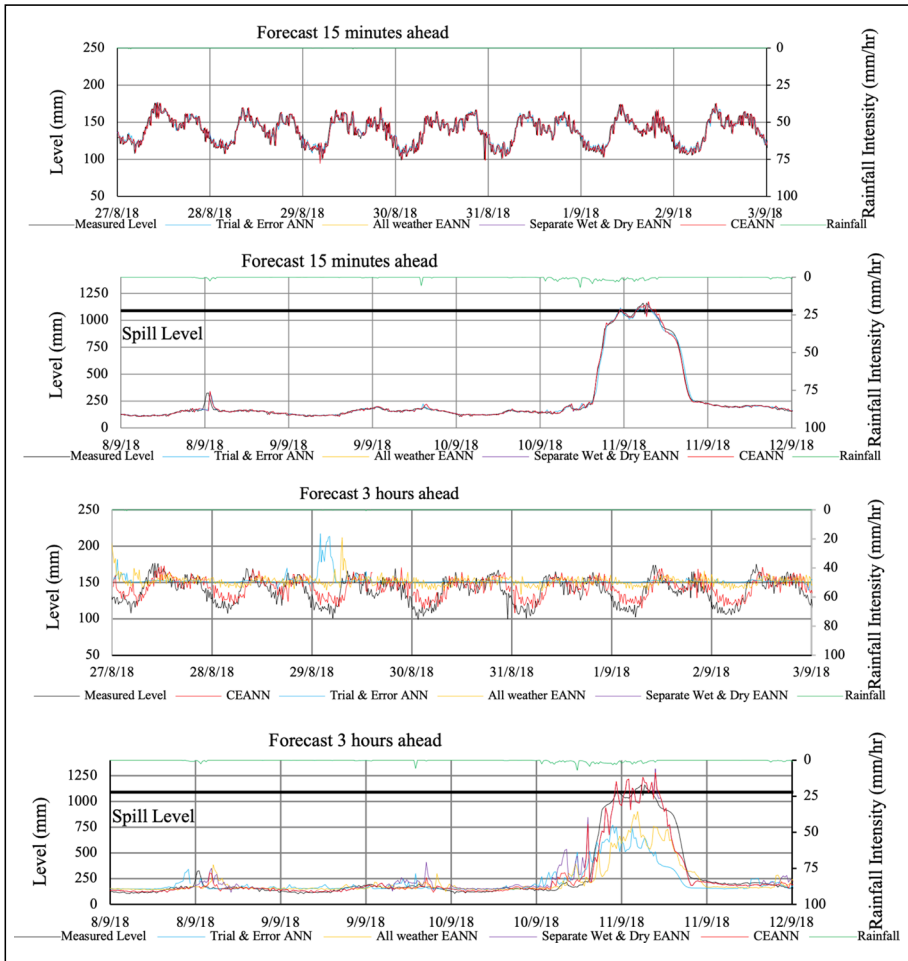
Fig. 2 Comparison of model forecasts for different ANN models during wet and dry weather for CSO site 1 15 min and 3 h ahead

– presumably again because this is a difficult timestep to predict and so requires additional information. No clear pattern was evident regarding the selection of the number of hidden neurons, the number of training cycles or the coefficient of weight decay regularisation and so are not presented.

A similar pattern was observed for all the case study sites. However, the actual values selected for each model varied considerably. Therefore, using a predetermined set of parameters based on these results, rather than using an EA to evolve the optimal parameters for a given location, will likely produce inferior results.

Table 1 Automatically selected EANN parameters & input structure for all weather, dry weather, and wet weather models for selected forecast horizons for CSO site 1

Forecast Horizon (Hours)	All Weather EANN					Dry Weather EANN					Wet Weather EANN				
	Day of Week	Time of Day	Level Data Lag Size	Rain Data Lag Size	Forecast Rain Data Lag Size	Day of Week	Time of Day	Level Data Lag Size	Rain Data Lag Size	Forecast Rain Data Lag Size	Day of Week	Time of Day	Level Data Lag Size	Rain Data Lag Size	Forecast Rain Data Lag Size
0.25	no	no	6	6	2	yes	no	4	4	22	no	no	4	18	1
0.5	no	no	18	4	2	no	no	12	4	4	no	no	4	12	2
0.75	no	no	18	4	8	no	no	6	4	4	no	no	6	12	2
1	no	no	12	12	10	no	no	6	6	4	no	no	4	12	4
2	no	yes	18	48	22	no	yes	6	6	10	no	no	4	4	1
3	no	yes	6	6	24	yes	yes	4	48	14	no	no	18	42	18
4	no	yes	18	30	24	no	yes	12	6	18	yes	no	24	24	22
5	no	yes	18	24	24	no	yes	6	18	22	yes	no	4	36	24
6	no	yes	6	42	24	no	yes	24	42	24	yes	no	18	30	18

4 Discussion

The CEANN model presented has been demonstrated to have the potential to be utilised by wastewater utilities to model CSO levels in near real time. Unlike physically based models commonly used by utilities (e.g., Infoworks), it does not require a physical understanding of the sewer system and is less expensive to build, calibrate and maintain. It is therefore anticipated that the system could be applied widely across an entire network, providing a real-time indication of sewer behaviour and enabling more informed and proactive decision-making. This aligns with the current trend in urban wastewater management of moving towards a more proactive 'smart' approach.

