Heliyon 10 (2024) e30892

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Influence of physical and water quality parameters on residual chlorine decay in water distribution network

Julius Caesar Kwio-Tamale *, Charles Onyutha

Department of Civil and Environmental Engineering, Kyambogo University, P.O. Box 1 Kyambogo, Kampala, Uganda

A R T I C L E I N F O

Keywords:
Chlorine decay
Physical parameters
Water parameters
Regression analysis
Water quality

A B S T R A C T

Chlorine is the most common disinfectant in drinking water distribution practice. World Health Organization recommends 0.2–5.0 mg/l of residual chlorine in drinking water. This paper analyzed influence of physical and water quality parameters on chlorine decay in drinking water distribution. Principal component analysis, directed tree and regression were used to investigate influence of these parameters on chlorine from water treatment plant to water consumption points. Results show that initial chlorine, electrical conductivity and distance explain 62 % of chlorine decay with estimated error of 0.045 mg/l. The decision-tree feature importance scores of initial chlorine and electrical conductivity were 0.47 and 0.23 respectively. The combined feature importance scores of physical parameters of distance (0.09), pipe diameter (0.06), flow velocity (0.03), pressure (0.02) and travel time (0.046) were less than that for initial chlorine concentration (0.47) alone. These results show that conventional chlorination at water treatment plants removes largely fast inorganic reactants leaving traces of slow organic reactants as the dominant secondary contaminants in water distribution system. The key policy recommendation is to use water quality parameters more than physical parameters in order to enable water utility managers maintain residual chlorine within safe public health standards.

1. Introduction

Residual chlorine is the most common disinfectant in drinking water treatment [1] because of its efficacy against pathogenic infections, low cost, ease of application, monitoring [2] and extended disinfectant durability compared to other disinfectants [3]. The World Health Organization (WHO) recommends a minimum residual chlorine concentration of 0.2–5 mg/l at water consumption points to safeguard public health from microbial secondary contamination in treated water supply [4]. During water borne disease outbreaks and emergencies, this minimum is increased to 1.0 mg/l at tap stands and 2.0 mg/l at water delivery trucks [5]. These residual chlorine specifications during times of no disease outbreaks and times of disease outbreaks and emergencies emphasize the importance of monitoring residual chlorine levels at all times of water supply. Item 6 of WHO 2011 WSP (water safety plan) that is "*Define monitoring of control measures—what limits define acceptable performance and how these are monitored*" [6] is relevant for monitoring of such residual chlorine levels in water distribution. This WHO 2011 WSP recommended by WHO in 2004 is mandatory in Australia, Iceland, New Zealand, Serbia, Switzerland, Uganda and the United Kingdom [6].

Both physical and water quality parameters influence residual chlorine decay during water distribution. Physical water parameters are pipe length [2,7], pipe diameter [7–9], pipe roughness [2,9], pipe age [2,9] and pipe material [10]. Water quality parameters

* Corresponding author.
E-mail address: tamalekwiojc@gmail.com (J.C. Kwio-Tamale).

https://doi.org/10.1016/j.heliyon.2024.e30892
Received 19 December 2023; Received in revised form 7 May 2024; Accepted 7 May 2024
Available online 10 May 2024
2405-8440/

include initial chlorine concentration [2,7,9,11–16], pH [4,7,12,15,17–22] and turbidity [2,4,7,11–16,18–22]. However, turbidity which is commonly used by water utilities as an indicator proxy surrogate for initial chlorine dosage for suspended and colloidal organic and inorganic impurities in water [16,21]. This therefore means that turbidity excludes dissolved chlorine reactants [16]. Instead, it is advised that a minimum Ct (concentration-time) factor with 30-min contact time is used instead of turbidity [16]. This also suggests that after break-point chlorination, the role of fast chlorine reactants which are largely organics is reduced after satisfaction of initial chlorine demand. This has the effect of leaving traces of slow reacting inorganics in water to exert continuous chlorine demand. The effect is continuous loss of residual chlorine as treated water is conveyed from treatment plant downstream to consumption points. Other water quality parameters are electrical conductivity [7,15,17,18,20,22,23] which together with other water quality parameters excluding turbidity account for about 75 % variability in free chlorine decay [16]. This statistic demonstrates the significance of electrical conductivity in free chlorine residual decay in water distribution. However, it is suggested that concentrations of inorganics less than 0.3 mg/l have insignificant effect on residual chlorine decay [24]. Temperature is another key water quality parameter that influences residual chlorine decay [2,4,8,9,12,13,15,16,18–20,25,26]. Temperature varies spatially and temporally with residual chlorine decay in water distribution networks [23].

The past studies mentioned above did not definitively and quantitatively evaluate influence of physical and water quality parameters of residual chlorine decay in water distribution system. This is a gap worth closing. Therefore, the aim of this paper was to investigate the influence of physical and water quality parameters on residual chlorine decay in water distribution system after initial treatment of raw water at water treatment plants. Proper understanding and appreciation of the effect of residual chlorine decay parameters by water supply utilities and practitioners is important for making appropriate decisions in control and management of these parameters to ensure residual chlorine remains within safe limits in drinking water distribution.

