

Prediction of Peracetic Acid Disinfection Performance for Secondary Municipal Wastewater Treatment Using Artificial Neural Networks

Kathryn B. Newhart,* Joshua E. Goldman-Torres, Daniel E. Freedman, K. Blair Wisdom, Amanda S. Hering,* and Tzahi Y. Cath



Cite This: <https://dx.doi.org/10.1021/acsestwater.0c00095>



Read Online

ACCESS |



Metrics & More



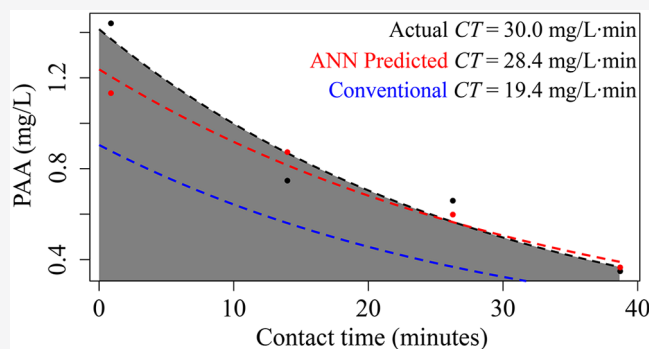
Article Recommendations



Supporting Information

ABSTRACT: Disinfection is one of the most critical processes for municipal wastewater treatment. However, traditional chemical dosing approaches do not consider how changes in water quality and process operation can alter disinfection performance. This work aims to develop novel disinfection models for precise prediction of peracetic acid (PAA) performance that considers real-time changes in water quality. Artificial and recurrent neural networks (ANN and RNN, respectively) are trained to predict PAA at various locations throughout the disinfection basin, CT (a function of the active concentration and contact time), and pre- and postdisinfection *Escherichia coli* using online and laboratory data. An ANN is found to predict PAA concentrations at an error rate comparable to that of an online analyzer. Additionally, an ANN can predict CT more accurately than a conventional first-principles method both with and without an online analyzer. An ANN with a lagged response variable can predict *E. coli* in a fraction of the time of an RNN, but with a slightly increased error. The integration of the models developed in this work into existing monitoring and control systems could provide treatment facilities with more robust and dynamic disinfection control without the need for costly analyzers.

KEYWORDS: wastewater treatment, machine learning, peracetic acid, disinfection kinetics, process modeling, performance prediction



1. INTRODUCTION

Municipal wastewater treatment plants (WWTPs) are critical in preventing the spread of waterborne disease caused by pathogenic organisms. As such, WWTPs have weekly and monthly permit limits on the maximum allowable concentration of *Escherichia coli* (*E. coli*), an indicator microorganism for fecal contamination, discharged to the environment. The primary method of destroying or inactivating pathogenic microorganisms for secondary treated municipal wastewater is chlorine-based disinfection (e.g., chlorine, chloramine, and chlorine dioxide) due to chlorine's widespread availability, low cost, effectiveness against bacteria and viruses, and acceptance among regulatory agencies. However, chlorine-based disinfectants have a high propensity to react with residual organic matter and form disinfection byproducts (DBPs), many of which are potent carcinogens.^{1,2} Peracetic acid (PAA, CH₃CO₃H) is a chemical disinfectant that can easily be utilized in most WWTPs, and it produces DBPs that are less toxic than those of chlorine-based disinfectants.^{3,4} PAA is widely used in the food, beverage, cosmetic, pharmaceutical, and agriculture industries^{5,6} and is capable of achieving the pathogen inactivation required to meet discharge permit limits in WWTPs.^{7,8} However, while the disinfection mechanism for commercial PAA is similar to those of other oxidizing agents,

quantifying the impacts of disinfection performance is still not fully understood at full scale.⁹

Hypothesized mechanisms of PAA disinfection include the release of "active" oxygen¹⁰ or hydroxyl radicals,¹¹ protein denaturing,¹² and dislocating or rupturing cell walls.¹³ Related to the method of *E. coli* inactivation with PAA, viable but nonculturable *E. coli* have been detected in WWTP effluent disinfected with PAA,¹⁴ potentially overestimating the measured effluent *E. coli* and underestimating the calculated log inactivation. The substantial variation in the modeling of *E. coli* inactivation by PAA could potentially be explained by the lack of a complete understanding of the disinfection mechanism and measurement bias.^{6,14–17} Therefore, models of PAA disinfection performance should consider the effects of changing water quality at full-scale WWTPs on the active PAA concentration and *E. coli* inactivation.

Received: August 6, 2020

Revised: September 27, 2020

Accepted: October 21, 2020

An additional component of modeling PAA concentration as a function of time is the substantial decline in concentration that occurs when PAA is initially dosed into a water matrix with oxidative demand [i.e., chemical oxygen demand (COD)]. Haas and Finch¹⁶ summarized demand and first-order decay PAA kinetics in eq 1:

$$C(t) = (C_0 - D)e^{-kt} \quad (1)$$

where $C(t)$ (milligrams per liter) is the residual concentration of PAA at time t , C_0 (milligrams per liter) is the initial applied PAA dose, D (milligrams per liter) is the initial oxidative PAA consumption, k (inverse minutes) is the first-order decay constant, and t (minutes) is the length of time the water is in contact with PAA. Equation 1 is used in this work to model PAA concentration as a function of time after first-order decay kinetics are verified. The literature reports a wide range of demand (D) and decay (k) values for various secondary treated wastewaters.^{8,15,18,19} The diversity of model parameter values implies that disinfection performance varies widely with water quality. Consequently, the assumption of static demand and decay model parameters could result in erroneous calculated PAA concentrations if the true demand and decay values change over time.

The potency of chemical disinfection is a function of the active concentration of the disinfectant [$C(t)$] and the time the disinfectant is in contact with the water (t), measured by CT. CT is a concept that measures disinfection dose relative to the time the disinfectant has to react. In traditional CT calculations for which the $C(t)$ function cannot be explicitly defined, the final concentration at time t is multiplied by t . However, this is a very conservative approach as it underestimates the actual active concentration of the chemical and thereby also underestimates CT. When $C(t)$ can be defined, CT is calculated by integrating eq 1 with respect to t , where t is interpreted as the hydraulic retention time (HRT) from the beginning to the end of the disinfection basin. Traditional chlorine-based disinfection processes frequently use CT control for design and validation purposes because of the well-established and predictable relationships among $C(t)$, CT, and pathogen inactivation. However, the relationship between PAA dose and pathogen activation of PAA is not nearly as well-understood.⁸ Manoli et al.¹⁷ proposed a novel CT-based dosing strategy for PAA derived from first principles. However, multiple fitted model parameters varied with each batch, further illustrating that a first-order model with constant model parameters may not fully describe PAA demand and decay kinetics in a real, constantly changing water matrix. Temperature, pH (>9; $pK_a = 8.2$), total suspended solids (TSS), and biochemical oxygen demand (BOD) have been shown to impact PAA decay and *E. coli* inactivation,^{9,14,19–21} but precise relationships have not yet been established such that the effect of changing water quality can be used as a predictor in a $C(t)$ model.

To incorporate information about real-time water quality and to better understand disinfection kinetics, nontraditional modeling approaches such as machine learning methods can be used. A particularly powerful nonlinear, data-driven modeling approach is an artificial neural network (ANN). ANNs map the relationship between predictor (i.e., input) and response (i.e., output) variables using linear combinations of multiple nonlinear functions (i.e., activation functions) arranged in layers. The ANN model is fit by iteratively adjusting the linear (e.g., bias) and nonlinear (e.g., weights) function parameters

internal to the ANN to minimize the training error. This is in contrast to mechanistic models that use complex formulations connected in simple (e.g., linear) ways. Additionally, mechanistic models are not constrained by the range of the training data. Data-driven modeling, such as an ANN, does not require knowledge of the underlying mechanisms driving changes in the response. Rather, data-driven modeling replicates patterns from a training data set, and conclusions or predictions based on the models are thus constrained to the range of features within that data set. In this work, two types of machine learning models are used to predict features in water quality data: ANNs and recurrent neural networks (RNNs). RNNs are a special case of ANNs that handle time dependency in data, which is a frequently observed property of water quality and treatment data,^{22,23} but have not yet been explored for disinfection performance prediction. Additional details about the structure and training of neural networks can be found in textbooks and the academic literature.^{24–26}

ANN modeling of chlorine dosing and residuals has been widely explored for drinking water²⁷ and wastewater treatment,²⁸ as well as other popular disinfection methods. Lin et al.²⁹ used an ANN to predict effluent total coliforms in a bench-scale ultraviolet (UV) disinfection system (combined with a linear model of UV dose and log inactivation), but they measured only the performance of the models on the training data set. Yu et al.³⁰ used five ANNs in a control scheme for a bench-scale chlorination–dechlorination system. Total coliform goals were achieved for a variety of reuse scenarios (*E. coli* was not measured), but the model was not able to fully describe the variation in residual chlorine. Carvajal et al.³¹ investigated the relationship among dose, time, pH, turbidity, and viral log inactivation on a bench-scale batch chlorination system for secondary treated wastewater using a Bayesian belief network; they seeded water with viruses and measured log inactivation to train Bayesian multilayer perceptron models to predict the required CT, one for each type of virus. The difficulty in predicting log inactivation arises when the pre-disinfection pathogen concentration is unknown and the kinetics of the disinfectant are difficult to model, as in the case of PAA and *E. coli* in WWTPs.

