Science of the Total Environment 799 (2021) 149398

Contents lists available at ScienceDirect

Science of the Total Environment

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

Assessing the effect of catchment characteristics to enhanced coagulation in drinking water treatment: RSM models and sensitivity analysis

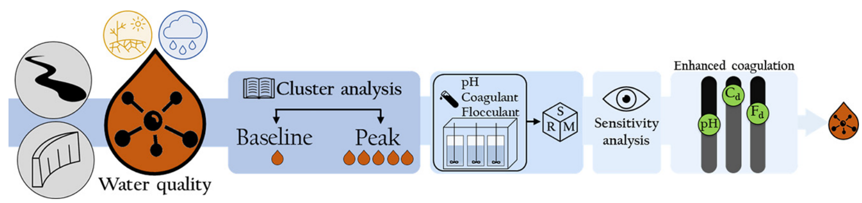
J. Suquet, Ll. Godo-Pla, M. Valentí, L. Ferràndez, M. Verdaguer, M. Poch, M.J. Martín, H. Monclús *

LEQUIA, Institute of the Environment, Universitat de Girona. C/Maria Aurèlia Capmany, 69, E-17003 Girona, Catalonia, Spain

HIGHLIGHTS

- RSM models optimise drinking water production by enhanced coagulation.
- Two surface water catchments were compared: river and reservoir.
- Cluster analysis determined baseline and peak organic matter loads at DWTPs.
- Sensitivity analysis allowed the interpretations of RSM models outputs.
- Peak scenarios were described as episodes particularly important for optimisation.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 4 May 2021
Received in revised form 9 July 2021
Accepted 28 July 2021
Available online 2 August 2021

Editor: Damià Barceló

Keywords:
Surface water
River
Reservoir
Enhanced coagulation
Natural organic matter

ABSTRACT

Coagulation is the main process for removing natural organic matter (NOM), considered to be the major disinfection by-products (DBPs) precursor in drinking water production. In this work, k-means clusters analysis were used to classify influent waters from two different surface drinking water treatment plants (DWTPs) located in the Mediterranean region. From this, enhanced coagulation models based on response surface methodology (RSM) were then developed to optimise coagulation at two water catchments (river and reservoir). The cluster analysis classified the water quality of the raw waters into two groups related to baseline and peak organic loads. The developed enhanced coagulation models were based on the turbidity, total organic carbon (TOC) and UV_{254} removals. Sensitivity analysis applied to the models (after predictors selection) determined the factors relative individual contributions for each DWTP scenario. Then, profile plots for enhanced coagulation were studied to identify the optimal levels for each case. Models mean R^2 were 0.85 and 0.86 in baseline and 0.85 and 0.84 in peak scenario for river and reservoir catchments, respectively. Results of this study indicate that the surface water quality variation in river DWTP is seasonal and is expressed by an increase of turbidity, while in the reservoir DWTP is related to extreme weather events showing high levels of dissolved organic load (TOC and UV_{254}). During baseline cases, where raw waters present low levels of organics, the three factors optimal adjustment should be ensured to optimise coagulation. Then, during peak scenarios, where influent waters present high organics, the optimal for enhanced coagulation relies on the correct adjustment of C_d. The presented work provides models for drinking water production aimed to propose the optimum conditions for enhanced coagulation, considering the influent water characteristics under different weather conditions.

1. Introduction

Drinking water treatment plants (DWTPs) deal with fluctuations in water quality and quantity. Surface water is plenty of organic

* Corresponding author.
E-mail address: hector.monclus@udg.edu (H. Monclús).

https://doi.org/10.1016/j.scitotenv.2021.149398

compounds, and its receipt is subjected to multiple factors including geography, geology, hydraulic regimes, land use, weather and the types of water catchments (Awad et al., 2018). Present and future scenarios related to climate change, where episodes of extreme events are becoming frequent and the challenges populations face, such as pandemics or the appearance of new contaminants, could induce some variations in water characteristics (Delpla et al., 2009; Poch et al., 2020; Sun et al., 2020). That said, natural organic matter (NOM) is ubiquitous and spread in water all over the Earth.

During water treatment processes, NOM is reduced because its direct impact on the formation of disinfection by-products (DBPs) (Chaukura et al., 2020; Godo-Pla et al., 2020b). DBPs are reported as being harmful compounds for humans and, as such, are regulated in both European (Directive (EU), 2020) and national directives (RD, 140/2003, 2003). NOM is a heterogeneous matrix composed of particulate, colloidal and dissolved organics. Inside the water-treatment train, coagulation is a key process and is usually located in the first stages of any water treatment. The typical parameter used to optimise coagulation in drinking water production is turbidity. However, enhanced coagulation aims to optimise coagulation for NOM removal, with the objective to reduce the residual organic water compounds (Sillanpää et al., 2018). These compounds react with the chemical disinfectants, added for water distribution, generating DBPs (Godo-Pla et al., 2019; Krzeminski et al., 2019). To avoid high concentrations of DBPs at the end of large distribution networks (i.e., just before consumption), the minimization of NOM compounds should be the main strategy in the initial stages of water treatment, especially during the coagulation process (Liu et al., 2012; Williams et al., 2019). In this sense, several water quality parameters are used to control and monitor the organics in full-scale facilities, some of which are: turbidity, Total Organic Carbon (TOC) or ultraviolet absorbance at 254 nm (UV_{254}) (Andersson et al., 2020; WHO, 2017). These quality parameters can be used to assess coagulation performance (Edzwald, 1993; Volk et al., 2000). TOC and UV_{254} are related to DBPs precursors and large organic pollutants (e.g., humic acids), respectively (Ates et al., 2007; Beauchamp et al., 2020).

Coagulation can be optimised through several pathways although there are various factors affecting coagulation performance. Because there are organic, inorganic, composite, hybrid coagulants and biocoagulants (Adesina et al., 2019; Harfouchi et al., 2016; Xia et al., 2018), the most efficient coagulant performance depends on the characteristics of the raw water. Operationally, pH and chemical dosages are the main factors influencing process performance and its optimisation (Arruda et al., 2018; Trinh and Kang, 2010; Xie et al., 2012; Yan et al., 2008). The optimal chemical dosages are usually determined from laboratory jar test experiments. Even though, coagulation operation is usually suboptimal due to other limitations related to full-scale operation and water quality fluctuations. To deal with that, several modelling approaches to optimise coagulation in water production, particularly response surface methodology (RSM), have been described in the literature.

RSM is a technique which allows planning for a set of minimum experiments to evaluate the effects and the interactions of coagulation factors and the respective responses (Zainal-Abideen et al., 2012). As a response for RSM, and in accordance with the abovementioned statements, it is important to select parameters which are monitored in a real facility. From this basis, RSM allows mathematical models that have the capacity to predict values focused on influent water characterization aimed to optimise coagulation process (Apostol et al., 2011; Suquet et al., 2020).

As DWTPs deal with influent water quality fluctuations, it is essential to adjust coagulation to cope with the different scenarios. This is especially relevant in Mediterranean regions where water provisioning is decreasing and seasonal changes and extreme events, such as heavy rains and droughts, are increasing in frequency (Jorda-Capdevila et al., 2019). Consequently, coagulation modelling requires not only knowledge of the process but also a broad range of available data. Numerous studies, fundamentally based on historical operational data, have attempted to develop models to optimise the coagulation step (Heddam et al., 2012; Kim and Parnichkun, 2017; Omran, 2018). Data-driven models such as artificial neural networks (ANNs) present high levels of predictability and accuracy with huge amounts of data but have some limitations when predicting responses outside the range of the training region (Baxter et al., 2001). On the other hand, models such as RSM are designed from experimental work, providing a detailed level of process understanding (Sadri Moghaddam et al., 2010). Hence, the development of a model integrating full-scale datasets and RSM models emerge as a useful tool for decision-making.

Within this framework, the motivation of this study arose from the assumption that the optimisation of coagulation based on the influent water characterization contributes to the minimization of NOM at the effluent of DWTPs. The main objective of this study is to understand how coagulation process is affected by different surface water catchments and their intrinsic fluctuations. In order to achieve that, this work proposes enhanced coagulation models based on RSM and influent DWTP characterization for the optimisation of coagulation process at two Mediterranean DWTPs. This study has been conducted with the following specific objectives: i) identify influent water quality classifications using clustering techniques, ii) develop RSM models for enhanced coagulation and iii) analyse the effect of operational parameters in enhanced coagulation through sensitivity analysis for the different water catchments.