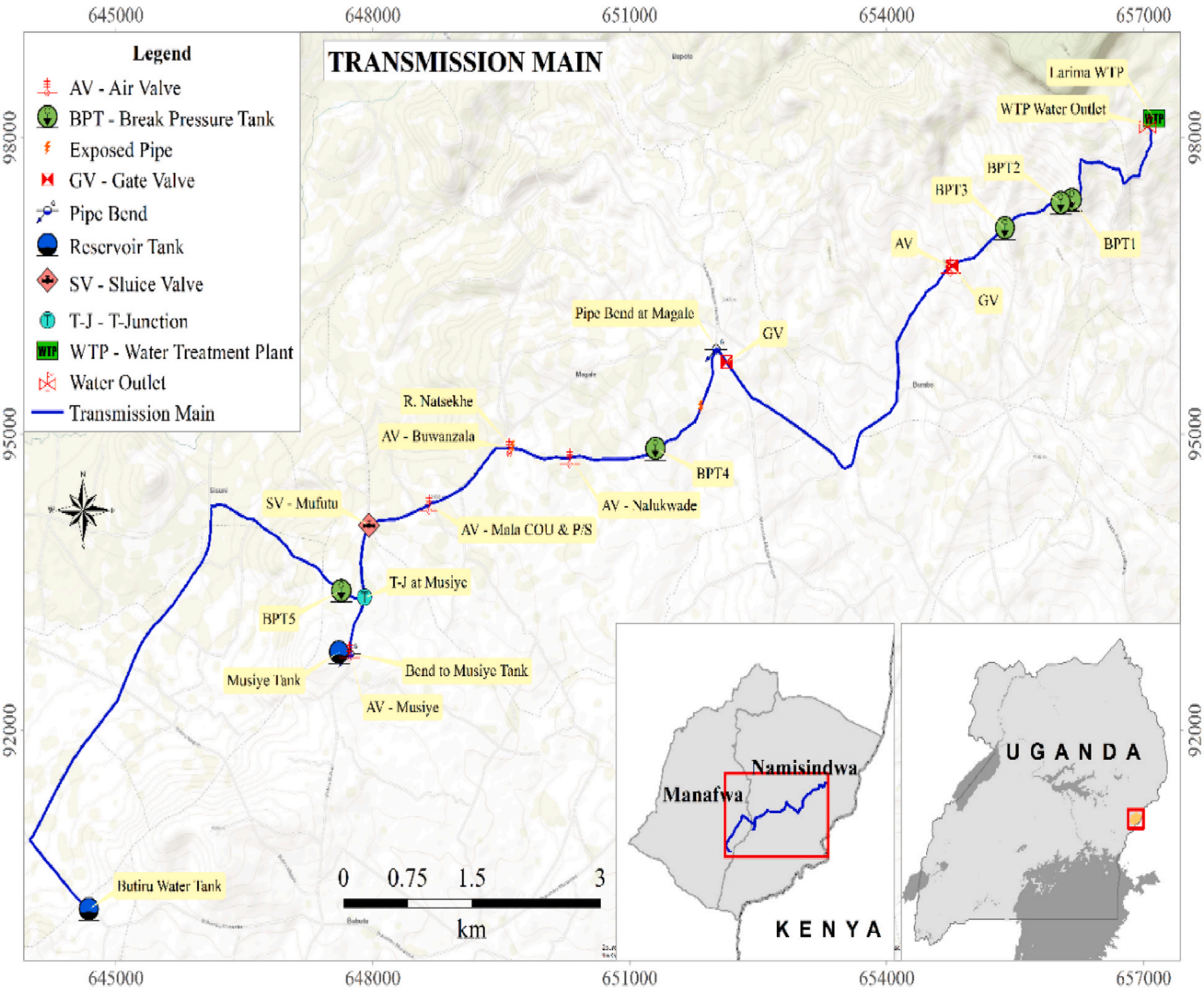


Fig. 1. Location of Lirima gravity water scheme in eastern Uganda.

2. Methods and materials

The study area, data collection strategy and procedure, data collection instruments and data analysis were as followed.

2.1. Study area

This research was conducted on Lirima Gravity Flow Scheme located in Manafwa and Namisinde districts in the Mount Elgon region in Eastern Uganda. This gravity scheme is owned and operated by National Water and Sewerage Corporation (NWSC) of Uganda which is a government parastatal. Fig. 1 shows the location and water transmission main from NWSC Lirima treatment plant.

The GPS (Geographical Positioning System) coordinates of the water source and treatment plant of this gravity flow scheme is 36 N (Latitude), 0657122 (Northing), 0098196 (Easting) at altitude of 1812 m above sea level. The scheme starts just inside Uganda at the Uganda-Kenya border and it traverses 90 Km in the hinterland of the study area.

2.2. Data collection

The strategy, procedure and instruments used to collect data for this study were as follows.

2.2.1. Data sample size and data collection strategy

Morning and afternoon runs were conducted each day on particular distribution mains. On each run, data was collected at sampling

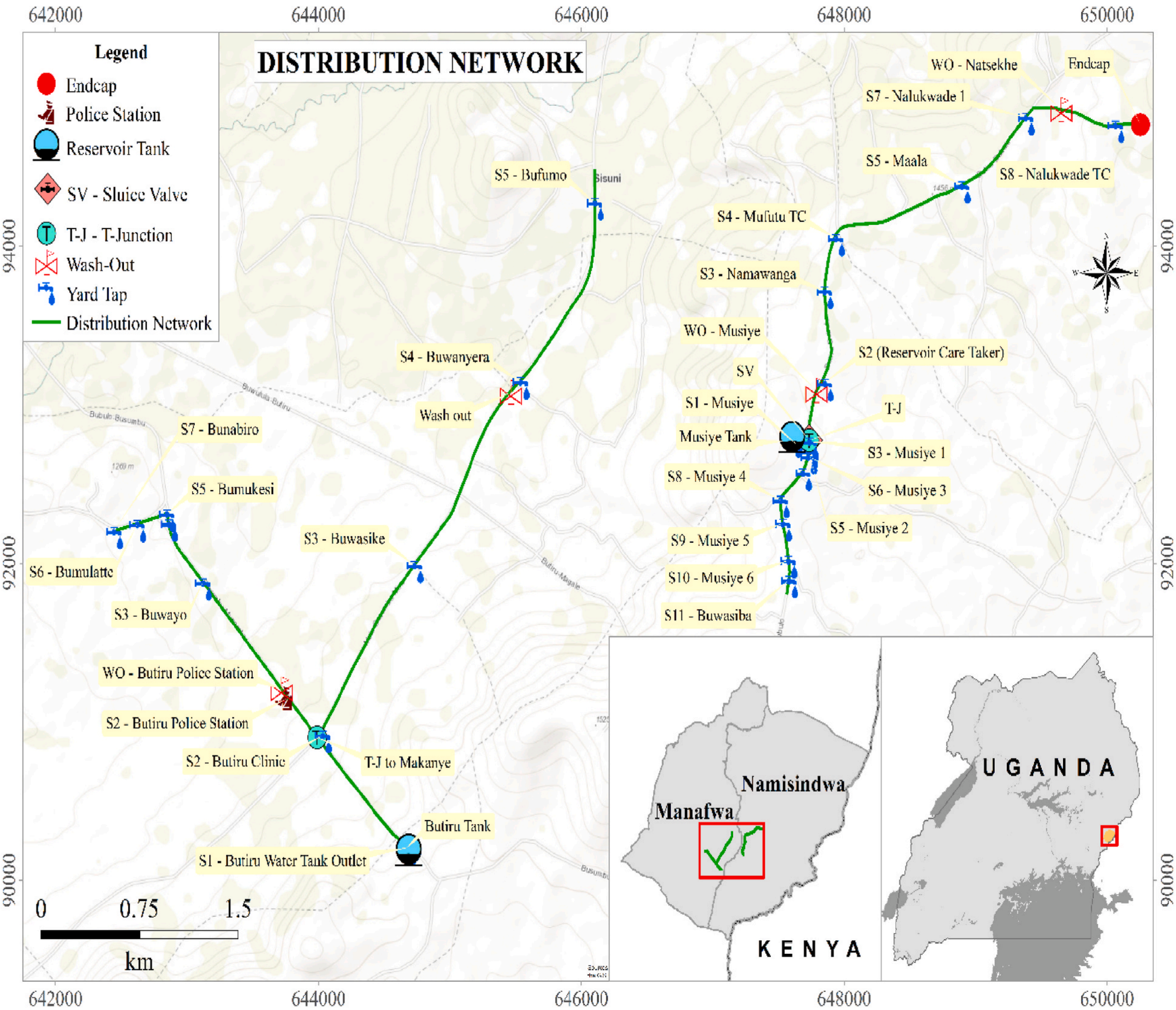


Fig. 2. Water demand draw-off points (yard taps) on Lirima gravity water scheme.

points at approximate intervals of 800 m-1000 m. This spacing interval was based on the advice of NWSC that closer intervals than this may not reveal significant variations in residual chlorine concentrations. This water sampling interval was based on the low initial chlorine dose of 0.73–1.00 mg/l at water treatment plant in order to minimize chlorination cost and also minimize formation of carcinogenic DBPs (disinfection by-products) that are associated with high chlorine dosage. Data collection was replicated on different days to simulate variations in study data. Replication of data on different days was also a strategy to increase sample size of study data. A total of 128 datasets were collected.

