Chapter 3

Effectiveness of PFA disinfection implemented at full scale (Seine Valenton WWTP)

© KEMIRA

3.1 INTRODUCTION

The industrial-scale trials performed at the Seine Valenton (SEV) wastewater treatment plant (WWTP) by SIAAP in 2018 were aimed at confirming the effectiveness of PFA on fecal bacteria and identifying potential optimization levers. These trials were conducted between August and October 2018 by disinfecting the totality of the WWTP discharge (nominal capacity: 550,000 m^3/day) using PFA on a week-on/week-off basis. A complete description of these industrial-scale trials is given in Part 2, Chapter 1. This third chapter will present the full set of results on PFA disinfection effectiveness generated from these trials. First, the fecal bacteria removal effectiveness will be discussed in

© 2021 The Editors. This is an Open Access eBook distributed under the terms of the Creative Commons Attribution Licence (CC BY-NC-ND 4.0), which permits copying and redistribution for non-commercial purposes with no derivatives, provided the original work is properly cited (https://creativecommons.org/licenses/by-nc-nd/4.0/). This does not affect the rights licensed or assigned from any third party in this book.
doi: 10.2166/9781789062106_0075

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

relation to the C × t (concentration × time) of PFA applied, along with the variability induced by daily variations in the SEV WWTP discharge and the impact of PFA on other pathogens (spores of sulfite-reducing anaerobes and F-specific RNA phages). Second, the impact of SEV WWTP discharge quality variations on fecal bacteria removal will be studied with a focus on the effect of a degradation in SEV WWTP operations and the potential for results to be normalized by initial water quality. Third, the prediction of *E. coli* and intestinal enterococci removal in SEV WWTP discharge by a statistical analysis of the industrial-scale results will be investigated. Fourth and last, the impact of PFA disinfection on the discharge quality will be quantified in terms of both conventional quality parameters and disinfection byproduct formation.

3.2 EFFECTIVENESS OF PFA DISINFECTION APPLIED TO SEV WWTP DISCHARGE

3.2.1 Fecal bacteria removal achieved during the industrial-scale trials at SEV WWTP

The overall quality and variability of the SEV WWTP discharges were in the normal range and are given in Table 13, along with the quality after disinfection. The reducing compounds capable of reacting with PFA (Ragazzo *et al.*, 2013), such as ammonia, nitrite and COD, were of very limited quantity in this effluent. In comparison, the average concentrations in samples used for laboratory-scale trials were: 8.7 ± 5.5 mg/L for TSS, 5.8 ± 0.8 mgC/L for DOC, 23 ± 8 mgO_2/L for COD, 2.0 ± 1.2 mgO_2/L for BOD_5, 1.1 ± 0.7 mgN/L for TKN, and 1.2 ± 0.7 mgP/L for TP. Similarly, the *E. coli* and intestinal enterococci concentrations in the samples were similar to what is normally encountered in this type of water, with median ± min-max log concentrations of: 4.27 ± 3.84–5.55 and 3.56 ± 3.03–4.93 MPN/100 mL, respectively, in the industrial-scale test samples (Rocher & Azimi, 2016). In addition, these results are highly comparable to log concentrations measured during laboratory-scale trials, that is,: 4.16 ± 3.23–5.83 (Mann–Whitney test, p-value = 0.180) and 3.67 ± 2.85–4.80 MPN/100 mL (Mann–Whitney test, p-value = 0.533), respectively, for *E. coli* and intestinal enterococci. The variability in water quality originates from different SEV WWTP operating conditions over the period (internal bypass, dry and rainy weather, etc.).

Figure 34 presents the results of the 43 campaigns performed within the PFA industrial-scale disinfection unit treating the SEV WWTP discharge. It represents the median concentration of *E. coli* or intestinal enterococci, with the 1st and 3rd quartile concentrations observed as error bars (upper part of the figure) and the average logarithmic removal of both fecal bacteria with the standard deviation shown as error bars (lower part). The quality limits (900 NPP/100 mL for *E. coli* and 330 NPP/100 mL for intestinal enterococci) to respect the European recreational bathing regulation (*Directive 2006/7/EC of 15 February 2006*

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

concerning the management of bathing water quality and repealing Directive 76/160/EEC, 2006) are provided as indications. During these campaigns, the actual PFA dose applied varied between 0.61 and 2.48 ppm of PFA (median value: 1.20 ppm), while depending on the SEV WWTP discharge flow, the hydraulic retention time in the discharge tunnel was 19.4 and 56.1 min (median value: 29.6 min), yielding C × t values ranging from 16 to 74 ppm.min (median value: 32 ppm.min). Four classes of C × t are depicted in Figure 34: 15–30, 30–50, 50–70, and 70–75 ppm.min. The water temperature is typical for the summer period, with an average and standard deviation of 23 ± 1°C.