2. Materials and methods

2.1. Case study

The study was carried out at two case-study DWTPs, namely the Llobregat and the Ter DWTPs (Fig. 1), located in Catalonia (NE Spain). These two DWTPs are managed by the Ens d'Abastament d'Aigua Ter-Llobregat (ATL) which supplies water to the Barcelona Metropolitan Area (≈4.5 m inhabitants). Treatment capacity are 3.2 and 8.0 $m^3 \cdot s^{-1}$ for the Llobregat and Ter DWTPs, respectively. The Llobregat DWTP catchment water is from the Llobregat River, which is the second longest river in Catalonia and historically characterized as being subjected to high levels of anthropogenic pressure, related to textile and pharmaceutical industry derived pollution (Kuster et al., 2008) and the exploitation of salt mines in the upper part of the basin (Postigo et al., 2018). On the other hand, the Ter DWTP takes water from the Sau-Susqueda-Pasteral reservoir system (Ter River basin). These reservoirs change river regimes and, consequently, water characterization (Espadaler et al., 1997). In both facilities, the first step of the conventional treatment chain consists of a primary oxidation (PO) step, followed by the

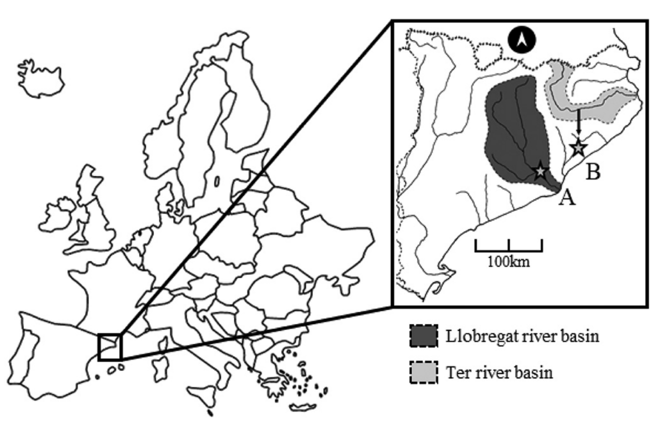


Fig. 1. Case-study DWTPs location (NE Spain). A and B correspond to the Llobregat and Ter DWTPs, respectively. At the Ter DWTP, raw water is pipe-conducted (55 km) from reservoirs to DWTP.

coagulation process. PO represents the first chemical barrier, located at the beginning of the treatment, attempting to oxidise a wide range of compounds present in raw waters (Godo-Pla et al., 2020a). For this purpose, potassium permanganate is added at the Llobregat DWTP, while chlorine-based oxidants are applied at the Ter DWTP before coagulation.

2.2. Cluster analysis

Influent water characteristics were evaluated using cluster analysis. In previous scientific literature, clustering has been stated as a suitable technique for water classification when detecting temporal changes in water characterizations (Celestino et al., 2018; Fathi et al., 2018; Hou et al., 2018). For this purpose, k-means clustering was applied in this work, with the aim to establish influent water quality classifications. K-means clustering is a partition method based on centroids aimed to classify large datasets into a pre-specified number of clusters. The first iteration of the cluster algorithm states randomly k-clusters along dataset and calculates the centroid of each cluster. Then, in the second iteration, each data point is assigned to the closest centroid. Centroids and their associated data constitute a cluster. The Euclidean distance was used to determine the similarity between clusters.

Historical datasets were obtained from case-study DWTPs daily laboratory analytics corresponding to the period 2017-2020. Using z-score mapping before clustering, input variables were scaled to have zero mean and a standard deviation (std) of 1. To classify data, the following influent water quality parameters were considered as features of the clustering algorithm: TOC, turbidity, UV_{254}, colour and the specific ultraviolet absorbance (SUVA). Based on the study of these parameters, a clustering algorithm with k-means = 2 was performed to classify the quality of the raw waters into two categories with the aim to identify cases with different raw waters composition through the seasons/year. Python programming language (Python Software Foundation, Wilmington, DE, USA) using Scikit-learn library (Pedregosa et al., 2011) was used to design and execute clustering algorithms. SUVA value was calculated resulting from the division of UV_{254} by TOC (USEPA, 2009).

2.3. Enhanced coagulation models

First, RSM was designed. Then, experimental laboratory jar tests were conducted to develop enhanced coagulation models. Subsequently, model validation was performed with sensitivity analysis.

2.3.1. Design

Central composite design (CCD) was selected as the RSM design for its efficiency and flexibility, and the capacity to provide detailed information in a minimum number of required runs for the region of interest/operability. In this case, CCD was performed for three study variables (factors) that influence the coagulation process: pH, coagulant and flocculant dose. These variables have been described in the literature as the most influential factors in the coagulation step (Trinh and Kang, 2010; Trinh and Kang, 2011). To assess coagulation performance, a set of water quality parameters were identified to determine process optimisation when accounting for all spectrum of water compounds (turbidity) and the specific NOM fractions (TOC and UV_{254}). Parameters were selected based on its nature and also the capacity to be monitored online at the full-scale facilities. As models developed for this work are aimed to aid decision-making, turbidity, TOC and UV_{254} were the chosen RSM responses. The total number of runs for CCD were 20, combining the conditions of various factors (see appendix A). The Design-Expert® (Stat-Ease, Inc., Minneapolis, MN, USA) software version 11.0 was used to perform RSMs.

The range of the factors for the RSMs in the two case-study DWTPs was 5.5 to 8.5 for pH level, 10 up to 70 $mg\cdot L^{-1}$ and 5.25 up to 70 $mg\cdot L^{-1}$ for the coagulant dose at Llobregat and Ter DWTP, respectively. Then, flocculant dose varied from 0.2 up to 1.5 $mg\cdot L^{-1}$ at Llobregat DWTP and 0.15 up to 1.74 at Ter DWTP. All RSM designs were planned by expanding operational full-scale ranges to cover the entire range of response (regions of interest). Two RSM were performed at each case study DWTP.

2.3.2. Experimental jar tests

Water samples were collected at each case-study DWTP to execute jar tests to develop the RSM models. At the Llobregat and Ter DWTPs, four RSM were conducted. Table 1 summarises the water characterization for the different sampling campaigns (SC).

For the laboratory analyses, turbidity, TOC and UV_{254} were measured with a Hach TU5200 turbidimeter, a Sievers M9 portable analyser and a Cary 3500 UV–Vis Agilent Tech spectrophotometer, respectively. For TOC and UV_{254} measurements, samples were filtered at 0.4 μm to ensure the analysis of dissolved NOM. For UV_{254}, a quartz cell with a 1 cm path length was used. Next, pH was determined using a Crison micro pH 2000 apparatus. ISO 5667-3:2018 requirements were ensured for the collection, storage, transport and pre-treatment of all the samples.

The jar test is a widespread methodology and for this study several jar test experiments were conducted based on the standardized DWTPs protocols to obtain feasible and comparable results according to full-scale application. A summary of the jar tests phases (times and speeds) is presented in appendix B. Jar tests were carried out using a Phipps & Bird (7790-910, Richmond, VA, USA) six paddle programmable jar tester and the chemical reagents employed were obtained from the DWTPs supporting this study. The case-study DWTPs use alum-based coagulant (Polyaluminium Chloride) to perform coagulation unit operation. There is a difference concerning flocculant type in that the Llobregat DWTP adds a cationic quaternary ammonium-based polymer (PolyDADMAC), while the Ter DWTP doses with a starch-based flocculant.

2.3.3. Predictors selection

From RSM experiments, water characterization (turbidity, TOC and UV_{254}) was obtained for the fixed coagulation conditions (see appendix A). Based on the selected factors (pH, coagulant and flocculant dose), the full quadratic equation is presented in Eq. (1).

$$Y = \beta_0 + \beta_1 pH + \beta_2 C_d + \beta_3 F_d + \beta_4 pHC_d + \beta_5 pHF_d + \beta_6 C_d F_d + \beta_7 pH^2 + \beta_8 {C_d}^2 + \beta_9 {F_d}^2 \quad (1)$$

where Y is the percentage of removal for responses, β_x are numerical model coefficients and pH, C_d and F_d the model factors. C_d and F_d are the coagulant and the flocculant dose, respectively. Related to the equation elements β_1, β_2 and β_3 are the one factor interactions; β_4, β_5 and β_6 are the two factor interactions and β_7, β_8 and β_9 are the quadratic effects.

Among all models' factors, it is necessary to state a procedure with the capacity to systematize the selection of the best features for each model. Hence, the *best subset selection method* was applied to the models obtained from jar test experimental data, allowing to determine the optimal number of predictors to ensure the best features under a

Table 1
DWTPs influent water characterization for the sampling campaigns (SC).

SC	DWTP	Date	Turbidity (NTU)	TOC ($mgC\cdot L^{-1}$)	UV_{254} (m^{-1})
SC 1[a]	Llobregat	03.2019	9.1	4.1	12.7
SC 2	Llobregat	03.2020	3.7	2.6	11
SC 3	Llobregat	08.2020	26.2	1.8	11.5
SC 4[a]	Llobregat	11.2020	63.3	3	12.9
SC 5[a]	Ter	02.2020	5.3	3.6	15.1
SC 6	Ter	03.2020	5.2	3.7	14.4
SC 7	Ter	07.2020	0.9	3.2	12.7
SC 8[a]	Ter	11.2020	4.8	2.8	11.9

[a] Sampling campaigns performed for RSM.

sufficient level of predictability. This method consists of fitting models considering each possible combination of the predictors candidates (p). In this case, the total number of p is listed in Eq. (1) ($p = 9$), being 2^p (2^9) the maximum number of combinations (Godo-Pla et al., 2020a). For this study, the identification of the best model was based on the determination of the following statistics: sum of squares error (SSE), R^2 and adjusted R^2 ($R^2_{adj.}$) values. From this, models can be selected by minimizing the prediction error or maximizing the $R^2 - R^2_{adj.}$ R^2 enables to identify the predictive accuracy while the $R^2_{adj.}$ value provides the coefficient of determination for each model pondered according to the number of predictors. Models coefficients significance was ensured (p-value <0.05) after the application of the best subset selection method. Plots of R^2 considering 2^p are presented in appendix C. This method is useful for a limited number of predictors due to computational limitations. The software used in this study was MATLAB 2019a (Mathworks®, Natick, MA, USA).