2.2.2. Data collection instruments and testing procedure

Water was sampled at clear water reservoir and break-pressure tank outlets, wash outs and nearest functional yard taps that were on direct supply lines from water distribution and transmission mains. The yard water taps from which water was sampled were those that were very close to distribution mains within off-sets of less than 5 m as shown in Fig. 2. Horizontal distances and altitudes of these physical infrastructure components i.e. clear water reservoirs, break-pressure tank outlets, wash-outs and nearest functional yard taps were captured using GARMIN GPSMAP64s hand-held GPI (Geographic positioning instrument). Internal pipe diameters (pipe bores) were measured using steel tape measures directly at break-pressure outlets when the outlets were empty. The GPS coordinates were used to track the hydraulic paths and gradients of water transmission and distribution pipelines for hydraulic modelling in EPANET. It was assumed that water quality parameters at yard taps close to distribution networks would not have varied significantly from the water in the nearby distribution lines. Therefore, water in yard taps was considered to be practically representative of water quality parameter values. Online tests of residual chlorine, turbidity, temperature and electrical conductivity were done on water samples drawn from each water sample point mentioned above. Standard 1 L bottles were used to draw water from break-pressure outlets, wash outs and yard taps. Within seconds of sampling water in standard 1 L bottles, a multifunctional Lovibond MD 600 digital meter was used to measure water quality parameters of residual chlorine in the range of 0–6 mg/l and turbidity (*NTU*), A pH and conductivity 901 digital meter was used to measure online temperature ($^{\circ}C$), pH and electrical conductivity ($\mu S/cm$). Test results of these water quality parameters were recorded in preprepared notebooks designed to record water qualities at geo-referenced positions in water distribution system.

2.2.3. Data analysis

EPANET 2.0 was used to develop hydrologic model from the GPS coordinates picked from water treatment plant, water outlets of break-pressure tanks, online wash-outs and other sections of transmission and distribution lines. EPANET 2.0 was further used to develop process model of residual chlorine decay from upstream to downstream points within water transmission and distribution lines. The influence of both water quality and water system parameters were investigated using three triangulated methods of: (1) Pearson's correlation coefficient at 95 % confidence interval (2) decision tree analysis and random forest ensemble importance scores and (3) (a) principal component analysis with Kaiser normalization equamax rotation method with KMO (Kaiser-Meyer-Olkin) measure of >0.5 and Bartlett's sphericity test at 5 % significance level and (b) p-values at 95 % confidence interval and standardized beta coefficients of independent variables in backward elimination in ordinary least squares regression models. Regression models were tested at 95 % confidence interval, multicollinearity of independent variables in regression models was tested at variable inflationary factor (VIFs) of less than 5 and Durbin-Watson statistic of 1.26 which falls within the acceptable range of 1–3. Tree-based modules of decision tree and random forest were used for feature importance of both physical and water quality parameters. Python and IBM SPSS V25 softwares were used to analyse correlation of both physical and water quality parameters with residual chlorine and also regression of these parameters on residual chlorine decay.

3. Results

We present three results for: (1) correlation between residual chlorine decay parameters with residual chlorine decay, (2) importance of these residual chlorine decay parameters in explaining residual chlorine decay and (3) significance test results of these residual chlorine decay parameters in explaining residual chlorine decay in drinking water distribution system.

Table 1
Descriptive statistics for physical, water quality and hydraulic parameters in water distribution network.

Water quality, physical and hydraulic parameters	count	mean	std	min	25 %	50 %	75 %	max
Residual chlorine (mg/l)	128	0.14	0.07	0.00	0.09	0.14	0.19	0.37
Distance (Km)	128	2.50	2.2	0.01	0.67	1.71	4.60	7.50
travel time (min)	128	46.13	42.63	5.00	15.00	30.00	65.00	190.00
Diameter (mm)	128	108.28	51.92	50.00	80.00	100.00	100.00	250.00
Turbidity (NTU)	128	0.96	0.77	0.00	0.75	1.07	1.07	5.00
Electrical Conductivity (μS/cm)	128	70.01	2.53	65.40	68.38	70.01	70.03	78.50
pH	128	7.53	0.17	6.71	7.48	7.53	7.60	7.83
Temperature (°C)	128	23.98	1.06	20.10	23.59	23.98	24.31	27.05
Pressure (Bar)	128	2.00	1.08	0.00	1.73	2.00	2.00	6.00
Velocity (m/s)	128	0.04	0.02	0.001	0.02	0.04	0.05	0.10

3.1. Univariate and bivariate aanalysis results

Table 1 shows the descriptive statistics for physical and water quality parameters in water distribution network. and Table 2 (correlation matrix) shows how physical and water quality parameters related with residual chlorine in water distribution network.

Table 1 shows that the mean residual chlorine of 0.14 mg/l was below the lower limit of 0.2–0.5 mg/l specified by WHO (2017). The pH that ranged from 6.71 to 7.83 were within the acceptable range of 6.5–8.5 as specified by US EAS 12 (Universal Standards of East African Standard 12) and Uganda National Bureau of Standards (UNBS, 2014). At water treatment plant and entry into water distribution system/network, chlorine dosage (initial chlorine concentration) was in the range of 0.73–1.00 mg/l, turbidity was in the range of 0.4–0.99 NTU, electrical conductivity ranged from 95.4 to 151.2 μS/cm, pH ranged from 7.41 to 7.80 and temperature ranged from 18.7 to 23.7 °C. It is evident from comparison of the range of values of each water quality parameter in water distribution network as in Table 1 against the corresponding range of values for each water quality parameter at water treatment plant and entry into water distribution system/network that there were reductions. These reductions show that conventional water treatment does not eliminate impurities in water fully. Traces of water impurities exert chlorine demand that consume residual chlorine (initial chlorine) as water is conveyed downstream to consumers.

3.2. Results of feature importance of residual chlorine decay parameters

Fig. 3 (a) and (b) show decision tree and random forest analysis results of importance of physical and water quality parameters in influencing residual chlorine decay in water transmission and distribution lines.

For both decision tree and random forest, initial chlorine and electrical conductivity both of which are water quality residual chlorine decay parameters influenced residual chlorine most. For the third most influential parameter, it was diameter (in decision tree) and distance in (in random forest). Diameter and distance are physical system residual chlorine decay parameters.

3.3. Statistical significance test results of residual chlorine decay parameters

Statistical importance of residual chlorine decay parameters was investigated by principal component analysis and multiple linear regression models. Results for each of these are presented as follows.

3.3.1. Principal component analysis results

Fig. 4 (scree plot) shows the eigenvalues that measures the variance of each principal component (PC).

Fig. 4 of the scree plot shows that all of the first three sharp curves are well above Kaiser minimum threshold criterion eigen value of one. This means that the first three principal components explain total variation substantially. From the fourth to tenth principal components, the gradient of the curve is less steep as the curve approaches the horizontal asymptotically indicating minor contribution to total variation in residual chlorine.