Similarly to the laboratory-scale trials, Figure 34 highlights the strong impact of PFA C × t applied to both the logarithmic removals and residual concentration of *E. coli* and intestinal enterococci. The logarithmic removal is stable around 2 log, between 15–30 and 30–50 ppm.min, but increases to 2.5 log above 50 ppm.min. Moreover, the residual concentrations of both pathogens decrease when increasing the PFA C × t, until reaching the limit of quantification at >70 ppm.min. The PFA C × t required to reach a sufficient quality for bathing is 10–30 ppm.min, while it rises to 50 ppm.min in order to reach residual concentrations below 100 MPN/100 mL.

The logarithmic removals of *E. coli* and intestinal enterococci are similar at a comparable PFA C × t; this similarity, for example, becomes statistically significant at a C × t of 30–50 ppm.min (Student test, p-value = 0.333). At the laboratory scale (Part 1, Chapter 1), the removal of intestinal enterococci was systematically 1 log lower than that of *E. coli*. Laboratory-scale removals were thus slightly higher than industrial-scale removals for *E. coli*, whereas they were comparable for intestinal enterococci. This finding proves to be significant for a C × t of 10–30 ppm.min since the average removal of *E. coli* is 2.3 ± 0.6 at the laboratory scale vs. 1.9 ± 0.7 at the industrial scale (Student test, p-value = 0.033); for the same C × t, this value is 1.7 ± 0.6 at the laboratory scale and 1.8 ± 0.7 at the industrial scale for intestinal enterococci (Student test, p-value = 0.996). Given that the exposure time was longer during industrial-scale testing (median value: 29.6 min) than the laboratory-scale trials (10 min), a comparable C × t is obtained for a lower PFA dose at the industrial scale. Even though C × t does drive fecal bacteria disinfection by PFA, this result would indicate that the PFA dose also exerts a major effect, in that an overly low dose can limit effectiveness.

3.2.2 Daily variations in fecal bacteria concentrations and impact on the PFA disinfection effectiveness

Figure 35 displays the elimination performance of *E. coli*; Figure 35a shows the residual concentrations obtained after disinfection, over a 24-hour cycle and for various PFA doses, while Figure 35b illustrates the removal rate calculated

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

during this cycle. In Figure 35c, the residual concentrations of *E. coli* are shown vs. the various C × t values. The correlation between removal rate and concentration level of the *E. coli* to be disinfected is shown in Figure 35d.

3.2.2.1 Evaluation of hourly PFA disinfection performance

The residual *E. coli* concentrations were on the whole less than 400 MPN/100 mL and less than 4 MPN/100 mL (limit of quantification) for the 2-ppm treatment dose (Figure 35a). Let us also note that these residual concentrations are low and less than 100 MPN/100 mL between 8pm and 4am (illustrated by a blue band in Figure 35a and b). The concentration levels of *E. coli* are nearly systematically below 900 MPN/100 mL (the bathing threshold), except in one case for which the concentrations of *E. coli* to be disinfected were very high (3.8×10^5 MPN/100 mL). Regarding the reductions in *E. coli*, these varied between 1.6 and 3.8 log according to Figure 35b, the maximum of which is observed at a treatment dose of 2 ppm. These removal rates are logically greater for the highest doses, but do remain high overall at night, starting at the 0.8 ppm dose rate. The low residuals observed during the night suggest this trend is being induced by other parameters.

Although these results are consistent with the monitoring carried out by the reference method with the microplate (see previous section), significantly higher reductions have been recorded here (a maximum of 3.8 log vs. 3.1 log for the reference method). This difference in removal rate is not due to the performance of the analytical methods but rather to the time elapsed between sampling and analysis, which induces the matrix change (bacterial mortality). In fact, this time is zero for the ALERT system, which carries out the sampling and analysis independently, while the reference method requires 24 hours between sampling and laboratory analysis.

3.2.2.2 Focus on key operating parameters

The daily monitoring of disinfection performance has enabled identifying two key parameters to be taken into account during operations: C × t, and the levels of *E. coli* concentrations to be disinfected. Figure 35c shows that the residual concentrations of *E. coli* after disinfection are clearly correlated with the C × t parameter. Independently of the treatment doses, the higher the C × t the greater the decrease in residual *E. coli* concentrations. These low residuals were systematically observed between 8pm and 4am (Figure 35a) when a lower water flow was considered, resulting in a higher C × t value. Above a C × t of 40 ppm.min, the residual concentrations are very low or even zero. Overall, a C × t of 20 ppm.mm seems sufficient to maintain concentrations below the swimming limit of 900 MPN/100 mL, except for a single peak at which a residual of 1175 MPN/100

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

mL in *E. coli* was obtained for a C × t of 22.2 ppm.min (Figure 35c). The variability in *E. coli* residuals observed for the same C × t value can be explained by the concentration levels of *E. coli* to be disinfected. The higher these levels, the more elevated becomes the removal rate. Figure 35d clearly shows this relationship by setting the treatment dose at 0.8 ppm and a C × t ranging from 18 to 26 ppm.min.