2.3.4. Sensitivity analysis

To study the enhanced coagulation models designed by RSM and performed by the best subset selection method, a sensitivity analysis was conducted to explore and to determine the impact of factors (pH, coagulant and flocculant dose) on the quality parameters (turbidity, TOC and UV_{254} removal efficiency). Equations were analysed through delta mean-squared sensitivity analysis to determine the contribution of the relative factors. The delta mean-squared (δ_i^{msqr}) non-dimensional sensitivity function was chosen to determine the significance of model factors as well as their interactions. Sensitivity analysis was used to verify the robustness of the models and their reliability for the different scenarios at each DWTP. Further details on the methodology can be found elsewhere (Godo-Pla et al., 2021; Sin and Gernaey, 2016). MATLAB 2019a (Mathworks®, Natick, MA, USA) was used to perform the sensitivity analyses.

3. Results and discussion

3.1. Cluster analysis

K-means clustering was developed to classify in two clusters influent water quality datasets (2017-2020 period) for the Llobregat and Ter DWTPs. Accounting for water characterization, the following parameters were selected: turbidity, TOC, UV_{254}, colour and SUVA value. Water colour is directly related to NOM content originated from wood and soil (Christman and Ghassemi, 1966; Dragon et al., 2018), turbidity accounts for particulate, colloidal and soluble water components (Gregor et al., 1997) while TOC and UV_{254} and SUVA values contribute to specific NOM fractions. Based on the study of these parameters, a clustering algorithm with k-means = 2 was performed to classify the quality of the DWTP raw waters into two categories with the aim to classify influent water quality. The initial datasets analysis coupled to expert knowledge provided by plant managers indicated that raw water quality is changing over the year and is subjected to scenarios where the quality is high and other which is low. Hence, this paper is focussed on enhanced coagulation models development, not to define the optimum number of clusters for the historical datasets. Studies focussed to stablish the optimum cluster number should evaluate metrics related to the separability and compactness of clusters (Rendón et al., 2011). Results obtained at the Llobregat and Ter DWTPs from the cluster analysis are presented in Fig. 2. There are five plots for each facility (left column Llobregat DWTP and right Ter DWTP) which classifies water quality for the chosen influent parameters in two groups. These groups were related to as baseline and peak water quality. In general, the Llobregat DWTP parameters fluctuations are higher than those of the Ter DWTP (see Fig. 2).

Observing the output from the cluster analysis performed at both DWTPs, basic differences between the plots could be attributed to the catchment characteristics. To explain this, it is necessary to examine the influent water quality fluctuations in depth (see Fig. 2). First, it is important to remark on the basic difference between the types of catchment the two case studies have. Reservoir system work as a massive water clarifier which, in turn, helps to maintain low fluctuations in the quality of the influent water. However, extreme events (heavy rains) could destabilize that system, leading to high changes in reservoir water quality. On the other hand, river catchment quality is more unstable, highly dependent on water flow (pollutants concentration) linked to weather (runoff effect) and other external factors related to human activities (Fernández-Turiel et al., 2003a; Gallart et al., 2011; Navarro et al., 2002).

Therefore, the outputs of the cluster analysis in both the Llobregat and Ter DWTPs classified influent water quality into two categories, which were attributed to baseline and peak organic content. The main difference was in the frequency of peak events: the river catchment water quality presented seasonal fluctuations, whereas water from the reservoir showed a lower frequency of peak events that were not strictly related to seasonal changes. At the Ter DWTP, data which comprises a major part of dataset corresponds to the baseline group and the punctual anomalies to peak scenario. These peak events can be visually identified in October 2018 and January 2020 (see Fig. 2, Ter DWTP). Both series of data are directly related to historical heavy storms events, rains with more than 180 L·m-2 (13th-15th October 2018) and more than 400 L·m-2; the latter was Storm Gloria, which provoked a great deal of damage and multiple issues in this part of Europe (19-23 January 2020) (Amores et al., 2020). The Mediterranean region is periodically affected by flash floods, characterized by local heavy rains in small surface regions (Cramer et al., 2018). These events consisting in heavy rains caused alterations in reservoirs stratifications and the quality of the influent of the Ter DWTP was affected by an increase of the water quality parameters. According to Casamitjana et al. (2003), this change in water composition is due to the resuspension of the organic content present in deep sediments towards epilimnetic waters (superficial waters). In summary, the time series show that the Ter River system of reservoirs act as a massive clarifier, thus maintaining the water quality in the influent of the Ter DWTP. However, there are some exceptional situations where reservoir stability is altered and then water quality recovery (reservoir stratification) is slow compared to the river regime fluctuations. In the Llobregat River, changes are strongly linked to seasonal events, and water content has higher fluctuations throughout the year. These results are aligned to previous local catchment studies (Fernández-Turiel et al., 2003a, 2003b). From this basis, DWTP should adapt coagulation performance to this influent water quality changes.

Influent water quality was classified depending on baseline and peak organic content for the two case-study DWTPs. Linking the information presented in Table 1 with cluster analysis, two SC were conducted at each DWTP (baseline and peak cases). At the Llobregat and Ter DWTPs, a SC for each influent water classification was selected to develop RSM models (see Fig. 2). SC 1 and SC 8 (see Table 2) belong to the baseline cluster. Then, SC 4 and SC 5 are related to the peak cluster. SC and their respective classification is indicated in Fig. 2. To link the SC with clustering analysis from here the SC 1, 4, 5, 8 are named Llobregat baseline (LB), Llobregat peak (LP), Ter peak (TP) and Ter baseline (TB), respectively.

3.2. Evaluation of enhanced coagulation models

The RSM experiments were conducted and supernatants from the jar test experiments were analysed. Model factors were pH, C_d and F_d, while responses were introduced as percentage of removal of turbidity,

Fig. 2. Influent water classifications resulting from cluster analysis at the Llobregat DWTP (left column) and Ter DWTP (right column) during the period 2017-2020. Y axis are the selected water quality parameters: turbidity, TOC, UV_{254}, colour and SUVA values. Black and grey colours indicate clusters: baseline and peak, respectively. The SCs are represented by vertical grey bars.

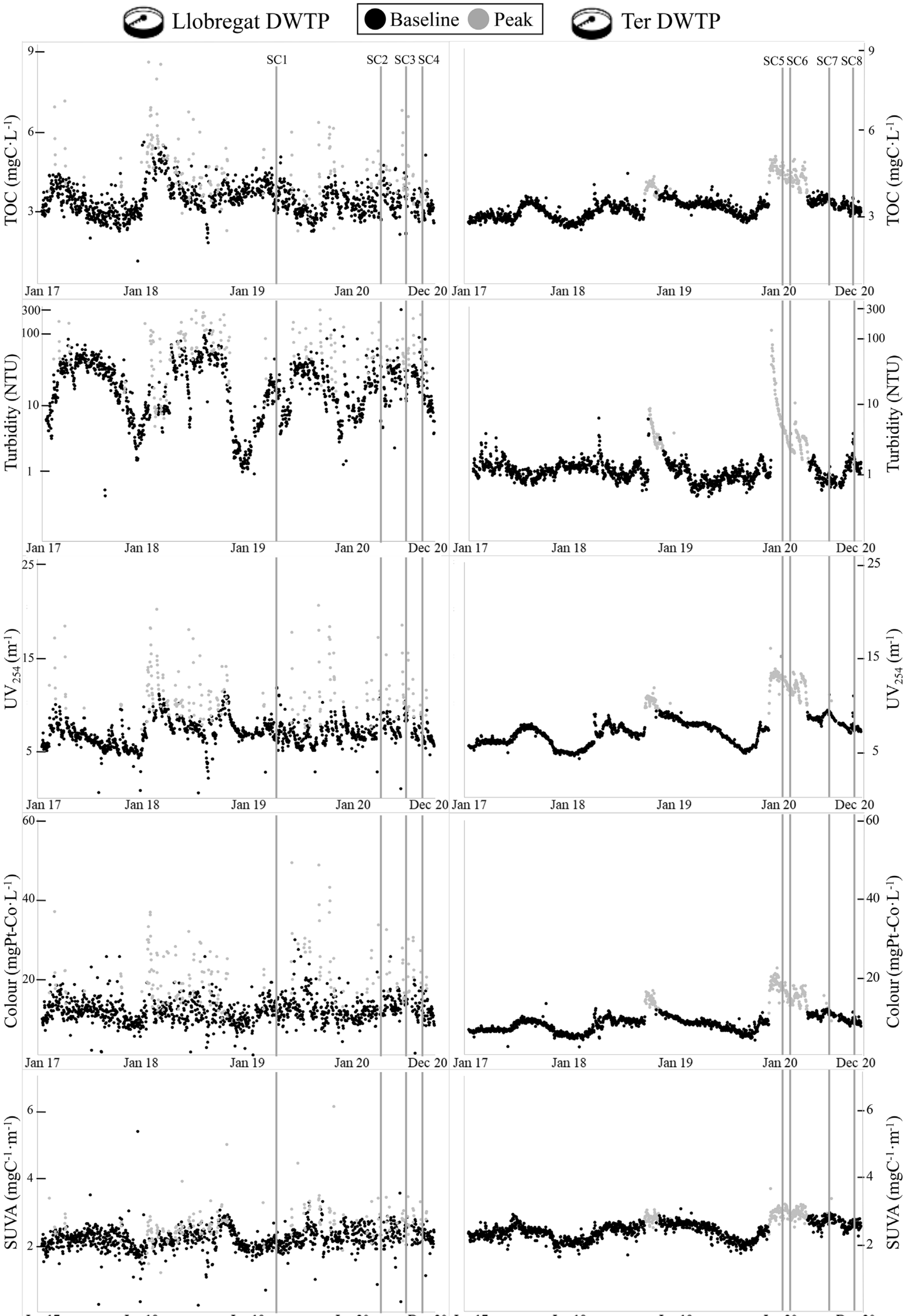

Llobregat DWTP
Baseline
Peak
Ter DWTP
TOC (mgC·L-1)
Turbidity (NTU)
UV254 (m-1)
Colour (mgPt-Co·L-1)
SUVA (mgC-1·m-1)
SC1
SC2
SC3
SC4
SC5
SC6
SC7
SC8
Jan 17
Jan 18
Jan 19
Jan 20
Dec 20

Table 2
Enhanced coagulation models for each DWTP. The number of predictors selected based on best subset selection method, coefficients for each factor and the coefficient of determination (R^2) are presented.