Table 3 shows the three principal component solution analysis results for residual chlorine decay parameters in water distribution system.

Table 3 shows that the KMO (Kaiser-Meyer-Olkin) Measure was 0.51 and Bartlett's test of sphericity (level of significance) was 0.005. Therefore, the three-principal component (PC) solution with principal components, PC 1 (initial chlorine), PC 2 (distance) and PC 3 (electrical conductivity) satisfy the two tests of KMO (Kaiser-Meyer-Olkin) Measure of >0.5 and Bartlett's test of sphericity (level of significance) of <0.05 required for credible principal component analysis solution.

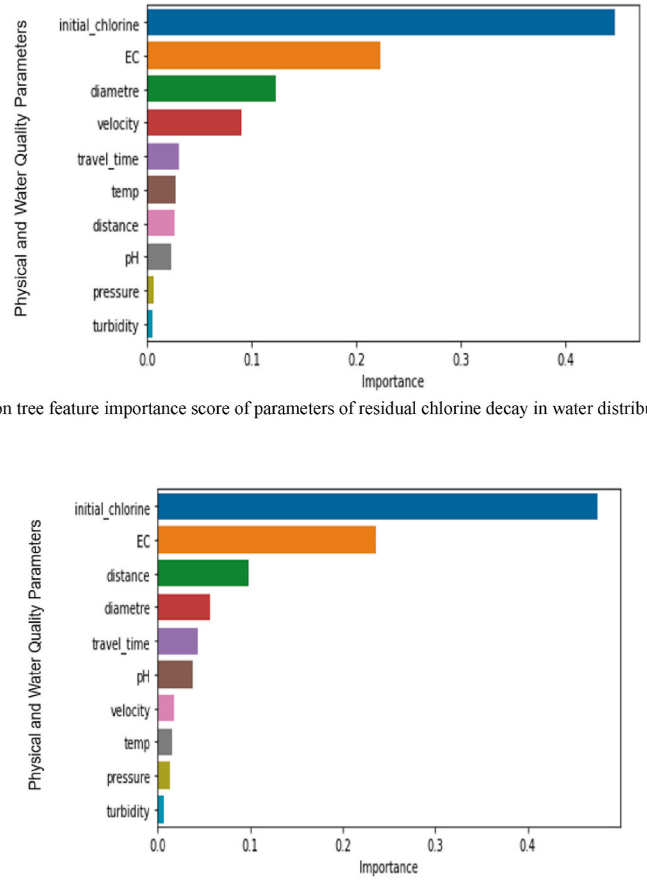
3.3.2. Regression models for statistically significant parameters

A multiple linear regression analysis for all the 10 physical and water quality parameters resulted into four (initial chlorine, electrical conductivity, travel time and pH) that were statistically significant. Distance at p-value of 0.089 was marginally statistically

Table 2
Correlation matrix of chlorine decay parameters.

Parameters	RC	IC	dist	tt	dia	tur	EC	pH	temp	pre	vel
RC (mg/l)	1										
IC (mg/l)	0.69	1									
Dist (Km)	−0.11	0.30	1								
tt (min)	−0.08	0.31	0.71	1							
Dia (mm)	−0.09	0.20	0.63	0.52	1						
tur (NTU)	−0.02	−0.07	−0.03	0.05	0.08	1					
EC (μSiem^{-1})	−0.20	−0.05	−0.05	−0.06	−0.10	−0.29	1				
pH	0.15	0.11	0.14	0.17	0.12	0.03	−0.11	1			
temp (°C)	−0.21	−0.21	−0.05	−0.002	0.05	−0.10	0.21	−0.06	1		
pres (Bar)	0.03	0.19	0.31	0.17	0.02	−0.11	−0.15	0.17	0.23	1	
vel (m/s)	−0.17	−0.07	−0.03	−0.004	−0.06	−0.01	−0.15	0.20	−0.001	0.32	1

Legend.
RC = residual chlorine, IC = initial chlorine, dist = distance, tur = turbidity, EC = electrical conductivity, temp = temperature, pre = pressure, vel = velocity, tt = travel time, dia = diameter.



(a) Decision tree feature importance score of parameters of residual chlorine decay in water distribution system

(b) Random forest feature importance score of parameters of residual chlorine decay in water distribution system

Fig. 3. Tree-based importance of parameters of residual chlorine decay in water distribution system.

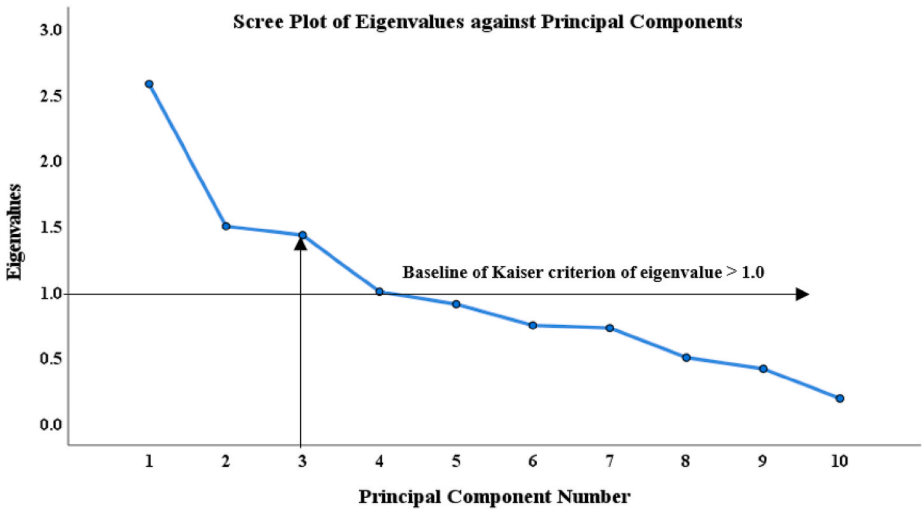


Fig. 4. Scree plot of eigenvalues against principal components.

insignificant. However, as shown in correlation matrix in Table 2, distance had a stronger Pearson's correlation coefficient $r = 0.11$ than travelling time with Pearson's correlation coefficient $r = 0.08$ in relation to final residual chlorine. However, distance and travelling time were strongly correlated at Pearson's correlation coefficient $r = 0.71$. This means both distance and travelling time

Table 3
Three principal component analysis statistics.

Item	No. of Principal Components (PCs)	KMO measure of sampling adequacy	Bartlett's test of sphericity	PC1 (%)	PC2 (%)	PC3 (%)	PC4 (%)	Explained total variance (%)
1	2	0.510	0.005	44.06	32.77	NA	NA	76.83
2	3	0.510	0.005	44.06	32.77	23.17	NA	100.00
3	4	0.547	0.000	23.15	25.26	14.45	12.28	65.14

should not be used together in regression analysis for residual chlorine decay. Therefore, elimination of travelling time that correlated weakly with final residual chlorine was necessary to reduce multicollinearity between distance and travel time. Applying this approach to all the original 10 parameters resulted into three statistically significant parameters as in Table 4.