Despite the growing adoption of PAA systems, ANNs have not been widely applied to PAA modeling. Wei et al.³² were able to replicate a computational fluid dynamics-based PAA disinfection model (i.e., a chemically reactive, time-dependent, three-dimensional turbulent flow model) for a conventional five-pass disinfection basin using physical mixing and chemical decay kinetic parameters as ANN model inputs. While Wei et al. demonstrated that an ANN could simulate turbulent flow, the model requires an existing calibrated computational fluid dynamics model to train the ANN. Second, their model does not consider real-time water quality or the inactivation of *E. coli*, each of which is explored here.

Given the difficulty in modeling CT and *E. coli* removal for full-scale PAA disinfection processes, the ability to predict real-time disinfection performance could improve the precision with which PAA is dosed and could reduce the number of unexpected permit exceedances of *E. coli* for municipal WWTPs due to more accurate monitoring. The objective of this work is to model full-scale PAA disinfection performance to predict PAA concentration, CT, pre-disinfection *E. coli* concentration, and post-disinfection *E. coli* concentration using machine learning models. Furthermore, this work identifies the environmental and operational conditions that are most

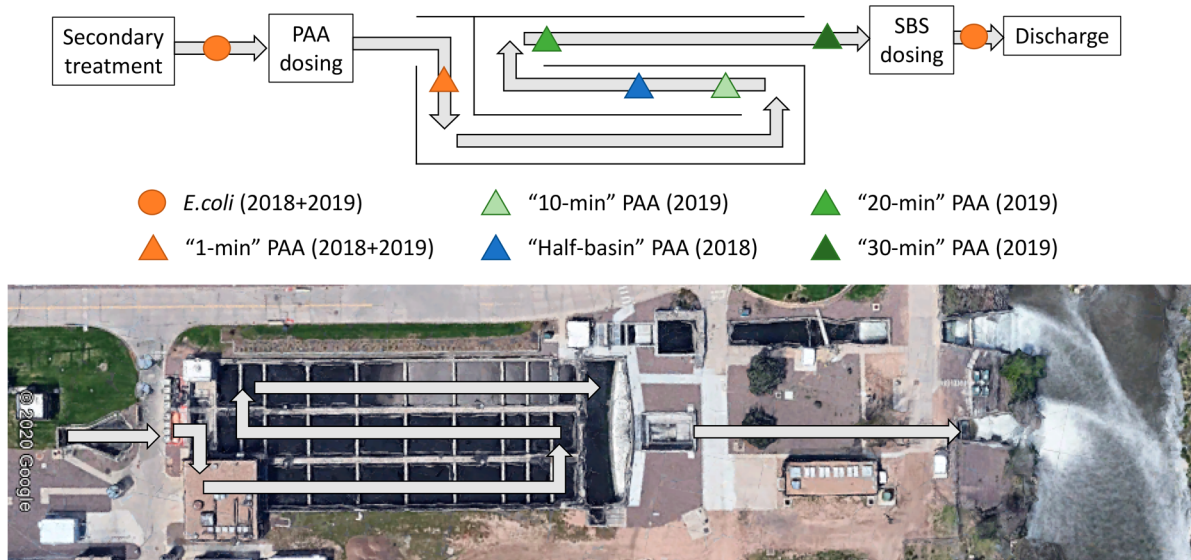


Figure 1. Process flow diagram of the PAA disinfection system at the RWHTF. Triangles indicate PAA samples taken at various locations in 2018 and/or 2019. Circles indicate *E. coli* sampling locations. The sample locations are named for the approximate HRT during peak flow from the PAA dosing location to the sampling location (e.g., “1 min”).

Table 1. Process Variables Used to Train ANN and RNN Models for PAA, CT, and *E. coli*^a

location	type	variable	short average	location	type	variable	sample type
S	O	influent flow rate	15 min	S	L	sludge volume index (SVI) ^b	grab
S	O	gravity thickener effluent flow rates ²	15 min	S	L	volatile suspended solids	24 h FC
S	O	basins in service ²	24 h	S, SE	L	alkalinity ²	24 h FC
S	O	aerated solid retention time (ASRT)	24 h	S, SE	L	BOD ²	24 h FC
S	O	TSS	60 min	SE	L	BOD (carbonaceous) ²	24 h FC
S	O	ammonia ²	60 min	S, SE	L	COD ²	24 h FC
S	O	nitrite ²	60 min	S, SE	L	total phosphorus ²	24 h FC
S	O	nitrate ²	60 min	S, SE	L	total Kjeldahl nitrogen ²	24 h FC
S	O	phosphate ²	60 min	SE	L	TSS ²	24 h FC
S	O	DO	60 min	S, SE	L	<i>o</i> -phosphorus ²	24 h FC
S	O	pH	60 min	SE	L	total inorganic nitrogen ²	24 h FC
S	O	clarifier blanket depth	15 min	SE	L	total nitrogen ²	24 h FC
S	O	waste-activated sludge flow rate ²	24 h	SE	L	nitrate + nitrite ²	24 h FC
S	O	return activated sludge (RAS) as a percent of influent flow	15 min	D	L	predisinfection <i>E. coli</i>	grab
S, SE, D	O	flow rate ^b	15 min	D	L	postdisinfection <i>E. coli</i>	grab
D	O	hydraulic retention time ^b	15 min	D	L	PAA	grab
D	O	PAA flow rate	15 min				
D	O	chemical oxygen demand (total)	5 min				
D	O	chemical oxygen demand (soluble)	5 min				
D	O	UV transmittance (UVT)	5 min				
S, SE, D	O	TSS ^b	5 min				
D	O	air temperature	60 min				

^aOnline sensor data (type code O, left columns) require averaging for some ANN and RNN configurations tested, and the “short-term” average is listed here (short average). Offline data (lab, type code L, right columns) are collected using either 24 h flow composites (24 h FC) or grab samples (sample type). Features include operational variables and water quality measurements from the main biological treatment process (secondary, location code S); water quality measurements as the treated water enters the disinfection basin (secondary effluent, location code SE), or operational variables that directly affect the dose of PAA (disinfection, location code D). Superscript numbers indicate process variables that are available for only the 2018 (1) or 2019 (2) data sets. ^bMultiple sampling locations for the same feature.

strongly linked to PAA disinfection performance from the ANN models using a variable selection process.

Methods includes a description of a full-scale PAA disinfection process at a large WWTP, the online instrumentation and laboratory data collected during two PAA sampling campaigns in 2018 and 2019 by the WWTP, and the ANN and RNN structures used to predict PAA and *E. coli* pre- and postdisinfection. Modeling results and implications for real-

time control are included in **Results and Discussion**, concluding with remarks on the potential for online PAA and *E. coli* monitoring using machine learning at full-scale WWTPs.

2. METHODS

2.1. Metro Wastewater Reclamation Facility. The Robert W. Hite Treatment Facility (RWHTF) in Denver,

CO, is owned and operated by the Metro Wastewater Reclamation District. The RWHTF treats both municipal and industrial wastewater for the Denver metro area and is rated for 220 million gallons per day (MGD). Secondary treated effluent is mixed with disinfectant using static mixers, flows through a three-pass serpentine contact basin prior to being quenched with sodium metabisulfite, and discharges to the South Platte River. In January 2018, the disinfection process at the RWHTF was changed from chloramination to PAA disinfection,³³ but the facility has since experienced challenges in both the monitoring and the optimization of the PAA dose. Due to the high unit cost of PAA, this has resulted in substantially higher than anticipated operating costs for disinfection to achieve the facility's effluent *E. coli* limit, which is 126 colony-forming units (cfu) per 100 mL for a 30-day geometric mean and 252 cfu/100 mL for a 7-day geometric mean. To better understand PAA disinfection kinetics, *E. coli* and PAA samples were taken throughout the disinfection basin in 2018 and 2019 at the locations denoted in Figure 1.

