

Article

ARIMA and TFARIMA Analysis of the Main Water Quality Parameters in the Initial Components of a Megacity's Drinking Water Supply System

Carlos Alfonso Zafra-Mejía ^{1,*}, Hugo Alexander Rondón-Quintana ² and Carlos Felipe Urazán-Bonells ³

¹ Grupo de Investigación en Ingeniería Ambiental-GIAUD, Facultad del Medio Ambiente y Recursos Naturales, Universidad Distrital Francisco José de Caldas, Bogotá E-110321, Colombia

² Facultad del Medio Ambiente y Recursos Naturales, Universidad Distrital Francisco José de Caldas, Bogotá E-110321, Colombia; harondonq@udistrital.edu.co

³ Facultad de Ingeniería, Programa de Ingeniería Civil, Universidad Militar Nueva Granada, Campus Cajicá E-250247, Colombia; carlos.urazan@unimilitar.edu.co

* Correspondence: czafra@udistrital.edu.co

Abstract: The objective of this paper is to use autoregressive, integrated, and moving average (ARIMA) and transfer function ARIMA (TFARIMA) models to analyze the behavior of the main water quality parameters in the initial components of a drinking water supply system (DWSS) of a megacity (Bogota, Colombia). The DWSS considered in this study consisted of the following components: a river, a reservoir, and a drinking water treatment plant (WTP). Water quality information was collected daily and over a period of 8 years. A comparative analysis was made between the components of the DWSS based on the structure of the ARIMA and TFARIMA models developed. The results show that the best water quality indicators are the following: turbidity > color > total iron. Increasing the time window of the ARIMA analysis (daily/weekly/monthly) suggests an increase in the magnitude of the AR term for each DWSS component (WTP > river > reservoir). This trend suggests that the turbidity behavior in the WTP is more influenced by past observations compared to the turbidity behavior in the river and reservoir, respectively. Smoothing of the data series (moving average) as the time window of the ARIMA analysis increases leads to a greater sensitivity of the model for outlier detection. TFARIMA models suggest that there is no significant influence of past river turbidity events on turbidity in the reservoir, and of reservoir turbidity on turbidity at the WTP outlet. Turbidity outlier events between the river and reservoir occur mainly in a single observation (additive outliers), and between the reservoir and WTP also have a permanent effect over time (level shift outliers). The AR term of the models is useful for studying the transfer of effects between DWSS components, and the MA term is useful for studying the influence of external factors on water quality in each DWSS component.

Keywords: ARIMA model; ARIMA transfer function model; drinking water supply system; forecast; time series analysis; water quality



Citation: Zafra-Mejía, C.A.; Rondón-Quintana, H.A.; Urazán-Bonells, C.F. ARIMA and TFARIMA Analysis of the Main Water Quality Parameters in the Initial Components of a Megacity's Drinking Water Supply System. *Hydrology* **2024**, *11*, 10. <https://doi.org/10.3390/hydrology11010010>

Academic Editor: Vasilis Kanakoudis

Received: 21 November 2023

Revised: 11 January 2024

Accepted: 13 January 2024

Published: 17 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Water resources are fundamental for life and the development of human activities in megacities [1]. The sustainable management of water resources and drinking water supply systems (DWSSs) is important to meet the economic, social, environmental, and public health demands of urban communities under rapidly growing conditions [2,3]. In recent decades, water scarcity due to climate variability and change and the deterioration of water quality due to the rapid growth of megacities have become some of the main problems of DWSSs [4,5]. Protection and control of water sources is necessary to comply water quality guidelines and reduce drinking water treatment costs [6]. Poor characterization of water quality results in a significant variability of the parameters involved during the operation

of DWSSs. In effect, this makes it difficult to manage DWSSs [7]. Water quality for human consumption is one of the most important determinants of public health in developing countries [8]. Despite the significant progress that has been made, the influence of natural and anthropogenic factors continue to affect DWSSs, creating significant challenges during their operation in urban areas [9]. Natural environmental factors such as water quantity and quality influence public health. Dynamic interplay of natural and anthropogenic factors can negatively influence DWSSs, leading to the emergence of infectious (biological agents) or noninfectious (chemical or physical agents) diseases [10]. This scenario is probably intensified by the particular social and economic conditions of developing countries.