3.3.2.1. Ordinary least squares multi-linear regression model. Ordinary least squares regression model based on initial chlorine, electrical conductivity and distance as the three most statistically independent and statistically significant predictors for residual chlorine decay in water distribution is summarized in Table 4.

The multicollinearity for all independent variables were low ranging from 1.005 to 1.105 and their *p-values* were also all below 0.05 for statistical significance at 5 % level of significance. The equation for final residual chlorine in water distribution network based on these statistically independent and statistically significant water quality and water distribution system parameters is as show1 in equation (1):

$$\textit{Final chlorine} = 0.415 + 0.548\ \textit{Initial chlorine} - 0.012\ \textit{distance} - 0.005\ \textit{EC}. \ldots\ldots \tag{1}$$

where final chlorine and initial chlorine are measured in mg/l, distance is measured in Km and EC (electrical conductivity) is measured in μS.

Table 4 further shows that in addition to the variable inflationary factor ranging from 1.005 to 1.105, the Durbin-Watson statistic was 1.262. The low VIFs (Variable Inflation Factors) all below 5 and the Durbin-Watson statistic of 1.26 which falls within the acceptable range of 1–3 all show low and acceptable (tolerable) multi-collinearity between these three parameters. This means that these three statistically significant parameters explain residual chlorine decay in water distribution system well.

3.3.2.2. Principal component analysis multi-linear regression model. The result for linear regression with principal components as predictors of residual chlorine decay in water distribution is summarized in Table 5.

The resulting principal components' linear regression model from the three principal components in Table 5 as predictors for final residual chlorine in water reticulation is as shown in equation (2).

$$\textit{Resdual chlorine} = 0.144 - 0.014\ \textit{EC} - 0.017\ \textit{length} + 0.053\ \textit{chlorine}_{dose} \ldots\ldots \tag{2}$$

3.3.3. Influence of physical and water quality parameters on residual chlorine decay

Table 6 shows the influence of the two broad categories of physical (diameter and distance) and water quality (initial chlorine, electrical conductivity, turbidity, pH and temperature) parameters on residual chlorine decay in water distribution system. Hydraulic parameters (travel time, pressure and velocity) are also included.

Table 6 shows that water quality contributes more between 87% and 93 % of residual chlorine decay than influence of physical water infrastructure at between 7% and 13 % of residual chlorine decay in water system contribution.

Table 4
Ordinary Least Squares linear regression model for statistically significant variables.

(a) Model summary

Model	R	R Squared	Adjusted R Squared	Std. Error of the Estimate
2	0.793[c]	0.628	0.619	0.0453

(b) Model coefficients

	Unstandardized coefficients		Standardized coefficients			Collinearity statistics	
Predictors	B	Std. Error	Beta	t	Sig.	Tolerance	VIF
Constant	0.415	0.112		3.692	0.0001		
1 initial chlorine	0.548	0.040	0.795	13.761	0.0001	0.906	1.104
2 distance	−0.012	0.002	−0.365	−6.250	0.0001	0.907	1.105
3 EC	−0.005	0.002	−0.175	−3.182	0.0020	0.995	1.005

Durbin-Watson = 1.262.
Dependent Variable: final chlorine.
EC = electrical conductivity.
Predictors: (1) Constant, (2) initial chlorine, (3) distance, (4) EC.

Table 5
Details of principal component analysis-based regression model.

(a) Model summary

Model	R	R squared	Adjusted R squared	Std. error of estimate
1	0.788[b]	0.620	0.611	0.04564

(b) Coefficients

Model	Unstandardized coefficients		Standardized coefficients	t	Sig.	Collinearity statistics	
	B	Std. error	Beta			Tolerance	VIF
(Constant)	0.144	0.004		35.597	0.001		
electrical_ conductivity	−0.014	0.004	−0.189	−3.423	0.001	1.000	1.000
distance	−0.017	0.004	−0.230	−4.155	0.001	1.000	1.000
chlorine dose	0.053	0.004	0.729	13.171	0.001	1.000	1.000

Legend: B = Unstandardized beta coefficient, t = test statistic = (B/std.error), VIF = Variable Inflationary Factor.
Dependent Variable: residual chlorine.
Predictors: (Constant), chlorine dose, length, electrical conductivity.

Table 6
Importance of physical and water quality parameters on residual chlorine decay in water distribution.

Item	Independent variable	Decision tree (Score)	Random forest (Score)	PCA (Loading)	OLS regression p-value, Standardized Beta coefficient
1	Initial chlorine	0.437	0.470	0.825	0.000, 0.816
2	distance	0.019	0.093	0.817	0.089, - 0.162
3	travel time	0.030	0.041	0.734	0.001, - 0.283
4	diameter	0.119	0.052	0.705	0.700, - 0.028
5	turbidity	0.004	0.007	0.639	0.849, - 0. 001
6	electrical conductivity	0.217	0.226	0.439	0.001, - 0. 019
7	pH	0.015	0.030	0.312	0.004, 0.160
8	temperature	0.022	0.013	0.748	0.873, 0.010
9	pressure	0.006	0.056	0.693	0.991, - 0.001
10	velocity	0.089	0.015	0.602	0.007, - 0.155

Legend: PCA = Principal Component Analysis, OLS = Ordinary Least Square.

4. Discussion

Results from section 3.1 and Table 2 (correlation), section 3.2 and Fig. 3 (a) (b) (feature importance) and section 3.3 (statistical significance) for residual chlorine decay parameters are discussed under each of the following sections.

4.1. *Correlation of residual chlorine parameters with residual chlorine decay*

This section discusses the relationship and control of physical and water quality parameters with residual chlorine decay in water distribution system.

4.1.1. *Correlation of physical and water quality parameters with residual chlorine decay*

Distance in the presence of travel time had p-value of 0.089 and diameter had p-value of 0.700 as physical parameters of residual chlorine decay in water transmission and distribution lines. However, Table 2 (correlation matrix) showed strong correlation between distance and travel time with Pearson's correlation coefficient $r = 0.71$. However, the correlation between distance and residual chlorine had Pearson's correlation coefficient $r =$ - 0.11 and that between travel time and residual chlorine had Pearson's correlation coefficient $r =$ - 0.08. This shows that distance that is better correlated with residual chlorine is a better predictor of residual chlorine decay than travel time after dropping travel time. This is illustrated in Table 2 (correlation matrix) in which distance with p-value of 0.000 was statistically significant as a physical parameter in explaining residual chlorine decay in water transmission and distribution lines.

Flow velocity correlated weakly with residual chlorine with Pearson's correlation coefficient $r = 0.17$ that is below the minimum threshold of 0.4 for consideration of covariance. This could have been due to the low velocities that ranged from 0.001 to 0.10 m/s as in Table 1 (descriptive statistics). This is similar to the velocities in the study of [27] that ranged from 0.0 to 0.12 m/s with Reynold's number <2000. The laminar flow could have resulted in steady flow because of little or no valve modulation and changes in pumping action in this gravity flow. For velocity to vary significantly to affect residual chlorine decay, there should be pressure variation and changes in pumping action [9]. This laminar flow ensured bulk chlorine decay reaction characterized by little or no chlorine mass transfer to pipe wall. Therefore, wash-off of wall organic biofilm and exposure of wall bacteria that are active chlorine reactants was most probably minimal [27]. Consequently, there was insignificant and negligible correlation between velocity and residual chlorine decay.