2.2. Data. Two PAA sampling campaigns, one in 2018 and one in 2019, were used in this work. From October 2 through October 15, 2018, three daily grab samples of secondary treated wastewater were collected (1) immediately downstream of the PAA dosing location ("1 min") and (2) halfway through the disinfection basin ("Half"), totaling 168 sampling events. A second sampling campaign was conducted in 2019 for four sampling locations: (1) immediately downstream of the PAA dosing location ("1 min"), (2) at the end of the first pass ("10 min"), (3) at the end of the second pass ("20 min"), and (4) at the end of the disinfection basin ("30 min") (totaling 66 sampling events). The 2019 sampling events occurred in the morning, at noon, and in the early afternoon for select weekdays between June 24 and August 1, 2019. PAA was measured at the RWHTF using Standard Method 4500.³⁴ For each sampling event, linear and first-order PAA concentration models were fit as a function of the HRT to the sampling location. Because the 2018 sampling events contain only two sampling locations, the 2019 data set was used to determine the order of the decay model for PAA concentration. From the fitted $C(t)$ model for each sampling event, the "actual CT" was calculated for each event by integrating eq 1. The actual CT was used as a response (i.e., output) for a set of neural network models and as a predictor (i.e., input) for the *E. coli* neural network models.

The inputs to the PAA concentration, CT, and *E. coli* models include online and offline process variables listed in Table 1. Online data included nutrients (ChemScan UV-6101, Waukesha, WI; s::can spectro::lyser and nitro::lyser, Vienna, Austria; Endress+Hauser ISEmax CAS40D, Reinach, BL, Switzerland), TSS (Cerlic ITX, Atlanta, GA), pH, and dissolved oxygen (DO) sensor measurements within and at the end (i.e., effluent) of the secondary treatment process. The frequency of collection ranged from multiple observations per minute (e.g., flows) to 10 min (e.g., nitrogen species). In 2018, an ultraviolet–visual spectrum instrument (YSI CarboVis, Yellow Springs, OH) recorded COD, UVT (UV transmittance), and TSS at the influent of the disinfection basin every 5 min. However, due to difficulty in calibrating the instrument, the raw measured voltage was used as input to the ANN. An online PAA analyzer (ChemScan Mini) was installed in 2019 immediately downstream of the 1 min sample location. The PAA analyzer data were not used to build the ANN and RNN models; instead, they were used to externally evaluate

and compare the performance of a costly, conventional CT calculation approach to the CT predicted by an ANN.

Given the wide range of data collection frequencies and delays inherent in laboratory methods, the process of data blending is the set of steps taken to organize and synthesize data such that they can be used to train machine learning models. To determine the data blending approach that best represents the water quality at the time of the PAA sampling events, online data were averaged using three different time horizons. Instantaneous data included only the last recorded observation (no averaging). Short-term average data were averaged over a time horizon estimated by RWHTF engineers both to approximate the HRT and to account for sensor noise. For example, a noisy sensor immediately upstream of the basin required a 5 min average to approximate the true value in the presence of high variability. Sensors present in the secondary treatment system were averaged for 60 min to account for HRT and to capture the variation in water quality. These time frames primarily range from 5 to 60 min and are listed in Table 1 under the short average column. The 24 h average data were averaged over a 24 h time horizon. If a sample was not taken within the averaging period, the last measured value was carried forward as the present value.

There are multiple complexities associated with selecting offline data at the time of the PAA sampling events, especially considering that the model is designed to simulate real-time prediction. For example, there are multiple PAA sampling events per day in 2019 but only one 24 h flow composite (24 h FC). The 24 h FC samples collect individual water quality samples over a 24 h period where the volume of each sample is proportional to the flow at the sampling location. Given the sparsity of the data (24 h between observations) and the unknown relationship between each observation, no interpolation is used in these cases to match minute-level time stamps exactly.

The final consideration for data blending is specific to the *E. coli* model development, specifically the comparison of the ANN and RNN. *E. coli* measurements require at least 42 h to collect, process, and upload the result at the RWHTF, and this is a common turnaround time for *E. coli* measured using Standard Method 9223-B.³⁵ Thus, the "most recent" *E. coli* observation when simulating real-time control is the *E. coli* measurement almost 2 days prior. For the neural network models that include *E. coli* concentrations as predictors (discussed in the next section), the *E. coli* observations lagged by 42 h to account for the lab processing delay.

2.3. Neural Network Prediction. Two types of neural networks were tested: ANNs and RNNs. The simplest and most common neural network used in this work is a feedforward ANN, which uses a single hidden layer to connect the input and output layers. Feedforward ANNs are also unidirectional and thus do not consider the order of observation. RNNs have internal memory nodes that train based on a sequence of observations rather than assuming each observation is independent of the others. However, RNNs are more efficient at discovering short-term dependencies than long-term. Long short-term memory (LSTM) nodes are one example of an RNN that selectively retains the optimum number of historical observations to avoid the problem of poor trend detection for longer training windows.^{36,37} Given the highly autocorrelated nature of microbial inactivation, we used an LSTM RNN to improve *E. coli* inactivation predictions over mechanistic or ANN modeling.

The neural network models evaluated in this work are listed in Table 2 and were constructed in R using the Keras API for

Table 2. Summary of Neural Network Models Built to Predict PAA Disinfection Performance in Simulated Real Time^a

response	neural network	data blending	testing
PAA concentration	ANN	instant, short-term average, 24 h average	leave one out
CT	ANN	instant	leave one out
predisinfection <i>E. coli</i>	ANN-lagged, RNN	instant, short-term average, 24 h average	90/10
postdisinfection <i>E. coli</i>	ANN-lagged, RNN	instant, short-term average, 24 h average	90/10

^aThe type of neural network tested depends on the response variable. For each neural network, multiple predictor data sets are used that differ in the length of time over which the variables are averaged. Each neural network set and each predictor set are trained and tested 10 times using either leave one out or a 90/10 split.

TensorFlow.^{38,39} We often use the term predictor to refer to input variables and response to refer to the output variable for a neural network model. Three thousand training epochs were used to fit each ANN and RNN model by minimizing mean square error. Each model contained an input layer with a node for each input variable using the softsign activation function (eq 2), a hidden layer with 3 times the number of nodes as the input layer using the softsign activation function, and an output layer with a single node and a linear activation function.

$$\text{softsign}(x) = x/(|x| + 1) \quad (2)$$

Fewer nodes in the hidden layer were explored with similar results, but the most complex architecture with the lowest testing error is presented here. The complexity that allows RNNs to retain information about the sequence of observations requires substantially more computational time and resources. For example, the ANN that we trained on the data from the 2018 sampling campaign includes 52 input variables, 156 nodes in the hidden layer, and one node in the output layer. This equates to 8425 weights and biases to estimate. For the analogous RNN, there are 130416 parameters. Thus, we tested an ANN that includes the last known measurement of *E. coli* as a predictor, termed “ANN-lagged” in Table 2, to attempt to account for the autocorrelation exhibited in *E. coli* measurements and to test if such an ANN can approximate the performance of an RNN with fewer fitting parameters in less time.

To build each model, data are blended from the data sources identified in Table 1 using the methods discussed in the previous subsection. Training data were prepared by min–max normalization scaling prior to model fitting and prediction and used to rescale the model’s prediction. Data were separated into training and testing sets (Table 2, testing), including “leave one out” (LOO) for an ANN and a 90% training/10% testing split (90/10) for an RNN. Neither PAA concentrations throughout the disinfection basin nor CT is strongly correlated over time; thus, the order of the observations in the training data set for these responses is irrelevant. The LOO approach takes advantage of this independence to train an ANN on all but a single observation and test on the held-out observation. In this way, the number of ANNs built for each trial was equal

to the number of observations in a data set. This was especially advantageous as it maximizes the “small” training data sets used in this work, relative to most machine learning applications in the literature. Additional training–testing variations were tested (e.g., 40%, 60%, 80%, and 90% training split), but the model error was found to be minimized by maximizing the number of training observations. RNN training observations are sequential and cannot be reordered as in LOO. Therefore, an RNN is trained on the first 90% of each data set and tested on the remaining 10% using a rolling window approach in which the number of observations in the training window is constant, and older observations are removed as the window “rolls” forward in time and includes new observations. The testing root-mean-square error (RMSE) from each iteration of model training and testing is calculated, and the process is repeated for a total of 10 iterations. The reported predictions are an average of the 10 trials for each model configuration.