In general, close monitoring of the dynamics of wastewater disinfection has made it possible to highlight the importance of taking the C × t parameter into account during industrial operations. However, it must be adapted to both the effluent flow rates and bacteria concentration levels to be disinfected.

3.2.3 Impact of PFA on other types of pathogens

Results obtained during the five sampling campaigns, in which additional pathogens were measured, are given in Table 14. Reductions in *E. coli* and intestinal enterococci are provided to control the representativeness of these campaigns, along with reductions observed for spores of sulfite-reducing anaerobes and F-specific RNA phages.

The five campaigns considered herein are representative of the full set of industrial-scale trials since the PFA dose applied covers the complete range tested (0.8–2.2 ppm), as well as the C × t (22.1–73.8 ppm.min). In addition, the TSS concentrations measured were low, except during the September 25, 2018 campaign, which represented the nominal operations at SEV WWTP. This set-up also led to fecal bacteria log concentrations of 4.0–4.7 for *E. coli* and 3.2–4.1 for intestinal enterococci. The log removals observed for both pathogens confirm the representativeness of these campaigns, with removals of 1.4–2.3 log for *E. coli* and 1.1–2.5 log for intestinal enterococci, which is consistent with the results shown in Figure 35. The impact of TSS presence in SEV WWTP discharge is observable on fecal bacteria, with higher initial concentrations for both bacteria, thus leading to higher residual concentrations after disinfection. This effect is also noticeable on both the other pathogens, as the SSR concentration is multiplied by six while the phages are only detected during this campaign.

Regarding SSR reduction by PFA, it can be concluded that such levels of PFA dose and C × t have a limited impact on this virus given that for two of the campaigns, a very low reduction of around 0.1 log is observed while for the three other campaigns no reduction was observed. Regarding F-specific RNA phages, it is difficult to conclude based on these results since this pathogen was quantified in just one campaign, at a low concentration of 170 PFU/100 mL, and totally eliminated after disinfection. These industrial-scale results are very consistent with those obtained at the laboratory scale (Part 1, Chapter 1).

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

3.2.4 Influence of WWTP effluent quality on fecal bacteria removal

3.2.4.1 Impact of SEV WWTP treatment degradation on disinfection effectiveness

Figure 36 presents a comparison of campaigns performed at the industrial scale at a comparable PFA C × t of 30–50 ppm.min under both nominal ($n = 8$) and degraded ($n = 9$) SEV WWTP operations. The PFA dose injected during these campaigns was similar, 1.0 or 1.2 ppm on nominal operating days and 1.2 ppm on days with degraded operations.

The degradation of SEV treatment mainly leads to significantly higher concentrations of TSS (Mann–Whitney test, p-value = 0.010), which was identified as a parameter influencing PFA effectiveness, along with slightly lower conductivity (Mann–Whitney test, p-value = 0.011). For these campaigns, the quality of wastewater observed during degraded operations was notably different from the quality observed under nominal operations for specific parameters, such as: TSS (4.1 ± 1.4 mg/L for nominal vs. 8.1 ± 3.4 mg/L for degraded operations), conductivity (1138 ± 26 vs. 1046 ± 34), COD (19 ± 2 vs. 27 ± 3 mgO_2/L), BOD_5 (1.1 ± 0.6 vs. 3.2 ± 1.1 mgO_2/L), or TP (0.6 ± 0.2 vs. 1.3 ± 0.8 mgO_2/L). On the other hand, soluble parameters such as TKN (1.0 ± 0.2 vs. 1.6 ± 0.3 mgO_2/L) and DOC (6.6 ± 0.5 vs. 6.3 ± 0.4 mgC/L) remained similar. The fecal bacteria concentrations are slightly higher under degraded operations, in particular for *E. coli*, but this increase is insignificant for both *E. coli* (Student test, p-value = 0.053) and intestinal enterococci (Student test, p-value = 0.130).