Temperature as measure of thermal energy influences chemical reactions [18,26] and affects all water quality processes [25,26]. However, there was little variation in temperature over the study area and during study time. The maximum temperature was ≈27 °C, minimum temperature was ≈20 °C and mean temperature was ≈24 °C over a small range of only ≈7 °C as shown in Table 1 (descriptive statistics). Therefore, the insignificant variation in temperature means that there was inadequate thermal energy variation to influence residual chlorine decay that is a temperature-dependent process [8]. This agrees with the view that for an independent variable (in this case temperature) to influence a dependent variable (in this case residual chlorine decay) in a stimulus-response type relation [28], there should be adequate and noticeable variability in the independent variable [28–30].

Turbidity and pH were also less influential. Similarly, the insignificant influence of pH that is temperature-dependent [18] could have been because of its small variance as shown by the small range of 1.12 in Table 1 (descriptive statistics). This result is consistent with finding of [12] that pH is not influential in residual chlorine decay.

4.1.2. Control of physical and water quality parameters in residual chlorine decay

Initial chlorine and electrical conductivity are controllable factors. These two water quality parameters can be controlled at water treatment plant and within water distribution network. In contrast, distance which is a physical water system parameter is hard to control in existing and operational water distribution system. Distance can be controlled and optimized largely during project planning, design and construction and controlled minimally during water infrastructure maintenance. It is hard to alter distance of constructed and commissioned water distribution scheme because of service disruption, reconstruction costs and related inconveniences. Therefore, water quality parameters that contribute as much as 67 % (two-thirds) of residual chlorine decay should be controlled more during water distribution operation. For a proposed new project, distance that is hard to control after construction but is responsible for a significant 33 % (one-third) of residual chlorine decay should be optimized by careful physical layout alternatives for the shortest possible alignment(s).

Both ordinary least squares regression and principal component analyses showed practically same results for these three statistically significant predictors that explain residual chlorine decay in water distribution, This consistence of results triangulated by regression and principal component analyses confirm the importance of initial chlorine, electrical conductivity and distance as the dominant parameters that influence residual chlorine decay in gravity water distribution. The consistency of regression and principal component analyses in identifying the same three influential parameters of residual chlorine decay in water distribution system emphasizes the need and urgency to review and refocus on water treatment strategy in treated water supply practice.

4.2. Importance of parameters in residual chlorine decay

In Fig. 4 (a) and (b), both decision tree and random forest respectively had initial chlorine and electrical conductivity as the first and second respectively most influential parameters of residual chlorine decay in water distribution network. Similarly, turbidity and pressure were the last two least influential parameters in both cases. However, there was difference in the third most influential parameter as diameter (in decision tree) and distance (in random forest). From Table 2 (correlation matrix) diameter correlated at Person's correlation coefficient $r = 0.09$ with final residual chlorine, less strongly than distance that had Person's correlation coefficient $r = 0.11$ with final residual chlorine. This means that decision tree was poor in ranking some of the parameters. In contrast, random forest was consistent with correlation matrix in relating diameter and distance with residual chlorine decay. This could have been because a single decision tree is prone to more randomness than random forest in prediction of outcomes. This result confirms the superiority of random forests over decision trees because random forests as ensembles of individual decision trees are robust to random errors than decision trees [29] especially for large data trees [29].

4.3. Statistical significance test results of residual chlorine decay parameters

We discuss in this section the statistical significance and dimensional reduction for ease of interpretation of parameters of residual chlorine decay in water distribution system.

4.3.1. Influence of physical and water quality parameters in residual chlorine decay

Table 3 also shows that three principal components explain practically all the total variation in residual chlorine decay. Of these three principal components, PC 1 (initial chlorine) and PC 3 (electrical conductivity) were water quality parameters which combined to explain approximately (44 % + 23 %) = 67 % (i.e. about two-thirds of total residual chlorine decay). PC 2 (distance) is a physical water system parameter that explained approximately 33 % (i.e. about one-third of total residual chlorine decay). This conforms to the descriptive and interpretative purpose of principal component analysis in classifying independent variables in terms of their relative importance in influencing dependent variables [30]. The disaggregation of residual chlorine parameters into highly influential and less influential parameters through sensitivity analysis is also consistent with the approach of [12] in modelling chlorine decay.

For inferential interpretation that depends on the actual relationships between water quality parameters and residual chlorine decay, Table 4 shows that majority influence of water quality parameters (initial chlorine and electrical conductivity) of between 87% and 93 % of residual chlorine decay requires more control of water quality during water treatment. The 7%–13 % influence of distance which is a physical parameter that is a one-off facility [31]. This means that the design, construction, rehabilitation and replacement of water infrastructure asset like water pipeline should be well optimized to minimize residual chlorine loss.

4.3.2. Dynamics between disinfectants and disinfectant reactants in water treatment

Inorganics are fast chlorine reactants compared to organics that are slow chlorine reactants [32]. However, chlorine demand for inorganics is smaller than that for organics [33] because up to 80 % of total suspended particles are organics, bacteria and pathogens [34]. Therefore, c-t (concentration-time) during contact time in water treatment removes mainly fast and fewer inorganic compounds. After chlorine oxidation of faster and fewer inorganic contaminants, oxidation of the slower and dominant organics starts. However, wide range of reactants remain in different waters even after treatment [35]. This means that at end of the common 30-min contact time, there are more organic contaminants that remain than inorganics to enter water distribution network. Therefore, chloramine which is a more stable, more persistent, less reactive and produces less carcinogenic DBPs (disinfection by-products) than chlorine is a better secondary disinfectant in water distribution networks [36]. Therefore, chloramine should be the preferred disinfectant to oxidize remnants of slow organic pollutants in drinking water in transit to water consumers.

4.3.3. Communalities for principal components of residual chlorine decay parameters

Although the extracted communalities for the three principal component solution were a perfect 1 (i.e., 100 %), in reality this is not the case. This is because PCA solution assumes total variance explained by indicator variables. However, variance attributed to specific and error sources invariably occurs. Besides, PCA also assumes orthogonality in transforming correlated variables into linearly uncorrelated variables [32,37,38]. Smaller sets of high variance variables are usually hypothesized as uncorrelated [39] although some scholars like [39] consider all original variables as mutually correlated before transformation into principal components. This may be because of high chances of multicollinearity in large dimensional multivariate data. Multicollinearity complicates analysis for causal inference between variables [38]. Therefore, the extracted communalities are near perfect instead of outright perfect. This is evidenced by Table 3 for the rotated component matrix of the three principal component solution that had near perfect component loadings of 0.999 (99.9 %) for electrical conductivity, 0.988 (98.8 %) for distance and 0.988 (98.8 %) for initial chlorine. This supports the interpretation of perfect extracted communalities for electrical conductivity, distance and initial chlorine as near perfect extracted communalities for residual chlorine decay.