There are several uncertainties inherent in the model. Forecast rainfall data decreases in accuracy for higher forecast horizons – in turn negatively effecting the model performance. The methodology also relies on accurate and correctly calibrated CSO level data. However, issues such as sensor drift and logger malfunctions can cause data anomalies. Automated data pre-processing is performed to identify and remove any erroneous data; however, anomalies can be challenging to detect, particularly gradual changes due to sensor drift.

The model also relies on several assumptions. Firstly, it is assumed that the future response of the sewer is similar to the historic data used to train the model. However, the behaviour of the system may change over time for many reasons, including planned changes to the sewer, the construction of new buildings connected to the network, and unintended issues such as blockages. This is a problem known as concept drift (Widmer and Kubat 1996), whereby the statistical properties of the target variable change over time in unforeseen ways. It is therefore recommended that the model is retrained periodically to accommodate any changes. Additionally, no work has yet been conducted to determine if seasonal variations need to be accounted for when training the model.

Future work will focus on validating the methodology on additional CSO sites and analysing the ability of the model to predict upcoming spills in a timely and reliable manner. The next step will then be the creation of a user-friendly online support tool, which can be integrated easily with the current working practices and technology utilised by wastewater utilities. Work should be conducted to quantify the costs and benefits of deploying the tool, to investigate the ability of the methodology to be applied on a large scale, and to understand the types of information and features wastewater utility personnel would find of use.

The model has been designed with a maximum forecast horizon of 6 h, as 6 h ahead rainfall nowcasts were utilised. However, future work could investigate if the model could predict further ahead if nowcasts with a higher lead time were available. The methodology also has the potential to be developed further by including additional EANN models as part of the committee machine. For example, multiple EANN models optimised for different rainfall event severities. Different types of ANN model could also be explored, for example recurrent neural networks. Additionally, as sensor technology becomes cheaper utilities may begin to install flow level monitors in addition to level sensors, thus there is the potential to apply the CEANN model to flow forecasting, providing the wastewater utility with further valuable information.

5 Conclusions

A Committee EANN (i.e. CEANN) model has been proposed here to forecast water level in a CSO chamber. The committee machine is composed of two feed-forward single hidden layer EANN models optimised for wet and dry weather data respectively and combined

using a non-linear weighted average approach. An evolutionary strategy algorithm is used to automatically select the ANN structure and parameter set.

The CEANN methodology was tested on four real case study CSO sites and compared with an ANN model developed manually through trial and error (i.e. Trail & Error ANN, approach adopted in existing literature), a single EANN model designed for all-weather conditions (i.e. All Weather EANN), and a joint wet and dry EANN model combined using a discontinuous transition (i.e. Combined Wet/Dry EANN model). The results obtained lead to the following conclusions:

1. The CEANN model produces more accurate CSO level predictions than the other three models. This is true for all forecast horizons (up to 6 h) clearly demonstrating the superiority of the CEANN methodology, especially over the Trail & Error ANN often used in the existing literature.
2. The CEANN and the Combined Wet/Dry EANN model accurately forecast the dry weather and wet weather level, predicting the timing and magnitude of, thus providing information that is of clear use to a wastewater utility in near real time. The same cannot be said for the other two models (Trial & Error ANN and All Weather EANN) that continually underestimate the CSO level during periods of heavy rainfall and hence cannot predict overflows.
3. The CEANN methodology is generic and thus requires minimal human effort to design and can automatically be applied to different catchments and forecast horizons. It is envisioned that the model could be used beneficially by utilities to model CSO levels in the wastewater network in near real-time and provide alerts for upcoming spills - enabling better decision making and proactive management of overflow events.