Regarding *E. coli*, no significant difference exists in the logarithmic removal value (Student test, p-value = 0.972) between the nominal and degraded WWTP operations. The average removals and standard deviations are: 2.07 ± 0.48 under nominal operations, and 2.06 ± 0.33 under degraded operations. In contrast, the degradation of SEV WWTP discharge quality, leading to higher TSS concentrations and lower conductivity, results in a significant decrease (1 log) of the intestinal enterococci removal (Student test, p-value = 0.001). These average removals and standard deviations are: 2.35 ± 0.30 under nominal operations, and 1.41 ± 0.59 under degraded operations. Since the initial intestinal enterococci concentrations are similar in both the nominal and degraded operations campaigns, higher residual concentrations were identified after disinfection, thus demonstrating the negative effect of TSS and/or conductivity on PFA disinfection effectiveness for this pathogen. This conclusion is consistent with laboratory-scale observations (Part 1, Chapter 1), even though the TSS effect was observable at higher TSS concentrations; it is also consistent with recent industrial-scale trials in Biarritz, which demonstrated a correlation between PFA disinfection effectiveness and both TSS content and water conductivity (Part 4, Chapter 1) (Pigot *et al.*, 2019). The particulate matter and conductivity

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

should be monitored in-line in PFA disinfection units and moreover used as a proxy to modulate the injected PFA dose.

3.2.4.2 Normalization of PFA effectiveness to the SEV WWTP discharge quality

In Part 1, Chapter 1, it was demonstrated with laboratory-scale results that normalization of the PFA C × t applied by initial water quality parameters offered promise. In addition, industrial-scale trials displayed the negative effect of TSS on intestinal enterococci removal. Consequently, similar normalizations were carried out with industrial-scale results and have been plotted alongside laboratory-scale results in Figure 37. This figure presents the residual concentrations of both types of bacteria after disinfection vs. the normalized PFA C × t. This normalization was performed by dividing the C × t applied by both initial bacterial concentrations, TSS or COD.

Figure 37 confirms with industrial-scale results the trends observed at the laboratory scale. Both *E. coli* and intestinal enterococci residual concentrations are significantly correlated with C × t normalized by the initial bacterial concentration (Spearman test, $r = -0.557$ and p-value $= 0.0002$ for *E. coli*, and $r = -0.353$ and p-value $= 0.024$ for intestinal enterococci), TSS (Spearman test, $r = -0.588$ and p-value $= 5 \times 10^{-5}$ for *E. coli*, and $r = -0.603$ and p-value $= 3 \times 10^{-5}$ for intestinal enterococci) or COD (Pearson test, $r = -0.618$ and p-value $= 2 \times 10^{-5}$ for *E. coli*, and $r = -0.537$ and p-value $= 0.0003$ for intestinal enterococci). Power law-type relationships can be determined in order to predict residual bacterial concentrations from C × t and the initial value of either bacteria, TSS or COD. This approach demonstrates that in-line measurements of quality parameters like TSS and DCO are consistent when regulating PFA injection, even if the approach does not totally eliminate the variability in results.

It is also interesting to observe that for intestinal enterococci, industrial-scale results are highly comparable to laboratory-scale results since both point clouds follow very similar power laws, regardless of the normalization. In contrast, an offset of 0.5–1.0 log is observed with *E. coli* industrial-scale results compared to the corresponding laboratory-scale results, thus leading to different power laws. The fact that the initial concentrations of fecal bacteria and conventional quality parameters are similar in both groups of water samples is explained by the slightly lower removals obtained at the industrial scale for this pathogen, originating from a dose limitation, than at the laboratory scale.

3.3 MATHEMATICAL CORRELATION

To further investigate the correlations observed in Section 2.1 between the disinfection efficiency observed on-site at the SEV WWTP discharge and the various operating parameters, a modeling study has been conducted. In the past,

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

several works have proposed models of varying complexity in order to simulate the disinfection of fecal indicators using different disinfectants (Fernando, 2009; Flores *et al.*, 2016; Hassen *et al.*, 2000). Such models typically involve application of the Chick–Watson or Hom Equation (Azzellino *et al.*, 2011), which can be summarized by Equation (3.1) below:

$$\ln \frac{N_t}{N_0} = -kC^n t^m \tag{3.1}$$

where N_0 and N_t are respectively the initial and current fecal indicator concentrations (typically expressed in CFU or MPN/100 mL), k is the disinfection kinetic constant ($L^3M^{-1}T^{-1}$), C the disinfectant concentration (ML^{-3}), and t is the time elapsed since the injection of C (T). n and m are fitting constants that account for the varying nature of disinfectants potentially used as well as other external conditions not directly taken into consideration in Equation (3.1), with m being set at 1 for the Chick–Watson equation. The initial value of C is usually considered to be constant as a means of model simplification. Other slightly more complex equations can also be used, such as Collins–Selleck. Regardless, these kinetics equations must be coupled with a description of the reactor hydraulics to properly simulate disinfection.