2.4. Neural Network Variable Selection. ANNs are “black box” systems, meaning that the internal model parameters do not directly represent the impact of an individual predictor variable on the response. To determine which predictor variables have the strongest impact on PAA disinfection prediction, we conducted a forward variable selection experiment to train and test ANNs.^{40,41} We used combinations of predictor variables and tested them sequentially to find the ANN model with the fewest predictors that minimizes the RMSE. Due to the number of iterations and to avoid overfitting, the number of training epochs for this experiment was reduced to 250. The procedures for variable selection of CT and *E. coli* ANN are as follows:

1. Train and test ANN using ν input variables where $\nu = 1$ to initiate the variable selection procedure. Calculate the mean RMSE of three train–test trials when $\nu = 1$. Repeat for each of the original ANN input variables, resulting in p mean RMSE values.
2. Identify the input variable with the lowest mean RMSE in step 1. This is considered the most influential predictor.
3. For the next iteration of variable selection ($\nu = \nu + 1$), create a list of predictors by combining all combinations of the most influential predictor with the remaining predictors, resulting in a $p - (\nu - 1)$ list of variable combinations.
4. Train and test the ANN using the combinations of predictors identified in step 3, and calculate the mean RMSE of three train–test trials. Identify the variable combination with the lowest mean RMSE as the most influential set of predictors.
5. Continue to iterate ($\nu = 3, 4, \dots$) steps 3 and 4 until the lowest mean RMSE using ν predictors is greater than the lowest mean RMSE from a model with $\nu - 1$ predictors; the optimum set of predictor variables results from iteration $\nu - 1$.

3. RESULTS AND DISCUSSION

3.1. ANN for Prediction of PAA Concentration. ANNs were used to predict PAA concentrations at each sampling location using online and offline process data. Three training data sets were compared for each location: instantaneous, short-term, or 24 h average. To compare ANN model performance among the three averaging horizons, testing

RMSE for each PAA sampling location and averaging horizon are plotted in Figure 2. Each box plot represents a set of 10

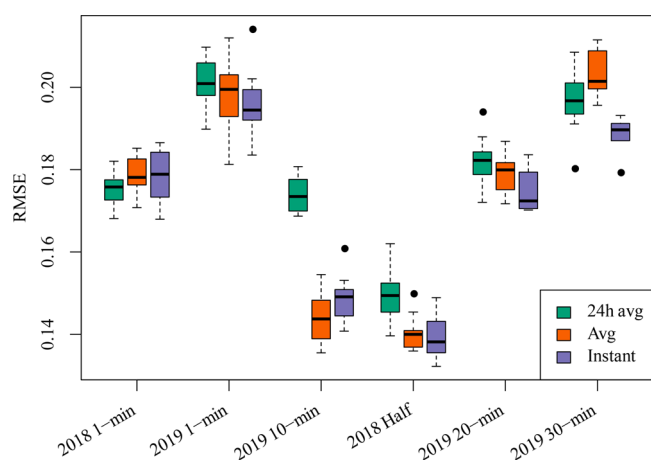


Figure 2. RMSE of ANN model predictions for three training data sets (each with a different averaging horizon) at each PAA sampling location shown in Figure 1.

testing RMSEs, one RMSE for each train–test iteration of an ANN model given random initialization. Overall, the PAA concentration was most accurately predicted when instantaneous data were used, and the difference is most profound by the third pass of the disinfection basin. This suggests that the instantaneous values are better for prediction than other types of averaging for representing water quality conditions, especially at distances farthest from the dosing site.

For the first sampling location downstream of PAA dosing (“1 min”), the use of real-time or averaged data did not substantially change the prediction error. This would suggest that the water quality and operational variables that vary over the length of the averaging period are not substantial contributors to PAA demand (D); these could include ASRT and laboratory measurements listed in Table 1 (e.g., alkalinity, BOD, COD, phosphorus species, and nitrogen species). However, the first sampling location in 2019 also exhibited the largest training variability, likely due to the largest range of PAA concentrations relative to the other sampling times and locations. The impact of training data selection is most important for predicting the PAA concentrations near the end of the disinfection basin (20 and 30 min, which are the end of the second and third passes, respectively).

As the sampling location moves downstream, the error declines (e.g., “2018 Half” or “2019 10 min”, along the second pass of the serpentine reactor). By the third pass (e.g., “2019 20 min” or “2019 30 min”), the error is similar to that of the “1 min” locations. This is most likely reflective of the increased degree of mixing due to non-ideal flow conditions. As is shown in ref 32, ANNs are capable of capturing the effect of flow dynamics on the ANN model response. A second important difference between the model accuracy at the “1 min” location compared to the downstream locations is the more prominent difference in RMSE among instantaneous, short-term, and 24 h averages. This could indicate that changes of an online process variable (Table 1) are not noise but instead are important controlling parameters that impact PAA decay.

ANNs are found to predict PAA concentrations throughout a full-scale disinfection basin within the range of 0.14–0.19 mg/L based on the standard error (SE) for the best model fit

of each sampling location. This is approximately twice the SE of a standard PAA online analyzer (0.07 mg/L SE for the PAA analyzer at the RWHTF) but is still a reasonable approximation for some applications. The benefit of an online analyzer is to respond to water quality changes in nearly real time; however, the next section illustrates that (i) the use of an online analyzer may not improve the precision with which PAA is dosed as substantially as previously assumed and (ii) the use of an ANN to predict CT is a better alternative for dose control.

3.2. ANN for Prediction of CT. From each sampling event in 2019, zero-order (i.e., linear) and first-order (i.e., exponential) decay models were fit to PAA concentrations as a function of HRT to determine if the decay kinetics were truly first-order, as is described in the majority of the literature, or are zero-order, as reported by Santoro et al.⁴² For all HRTs, the R^2 of the zero-order model was significantly less than the R^2 of the first-order model (one-sided Wilcoxon signed-rank test p value = 0.009; median zero-order R^2 = 0.973; median first-order R^2 = 0.983). When the HRT was >45 min, the difference in model fit between zero- and first-order models became more significant with a smaller p value (p value = 8.1×10^{-5} ; median zero-order R^2 = 0.927; median first-order R^2 = 0.993). While PAA decay at short HRTs (<45 min) may appear to be approximated by zero-order decay, the true function is likely first-order exponential decay, and thus, CT can be calculated using eq 1.

By integration of eq 1, the actual CT was calculated for each sampling event. In Figure 3, the dashed black line represents

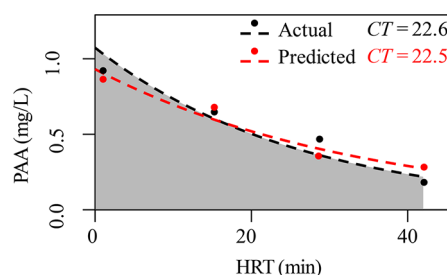


Figure 3. Observed (black) and average predicted (red) CT from the 2019 sampling campaign. The black dots are the actual sampled PAA concentrations; the black dashed line is the first-order model fit to the actual sampled observations, and the gray area beneath the black curve is the actual CT. The red dots represent the average of 10 ANN model predictions of PAA concentrations at four sampling locations, trained on instantaneous process data. The red dashed line is the first-order model fit to the predicted observations.

$C(t)$ and was fit to the PAA concentration at each sampling location (where there are two observations per sampling event in 2018 and four observations per event in 2019). The shaded gray area under the curve sums to CT and is labeled as “actual CT” in the top right of each plot. Compared to the actual PAA disinfection kinetics, the average ANN prediction using instantaneous data is plotted in red in Figure 3. $C(t)$ fit to the ANN-predicted concentrations is plotted as the dashed red line. The area under the predicted PAA concentration model fit equals the “predicted CT” listed in the top right corner.

Real-time CT was calculated by ANN predictions of PAA concentrations at designated sampling locations followed by a first-order kinetic model fit. No significant difference from zero was found between the median difference in actual and predicted CT for all sampling events (the two-sided paired

Wilcoxon signed-rank test fails to find a significant difference; $p = 0.41$). A simpler approach was also tested in which CT was directly predicted by an ANN (same set of predictors as PAA concentration ANN models). Comparatively, the direct prediction approach predicted CT with almost equal accuracy as the PAA concentration approach (Figure 4) (the two-sided paired Wilcoxon signed-rank test fails to find a significant difference in median CT; p value = 0.61).