References

Al-Ani RRA, Al-Obaidi BHK (2019) Prediction of sediment accumulation model for trunk sewer using multiple linear regression and neural network techniques. Civil Eng J 5(1):82

Benbassat A, Sipper M (2013) Evolving Artificial Neural Networks with FINCH. Pp. 1719–20 in GECCO 2013 - Proceedings of the 2013 Genetic and Evolutionary Computation Conference Companion. Vol. 87

Berkhahn S, Fuchs L, Neuweiler I (2019) An ensemble neural network model for real-time prediction of urban floods. J Hydrol 575:743–754

Breakey D, Meskell C (2013) Comparison of metrics for the evaluation of similarity in acoustic pressure signals. J Sound Vib 332(15):3605–3609

Butler D, Friedler E, Gatt K (1995) Characterising the quantity and quality of domestic wastewater inflows. Water Sci Technol 31(7):13–24

Chen YH, Chang FJ (2009) Evolutionary artificial neural networks for hydrological systems forecasting. Journal of Hydrol 367(1–2):125–137. https://doi.org/10.1016/j.jhydrol.2009.01.009

Ebtehaj I, Bonakdari H (2014) Performance evaluation of adaptive neural fuzzy inference system for sediment transport in sewers. Water Resour Manag 28(13):4765–4779

Fernando AK, Zhang X, Kinley PF (2007) Combined sewer overflow forecasting with Feed-forward Back-propagation Artificial Neural Network. Journal of Civil and Environmental Engineering 1(12):128–134. https://doi.org/10.5281/zenodo.1330365

He H, Ma Y (Eds) (2013) Imbalanced learning: foundations, algorithms, and applications. Wiley-IEEE Press

Hines A, Harte N (2012) Speech intelligibility prediction using a Neurogram similarity index measure. Speech Comm 54(2):306–320

Hines A, Skoglund J, Kokaram A, Harte N (2012) ViSQOL: The Virtual Speech Quality Objective Listener. *Proceedings of the International Workshop on Acoustic Signal Enhancement,* Aachen, Germany 1–4

Joseph-Duran B, Jung MN, Ocampo-Martinez C, Sager S, Cembrano G (2014) Minimization of sewage network overflow. Water Resour Manag 28(1):41–63

Kim HI, & Han KY (2020) Urban flood prediction using deep neural network with data augmentation. Water 12(3). https://doi.org/10.3390/w12030899

Kurth A, Saul A, Mounce S, Shepherd W, Hanson D (2008) Application of Artificial Neural Networks (ANNs) for the Prediction of CSO Discharges. Proceedings of 11th International Conference on Urban Drainage 1–10

Mo R, Ye C, Whitfield P (2013) Some similarity indices with potential meteorological applications (pp. 201–401). Retrieved from http://www3.telus.net/ruping/lib/reports/2013-002-CMML-TR_Mo_etal.pdf

Moradi AM, Dariane AB (2017) Evolving neural networks and fuzzy clustering for multireservoir operations. Neural Comput & Applic 28(5):1149–1162

Mounce SR, Shepherd W, Sailor G, Shucksmith J, Saul AJ (2014) Predicting combined sewer overflows chamber depth using artificial neural networks with rainfall radar data. Water Science and Technology 69(6):1326–1333. https://doi.org/10.2166/wst.2014.024

Norazian MN, Shukri YA, Azam RN, Al Bakri AMM (2008) Estimation of missing values in air pollution data using single imputation techniques. ScienceAsia 34(3):341–345

Romano M, Kapelan Z (2014) Adaptive water demand forecasting for near real-time Management of Smart Water Distribution Systems. Environ Model Softw 60:265–276

Rosin T, Romano M, Kapelan Z, Keedwell E, Woodward K (2018) Prediction of CSO Chamber Level Using Evolutionary Artificial Neural Networks. In 13th International Conference on Hydroinformatics

Salerno F, Viviano G, Tartari G (2018) Urbanization and climate change impacts on surface water quality: enhancing the resilience by reducing impervious surfaces. Water Res 144:491–502

Schwefel H-P (1998) Evolution and optimum seeking, vol 27. Wiley, New York

She L, You X y (2019) A dynamic flow forecast model for urban drainage using the coupled artificial neural network. Water Resour Manag 33(9):3143–3153

Sola J, Sevilla J (1997) Importance of input data normalization for the application of neural networks to complex industrial problems. IEEE Trans Nucl Sci 44(3 PART 3):1464–1468

Sumer D, Gonzalez J, Lansey K (2007) Real-time detection of sanitary sewer overflows using neural networks and time series analysis. J Environ Eng 133(4):353–363

Tadeusiewicz R (1995) Neural networks: a Comprehensive Foundation, vol 3. Pearson Prentice Hall, Upper Saddle River

Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612

Widmer G, Kubat M (1996) Learning in the presence of concept drift and hidden contexts. Mach Learn 23(1):69–101

Yao X (1993) A review of evolutionary artificial neural networks. Int J Intell Syst 8(4):539–567

Yao X (1999) Evolving artificial neural networks. Proc IEEE 87(9):1423–1447

Zhang D, Lindholm G, Ratnaweera H (2018) Use long short-term memory to enhance internet of things for combined sewer overflow monitoring. J Hydrol 556:409–418

Zhang Q, Li Z, Snowling S, Siam A, El-Dakhakhni W (2019) Predictive models for wastewater flow forecasting based on time series analysis and artificial neural network. Water Sci Technol 80(2):243–253