While reportedly appropriate for disinfectants such as chlorine or ozone, this approach is less relevant when modeling the performance of peracids (Azzellino *et al.*, 2011). Since it is known that peracid concentration decreases upon injection and slowly during disinfection, the constant Ct approach of the basic Chick–Watson equation is often considered as insufficient to properly describe the varying rates of fecal indicator removal. An integral Ct approach has thus been proposed instead (Santoro *et al.*, 2007, 2015). The disinfection kinetics of peracetic acid (PAA) have been more widely studied in the literature than those of performic acid (PFA). Various authors have proposed using a second equation to describe the variation of C over time. This variation includes an instantaneous decrease due to an immediate initial demand followed by a first-order decay (Domínguez Henao *et al.*, 2018; Murray *et al.*, 2016). In particular, Domínguez Henao *et al.*, (2018) investigated the impact of wastewater composition on the kinetics of PAA decay and *E. coli* disinfection at the laboratory scale; they observed that both the initial demand and decay rate of PAA increased to varying extents once a threshold TSS or soluble COD concentration had been reached. Likewise, the disinfection efficiency decreased at a similar equivalent PAA dose exposure when comparing samples containing, respectively, 0 and 40 mg/L of TSS.

Such modeling efforts can be beneficial not only to better assess the impacts of various factors on disinfection efficiency, but also for on-site process control. The appropriate dose of disinfectant needs to be estimated and controlled in order to maintain a sufficient level of disinfection. These efforts also help minimize both the injection cost and remaining disinfectant concentration at the plant outfall. A suitable model can thus be used either as an optimization benchmark platform or

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

in a feedforward-type control scheme for disinfectant injection (Manoli *et al.*, 2019; Murray *et al.*, 2016).

Despite their intrinsic interest, conceptual approaches to disinfection modeling commonly found in the literature would be difficult to apply in the case study performed at the SEV WWTP. Even though PFA is not exactly the same disinfectant as PAA, it is still expected to be similarly consumed at the time of injection and during the disinfection process itself. Since PFA is more difficult to measure in samples and online than PAA, no values are available for either the laboratory experiments or the on-site case study campaigns. Likewise, the authors could not identify any modeling work regarding the modeling of PFA disinfection. No default or reference parameter values are therefore available to use with the Chick–Watson or peracid decay equations typically employed. In this work, a multiple linear regression approach has instead been used to simulate the effluent *E. coli* concentrations at the SEV WWTP; this approach takes into consideration the impact of both operation variables (wastewater flow rate, PFA dose) as well as wastewater quality at the injection point, despite the lack of PFA concentration measurements. Different criteria were used to select the optimal model input variables from a partial measurement dataset representing what could easily be measured online. The selected models were then compared using basic validation checks, plus leave-one-out and k-fold cross validation.

3.3.1 Model construction and validation

The dataset used consists of the 45 influent and effluent grab samples collected for analysis during the weeks of PFA injection, along with the six samples collected during non-injection week 42 (previously described in Part 2, Chapter 1). This dataset includes operation data at the time the grab sample was collected (wastewater flow rate, water height in the discharge channel, temperature, PFA dose, equivalent initial Ct) as well as measured concentrations. Some data are also available from more than one source. Conductivity and pH were measured on-site at the sampling point and once again after the arrival of samples at the laboratory. Temperature was also measured during sampling and by a sensor used by plant operators permanently installed in the discharge channel. Samples for which the effluent fecal indicator bacteria (FIB) concentrations were at the method quantification limit were removed. The effluent DOC concentrations were not considered since these data were unavailable for a small fraction of the dataset. Likewise, given that the purpose of this study was to estimate the effluent concentrations of each type of FIB, the measured intestinal enterococci concentrations were discarded as input variables for the *E. coli* model, and vice versa. In addition, since the correlation between influent and effluent concentrations for each measured variable was rather high (Pearson coefficient values ranging from 0.777 to 0.995), only the influent concentrations were retained as potential explanatory variables. This decision also served to decrease

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

the number of potential explanatory variables, which was initially large compared to the number of experiments available for model fitting. The final datasets for each model thus contained 22 explanatory variables, for 48 valid measurements of EC (Table 15) and only 34 of IE. For this reason, only those results concerning the EC models will be presented herein.

A second dataset, containing the same number of valid measurements, was also prepared. For this case, only the explanatory variables considered to be easily measurable at the discharge channel intake with either sensors or *in situ* analyzers were kept. This dataset included the operational data, as well as pH, temperature, conductivity, EC concentration, DOC, turbidity and ammonia nitrogen; TSS were not included since TSS sensors actually often measure turbidity, which has already been included. Likewise, COD and BOD were excluded since they often constitute a composite estimation based on turbidity (particulate fraction) and UV-254 (soluble fraction). DOC was chosen instead as it tends to be estimated from sensors solely on the basis of UV-254. The EC concentration at the channel intake was also retained given that it could potentially be measured online, although probably with some delay between sampling and result. The secondary dataset thus contained 48 measurements and 14 explanatory variables, including similar measurements from the various sources detailed above. The objective of the first dataset was to obtain an 'ideal' model using every piece of data available, whereas the second one was considered in order to evaluate the possibility of using such a model for process control or a simplified on-site performance estimation.