4.3.4. Parsimony of residual chlorine decay parameters

The three principal component solution in this study achieved parsimony by data dimensionality compression from the original multivariate dataset of 10 variables into three easily interpretable variables that explain almost all variability. This enhances interpretation of total variance through graphing and visualization in 3D (three-dimensional) space as advanced by Ref. [40]. Parsimony in this study was achieved by the three key variables of (1) initial chlorine with 44 % feature importance, (2) electrical conductivity with 23 % feature importance and (3) distance from upstream water distribution pipeline point to consumption point with 33 % feature importance. Together, these three key variables (initial chlorine, electrical conductivity and distance) explain most of residual chlorine decay in water distribution network.

5. Conclusions

This study investigated the influence of physical and water quality parameters on residual chlorine decay in gravity water flow scheme. Variation of residual chlorine was tracked from entry of treated water into gravity flow distribution system starting at water treatment plant to water consumption points downstream. All three methods of random forest, principal component analysis and multi-linear regression showed that initial chlorine, electrical conductivity and distance were the three parameters that influence residual chlorine decay most in gravity water distribution systems. Water quality parameters (initial chlorine and electrical conductivity) which are controllable factors dominate residual chlorine decay with 67 % contribution. Distance which is controllable only during design and construction but uncontrollable after construction of a water supply scheme contributes residual chlorine decay at 33 % contribution. Initial chlorine, electrical conductivity and distance together explain 62 % of residual chlorine decay with of estimated error of 0.045 mg/l in gravity water distribution network. After breakpoint chlorination, fast chlorine reactants which are mainly dissolved salts should be checked and reduced below their allowed maximum limits to ensure that electrical conductivity of treated water that significantly reduces residual chlorine is maintained at minimum during water distribution. Future research on analysing water quality after breakpoint chlorination at water treatment plants to assess water quality parameter profile is needed. This is important for identifying which controllable water quality parameters remain in treated water *after* primary treatment at water treatment plant *before* entry into drinking water distribution network. This calls for two-stage analysis in water treatment of: (a) conventional water treatment at water treatment plant and (b) pipeline distribution/transmission water quality analyses in drinking water supply practice. However, the application of the results of this study is limited to HDPE (high density polyethylene) gravity water distribution systems. This is because gravity flow schemes have different hydraulic pressure and velocity distributions compared to pressurised water supply systems. Secondly, pipe material and age which are physical parameters of residual chlorine were not included because the pipelines of this studied gravity flow scheme were all HPDE and constructed around the same time. Lastly, the study was carried out during the months of February and March which was a dry spell. Temperature which is a water quality parameter varies seasonally. Therefore, to apply the results all year round even on gravity flow schemes, there is need to carry out similar study during rain season.

Funding

This study was financed through an award of the third competitive research grant from Kyambogo University, Uganda, under the support from the Government of Uganda.

Data availability statement

Data used in this study will be made available on formal request to the corresponding author.

CRediT authorship contribution statement

Julius Caesar Kwio-Tamale: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Charles Onyutha:** Writing – review & editing, Visualization, Supervision, Resources, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

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.

Acknowledgment

We are grateful to National Water and Sewerage Corporation for their kind permission and staff support in conducting this study on their Lirima Gravity Flow Scheme. The authors acknowledge that this paper was partially based on the first author's M.Sc. dissertation "Comparison of physical and statistical models in predicting space-time decay of residual chlorine in water distribution system" submitted in 2022 for award of the M.Sc. in Water and Sanitation Engineering of Kyambogo University.

References

[1] M.A. Bensoltane, L. Zeghadia, L. Djemili, A. Gheid, Y. Djebbar, Enhancement of the free residual chlorine concentration at the ends of the water supply network: case study of Souk Ahras City-Algeria, J. Water Land Dev. 38 (VII-IX) (2018) 3–9, https://doi.org/10.2478/jwld-2018-0036.
[2] A.T. Tiruneh, T.Y. Debessai, G.C. Bwembya, S.J. Nkambule, L. Zwane, Variable chlorine decay rate modelling of the Matsapha town water network using EPANET program, J. Water Resour. Protect. 11 (2019) 37–52, https://doi.org/10.4236/jwarp.2019.111003.
[3] A.T. Sharif, A. Farahat, H. Haider, M.A. Al-Zahrrani, M.J. Rodriguez, R. Sadiq, Risk-based framework for optimizing residual chlorine in large water distribution systems, Environ. Monit. Assess. 189 (307) (2017), https://doi.org/10.1007/s10661-017-5989-0.
[4] World Health Organization, Principles and Practices of Drinking-Water Chlorination: a Guide to Strengthening Chlorination Practices in Small-To Medium Sized Water Supplies, 2017. https://creativecommons.org/licenses/by-nc-sa/3.0/igo.
[5] A. Rajasingham, B. Harvey, Y. Taye, S. Kamwaga, A. Martinsen, M. Sirad, M. Aden, K. Gallagher, H. Thomas, Improved chlorination and rapid water quality assessment in response to an outbreak of acute watery diarrhea in Somali region, Ethiopia, J. Water, Sanit. Hyg. Dev. 10 (3) (2020) 596–602, https://doi.org/10.2166/washdev.2020.146.
[6] T. Bereskie, J.R. Manuel, R. Sadiq, Drinking water management and governance in Canada: an innovative plan-do-check-act framework for a safe drinking water supply, Environ. Manag. 60 (2017) 243–262, https://doi.org/10.1007/s00267-017-0873-9.
[7] V. Torretta, A.K. Tolkou, I.A. Katsoyoyiannis, A. Katsoyoyiannis, E. Trulli, E. Magaril, E.C. Rada, Consumption of free chlorine in an aqueduct with low protection: case study of the new aqueduct Simbrivio-Castelli (NASC), Italy, Water 2018 MDPI 10 (127) (2019), https://doi.org/10.3390/w10020127www.mdpi.com/journal/water.
[8] F. Garcia-Avila, A.A. Anazco, J. Ordonez-J, C. Guanuchi-Quezada, L. Flores del Pino, L. Ramos-Fernandez, Modeling of residual chlorine in a drinking water network in times of pandemic of the SARS-COV-2 (COVID-19), Sustainable Environmental Research 31 (12) (2021), https://doi.org/10.1186/s42834-021-00084-w, 2021.
[9] K. Hyunjun, K. Sanghyun, Evaluation of chlorine decay models under transient conditions in a water distribution system, Curr. Sci. 3 (8) (2017) 522–537, https://doi.org/10.2166/hydro.2017.082.
[10] C. Zhang, C. Li, X. Zheng, J. Zhao, G. He, T. Zhang, Effect of pipe materials on chlorine decay, trihalomethanes formation and bacterial communities in pilot-scale water distribution systems, Int. J. Environ. Sci. Technol. 14 (2016) 84–94, https://doi.org/10.1007/s13762-016-1104-2.
[11] E. Chirwa, V. Madzivahandil, Modelling chlorine decay in water distribution systems using Aquasim, Chemical Engineering Transactions (CET) 57 (2017) aidiic.it/cet.
[12] C.G.A. Cuesta, L. Tuhovcak, M. Taus, Using artificial neural networks to assess water quality in water distribution networks, in: 12th International Conference on Computing and Control for the Water Industry, CCWI2013, vol. 70, 2014, pp. 399–408, https://doi.org/10.1016/j.proeng.2014.02.045.
[13] M.S. Gibbs, N. Morgan, H.R. Maier, G.C. Dandy, M. Holmes, J.B. Nixon, Use of Artificial Neural Networks for Modelling Chlorine Residuals in Water Distribution Networks, United Water International Pty Ltd, Adelaide-Australia, 2019. https://www.mssanz.org.au.
[14] A.T. Tiruneh, T.Y. Debessai, G.C. Bwembya, S.J. Nkambule, A mathematical model for variable chlorine decay rates in water distribution systems, Hindawi Modelling and Simulation in Engineering (2019), https://doi.org/10.1155/2019/5863905, 2019.
[15] T.F. Vargas, C.C. Baia, T.L. Machado, C.C. Dorea, W.R. Bastos, Decay of free residual chlorine in wells water of northern Brazil, Water 2021 13 (7) (2021) 992, https://doi.org/10.3390/w13070992.
[16] H. Wu, C.C. Dorea, Towards a predictive model for initial chlorine dose in humanitarian emergencies, Water 12 (2020) 1506, https://doi.org/10.3390/w12051506, 2020.
[17] S. Tiwari, R. Babbar, G. Kaur, Performance evaluation of two ANFIS models for predicting water quality index of River Satluj (India), Hindawi Advances in Civil in Engineering (2018), https://doi.org/10.1155/2018/8971079, 2018.
[18] Environmental Protection Agency, *Online Water Quality Monitoring Systems in Distribution Systems: for Water Quality Surveillance and Response Systems*, MC 140, Office of Water, Cincinnati, USA, 2018.
[19] G.J. Bowden, J.B. Nixon, G.C. Dandy, H.R. Maier, M. Holmes, Forecasting Chlorine Residuals in a Water Distribution System Using Generalized Neural Network, 2019, 10.1.1.519.8859.pdf.
[20] S.A. Baig, Z. Lou, M.A. Baig, M. Qasim, D.F. Shams, Q. Mahmood, X. Xu, Assessment of tap water quality and corrosion scales from the selected distribution systems in northern Pakistan, Environ. Monit. Assess. 189 (194) (2017), https://doi.org/10.1007/s10661-017-5907-5.