The Latin American megacity where the DWSS under study is located (Bogota, Colombia) is undergoing massive urban expansion processes [11]. Urban expansion shows strong pressures on the territory due to the construction of rural housing and the development of tourism, industrial, agro-industrial, and agricultural projects (e.g., potato crops, pig farms, and cattle ranching) [12]. This scenario generates a high water demand and consequently a high wastewater discharge into the water supply sources [13]. In the sanitation and discharge management plans of the basins, deficiencies are reported in the wastewater treatment generated by the different residential and industrial sectors settled in these areas [14]. There are also difficulties on the part of environmental institutions to implement effective control and surveillance mechanisms to monitor compliance with local water quality guidelines [15]. Moreover, adverse effects on the quality and quantity of water supply sources associated with the El Niño–Southern Oscillation (ENSO) climate variability phenomenon have been reported. On the one hand, there is an increase in water turbidity in water supply sources due to sediment entrainment from increased rainfall and runoff. On the other hand, there is a decrease in the water flow in the supply sources during periods of decreased rainfall [16]. Another influential factor for the DWSS under study is the destruction of vegetation cover in the upper part of the basin, thus disturbing the hydrological cycle and reducing water availability [17].

Anthropic and natural factors influence the operation of DWSSs. This influence generates an increase in the consumption of reagents and hinders the operation of the drinking water treatment plants (WTPs) [18]. Under this scenario, there is an increase in the operating costs of the DWSS in order to comply with water quality guidelines [10]. The organizations in charge of the DWSS operation have visualized new management tools for the monitoring and control of water quality. Autoregressive, integrated, and moving average (ARIMA) models are then presented as a suitable analysis tool for decision making related to the management of DWSSs in developing countries [19]. Multivariate statistical analysis methods (e.g., cluster and principal component analysis) in combination with ARIMA models may provide a comprehensive view to study water quality parameters in each of the main components of a DWSS (water source, reservoir, and WTP) [20]. In addition, the use of ARIMA transfer function (TFARIMA) models is useful in this type of analysis in DWSSs. ARIMA transfer function models are used to model the relationship between two time series [21]. These models are built from an ARIMA model applied to both series and use the transfer function to describe how changes in one series directly affected the other. This approach is useful in situations where there is a known or theoretical relationship between the two series and the aim is to model and understand this relationship for forecasting or analytical purposes (effect of one DWSS component on another) [22].

Time series analysis of water quality parameters using ARIMA models can provide accurate short-term forecasts from a significant amount of information [23]. ARIMA modeling is the combination of three processes: autoregressive (time series memory), differencing (time series trend), and moving average (time series variability) [24]. In time series analysis, ARIMA models are flexible and widely used in the water quality context. For example, [25] adequately modeled the water quality of a river (Johor River) in Malaysia using ARIMA models. These researchers studied the possible relationship between river water quality parameters (pH, color, and turbidity) and hydrological variables of the catchment (rainfall and river flow). The authors of [26] reported that by using

ARIMA models improved the accuracy of water quality prediction (total phosphorus and total nitrogen) in a reservoir by 97.5%. The authors of [27] used ARIMA models to study the daily performance of a DWSS. These researchers used turbidity (ARIMA: 3, 1, 0) and pH (ARIMA: 1, 1, 1) as indicator parameters of WTP performance from the use of ARIMA models. In addition, [28] reported the need to use ARIMA models in combination with multivariate analysis to study in more detail the information of water quality (electrical conductivity, temperature, dissolved oxygen, ammonium, nitrates, pH, and total phosphorus) of the Pinios River (Greece). These researchers were able to visualize decision-making strategies for river pollution control. Multivariate statistical analyses were also used in the analysis of physicochemical water quality information and these techniques were an effective tool for their evaluation [29]. For example, [30] used hierarchical cluster analysis to detect the best water quality indicator parameters in three WTPs in Iraq. Of the 32 parameters analyzed, they detected the following as water quality indicators: Turbidity, electrical conductivity, total alkalinity, total hardness, total coliforms, and fecal coliforms. The authors of [31] used principal component analysis (PCA) and cluster analysis (CA) to assess the water quality (Ph, electrical conductivity, total dissolved solids, hardness, salinity, and alkalinity) of Bannu district (Pakistan). The authors of [32] also used multivariate techniques (PCA and CA) to analyze water quality (alkalinity, calcium, chloride, pH, conductivity, hardness, nitrate, and sulfate) in DWSSs of 164 municipalities in Italy.