For each dataset, a similar model building and evaluation protocol was followed. The effluent EC concentrations were natural log-transformed before use, as often takes place in FIB modeling (Eleria & Vogel, 2005). An initial Box–Cox power transform ($\lambda_{opt} = -0.16$) confirmed the choice of a logarithmic form (Herrig *et al.*, 2015). The regressions were performed according to a stepwise procedure, using the *stepwiselm* function available in Matlab® R2018b (MathWorks). Since this procedure does not necessarily guarantee finding an optimal model on the first try, the explanatory variables are identified using the Sum of Squared Error (SSE) criteria, the Aikake Information Criteria (AIC) and the Bayesian Information Criteria (BIC), by employing both a forward and backward procedure in each case. The identified models were further simplified by subsequently removing those explanatory variables for which the 95% confidence interval crossed zero and then refitting the model with the remaining variables. For each dataset, the stepwise procedures were first performed by only considering the direct effects of explanatory variables (no interactions) and including two-way interactions for all available variables during a second run (Table 15). All models identified by this stepwise procedure were verified for normal error distribution as well as for obvious trends between errors and each retained explanatory variable. The Variance Inflation Factor (VIF), normal and

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

adjusted coefficients of determination (R^2 and adj-R^2) and the mean square error (MSE) were computed in all cases (Montgomery *et al.*, 2012).

Lastly, a cross-validation was performed for each model identified as a means of further optimal model selection, using the *crossval* function in Matlab. A leave-one-out (LOO) and a series of k-fold (for k = 2–10) procedures were applied. The single value of the LOO validation MSE (MSE_{val}) was considered, while the median value from a 1000 random sampling was considered for k-fold. The MSE_{val} values were compared both between models and to their construction MSE results in order to identify those models for which a large loss of accuracy was to be expected outside of the specific training dataset used.

3.3.2 Main results

The main results obtained for models found on the full dataset are summarized in Table 16. Four distinct combinations of optimal explanatory variables were found, as labeled F1 to F4. As is regularly the case, the forward stepwise selections (F1 and F2) identified models with fewer variables than the backward selections (F3 and F4). In both cases, the two identified models had a very similar structure; models F1 and F2 include a very small number of variables, based on the PFA dose, Ct, temperature and, potentially, the influent ammonia concentration.

The resulting maximum VIF value is high in both cases (23–23.1) and corresponds to the significant correlation between PFA dose and Ct value. For purposes of comparison, optimum VIF values under 5 or 10 are often reported in the literature (Herrig *et al.*, 2015; Montgomery *et al.*, 2012). The R^2 and adjusted R^2 values are relatively high in both cases, although the removal of [NH_4^+] in F2 results in a decrease in these scores compared to F1. Models F3 and F4 contain more variables (eight) and are based instead on a combination of water flow rate and height, temperature, PFA dose, influent carbon concentrations and either ammonia (F3) or pH (F4). Both models exhibit very similar values of each selection criterion applied. The maximum VIF value equals 11.6, which stems from the rather strong correlation between water flow rate and channel height (Pearson correlation of 0.95). An increase in the adjusted R^2 values is observed when compared to F1 and F2, thus suggesting that the additional explanatory variables do raise model accuracy. Similarly, models F3 and F4 produce a much lower cross-validation MSE value than the 'simpler' models F1 and F2. The increase between the MSE of model construction and validation is also much lower (1.4–1.6 vs. 2.0–2.1), suggesting that the simpler models are less accurate when applied to 'new' conditions. For all these reasons, models F3 and F4 seem more appropriate for general use. F3 is, at first glance, preferable due to the slight increase in cross-validation MSE between F3 and F4; however, it could be argued that ammonia is not as easily measurable as pH.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

Table 17 shows the main results obtained for models found on the 'online dataset' without the inclusion of any variable interactions.

In this case, only two distinct models were identified. Every forward stepwise criterion identified model S1, which is exactly the same as model F1 based on the full dataset. Since the 'online' dataset contains a smaller number of potential explanatory variables while conserving every variable identified in model F1, this result is not unexpected. Moreover, every backward stepwise procedure converged on the same model (S2, Table 17). Model S2 is also quite similar to model F3 built on the full dataset, only excluding the BOD and COD concentrations, which results in a small decrease in R^2 (0.036) and adjusted R^2 (0.024) as well as an increase in the construction and cross-validation MSE results when compared to F3. The maximum VIF value is also slightly lower (10.5 vs. 11.6), which suggests that most of the key information needed to predict the effluent EC concentration could, in theory, be made available by online measurements. One important caveat of the work on this subject, however, is that the same measurements are used as potential model inputs for both the full and online datasets. The concentration values were obtained by laboratory analyses (except pH and conductivity) without any duplicate measurements using sensors. Both laboratory and *in situ* measurement performance may differ depending on many factors. Actual model performance using online sensors could thus differ from that obtained in this case. A proper data acquisition campaign would be needed in order to fully validate these findings. Once again, however, the slightly more complex model S2 can be considered as a preferable compromise between complexity and performance.