DWTP	Response	No. predictors	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	R^2
LB	Turbidity	4	34.5	–	1.4	18.1	–	–	-0.3	–	-0.01	–	0.91
	TOC	3	6.8	–	0.3	21.3	–	-2.3	–	–	–	–	0.91
	UV_{254}	5	223.3	-60.4	0.6	–	–	3.9	-0.4	4.03	–	–	0.73
LP	Turbidity	4	77.7	1.6	0.5	–	-0.02	–	–	–	-0.01	–	0.84
	TOC	3	9.5	–	1.2	–	-0.1	–	–	–	-0.01	–	0.84
	UV_{254}	2	33.5	–	0.3	–	–	–	–	-0.3	–	–	0.91
TB	Turbidity	5	−16.1	7.9	2	43.1	–	-6.3	–	–	-0.02	–	0.94
	TOC	5	30.2	-2.9	0.6	-19	–	–	–	–	-0.01	9.42	0.72
	UV_{254}	3	23.9	–	1.1	–	–	–	–	-0.2	-0.01	–	0.9
TP	Turbidity	4	86.1	–	0.4	–	–	-0.8	0.1	–	-0.01	–	0.79
	TOC	4	−246	93.3	–	–	0.17	–	–	-7.6	-0.02	–	0.83
	UV_{254}	4	74.5	–	-0.8	–	0.3	–	–	-0.7	-0.01	–	0.9

TOC and UV_{254}. All runs and different standards for each RSM are presented in appendix A section.

Models were obtained from the RSM experiments and after the predictors selection. Prior to determine the optimum number of predictors an evaluation task was carried out to pre-validate models. Several analyses were checked such as diagnostics related to normalized plots (models residual for experimental runs, predicted and actual values, among others) to ensure models' fitting and detect outliers. Also, 3D surface plots were analysed in both facilities under the two scenarios to interpret models, which are presented in appendix D section. These plots are useful to interpret visually (colour legend) factors interactions and responses variations inside the response surface. The final equations for each DWTP scenario with the number of predictors selected, coefficients and $R^2_{adj.}$ values are presented in Table 2.

Results from laboratory experiments (see appendix A) revealed that turbidity experimented the highest mean removals in all enhanced coagulation models. The unified mean turbidity removals for all models was 85%. This effect can be explained by considering that turbidity accounts for the whole spectrum of water compounds: organic, inorganic, particulate, colloidal and dissolved. Concerning TOC and UV_{254}, which are related to water dissolved organic fractions, were less removed during coagulation. From this results, it is appreciable that turbidity removal is achieved in coagulation for the performed scenarios. The general trend in all RSMs were that the percentage of removal was higher in UV_{254} with respect to TOC. Examining all the RSMs, the average of removal for TOC and UV_{254} were 23% and 38%, respectively. TOC and UV_{254} removals were low compared to turbidity, indicating that the optimum conditions needs to be ensured in accordance to these parameters to achieve an optimal removal of dissolved pollutants during coagulation process.

Going into greater detail for both DWTPs, the maximum removals (three responses combined) were obtained at neutral pH and medium C_d and F_d. Otherwise, minimum removals were shown at high pH values above 8 and low C_d combined with high F_d. Regarding the latter, this is due to the fact that a flocculant overdose induces a decrease of sedimentation coagulation effect and, consequently, less efficiency in the process (Katrivesis et al., 2019). The highest turbidity removals >80% were observed at neutral pH and medium C_d and F_d, while the higher removals for TOC were obtained at depressed pH levels. According to Bell-Ajy et al. (2000) and Edwards (1997), this is attributed to the increase of floc precipitation originated by the entrapment of sorbable TOC fraction combined with alum hydroxide from the coagulant. The highest removals of UV_{254} were shown at low to neutral pH levels, linked with the removal of the organic compounds in these conditions (Altmann et al., 2016). The C_d and F_d affected the removals in a different way for each RSM. More information about the individual experiments, including the RSM models' runs and responses analyses is provided in the appendix A section.

At the Llobregat DWTP, mean RSMs percentage of removals were 86% ± 23%, 21% ± 9%, 32% ± 13% for turbidity, TOC and UV_{254}, respectively. On the other hand, the results obtained reflected that the Ter DWTP responses mean percentage of removal were 84% ± 15% for turbidity, 27% ± 20% for TOC and 45% ± 13% for UV_{254}. For baseline clusters (LB and TB), in both facilities RSM outputs presented similar percentage of removals, especially for turbidity and UV_{254} removal. However, influent values of turbidity at the Llobregat DWTP were higher than those at the Ter DWTP. The reason for this is because Llobregat River water is affected by weather (periods of rains/droughts), runoff and some anthropogenic discharges of industrial origins, while reservoir remains stable. Turbidity removals >90% at Ter DWTP were difficult to observe during RSMs due to the water quality from reservoir, expressed in low turbidity values at the influent of the DWTP (5.23 and 4.76 NTU, respectively).

LP results presented turbidity removals ≥90% for all RSM experiments. This is due to the high initial value (63 NTUs), propitiating elevated removal values for this parameter. At the Ter DWTP, TP showed high influent values of TOC and UV_{254}, indicating that during peak scenarios the dissolved NOM fraction needs to be removed for the optimal coagulation at this facility. The percentage of removal of TOC and UV_{254} comparing TB and TP is significant, where TOC has a mean removal of 11% in TB and 43% in TP and UV_{254} has a mean removal of 37% for TB and 53% for TP.

All enhanced coagulation models are presented in Table 2. Related to predictors selection, no single model with $p = 9$ was selected after the application of the best subset selection method. The maximum number of predictors was located at $p = 5$. Despite the total number of predictors, it is important to identify which are the selected ones in order to proceed with the sensitivity analysis. R^2 predictors selection for all DWTPs scenarios and coefficients combination (2^p) are presented in appendix C. Enhanced coagulation models considering turbidity, TOC and UV_{254} were performed with a mean R^2 of 0.85. The mean responses R^2 were 0.87, 0.83 and 0.86 for turbidity, TOC and UV_{254} removals. The equation coefficients are not normalized, therefore some of them are negative. As a consequence, the following step is to perform a sensitivity analysis to identify the relative weight of each individual factor/predictor to understand the enhanced coagulation. Predictors selection based on R^2, $R^2_{adj.}$ aid to ensure predictability. Also, the fact to consider SSE (residuals) to choose the best model works as a prevention barrier for avoiding biased models (James et al., 2013).

Hence, depending on the nature of the influent waters, coagulation can be optimised following models developed with RSM in a baseline or peak scenario. These model outputs linked to Table 1 suggest that during LP an increase of particulate water fraction at the influent is detected (high values of turbidity) while TP shows high values of dissolved NOM fraction at the influent (high TOC and UV_{254}).

3.3. Sensitivity analysis

Sensitivity analysis was performed to discuss the individual factors influence for each response, based on the equations resulting from

Table 3
Delta mean-squared (δ^{msqr}) values for individual factors' coefficients for each enhanced coagulation model. Factors' relative impact was simplified for pH, Cd and Fd. Hyphenated cells correspond to the coefficients dismissed after predictors selection.

DWTP	Response	$\delta^{msqr}\beta_1$	$\delta^{msqr}\beta_2$	$\delta^{msqr}\beta_3$	$\delta^{msqr}\beta_4$	$\delta^{msqr}\beta_5$	$\delta^{msqr}\beta_6$	$\delta^{msqr}\beta_7$	$\delta^{msqr}\beta_8$	$\delta^{msqr}\beta_9$	Simplified relative impact
LB	Turbidity	–	0.5	1.3	–	–	0.6	–	0.1	–	Fd[a] > Cd
	TOC	–	0.02	0.3	–	0.7	–	–	–	–	Fd[a] > Cd
	UV_{254}	0.3	0.002	–	–	0.09	0.01	0.4	–	–	pH[a] > Cd
LP	Turbidity	0.4	0.1	–	0.1	–	–	–	0.03	–	pH[a] > Cd
	TOC	–	0.2	–	0.5	–	–	–	0.04	–	Cd[a]
	UV_{254}	–	0.01	–	–	–	–	0.5	–	–	pH[a] > Cd
TB	Turbidity	0.5	0.1	0.7	–	1.6	–	–	0.02	–	Fd[a] > pH > Cd
	TOC	0.5	0.02	0.9	–	–	–	–	0.01	1.8	Fd[a] > pH > Cd
	UV_{254}	–	0.4	–	–	–	–	0.9	0.1	–	pH[a] > Cd
TP	Turbidity	–	0.1	–	–	0.2	0.1	–	0.01	–	Cd[a]
	TOC	7.4	–	–	0.2	–	–	2.2	0.1	–	pH[a] > Cd
	UV_{254}	–	0.2	–	0.2	–	–	0.2	0.1	–	pH[a] ≈ Cd

[a] Factors selected for profile plots visualization.