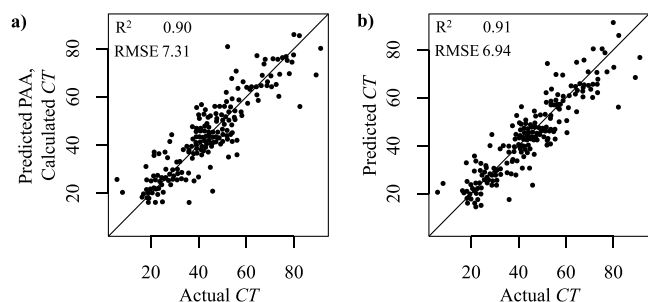


Figure 4. Accuracy of the prediction of CT from a first-order model integration of (a) ANN-predicted PAA concentrations and (b) ANN-predicted CT from the same set of predictor variables. No significant difference was found between the median CT of the model approaches.

In conventional CT dose control, the decay parameter (k in eq 1) and the demand parameter (D in eq 1) are assumed to be constant, an artifact of chlorine-based CT dosing wherein such an assumption is valid. If an online analyzer is available and the HRT to the analyzer is known, then D can be calculated assuming a constant k . The impact of these assumptions on conventionally calculated CT relative to actual CT for PAA disinfection is illustrated in Figure 5. Our real-time ANN CT prediction is also plotted for comparison. The 2018 data set (left panel of Figure 5) replicates the constant D and k control strategy. The 2019 data set (right panel of Figure 5) replicates a constant k control strategy because an online analyzer was installed prior to the sampling campaign for that purpose. The top two rows of Figure 5 show the actual k and D parameters as black triangles. The red dashed line is the mean parameter value and illustrates the parameter being held constant. For the 2019 trial, the D parameter was calculated by solving for the initial concentration term in $C(t)$ given the concentration at the analyzer (known t), initial dose, and mean k .

There is substantial variation around the mean k and D for both the 2018 and 2019 data sets (Figure 5). A slightly larger error for the 2019 k RMSE relative to 2018 is expected, given the smaller sample size over a longer period of time. However, the increased error in D with the addition of the PAA analyzer is surprising. The use of a mean parameter value is not a true real-time simulation; thus, it is important to note the increased error in the conventional D calculation between 2018 and 2019 is not necessarily indicative of a poor analyzer. Rather, the analyzer can reasonably approximate the mean D when the true mean is unknown, as in a real-time scenario. The observed error (as RMSE) between the conventional CT and actual CT is higher in 2018 than in 2019, most likely due to the significantly larger range of CT values observed in 2018. However, the RMSE for all CT predictions is still relatively low (6.94 mg/L from Figure 4b). As a general consequence of the static parameter assumptions, the conventional CT consis-

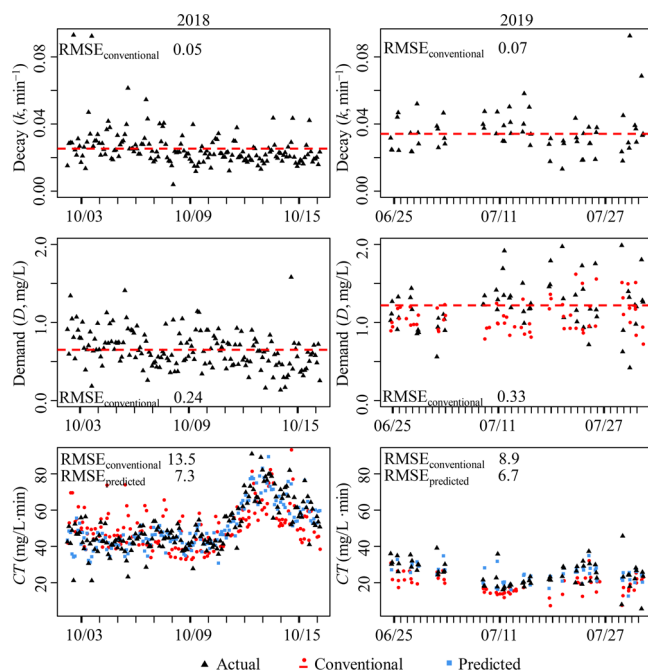


Figure 5. Actual decay (k) (top row), demand (D) (middle row), and CT (bottom row) are calculated for each sampling event in 2018 (left column) and 2019 (right column) (black triangles). Two common disinfection control methods (static k and D ; static k and variable D calculated from an online analyzer) are simulated and labeled “conventional” (red lines and dots). The ANN CT prediction (blue squares) has a lower prediction error ($RMSE_{\text{predicted}}$) compared to conventional CT ($RMSE_{\text{conventional}}$).

tently over- or underestimated the actual CT, depending on whether the actual D was above or below the conventional value. For all cases, ANN CT prediction was more a more accurate approximation of actual CT compared to conventional CT ($RMSE_{\text{predicted}} < RMSE_{\text{conventional}}$ for 2018 and 2019).

CT control is considered a more precise disinfectant dosing strategy than concentration-based control because of the increased microbial inactivation observed when t increases at the same initial PAA concentration. However, CT control requires knowledge of how the PAA concentration changes as a function of time to construct a model. In conventional CT control, this would require approximations of PAA kinetics, which have a large error, or the installation of multiple analyzers to fit eq 1 in real time, which is cost-prohibitive.

From Figure 5, approximations of D and/or k cannot accurately estimate the true behavior of PAA in secondary treated wastewater. The use of an ANN-predicted CT in a cascade control strategy could substantially reduce the amount of excess PAA dosed. For example, in 2019 CT was frequently underpredicted using an online analyzer. The control system assumes a CT set point by the conventional CT even though the actual CT is substantially above the set point, meaning excess PAA is dosed. If the ANN-predicted CT is used in this case, the error between the calculated CT in the set point control loop and the actual CT would be smaller. Considering the percent difference in the RMSE, the ANN-predicted CT is 46% more accurate than the conventional CT and 25% more accurate compared to the conventional CT with an analyzer. Thus, PAA dose could have been decreased, thereby reducing chemical costs.

3.3. CT ANN Variable Selection. ANNs used to predict CT were also constructed using forward variable selection: one predictor at a time is used to predict CT using an ANN, the predictor with the lowest testing error is selected as the most influential model input, and the ANN is retrained using the most influential predictor and one of the remaining model predictors at a time. The sequential addition of process variables allows the error of the ANN models to be compared, and the set of process variables that produces the lowest RMSE is considered to be the most influential or important set of model predictors. For CT, six process variables were identified that sequentially produced lower RMSEs of the variable selection set of ANN. However, possibly due to the smaller number of training epochs to avoid overfitting in this set of ANN experiments, the lowest RMSE for six iterations of variable selection (7.67) is slightly higher than the ANN CT prediction with all predictors (RMSE of 6.95 for 2018 and 2019). The variables selected for the CT ANN model for each step are plotted as a function of the average ANN testing error in Figure 6.

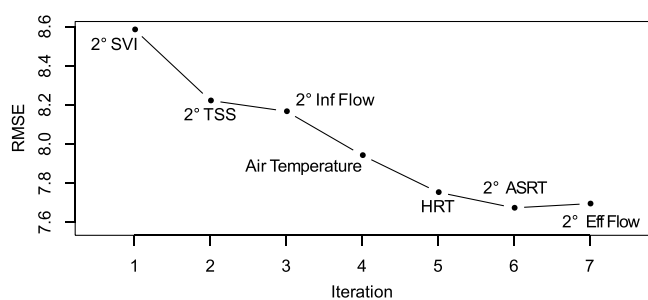


Figure 6. Forward variable selection for a CT ANN. The number of iterations is equal to the number of inputs for the ANN model, and iterations of more than one include all preceding predictors in the ANN model.

Predictors are selected on the basis of the minimum RMSE for a set of predictor combinations. In this case (Figure 6), the addition of secondary (2°) effluent flow increased the testing RMSE on the seventh iteration of model building. Thus, only variables used up to and including the sixth iteration are selected as highly influential in predicting CT. It is expected that HRT is included in the model, as it was used to calculate the values of CT used for training and testing (eq 1). The inclusion of temperature is also expected, as most chemical disinfectants are more effective at higher temperatures due to speciation. The most surprising predictor variables selected for the CT ANN describe the properties of the sludge in the secondary (2°) treatment basins. The concentration of solids (TSS), the age of the solids (aerobic solid retention time or ASRT), and the settleability of the solids (sludge volume index or SVI) have been understood for decades to impact final water quality variables such as TSS and BOD. Excess TSS and BOD in the final treated water could consume active PAA, thereby accelerating PAA decay and preventing full disinfection. While the relationship of TSS or BOD with PAA decay has been identified in the literature,^{9,14,19–21} this is the first time that it has been synergistically and quantifiably included in a model of CT.