Table 18 summarizes the positive results obtained when considering two-way interactions during the model identification procedures.

Most attempts led to a very high number of variables and interactions being included, compared to the amount of data available, resulting in overfitted models with poor cross-validation MSE values. Only the forward AIC and backward SSE criteria on the 'online' dataset found models containing a reasonable number of interactions that apparently increased model accuracy when compared to previous results. Both models in Table 3.18 are similar to model S2. SI1 replaces the PFA dose by interactions of PFA dose with temperature and DOC concentration. SI2, on the other hand, replaces the water height by an interaction between height and temperature. In both cases, the temperature recorded during sampling replaces that taken by the WWTP sensor. SI1 yields an improvement in overall model performance (both construction and validation) while SI2 maintains similar construction results, albeit with an increase in cross-validation MSE values.

Figure 38 shows the model results over the entire dataset for the optimal models F3 and S2. As can be observed, restricting the available variables to 'online' measurements leads to greater model dispersion for the EC data around 5–7 log;

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

it also leads to underestimating the high effluent readings recorded during the week of measurement without any PFA injection. Also noticed in Figure 38 are the few concentration measurements available in the log 7–9 region (i.e., a low PFA dose and/or high water flow rate). Additional measurements in this region would be beneficial to ensure that the optimal models are suited for the whole range of possible conditions.

Although care must be taken when interpreting the physical meaning of variables identified by stepwise procedures in linear regression, a comparison of these results with factors incorporated in knowledge-based models can be useful. As highlighted in previous sections, the typical factors included in EC disinfection knowledge-based models are reaction time, initial EC concentration, and integrated Ct value, which itself is dependent on the initial acid dose and any factors affecting its speed of consumption (e.g., TSS, soluble COD, temperature). Domínguez Henao *et al.*, (2018) also considered TSS as a direct influential factor, by acting like a protection for fecal bacteria and not only through an increase in the disinfectant consumption speeds. The optimal models obtained in this study all have a similar structure; they are based on the water flow rate and discharge channel height (correlated with the hydraulic residence and hence the reaction time), the PFA dose introduced, the water temperature, and different concentrations of remaining pollution in the WWTP effluent, for the most part carbon-based. This finding is partly similar to the observations highlighted in the previous sections (Part 1, Chapter 1), directly for DOC and indirectly for TSS and the initial Ct.

The influent EC concentration, however, is missing from every identified model; this situation could be due to the relatively strong correlation in available data between the influent EC concentration and several of the explanatory variables found. Such is the case for Q (Pearson $\rho = 0.61$), H (0.58), COD (0.74) and BOD (0.84). At the effluent of a WWTP, the major increases in EC concentration are in fact typically caused by a partial bypass of the normally complete treatment lane. These bypasses are often due to either treatment process unavailability (maintenance, repair) or high flow events, such as rain periods. Partial bypasses also often result in an increase in pollution concentrations in the discharge channel. A combination of these factors could be sufficient, in this specific case, to estimate the order of magnitude of the influent EC concentration. This approach, however, also leads to multicollinearity amongst the various concentration explanatory variables. As such, the preference of COD, BOD and DOC over other variables, like TSS or turbidity, by the model selection procedure employed should not be over-interpreted. For similar reasons, the maximum VIF value for all optimal models lies above the recommended maximum of 10. As mentioned above, in most cases this situation is due to the strong correlation between water flow rate and channel height. Although additional data acquisition could potentially help reduce this VIF value, it is unlikely to be of much improvement.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

3.4 INTERACTIONS BETWEEN PFA AND THE PHYSICOCHEMICAL QUALITY OF SEV WWTP DISCHARGE

3.4.1 Impact of PFA on conventional quality parameters

Together with fecal bacteria measurements, conventional quality parameters were monitored in both SEV WWTP discharge and disinfected water to determine whether PFA has a significant impact on water quality. The average concentrations and standard deviations are given in Table 13. A stability in conductivity, TSS, TKN, $N\text{-}NH_4^+$, $N\text{-}NO_3^-$, $N\text{-}NO_2^-$, TP and $P\text{-}PO_4^{3-}$, along with a sizable increase in DOC, COD and pH, can be observed. More specifically, the lack of a PFA disinfection effect on TSS (Mann–Whitney test, p-value = 0.192) and conductivity (Mann–Whitney test, p-value = 0.910) is statistically significant. The increases in DOC (Student test, p-value = 8×10^{-19}), COD (Mann–Whitney test, p-value = 0.001) and pH (Mann–Whitney test, p-value = 0.0001) are also significant. The evolutions of DOC, COD and pH are depicted in Figure 39 for the entire industrial-scale trial period. An average increase of 1.2 mgC/L, 2 mgO_2/L and 0.19 pH unit is observed, with an average applied PFA dose of 1.2 ± 0.5 ppm. A slight change in effluent pH due to PFA was also detected by Ragazzo *et al.*, (2013) at the industrial scale even though it was attributed to acidification. Similarly, some cases of slight acidification have been reported in the literature with peracetic acid, with in most instances an insignificant evolution (Luukkonen & Pehkonen, 2017). The increase in DOC or COD was reported in the literature (Luukkonen & Pehkonen, 2017), along with a slight increase in dissolved oxygen due to peracid decomposition.