RSM experiments and the best subset selection method. For this purpose, delta mean-squared analysis was applied to the models with the aim to identify the relative weights of model factors. For each scenario three enhanced coagulation models were obtained, one for each selected response: turbidity, TOC and UV_{254}; giving a total number of twelve equations (Table 2, β coefficients).

Delta mean-squared values for individual factors (β_1, β_2, β_3), combined interactions (β_4, β_5, β_6) and quadratic effects (β_7, β_8, β_9) are presented in Table 3. To proceed with the discussion, the relative impact of individual factors was included to Table 3, considering only the single individual factors contribution. Despite this simplification, factors interactions expressed by β_4, β_5, β_6 are significant for some of the enhanced coagulation models. For almost all scenarios (LB, LP, TB and TP) pH and C_d emerged as important factors. According to Bell-Ajy et al. (2000), pH is the most important factor for NOM removal during coagulation process. When coagulation is adjusted at the optimum pH, removals are improved because of major alum-NOM complexation and less coagulant demand. At this point, it is important to note that pH is legislated for water consumption and it is not feasible to optimise the process in the range of optimum pH, because this can be located outside of the threshold limits. From this, C_d and F_d adjustment play a key role to achieve enhanced coagulation in drinking water production.

Prior to comparing the models, it is important to remark that results from sensitivity analysis comparison is applicable for models developed at the same DWTP, because each RSM was designed under a specific treatment train and operational ranges. Furthermore, nature and regime of the sources as well as water characterization differs from the Llobregat river in comparison to the Ter river reservoirs system, details of which can be found in the cluster analysis section.

Starting with the comparison between baseline and peak events, there are similar behaviours at the two DWTPs. Regarding LP and TP, an increase of organic load is expected at the influent, and the sensitivity analysis reveals that pH and C_d are the key factors to ensure enhanced coagulation (Table 3). For these cases, when influent waters belong to peak scenarios, the optimum pH range is wider and the Cd needs to be carefully adjusted to ensure NOM adsorption and chemical bridging (Gaikwad and Munavalli, 2019). However, during LB and TB, it is necessary to carefully adjust pH, C_d and F_d to achieve high levels of pollutants removals during coagulation. In this cases F_d is also considered a key parameter, presenting high delta mean-squared, hence its importance on process performance.

Accounting for individual responses, C_d has a great influence on turbidity and UV_{254} removals efficiency and was selected as a predictor for all models accounting for these responses, ensuring high levels of removals when C_d is correctly optimised (Rocha et al., 2020). Regarding TOC value, there are different relevant factors depending on the scenario. This basically can be explained from the assumption that TOC value represents a great variety of dissolved compounds, depending on the predominant group of water pollutants the key conditions for coagulation can be modified.

Stablishing the comparison from the different water catchments, as a result of clustering analysis section Llobregat river peak events (LP) are linked clearly to an increase of turbidity. This is mainly associated to rains and its derived runoff effect. Based on that, a major part of water pollutants present in water are linked to particles accounting for turbidity more than the dissolved fraction (TOC and UV_{254}). According to Aboubaraka et al. (2017), turbidity is representative for coloured organic compounds, changing water colour. Within this context and according to the sensitivity analysis, during LP C_d requires to be carefully optimised to ensure the particles aggregation and its sedimentation during coagulation process.

At the Ter DWTP, turbidity is usually stable thorough the year. Clustering analysis denoted that during peak scenarios (TP) TOC and UV_{254} are high, then pH adjustment gains importance, especially for TOC removal. TOC removal is usually related to a decrease in operational pH, resulting in NOM protonation by acid addition, favouring complexation reactions between metal-based coagulants and carbon-based compounds (Bell-Ajy et al., 2000). On the other hand, in baseline scenario (TB) F_d gains importance (see Table 3) and requires being precisely adjusted because water contains less colloidal and suspended compounds, and the optimum range of F_d becomes more specific. An overdose of flocculant could induce an increase in turbidity at the effluent of coagulation process.

Profile plots for enhanced coagulation were performed to complement the information provided by the sensitivity analysis (Fig. 3). Fig. 3 presents the profile plots for the enhanced coagulation models for turbidity, TOC and UV_{254} removals at each DWTP under the different scenarios. In each plot, the entire range of pH, C_d or F_d is presented for the low, medium or high values of the other factors. The pH was ranged in a feasible full-scale values of operation from 6 to 8. The selected profile plots (see Fig. 3) were the cases where TOC or UV_{254} emerge as significant factors (see Table 3). Profile plots (see Fig. 3) for each response (turbidity, TOC and UV_{254}) were levelled according to the other significant factors (see Table 3).

According to Fig. 3, in LB the pH should be adjusted at a neutral levels and F_d has a positive relation for turbidity, TOC and UV_{254} removals. In the case of turbidity removal, high F_d increases significantly turbidity removal with low C_d dose, being the medium dose of C_d and F_d the best for an optimal operation. Regarding TOC, pH is not critical (≈7) and F_d and C_d have a strong positive impact in the percentage of removal. For LB-UV_{254}, pH is significant for the model (sensitivity analysis results) but not critical between 6 and 8. However, C_d has clear positive impact to remove UV_{254}. In summary, for LB at a neutral pH, high C_d and F_d TOC and UV_{254} removals improve. Compared to LP, medium level of C_d and low-neutral pH are required for turbidity, TOC and UV_{254} removals. In accordance to sensitivity analysis, F_d is not relevant for enhanced coagulation

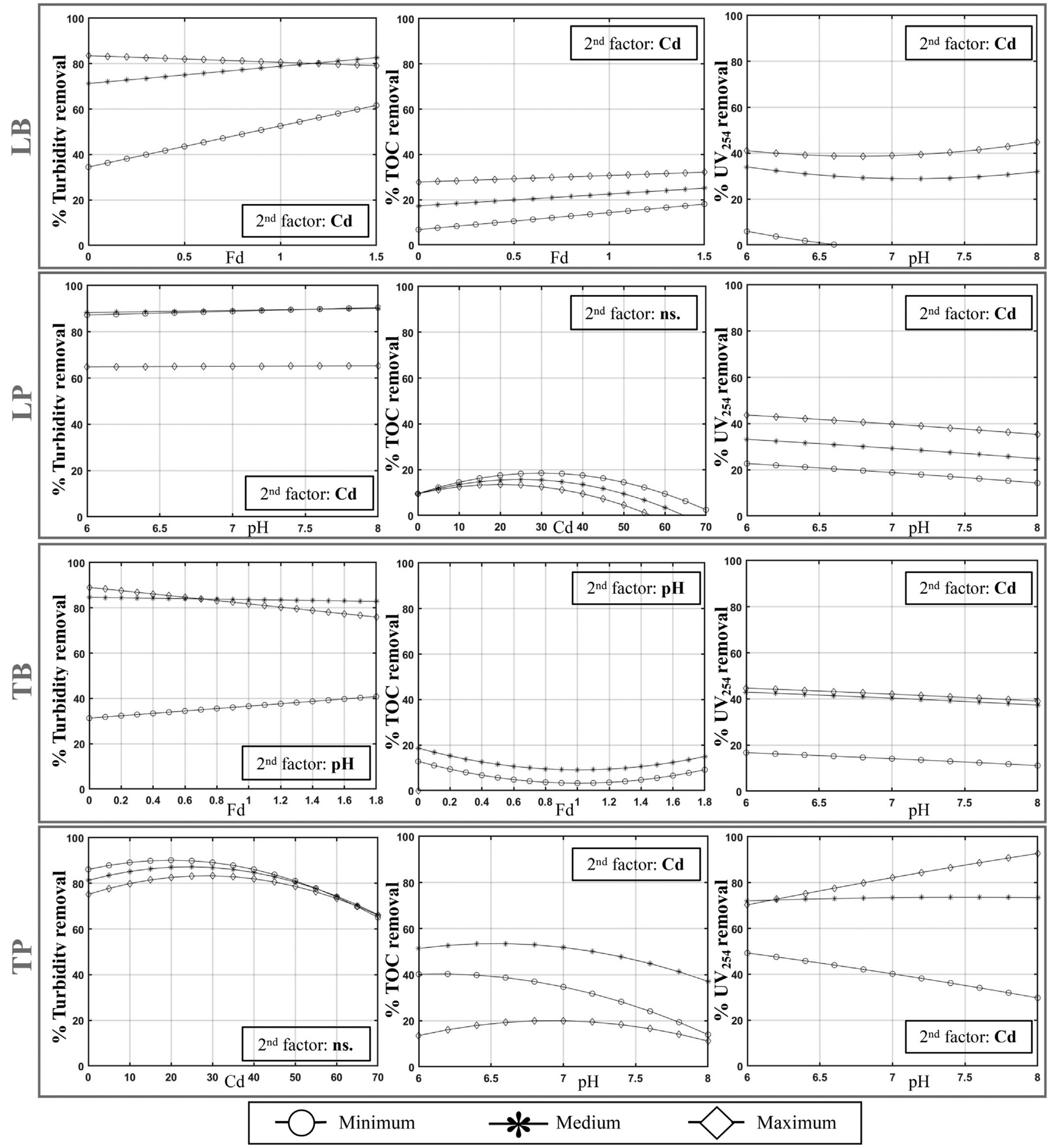


Fig. 3. Profile plots for Turbidity, TOC and UV_{254} percentage of removals for LB, LP, TB and TP (for the factors presenting the highest δ^{msqr}). The X axis are this factors presenting the highest δ^{msqr} (pH, C_d and F_d) for each response DWTP scenario, located at Y axis as % of removal. Then, the other factors presenting lower relative impact were levelled for the minimum, medium and maximum values according to the operational DWTP ranges.

in LP case. Regarding TB, the optimal adjustment of F_d is located at medium pH and C_d. Concerning the TP scenarios, the optimum removals for dissolved NOM are located at medium pH and C_d levels (F_d is not significant). In those cases, the medium range of pH-C_d is highlighted as a proper option for coagulation removals. Profile plots are useful to complement the information obtained through the sensitivity analysis, increasing model understanding for each specific scenario.