[21] A. Branz, M. Levine, L. Lehmann, A. Bastable, S.I. Ali, K. Kadir, T. Yates, D. Bloom, D. Latagne, Chlorination of drinking water in emergencies: a review of knowledge to develop recommendations and implementation for research needed, Waterlines 36 (1) (2017), 10.336/1756-333488.2017.002.

[22] S. Geetha, S. Gouthami, Internet of things enabled real time water quality monitoring system, Smart Water 2 (1) (2017) 1–19, https://doi.org/10.1186//s40713-0005-y.

[23] L. Monteiro, D. Figueiredo, D. Covas, J. Menaia, Integrating water temperature in chlorine decay modelling: a case study, Urban Water J. 19 (4) (2017) 1–5, https://doi.org/10.1080/1573062X.2017.1363249.

[24] D. Nono, P.T. Odirile, I. Basupi, P.T. Parida, Assessment of probable causes of chlorine decay in water distribution systems of Gaborone city-Botswana, WaterSA 45 (2) (2019) 190–198, https://doi.org/10.4314/wsa.v45i2.05.

[25] D.P. Loucks, E. van Beek, Water Resource Systems Management and Planning: an Introduction to Methods, Models and Applications, second ed., Springer International Publishing AG, Cham, Switzerland, 2017 https://doi.org/10.1007/978-3-319-44234-1.

[26] L. Zlatanovic, A. Moerman, J.P. van der Hoek, J. Vreeburg, M. Blokker, Development and validation of drinking water temperature model in domestic drinking water supply systems, Urban Water J. 14 (10) (2017) 1031–1037. https://doi:10.1080/1573062X.2017.132550.

[27] P. Jamwal, M.S.M. Kumar, Effect of flow velocity on chlorine decay in a water distribution network: a pilot study, J. Hydroinf. 19 (4) (2016) 1349–1354, https://doi.org/10.18520/cs/v111/i8/1349-1354.

[28] K. Jones, An introduction to statistical modelling, Chemical Engineering Transactions (CET) (2014) 241–255, researchgate.net/publication/256801601.

[29] C.A. Ali, Application of principal component analysis in monitoring educational prospects in mathematics related careers in the upper east region of Ghana, Education and Science Journal of Policy Review and Curriculum (ESJPRCD) 4 (1) (2014) 1741–8763, internationalpolicybrief.org.

[30] I.T. Jollife, J. Cadima, Principal component analysis: a review and recent developments, Phil. Trans. R. Soc. A374 (2016) 1–16, https://doi.org/10.1098/rsta.2015.0202.

[31] Water Research Australia Limited, *Good Practice Guide to the Operation of Drinking Water Supply Systems for the Management of Microbial Risk*, Final Report Project 1074, World Health Organization Regional Office for South-East Asia, Adelaide-Australia, 2015. https://www.waterra.com.au.

[32] S. Karamizadeh, S.M. Abdullah, A.A. Manaf, M. Zamani, A. Hooman, 'An overview of principal component analysis, J. Signal Inf. Process. 4 (2013) 173–175, https://doi.org/10.4236/jsip.2013.43B031. http://www.scirp.org/journal/jsip.

[33] L. Vuta, G.E. Dumitran, 'Some aspects regarding chlorine decay in water distribution networks,' [Online], Aerul si Compaonente ale Mediului (2019) 253–259. researchgate.net>2284….

[34] I. Stoianov, A. Aisopou, 'Chlorine decay under steady and unsteady state hydraulic conditions,', in: 12th International Conference on Computing and Control for the Water Industry, CCW12013, Procedia Engineering, vol. 70, 2014, pp. 1592–1601, https://doi.org/10.1016/j.proeng.2014.02.176, 2014.

[35] I. Fisher, G. Clark Kastl, A. Sathasivan, Evaluation of suitable chlorine bulk-decay models for water distribution systems, Water Res. 45 (2011) 4896–4908, https://doi.org/10.1016/j.watres.2011.06.032, 2011.

[36] H. Ricca, V. Aravinthan, G. Mahinthakumar, Modelling chloramine decay in full-scale drinking water supply systems, Water Environment Federation (2019) 441–454, https://doi.org/10.1002/wer.1046.

[37] A.N. Dar, 'Principal component analysis (PCA): using eigen decomposition, Global Scientific Journals (GSJ) 9 (7) (2021). www.globalscientificjournal.com.

[38] Y. Meng, Application of principal component analysis in teaching and evaluation, Frontiers in Sports Research 1 (1) (2017) 34–42, https://doi.org/10.25236/FSST.080106.

[39] Z. Gniazdowski, New interpretation of principal component analysis, Zeszyty Naukowe WWSI 16 (11) (2017) 43–65.

[40] K. Labid, R. Vemuri, An Application of Principal Component Analysis to the Detection and Visualization of Computer Network Attacks, Department of Applied Science, Faculty of Applied Sciences, University of California, Davis, USA, 2004.