3.4. ANN and RNN Prediction of *E. coli* Inactivation. To compare conventional *E. coli* inactivation models to full-scale PAA disinfection performance, traditional microbial

inactivation models from the literature were fit to the log inactivation observed in this work. Multiple models were fit, including linear, exponential, power, double-exponential,⁴³ Chick–Watson,¹⁶ and Holm model,¹⁶ and the results are plotted in Figure S5. However, these mechanistic model results demonstrated low R^2 (<0.38) and high RMSE (>2 log). There are two possible sources of error that contribute to the poor model fit, including (1) measurement error and (2) variation in *E. coli* inactivation by PAA that is not solely described by CT.

The calculation for log inactivation is dependent on the predisinfection and postdisinfection concentrations of *E. coli*. When the predisinfection *E. coli* concentration is low, log removal appears to be a poor indicator of disinfection performance as the postdisinfection *E. coli* concentration does not change significantly relative to the predisinfection *E. coli* concentration. This concept is explored in Figure 7, where

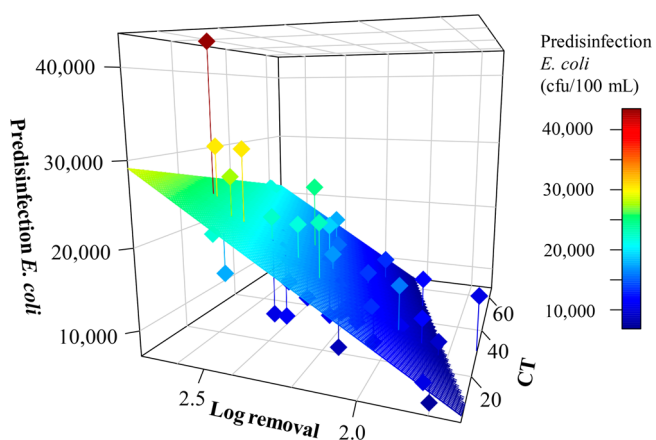


Figure 7. Plots of PAA CT, *E. coli* log removal, and predisinfection *E. coli* concentration for sampling events (diamonds) with a linear plane to model predisinfection *E. coli* concentration as a function of CT and log removal.

predisinfection *E. coli*, log inactivation, and CT are plotted to better understand the relationship between PAA dose and *E. coli* concentration. Despite a large range of actual CT (20–60 $\text{mg L}^{-1} \text{min}^{-1}$), there is a small range of log removal (2.0–2.5). This is counter to the conventional understanding of disinfection performance that correlates a CT value to a desired log removal. In this case, log removal is more sensitive to a high predisinfection *E. coli* concentration than to a low postdisinfection *E. coli* concentration. This may be due to the measurement bias discussed in the introduction that prevents the lowest *E. coli* concentrations from being accurately measured (see Introduction). Given that the variation in log removal is influenced more by the predisinfection *E. coli* concentration than CT, increasing the applied PAA dose when the predisinfection *E. coli* concentration is low may not substantially increase log removal. Thus, we suggest evaluating the performance of PAA disinfection systems by the postdisinfection *E. coli* concentration as opposed to the log removal metric. Next, we continue the *E. coli* exploration by predicting pre- and postdisinfection *E. coli* concentrations using an ANN and an RNN to capture the complex dynamics of *E. coli* inactivation.

Predisinfection and postdisinfection *E. coli* concentrations are highly correlated over time (see Figure S4); however, it is impractical to obtain a measurement of *E. coli* for a water

quality sample in <24 h, and results are often not available until approximately 42 h after the sample is taken (on average at the RWHTF). Therefore, the *E. coli* ANN and RNN models included a lagged *E. coli* predictor to address autocorrelation. To simulate real-time prediction, the lag was 42 h. The *E. coli* concentrations were also $\log_{10}$ -transformed prior to normalization and training to improve performance.

The distributions of RMSE for ANN and RNN predictions (based on 10 iterations of randomly initiated sets of hyperparameters for model training and testing per observation) of predisinfection and postdisinfection *E. coli* concentrations using instantaneous, short-term averaged, and 24 h averaged process data are shown in Figure 8. For comparison,

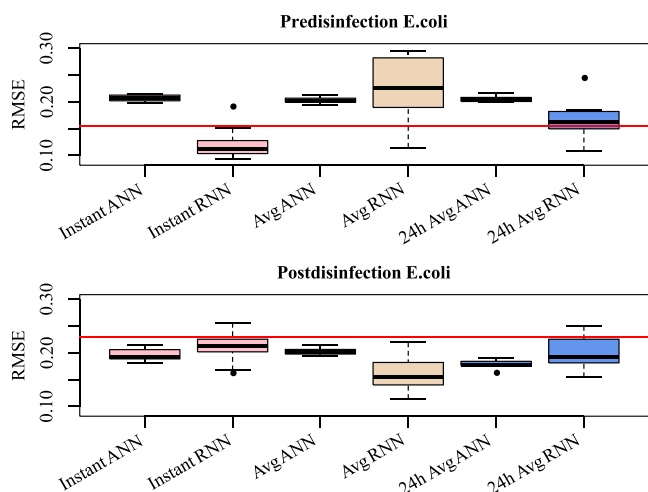


Figure 8. Box plots of RMSE for $\log_{10}$ -transformed predisinfection (top) and postdisinfection (bottom) *E. coli* model predictions. The red line is the average RMSE of persistence, i.e., the *E. coli* concentration 42 h prior.

the ANN predisinfection *E. coli* model averages 4.9 s to train and test an observation whereas the RNN model averages 10.4 min per observation; in other words, training and testing the ANN is 127 times faster than the RNN. The prediction errors are compared to the error of a model that uses the previously known value, the *E. coli* measurement 42 h prior (i.e., persistence). Only the RNN model trained on instantaneous data can predict predisinfection *E. coli* concentration better than persistence. Therefore, it is suspected that predisinfection *E. coli* ANN prediction accuracy depends on (1) the most recently measured *E. coli* and (2) water quality variables in the model that have a data collection frequency of ≥ 24 h. A previous study showed that ammonia, orthophosphate, and TSS concentrations were significantly correlated with *E. coli* removal in WWTPs.⁴⁴ Therefore, predisinfection *E. coli* concentrations could be more heavily influenced by the performance of nutrient removal processes. More work is needed to identify additional WWTP features that impact *E. coli* removal upstream of the disinfection process.

ANN predictions of postdisinfection *E. coli* are more accurate than persistence, but the lowest mean error is the RNN using short-term averaged data. Upon inclusion of a lagged response variable, the accuracy of the postdisinfection *E. coli* ANN models was comparable to those of the instantaneous and 24 h average RNN models. This could save a substantial amount of time and make the prediction framework possible for real-time control, as RNNs require hours to train and test

whereas ANNs take minutes. In the case of both CT and *E. coli*, the real-time model update frequencies are 24 and 42 h, respectively. Therefore, the additional hour required to build an RNN as opposed to an ANN may be worth the delay if the prediction is used for process monitoring.

For implementation of a real-time *E. coli* prediction in process control, we suggest that the optimum postdisinfection *E. coli* concentration model be selected as opposed to a log inactivation model or predisinfection *E. coli*. The ability to predict the predisinfection *E. coli* concentration is only valuable when there is a definitive relationship between dose and inactivation. Dose could be more effectively controlled if the postdisinfection *E. coli* concentration is accurately predicted, and the dose could then be increased or decreased relative to an *E. coli* set point. A factor of safety accounting for the RMSE would accommodate model error. For example, the best performing postdisinfection *E. coli* model used an RNN and instantaneous process data for 0.13 log. To maintain 126 cfu/100 mL, a conservative postdisinfection *E. coli* set point would be 93 cfu/100 mL $[=10^{\log(126)-0.13}]$.