Results from sensitivity analysis indicate that enhanced coagulation for river catchment in baseline scenario is subjected to the optimal adjustment of the three factors influencing enhanced coagulation while during peaks, which are related to the increase of particulate compounds in water resulting from rain runoff effects, pH and C_d are crucial for enhanced coagulation. On the other hand, reservoir catchment is stable all over the time series where enhance coagulation is controlled by high C_d and F_d, but during peaks (extreme events) an increase of dissolved NOM is expected at the influent waters, resulting from the resuspension of reservoir deep sediments. For these cases, coagulation pH should be carefully levelled to ensure the optimum TOC removal and C_d emerges as a crucial factor for turbidity and UV_{254}.

3.4. Practical implications

In this section, practical implications as well as the limitations for models implementation are stated. Enhanced coagulation models implementation is based on influent waters classification, thus determining which model to propose for coagulation. Then, depending on the fixed enhanced coagulation optimisation criteria for the selected quality standards (responses removal), a specific pH, C_d and F_d can be proposed by models for the desired operation (Fig. 4).

This study, which has been developed within the context of two water treatment facilities, is adaptable to other full scale DWTPs, but some requirements should be taken in consideration. Firstly, the type of catchment. Results demonstrate that influent water quality is subjected to the type of surface water catchment (river or reservoir) and these has an effect on the optimisation of coagulation process. The water regimes differ between them and this affects water quality, quantity, as well as the frequency of these fluctuations. Then, the installation and the capacity to monitor and track influent water quality is crucial when implementing the models. To characterise waters, a number of minimum parameters should be analysed, i.e., at least the three basics for enhanced coagulation models: turbidity, TOC and UV_{254}. This step, in some cases, could imply capital investment and derived operational costs (sensor maintenance and replacement). Also, all these data generated by influent water sensors should be upload to supervisory control and data acquisition (SCADA), a control system architecture with the capacity to register and display data. Moreover, a flexible operation for coagulation would be required, allowing water treatment to be adapted to the proposed model outputs (operational factors). In other words, DWTPs should have the capacity to easily change pH, C_d and F_d. It is important to considerer coagulation location in water treatment train, at the two case studies raw water parameters were used because no other process was affecting the water characterization. Should there be other steps before coagulation, the quality of influent waters would not be representative for the optimisation.

There are some practical operational implications which could be taken in consideration. In the case of extreme events, detecting high levels of turbidity in the influent (>40NTUs), coagulation optimisation criteria should be readjusted by, at least, >95% to ensure the removal during coagulation and avoid pore blockage in the subsequent processes if there are some filtration-based technologies involved. Also, taking in consideration the results obtained in this study, in these cases the correct adjustment of C_d becomes crucial for enhanced coagulation because pH level should be maintained within the legislated ranges. This means that can exist cases where model-propose pH levels that are not applicable to full-scale operations. For example, if pH = 5.5 or 8.5 are suggested to ensure optimum conditions for a specific removal criterion, this may will be the ideal for enhanced coagulation, but is not feasible for drinking water production and the later consumption. Then, some restrictions can be applied to the models by considering a readjustment of C_d and F_d instead of the decrease in pH by fixing some limiting thresholds (e.g., pH > 7 and pH < 8). During baseline scenarios, where water quality at the influent is not considered poor and water characterization levels is low, enhanced coagulation depends on the correct adjustment of pH, C_d and F_d, being the latter a determinant parameter for process performance.

The mathematical models developed here have some design limitations and it is important to state them for the future applications. First, RSM experiments and models should be performed for individual catchments. Moreover, to develop RSM model replicates (jar test experiments) during the year within specific catchment/weather situations would bring additional information for the existing models and increase their reliability in all scenarios. Under the supervision of an experienced user, the proposed models also have the potential to act as decision-making support tools with which to check the viability of any proposed values (see Fig. 4). That said, a user interface for that task would need to be developed.

As a point of insight into (and related to) prevention tasks, increasing the capacity to be able to monitor and predict weather forecast in the drinking-water sector by controlling hydraulic regimes, retention times and catchment basins, is important. As present and near-future predictions anticipate, the frequency of extreme weather events will be reduced in time and therefore, regular floods will have to be taken in consideration for decision-making purposes. This study highlights the importance of meteorology to water production/management sector. In the case of river (seasonality) and reservoir (extreme events) catchments the relationship between influent water quality and rains/storms was highlighted. These exceptional circumstances will become habitual and adapting water treatment to them will be required to safeguard the water supply during these stages.

4. Conclusions

Enhanced coagulation models were developed to optimise coagulation processes at two Mediterranean DWTPs. Specifically, these models were designed to remove turbidity, TOC and UV_{254}, which are stated as water quality parameters and NOM surrogates. Cluster analysis based on K-means algorithm was applied to influent water characterization databases to classify waters into baseline and peak organic content. For each cluster, a coagulation RSM accounting for turbidity, TOC and UV_{254} percentage of removals was designed and developed. Then, after the predictors selection the models outputs (equations) were validated with a sensitivity analysis based on delta mean-squared (δ^{msqr}) to quantify model factors relative impact for the previously-mentioned scenarios. The models mean R^2 value for the three responses at Llobregat DWTP were 0.85 and 0.86 while in Ter DWTP were 0.85 and 0.84, in both cases for baseline and peak scenarios, respectively.

The study of these models was conducted to determine that the differences between water catchments alter the quality of the influent water at the DWTPs, thus affecting the optimum for enhanced coagulation. Results from clusters analysis revealed that the water catchment determines drinking water quality because of the temporal fluctuations of influent organic load. Clustering analysis provided information about the intensity, the frequency and the water characterization during baseline and peak scenarios. Then, sensitivity analysis allowed to find out which are the key factors for enhanced coagulation depending on the scenario.

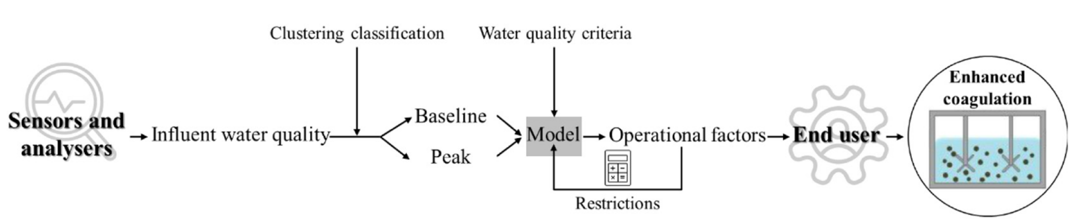


Fig. 4. Roadmap for the implementation of enhanced coagulation models at a full-scale DWTP.

The Llobregat River is a challenging case due to seasonal fluctuations and the sudden high organic loads in the waters caused by anthropogenic pressure and rains runoff. On the other hand, rather than the seasonal change, the influent waters at the Ter DWTP (reservoir catchment) are altered through extreme weather events. Results of cluster analysis determined that peak events at Llobregat DWTP are seasonal and related to an increase of particles and coloured compounds, expressed by high levels of turbidity at the influent. On the other hand, Ter DWTP peak scenarios are linked to extreme weather events and are challenging due to the increase of dissolved NOM, which is expressed by higher values of TOC and UV_{254} more than turbidity. From this and considering the sensitivity analysis, in baseline scenarios it is important to adjust at the optimum levels (which are not the highest) the three influential factors (pH, C_d and F_d) in order to ensure enhanced coagulation, resulting obvious if low pollutants load is located at the influent waters. However, during peak scenarios pH and C_d become the key factors for enhanced coagulation and F_d is not relevant for the process itself due to reduce the high levels of pollutants present at the influent. In this cases C_d is the key factor to ensure the water quality.

CRediT authorship contribution statement

J. Suquet: Conceptualization, Methodology, Investigation, Software, Writing – original draft. **Ll. Godo-Pla:** Conceptualization, Software, Supervision, Writing – review & editing. **M. Valentí:** Methodology. **L. Ferrandez:** Writing – review & editing. **M. Verdaguer:** Data curation. **M. Poch:** Supervision. **M.J. Martin:** Funding acquisition, Resources. **H. Monclús:** Conceptualization, Project administration, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the Ministerio de Economía, Industria y Competitividad, Gobierno de España: Retos de la Sociedad Projects (CTM2017-83598-R & PID2020-112615RA-I00). In addition, this study was supported by the Universitat de Girona through the PhD student grants IFUdG2018-69 (Jordi Suquet) and 2020 FISDU 00330 (Meritxell Valentí), and by the Spanish Ministry of Science and Innovation (MICINN) for the Ramon y Cajal research fellowship (RYC2019-026434-I) (Hèctor Monclús). LEQUIA has been recognized as consolidated research group by the Catalan Government (2017-SGR-1552). All authors would like to thank the Ens d'Abastament d'Aigua Ter-Llobregat (ATL) and, in particular, Fernando Valero and Pere Emiliano from the ATL R&D department for their cooperation and implication during this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2021.149398.