This discussion confirms, at the industrial scale, that PFA disinfection has a limited impact on conventional quality parameters in SEV WWTP discharge, with the only significant impact being on carbon due to the injection of a non-negligible quantity of carbon via PFA (Figure 39). This finding is in complete accordance with laboratory results (Part 1, Chapter 3), which measured a DOC input of 0.78 ppm C/ppm PFA injected, vs. a theoretical input of 0.79 ppm C/ppm PFA injected. The average PFA dose injected was in fact 1.2 ± 0.5 ppm, thus corresponding to a theoretical input of 1.0 ± 0.4 mgC/L. Ragazzo *et al.*, (2013) observed an increase of 0.68 mgC/L per ppm of PFA applied.

3.4.2 Production of disinfection byproducts caused by PFA

The potential for disinfection byproduct formation by PFA has been studied in detail in Part 1, Chapter 4, for both the laboratory-scale and industrial-scale trials. Table 19 displays the concentrations of adsorbable organic halogens (AOX), bromide and bromate measured in SAV WWTP discharge before and after PFA injection during the five sampling campaigns performed within the industrial-scale trials. Bromide is a precursor of both brominated oxidation byproducts and bromate.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

An increase of 10–41% of AOX in SEV WWTP discharge with PFA injection is systematically observed at the industrial scale. Slight increases of AOX with PFA disinfection have also been reported in the literature (Karpova *et al.*, 2013; Luukkonen & Pehkonen, 2017), but these were notably lower when compared with other peracids or oxidants. The PFA dose and C × t applied seem to produce an effect on AOX formation since the largest AOX increase (+41%) is observed with the highest PFA dose (2.2 ppm) and applied C × t value (73.8 ppm.min). In contrast, the bromide concentration is reduced by PFA disinfection from 0 to −58%, even though the initial bromide concentrations are low (100–250 µg/L). No bromate was detected either before or after disinfection despite bromide levels being above 100 µg/L, which was defined as the concentration range where bromate can be formed by oxidation (von Gunten, 2003). This finding confirms that no bromate formation is expected with PFA at such low doses.

Key points

- Industrial-scale trials performed at the SEV WWTP confirmed the performance obtained at the laboratory scale with logarithmic removals of around 2–2.5 log for both *E. coli* and intestinal enterococci at PFA doses of 0.8–2.2 ppm and a C × t of 15–75 ppm.min.
- High-frequency measurements of *E. coli* with the ALERT apparatus highlighted the daily variability in fecal bacteria concentrations and its impact on logarithmic removal with PFA, thus leading to the conclusion that the PFA dose should be modulated between daytime and nighttime to optimize disinfection.
- PFA has a limited disinfection effectiveness on SSR and on totally inhibited F-specific RNA phages, for the only campaign during which this pathogen was detected in SEV WWTP discharge.
- The influence of SEV WWTP discharge quality variations on the PFA disinfection effectiveness observed at the laboratory scale was confirmed at the industrial scale, with better fecal bacteria removals during nominal SEV WWTP operations, corresponding to lower TSS concentrations and higher conductivity. Residual concentrations of fecal bacteria are correlated with the applied C × t per unit of bacteria in the water. Both TSS and COD serve as good proxies when employed to modulate PFA injection.
- PFA disinfection at the industrial scale led to a slight increase in DOC, COD and pH; however, no bromate formation was observed despite a decrease in bromide. A slight increase in AOX indicated the formation of disinfection byproducts, but specific results on this point (see Part 1, Chapter 4) did not display any formation of problematic byproducts.
- Two series of optimal linear regression models were developed to estimate EC concentration at the discharge channel outfall during the disinfection

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest

trials at SEV WWTP. These optimal models contained relatively few explanatory variables and featured both high R^2 and adjusted R^2 values, thus indicating the potential interest of such a model for the online on-site estimation of disinfection efficiency. Also, the use of actual sensors to measure the concentration values identified as explanatory variables would be necessary to draw a complete conclusion on the online model's potential.

Downloaded from http://iwaponline.com/ebooks/book/chapter-pdf/907042/9781789062106_0075.pdf
by guest