3.5. *E. coli* ANN Variable Selection. The ANN models used to predict pre- and postdisinfection *E. coli* concentrations were used for the feedforward variable selection (Figure 9),

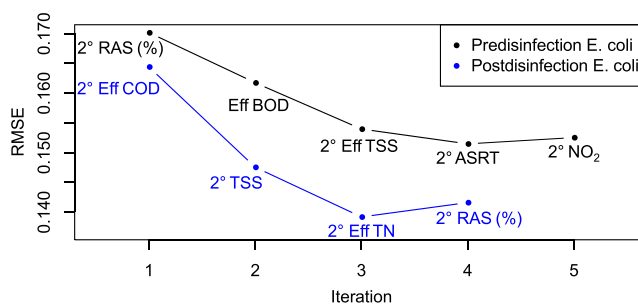


Figure 9. Forward variable selection for *E. coli* ANN. The number of iterations is equal to the number of inputs for the ANN model, and iterations greater than one include all preceding predictors. Predictors are selected on the basis of the minimum RMSE, and model selection stops when the minimum RMSE is higher than that of the previous iteration.

similar to CT variable selection discussed in section 3.3. The predisinfection *E. coli* concentration is most heavily influenced by the operating conditions of the 2° treatment process immediately upstream of disinfection. The most influential model predictor was the recirculated activated sludge (RAS) flow as a percentage of the influent flow to the secondary treatment system. This recirculation rate is an important control parameter for nutrient removal and is increased or decreased to achieve the desired nitrogen and phosphorus removal. The RAS flow is related to the average sludge age of the secondary treatment system, as flow that is not removed from the secondary process for solids treatment to limit the ASRT is returned as the RAS. Similar to CT, the properties of the biologically active solids in the secondary process and the performance of carbon and nitrogen removal [e.g., COD, BOD, and total nitrogen (TN)] appear to play an important role in pre- and postdisinfection *E. coli* concentration.

4. CONCLUSION

This work illustrates the potential of machine learning to model the near real-time PAA disinfection performance of a

full-scale WWTP with considerations for water quality and operational changes. We have shown that the full-scale disinfection performance of PAA for secondary wastewater effluent at the RWHTF is not accurately described using traditional and batch literature models for PAA concentration, CT, and *E. coli* removal, and that a machine learning modeling approach could supplement costly analyzers in process control. Further study of the integration of model predictions with existing process control strategies is needed to fully test the real-world applicability of the models developed here. Given the ability of machine learning methods to map predictors to response variables in process engineering, ANNs and RNNs can predict the CT (using PAA concentration as a function of time and integrated over the HRT) and *E. coli* concentrations using online and laboratory data at a full-scale WWTP.

The wide ranges of demand (*D*) and decay (*k*) kinetic parameters observed for PAA at RWHTF suggest that both are directly related to changing water quality, and static values should not be assumed for either when designing a CT-based control strategy. A suitable, more reliable, and less expensive alternative to analyzer-based CT control is the use of an ANN to predict CT using a mixture of online and lab process data. While the secondary treatment data used in this study do capture information about water quality that measurably changes CT, the variables included in the training data do not fully describe the variation in predisinfection *E. coli* concentration. This error is propagated to the postdisinfection *E. coli* concentration, although the accuracy of the RNN does improve over the former.

The time required to train and test a neural network is proportional to the number of parameters in ANNs and RNNs and governs how quickly new laboratory observations can be incorporated into control systems. If the prediction is used in the control, ANNs may be preferential due to their comparable error rates, reduced complexity, and consistency, but the application may require an accuracy that only a more complex network, like an RNN, can achieve. For example, if an RNN is used to predict pre- and postdisinfection *E. coli* concentrations, then deviation from the predictions could provide WWTP operations with indications of process upsets or inefficiencies. For direct control, an ANN could be used to accurately predict real-time CT. By adjusting PAA dosing to achieve a target CT, a WWTP may be able to prevent overdosing with a costly disinfectant without a comparably costly analyzer.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acsestwater.0c00095>.

Figures that support the use of an exponential model and illustrate the variation in first-order decay parameters, autocorrelation of *E. coli*, and conventional modeling of *E. coli* inactivation as a function of CT (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Kathryn B. Newhart — Colorado School of Mines, Golden, Colorado 80401, United States; Metro Wastewater Reclamation District, Denver, Colorado 80229, United States; orcid.org/0000-0002-7071-6536; Email: kbnewhart@mines.edu

Amanda S. Hering — Baylor University, Waco, Texas 76706, United States; Email: mandy_hering@baylor.edu

Authors

Joshua E. Goldman-Torres — Metro Wastewater Reclamation District, Denver, Colorado 80229, United States

Daniel E. Freedman — Metro Wastewater Reclamation District, Denver, Colorado 80229, United States

K. Blair Wisdom — Metro Wastewater Reclamation District, Denver, Colorado 80229, United States

Tzahi Y. Cath — Colorado School of Mines, Golden, Colorado 80401, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acsestwater.0c00095>

Author Contributions

J.E.G.-T., D.E.F., and K.B.W. operated the PAA system and modeled PAA disinfection. A.S.H. and T.Y.C. reviewed and evaluated the machine learning methods and performance.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation Partnership for Innovation: Building Innovation Capacity (Project 1632227); the National Science Foundation Engineering Research Center program, ReNUWit (Cooperative Agreement EEC-1028968); a grant from the Colorado Higher Education Competitive Research Authority (CHECRA); and Metro Wastewater Reclamation District's operations, laboratory, and technology and innovation staff. Engineers at the Metro Wastewater Reclamation District contributed through operation of the PAA system and experience modeling PAA disinfection. Academic partners contributed through the review and evaluation of machine learning methods and performance.

■ REFERENCES

- (1) Krasner, S. W.; Westerhoff, P.; Chen, B.; Rittmann, B. E.; Amy, G. Occurrence of Disinfection Byproducts in United States Wastewater Treatment Plant Effluents. *Environ. Sci. Technol.* **2009**, *43* (21), 8320–8325.
- (2) Comprehensive Disinfectants and Disinfection Byproducts Rules (Stage 1 and Stage 2): Quick Reference Guide. EPA 816-F-10-080; Office of Water, U.S. Environmental Protection Agency, 2010.
- (3) Crebelli, R.; Conti, L.; Monarca, S.; Feretti, D.; Zerbini, I.; Zani, C.; Veschetti, E.; Cutilli, D.; Ottaviani, M. Genotoxicity of the Disinfection By-Products Resulting from Peracetic Acid- or Hypochlorite-Disinfected Sewage Wastewater. *Water Res.* **2005**, *39* (6), 1105–1113.
- (4) Monarca, S.; Richardso, S. D.; Feretti, D.; Grottolo, M.; Thruston, A. D.; Zani, C.; Navazio, G.; Ragazzo, P.; Zerbini, I.; Alberti, A. Mutagenicity and Disinfection By-Products in Surface Drinking Water Disinfected with Peracetic Acid. *Environ. Toxicol. Chem.* **2002**, *21* (2), 309–318.
- (5) da Silva, W. P.; Carlos, T. D.; Cavallini, G. S.; Pereira, D. H. Peracetic Acid: Structural Elucidation for Applications in Wastewater Treatment. *Water Res.* **2020**, *168*, 115143.
- (6) Kitis, M. Disinfection of Wastewater with Peracetic Acid: A Review. *Environ. Int.* **2004**, *30* (1), 47–55.
- (7) Gehr, R.; Wagner, M.; Veerasubramanian, P.; Payment, P. Disinfection Efficiency of Peracetic Acid, UV and Ozone after Enhanced Primary Treatment of Municipal Wastewater. *Water Res.* **2003**, *37* (19), 4573–4586.