References

Aboubaraka, A.E., Aboelfetoh, E.F., Ebeid, E.Z.M., 2017. Coagulation effectiveness of graphene oxide for the removal of turbidity from raw surface water. Chemosphere 181, 738–746. https://doi.org/10.1016/j.chemosphere.2017.04.137.

Adesina, O.A., Abdulkareem, F., Yusuff, A.S., Lala, M., Okewale, A., 2019. Response surface methodology approach to optimization of process parameter for coagulation process of surface water using Moringa oleifera seed. S. Afr. J. Chem. Eng. 28, 46–51. https://doi.org/10.1016/j.sajce.2019.02.002.

Altmann, J., Massa, L., Sperlich, A., Gnirss, R., Jekel, M., 2016. UV254 absorbance as real-time monitoring and control parameter for micropollutant removal in advanced wastewater treatment with powdered activated carbon. Water Res. 94, 240–245. https://doi.org/10.1016/j.watres.2016.03.001.

Amores, A., Marcos, M., Carrió, Di S., Gomez-Pujol, L., 2020. Coastal impacts of Storm Gloria (January 2020) over the north-western Mediterranean. Nat. Hazards Earth Syst. Sci. 20, 1955–1968. https://doi.org/10.5194/nhess-20-1955-2020.

Andersson, A., Lavonen, E., Harir, M., Gonsior, M., Hertkorn, N., Schmitt-Kopplin, P., Kylin, H., Bastviken, D., 2020. Selective removal of natural organic matter during drinking water production changes the composition of disinfection by-products. Environ. Sci. Water Res. Technol. 6, 779–794. https://doi.org/10.1039/c9ew00931k.

Apostol, G., Kouachi, R., Constantinescu, I., 2011. Optimization of coagulation-flocculation process with aluminum sulfate based on response surface methodology. UPB Sci. Bull. Ser. B Chem. Mater. Sci. 73, 77–84.

Arruda, P.M., Pereira-Filho, E.R., Libânio, M., Fagnani, E., 2018. Response surface methodology applied to tropical freshwater treatment. Environ. Technol. (United Kingdom), 1–11. https://doi.org/10.1080/09593330.2018.1514072.

Ates, N., Kitis, M., Yetis, U., 2007. Formation of chlorination by-products in waters with low SUVA-correlations with SUVA and differential UV spectroscopy. Water Res. 41, 4139–4148. https://doi.org/10.1016/j.watres.2007.05.042.

Awad, J., Fisk, C.A., Cox, J.W., Anderson, S.J., Van Leeuwen, J., 2018. Modelling of THM formation potential and DOM removal based on drinking water catchment characteristics. Sci. Total Environ. 635, 761–768. https://doi.org/10.1016/j.scitotenv.2018.04.149.

Baxter, C.W., Zhang, Q., Stanley, S.J., Shariff, R., Tupas, R.-R.T., Stark, H.L., 2001. Drinking water quality and treatment: the use of artificial neural networks. Can. J. Civ. Eng. 28, 26–35. https://doi.org/10.1139/cjce-28-s1-26.

Beauchamp, N., Bouchard, C., Dorea, C., Rodriguez, M., 2020. Ultraviolet absorbance monitoring for removal of DBP-precursor in waters with variable quality: enhanced coagulation revisited. Sci. Total Environ. 717, 137225. https://doi.org/10.1016/j.scitotenv.2020.137225.

Bell-Ajy, K., Abbaszadegan, M., Ibrahim, E., Verges, D., LeChevallier, M., 2000. Conventional and optimized coagulation for NOM removal. J. Am. Water Work. Assoc. 92, 44–58. https://doi.org/10.1002/j.1551-8833.2000.tb09023.x.

Casamitjana, X., Serra, T., Colomer, J., Baserba, C., Pérez-Losada, J., 2003. Effects of the water withdrawal in the stratification patterns of a reservoir. Hydrobiologia 504, 21–28. https://doi.org/10.1023/B:HYDR.0000008504.61773.77.

Celestino, A.E.M., Cruz, D.A.M., Sánchez, E.M.O., Reyes, F.G., Soto, D.V., 2018. Groundwater quality assessment: an improved approach to K-means clustering, principal component analysis and spatial analysis: a case study. Water (Switzerland) 10, 1–21. https://doi.org/10.3390/w10040437.

Chaukura, N., Marais, S.S., Moyo, W., Mbali, N., Thakalekoala, L.C., Ingwani, T., Mamba, B.B., Jarvis, P., Nkambule, T.T.I., 2020. Contemporary issues on the occurrence and removal of disinfection byproducts in drinking water - a review. J. Environ. Chem. Eng. 8, 103659. https://doi.org/10.1016/j.jece.2020.103659.

Christman, R.F., Ghassemi, M., 1966. Chemical nature of organic color in water. J. Am. Water Works Assoc. 58, 723–741. https://doi.org/10.1002/j.1551-8833.1966.tb01631.x.

Cramer, W., Guiot, J., Fader, M., Garrabou, J., Gattuso, J.P., Iglesias, A., Lange, M.A., Lionello, P., Llasat, M.C., Paz, S., Peñuelas, J., Snoussi, M., Toreti, A., Tsimplis, M.N., Xoplaki, E., 2018. Climate change and interconnected risks to sustainable development in the Mediterranean. Nat. Clim. Chang. 8, 972–980. https://doi.org/10.1038/s41558-018-0299-2.

Directive (EU), 2020. 2020/2184 of the European parliament and of the council of 16 December 2020 on the quality of water intended for human consumption (recast). Off. J. Eur. Union L435, 1–62.

Delpla, I., Jung, A.V., Baures, E., Clement, M., Thomas, O., 2009. Impacts of climate change on surface water quality in relation to drinking water production. Environ. Int. 35, 1225–1233. https://doi.org/10.1016/j.envint.2009.07.001.

Dragon, K., Górski, J., Kruc, R., Drozdzynski, D., Grischek, T., 2018. Removal of natural organic matter and organic micropollutants during riverbank filtration in Krajkowo, Poland. Water (Switzerland) 10, 1457. https://doi.org/10.3390/w10101457.

Edwards, M., 1997. Predicting DOC removal during enhanced coagulation. J. Am. Water Work. Assoc. 89, 78–89. https://doi.org/10.1002/j.1551-8833.1997.tb08229.x.

Edzwald, J.K., 1993. Coagulation in drinking water treatment: particles, organics and coagulants. Water Sci. Technol. 27, 21–35. https://doi.org/10.2166/wst.1993.0261.

Espadaler, I., Caixach, J., Om, J., Ventura, F., Cortina, M., Pauné, F., Rivera, J., 1997. Identification of organic pollutants in Ter river and its system of reservoirs supplying water to Barcelona (Catalonia, Spain): a study by GC/MS and FAB/MS. Water Res. 31, 1996–2004. https://doi.org/10.1016/S0043-1354(97)00003-1.

Fathi, E., Zamani-Ahmadmahmoodi, R., Zare-Bidaki, R., 2018. Water quality evaluation using water quality index and multivariate methods, Beheshtabad River, Iran. Appl Water Sci 8, 1–6. https://doi.org/10.1007/s13201-018-0859-7.

Fernández-Turiel, J.L., Gimeno, D., Rodriguez, J.J., Carnicero, M., Valero, F., 2003a. Spatial and seasonal variations of water quality in a Mediterranean catchment: the Llobregat River (NE Spain). Environ. Geochem. Health 25, 453–474. https://doi.org/10.1023/B:EGAH.0000004566.75757.98.

Fernández-Turiel, J.L., Gimeno, D., Rodriguez, J.J., Carnicero, M., Valero, F., 2003b. Factors influencing the quality of a surface water supply system: the Ter river, northeastern Spain. Fresenius Environ. Bull. 12, 67–75.

Gaikwad, V.T., Munavalli, G.R., 2019. Turbidity removal by conventional and ballasted coagulation with natural coagulants. Appl Water Sci 9, 1–9. https://doi.org/10.1007/s13201-019-1009-6.

Gallart, F., Delgado, J., Beatson, S.J.V., Posner, H., Llorens, P., Marcé, R., 2011. Analysing the effect of global change on the historical trends of water resources in the headwaters of the llobregat and Ter river basins (Catalonia, Spain). Phys. Chem. Earth 36, 655–661. https://doi.org/10.1016/j.pce.2011.04.009.