- (8) Rossi, S.; Antonelli, M.; Mezzanotte, V.; Nurizzo, C. Peracetic Acid Disinfection: A Feasible Alternative to Wastewater Chlorination. *Water Environ. Res.* **2007**, 79 (4), 341–350.
- (9) Block, S. S. *Disinfection, Sterilization, and Preservation*, 5th ed.; LWW: Philadelphia, 2000.
- (10) Lefevre, F.; Audic, J.; Ferrand, F. Peracetic Acid Disinfection of Secondary Effluents Discharged off Coastal Seawater. *Water Sci. Technol.* **1992**, 25 (12), 155–164.
- (11) Bianchini, R.; Calucci, L.; Pinzino, C.; Piscicelli, M. An EPR Study on Wastewater Disinfection by Peracetic Acid, Hydrogen Peroxide and UV Irradiation. *Ann. Chim.* **2002**, 92 (9), 783–793.
- (12) Liberti, L.; Lopez, A.; Notarnicola, M. Disinfection with Peracetic Acid for Domestic Sewage Re-Use in Agriculture. *Water Environ. J.* **1999**, 13 (4), 262–269.
- (13) Baldry, M.; French, M. Disinfection of Sewage Effluent with Peracetic Acid. *Water Sci. Technol.* **1989**, 21 (3), 203–206.
- (14) Sánchez-Ruiz, C.; Martínez-Royano, S.; Tejero-Monzón, I. An Evaluation of the Efficiency and Impact of Raw Wastewater Disinfection with Peracetic Acid Prior to Ocean Discharge. *Water Sci. Technol.* **1995**, 32 (7), 159–166.
- (15) Dell'Erba, A.; Falsanisi, D.; Liberti, L.; Notarnicola, M.; Santoro, D. Disinfecting Behaviour of Peracetic Acid for Municipal Wastewater Reuse. *Desalination* **2004**, 168, 435–442.
- (16) Haas, C. N.; Joffe, J.; Jacangelo, J. G.; Anmangandla, U.; Heath, M. Water Quality and Disinfection Kinetics. *J. - Am. Water Works Assoc.* **1996**, 88 (3), 95.
- (17) Manoli, K.; Sarathy, S.; Maffettone, R.; Santoro, D. Detailed Modeling and Advanced Control for Chemical Disinfection of Secondary Effluent Wastewater by Peracetic Acid. *Water Res.* **2019**, 153, 251–262.
- (18) Block, P. The Decomposition Kinetics of Peracetic Acid and Hydrogen Peroxide in Municipal Wastewaters. *Proceedings of the Water Environment Federation* **2016**, 2016 (10), 555–563.
- (19) Falsanisi, D.; Gehr, R.; Santoro, D.; Dell'Erba, A.; Notarnicola, M.; Liberti, L. Kinetics of PAA Demand and Its Implications on Disinfection of Wastewaters. *Water Qual. Res. J. Can.* **2006**, 41 (4), 398–409.
- (20) Baldry, M. G. C.; French, M. S.; Slater, D. The Activity of Peracetic Acid on Sewage Indicator Bacteria and Viruses. *Water Sci. Technol.* **1991**, 24 (2), 353–357.
- (21) Stampi, S.; De Luca, G.; Zanetti, F. Evaluation of the Efficiency of Peracetic Acid in the Disinfection of Sewage Effluents. *J. Appl. Microbiol.* **2001**, 91 (5), 833–838.
- (22) Harrou, F.; Sun, Y.; Hering, A. S.; Madakyaru, M.; Dairi, A. Chapter 3 - Fault Isolation. In *Statistical Process Monitoring Using Advanced Data-Driven and Deep Learning Approaches*; Harrou, F., Sun, Y., Hering, A. S., Madakyaru, M., Dairi, A., Eds.; Elsevier, 2021; pp 71–117. DOI: 10.1016/B978-0-12-819365-5.00009-7
- (23) Harrou, F.; Sun, Y.; Hering, A. S.; Madakyaru, M.; Dairi, A. Chapter 8 - Case Studies. In *Statistical Process Monitoring Using Advanced Data-Driven and Deep Learning Approaches*; Harrou, F., Sun, Y., Hering, A. S., Madakyaru, M., Dairi, A., Eds.; Elsevier, 2021; pp 255–303. DOI: 10.1016/B978-0-12-819365-5.00014-0
- (24) Dreyfus, G. Neural Networks: An Overview. *Neural Networks SE - 1* **2005**, 1–83.
- (25) Nielsen, M. A. *Neural Networks and Deep Learning*; Determination Press, 2015.
- (26) Schmidhuber, J. Deep Learning in Neural Networks: An Overview. *Neural Networks* **2015**, 61, 85–117.
- (27) Sérodes, J.-B.; Rodriguez, M. J.; Ponton, A. Chlorcast©: A Methodology for Developing Decision-Making Tools for Chlorine Disinfection Control. *Environmental Modelling & Software* **2001**, 16 (1), 53–62.
- (28) Gibbs, M. S.; Morgan, N.; Maier, H. R.; Dandy, G. C.; Nixon, J. B.; Holmes, M. Investigation into the Relationship between Chlorine Decay and Water Distribution Parameters Using Data Driven Methods. *Mathematical and Computer Modelling* **2006**, 44 (5), 485–498.
- (29) Lin, C.-H.; Yu, R.-F.; Cheng, W.-P.; Liu, C.-R. Monitoring and Control of UV and UV-TiO₂ Disinfections for Municipal Wastewater Reclamation Using Artificial Neural Networks. *J. Hazard. Mater.* **2012**, 209–210, 348–354.
- (30) Yu, R.-F.; Chen, H.-W.; Cheng, W.-P.; Shen, Y.-C. Dynamic Control of Disinfection for Wastewater Reuse Applying ORP/PH Monitoring and Artificial Neural Networks. *Resources, Conservation and Recycling* **2008**, 52 (8), 1015–1021.
- (31) Carvajal, G.; Roser, D. J.; Sisson, S. A.; Keegan, A.; Khan, S. J. Bayesian Belief Network Modelling of Chlorine Disinfection for Human Pathogenic Viruses in Municipal Wastewater. *Water Res.* **2017**, 109, 144–154.
- (32) Wei, W.; Haas, C. N.; Farouk, B. *Optimized Design of Wastewater Disinfection Reactors Based on an Artificial Neural Network Metamodel*; American Society of Mechanical Engineers Digital Collection, 2017. DOI: 10.1115/IMECE2016-65139
- (33) Wisdom, B.; Werth, E.; Stephens, N.; Da Silva, A. Peracetic Acid Quenching with Sodium Bisulfite. *Proceedings of the Water Environment Federation* **2018**, 2018 (17), 469–477.
- (34) Standard methods for the examination of water and wastewater. 4500 Peracetic Acid (Residual) (Proposed); 2019.
- (35) Standard methods for the examination of water and wastewater. 9223 Enzyme Substrate Coliform Test; 2012. DOI: 10.2105/SMWW.2882.194
- (36) Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Computation* **1997**, 9 (8), 1735–1780.
- (37) Harrou, F.; Sun, Y.; Hering, A. S.; Madakyaru, M.; Dairi, A. Chapter 7 - Unsupervised Recurrent Deep Learning Scheme for Process Monitoring. In *Statistical Process Monitoring Using Advanced Data-Driven and Deep Learning Approaches*; Harrou, F., Sun, Y., Hering, A. S., Madakyaru, M., Dairi, A., Eds.; Elsevier, 2021; pp 225–253. DOI: 10.1016/B978-0-12-819365-5.00013-9
- (38) Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G. S.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Goodfellow, I.; Harp, A.; Irving, G.; Isard, M.; Jia, Y.; Jozefowicz, R.; Kaiser, L.; Kudlur, M.; Levenberg, J.; Mané, D.; Monga, R.; Moore, S.; Murray, D.; Olah, C.; Schuster, M.; Shlens, J.; Steiner, B.; Sutskever, I.; Talwar, K.; Tucker, P.; Vanhoucke, V.; Vasudevan, V.; Viégas, F.; Vinyals, O.; Warden, P.; Wattenberg, M.; Wicke, M.; Yu, Y.; Zheng, X. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. *arXiv* **2015**, 1603.04467v2.
- (39) Chollet, F. Keras; <https://github.com/fchollet/keras>, 2015.
- (40) Guyon, I.; Elisseeff, A. An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research* **2003**, 3 (Mar), 1157–1182.
- (41) May, R.; Dandy, G.; Maier, H. Review of Input Variable Selection Methods for Artificial Neural Networks. *Artificial Neural Networks-Methodological Advances and Biomedical Applications* **2011**, 10, 16004.
- (42) Santoro, D.; Gehr, R.; Bertrand, T. A.; Liberti, L.; Notarnicola, M.; Dell'Erba, A.; Falsanisi, D.; Haas, C. N. Wastewater Disinfection by Peracetic Acid: Assessment of Models for Tracking Residual Measurements and Inactivation. *Water Environ. Res.* **2007**, 79 (7), 775–787.
- (43) Santoro, D.; Crapulli, F.; Raisee, M.; Raspa, G.; Haas, C. N. Nondeterministic Computational Fluid Dynamics Modeling of Escherichia Coli Inactivation by Peracetic Acid in Municipal Wastewater Contact Tanks. *Environ. Sci. Technol.* **2015**, 49 (12), 7265–7275.
- (44) Barrios-Hernández, M. L.; Pronk, M.; Garcia, H.; Boersma, A.; Brdjanovic, D.; van Loosdrecht, M. C. M.; Hooijmans, C. M. Removal of Bacterial and Viral Indicator Organisms in Full-Scale Aerobic Granular Sludge and Conventional Activated Sludge Systems. *Water Research X* **2020**, 6, 100040.