Godo-Pla, L., Emiliano, P., Valero, F., Poch, M., Sin, G., Monclús, H., 2019. Predicting the oxidant demand in full-scale drinking water treatment using an artificial neural network: uncertainty and sensitivity analysis. Process Saf. Environ. Prot. 125, 317–327. https://doi.org/10.1016/j.psep.2019.03.017.
Godo-Pla, L., Emiliano, P., González, S., Poch, M., Valero, F., Monclús, H., 2020a. Implementation of an environmental decision support system for controlling the pre-oxidation step at a full-scale drinking water treatment plant. Water Sci. Technol. 81, 1778–1785. https://doi.org/10.2166/wst.2020.142.
Godo-Pla, L., Rodríguez, J.J., Suquet, J., Emiliano, P., Valero, F., Poch, M., Monclús, H., 2020b. Control of primary disinfection in a drinking water treatment plant based on a fuzzy inference system. Process Saf. Environ. Prot. 145, 63–70. https://doi.org/10.1016/j.psep.2020.07.037.
Godo-Pla, L., Emiliano, P., Poch, M., Valero, F., Monclús, H., 2021. Benchmarking empirical models for THMs formation in drinking water systems: an application for decision support in Barcelona, Spain. Sci. Total Environ. 763, 144197. https://doi.org/10.1016/j.scitotenv.2020.144197.
Gregor, J.E., Nokes, C.J., Fenton, E., 1997. Optimising natural organic matter removal from low turbidity waters by controlled pH adjustment of aluminium coagulation. Water Res. 31, 2949–2958. https://doi.org/10.1016/S0043-1354(97)00154-1.
Harfouchi, H., Hank, D., Hellal, A., 2016. Response surface methodology for the elimination of humic substances from water by coagulation using powdered Saddled Sea bream scale as coagulant-aid. Process Saf. Environ. Prot. 99, 216–226. https://doi.org/10.1016/j.psep.2015.10.019.
Heddam, S., Bermad, A., Dechemi, N., 2012. ANFIS-based modelling for coagulant dosage in drinking water treatment plant: a case study. Environ. Monit. Assess. 184, 1953–1971. https://doi.org/10.1007/s10661-011-2091-x.
Hou, L., Zhou, Q., Wu, Q., Gu, Q., Sun, M., Zhang, J., 2018. Spatiotemporal changes in bacterial community and microbial activity in a full-scale drinking water treatment plant. Sci. Total Environ. 625, 449–459. https://doi.org/10.1016/j.scitotenv.2017.12.301.
James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning with Applications in R. Springer, New York.
Jorda-Capdevila, D., Gampe, D., Huber García, V., Ludwig, R., Sabater, S., Vergoñós, L., Acuña, V., 2019. Impact and mitigation of global change on freshwater-related ecosystem services in southern Europe. Sci. Total Environ. 651, 895–908. https://doi.org/10.1016/j.scitotenv.2018.09.228.
Katrivesis, F.K., Karela, A.D., Papadakis, V.G., Paraskeva, C.A., 2019. Revisiting of coagulation-flocculation processes in the production of potable water. J. Water Process Eng. 27, 193–204. https://doi.org/10.1016/j.jwpe.2018.12.007.
Kim, C.M., Parnichkun, M., 2017. Prediction of settled water turbidity and optimal coagulant dosage in drinking water treatment plant using a hybrid model of k-means clustering and adaptive neuro-fuzzy inference system. Appl Water Sci 7, 3885–3902. https://doi.org/10.1007/s13201-017-0541-5.
Krzeminski, P., Vogelsang, C., Meyn, T., Köhler, S.J., Poutanen, H., de Wit, H.A., Uhl, W., 2019. Natural organic matter fractions and their removal in full-scale drinking water treatment under cold climate conditions in nordic capitals. J. Environ. Manag. 241, 427–438. https://doi.org/10.1016/j.jenvman.2019.02.024.
Kuster, M., Martín-Alonso, J., Petrovic, M., Barceló, D., Hernando, M.D., López de Alda, M.J., 2008. Analysis and occurrence of pharmaceuticals, estrogens, progestogens and polar pesticides in sewage treatment plant effluents, river water and drinking water in the llobregat river basin (Barcelona, Spain). J. Hydrol. 358, 112–123. https://doi.org/10.1016/j.jhydrol.2008.05.030.
Liu, H., Liu, R., Tian, C., Jiang, H., Liu, X., Zhang, R., Qu, J., 2012. Removal of natural organic matter for controlling disinfection by-products formation by enhanced coagulation: a case study. Sep. Purif. Technol. 84, 41–45. https://doi.org/10.1016/j.seppur.2011.07.009.
Navarro, E., Guasch, H., Sabater, S., 2002. Use of microbenthic algal communities in ecotoxicological tests for the assessment of water quality: the Ter river case study. J. Appl. Phycol. 14, 41–48. https://doi.org/10.1023/A:1015242301451.
Omran, H.A., 2018. A mathematical model to determination of alum amount added for the purpose of coagulation in water purification plants. Eng. Technol. J. 36, 456–460. https://doi.org/10.30684/etj.36.4a.13.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830. https://doi.org/10.1289/EHP4713.
Poch, M., Garrido-Baserba, M., Corominas, L., Perelló-Moragues, A., Monclús, H., Cermerón-Romero, M., Melitas, N., Jiang, S.C., Rosso, D., 2020. When the fourth water and digital revolution encountered COVID-19. Sci. Total Environ. 744, 140980. https://doi.org/10.1016/j.scitotenv.2020.140980.
Postigo, C., Emiliano, P., Barceló, D., Valero, F., 2018. Chemical characterization and relative toxicity assessment of disinfection byproduct mixtures in a large drinking water supply network. J. Hazard. Mater. 359, 166–173. https://doi.org/10.1016/j.jhazmat.2018.07.022.
RD, 140/2003, 2003. Royal Decree 140/2003 of 7th February, sanitary criteria for the quality of water for human consumption. Bol. Estado 1–41.
Rendón, E., Abundez, I., Arizmendi, A., Quiroz, E.M., 2011. Internal Versus External Cluster Validation Indexes, p. 5.
Rocha, V.V.F., dos Santos, I.F.S., Silva, A.M.L., Sant'Anna, D.O., Junho, A.L., Barros, R.M., 2020. Clarification of high-turbidity waters: a comparison of Moringa oleifera and virgin and recovered aluminum sulfate-based coagulants. Environ. Dev. Sustain. 22, 4551–4562. https://doi.org/10.1007/s10668-019-00397-2.
Sadri Moghaddam, S., Alavi Moghaddam, M.R., Arami, M., 2010. Coagulation/flocculation process for dye removal using sludge from water treatment plant: optimization through response surface methodology. J. Hazard. Mater. 175, 651–657. https://doi.org/10.1016/j.jhazmat.2009.10.058.
Sillanpää, M., Ncibi, M.C., Matilainen, A., Vepsäläinen, M., 2018. Removal of natural organic matter in drinking water treatment by coagulation: a comprehensive review. Chemosphere 190, 54–71. https://doi.org/10.1016/j.chemosphere.2017.09.113.
Sin, G., Gernaey, K., 2016. Data handling and parameter estimation. Exp. Methods Wastewater Treat. 201–234.
Sun, S., Jiang, T., Lin, Y., Song, J., Zheng, Y., An, D., 2020. Characteristics of organic pollutants in source water and purification evaluations in drinking water treatment plants. Sci. Total Environ. 733, 139277. https://doi.org/10.1016/j.scitotenv.2020.139277.
Suquet, J., Godo-Pla, L., Valentí, M., Verdaguer, M., Martin, M.J., Poch, M., Monclús, H., 2020. Development of an environmental decision support system for enhanced coagulation in drinking water production. Water (Switzerland) 12. https://doi.org/10.3390/W12082115.
Trinh, T.K., Kang, L.-S., 2010. Application of response surface method as an experimental design to optimize coagulation tests. Environ. Eng. Res. 15, 63–70. https://doi.org/10.4491/eer.2010.15.2.063.
Trinh, T.K., Kang, L.S., 2011. Response surface methodological approach to optimize the coagulation-flocculation process in drinking water treatment. Chem. Eng. Res. Des. 89, 1126–1135. https://doi.org/10.1016/j.cherd.2010.12.004.
USEPA, 2009. Method 415.3 determination of total organic carbon and specific uv absorbance at 254 nm in source water and drinking water. EPA Doc.
Volk, C., Bell, K., Ibrahim, E., Verges, D., Amy, G., Lechevallier, M., 2000. Impact of enhanced and optimized coagulation on removal of organic matter and its biodegradable fraction in drinking water. Water Res. 34, 3247–3257. https://doi.org/10.1016/S0043-1354(00)00033-6.
WHO, 2017. Guidelines for Drinking-water Quality: Fourth Edition Incorporating the First Addendum. World Health Organization, Geneva.
Williams, C.J., Conrad, D., Kothawala, D.N., Baulch, H.M., 2019. Selective removal of dissolved organic matter affects the production and speciation of disinfection byproducts. Sci. Total Environ. 652, 75–84. https://doi.org/10.1016/j.scitotenv.2018.10.184.
Xia, X., Lan, S., Li, X., Xie, Y., Liang, Y., Yan, P., Chen, Z., Xing, Y., 2018. Characterization and coagulation-flocculation performance of a composite flocculant in high-turbidity drinking water treatment. Chemosphere 206, 701–708. https://doi.org/10.1016/j.chemosphere.2018.04.159.
Xie, J., Wang, D., Van Leeuwen, J., Zhao, Y., Xing, L., Chow, C.W.K., 2012. PH modeling for maximum dissolved organic matter removal by enhanced coagulation. J. Environ. Sci. 24, 276–283. https://doi.org/10.1016/S1001-0742(11)60717-1.
Yan, M., Wang, D., Yu, J., Ni, J., Edwards, M., Qu, J., 2008. Enhanced coagulation with polyaluminum chlorides: role of pH/Alkalinity and speciation. Chemosphere 71, 1665–1673. https://doi.org/10.1016/j.chemosphere.2008.01.019.
Zainal-Abideen, M., Aris, A., Yusof, F., Abdul-Majid, Z., Selamat, A., Omar, S.I., 2012. Optimizing the coagulation process in a drinking water treatment plant - comparison between traditional and statistical experimental design jar tests. Water Sci. Technol. 65, 496–503. https://doi.org/10.2166/wst.2012.561.