The main objective of this paper is to use ARIMA and TFARIMA models to analyze the temporal behavior of the main water quality parameters in the initial components of a DWSS in a Latin American megacity (Bogota, Colombia). The components of the DWSS considered in this study are the following: a natural source of water (Teusacá River), a reservoir (San Rafael), and a WTP (Francisco Wiesner). In the context of the management of DWSSs, this study is relevant for the following practical aspects: (1) the applicability of ARIMA and TFARIMA models to study the performance of DWSSs; (2) the combined use of ARIMA and TFARIMA models, and multivariate statistical methods as an integral decision-making tool for the management of these DWSSs; and (3) to study the occurrence of atypical events (outliers) and their transfer over time on the different components of a DWSS.

2. Materials and Methods

2.1. Research Site

The research site corresponded to the northern DWSS of the megacity of Bogota, Colombia (4°41'23" N—73°59'44" W). The DWSS considered in this study consisted of the Teusacá river, San Rafael reservoir, and Francisco Wiesner treatment plant (Figure 1). The climate around the DWSS was characterized as equatorial mountain (Csbi) or tropical mountain climate, with moderate rainfall and drought. This is according to the Köppen-Geiger climate classification [33]. On average, on an annual basis, the temperature was 12.4 °C (daily oscillation between 8–17 °C) and rainfall was 1520 mm. Due to the influence of the Intertropical Convergence Zone (ITCZ), the annual rainfall regime was bimodal. That is, two seasons of increased rainfall (April–May and October–November) and two seasons of decreased rainfall (January–February and July–August) were evident. The average elevation of the study sites was between 2776–2813 masl.

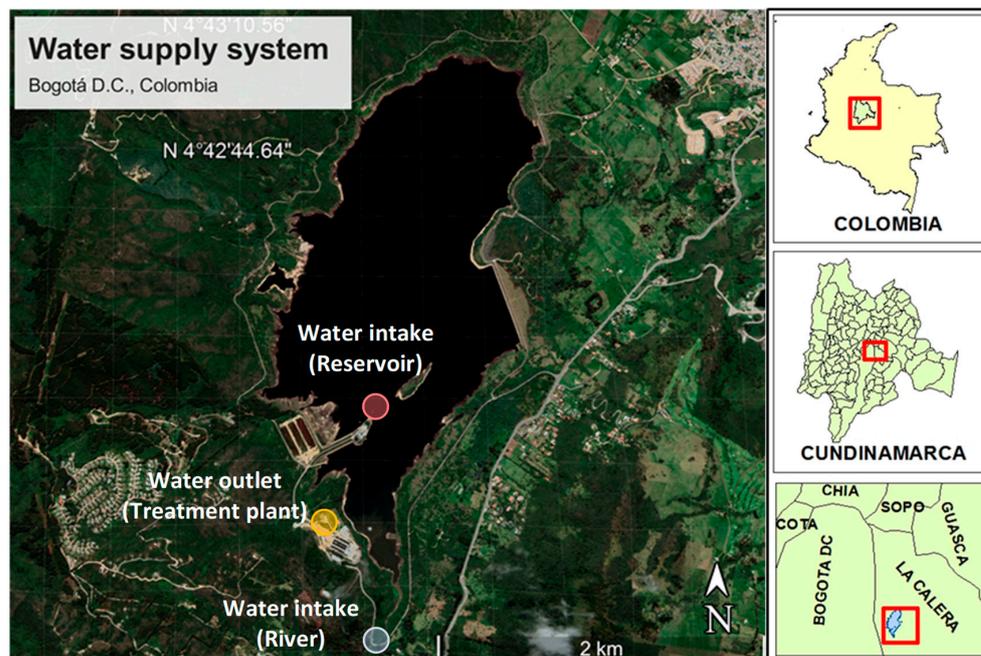


Figure 1. Components of the DWSS under study (river, reservoir, and WTP).

The mean flow from multiperiod (1 January 2008–31 December 2015) of the Teusacá river was $2.55 \text{ m}^3/\text{s}$. This river flow record corresponded to the entry point into the San Rafael reservoir (Figure 1; water intake, river). On average, the water volume contributed by the river corresponded to 5% of the total water volume of the reservoir (75 million m^3). The remaining 95% of the water volume was provided by a transfer from the Chuza reservoir, which was located 33.7 km away. The water intake at the reservoir was located 1.05 km from the river discharge. Francisco Wiesner's treatment plant was a direct filtration plant. In other words, the water was treated without the conventional flocculation and sedimentation processes. The water to be treated was received directly from the San Rafael reservoir from a pumping station made up of four units, each of which had a capacity of $5 \text{ m}^3/\text{s}$. The maximum treatment capacity of the WTP was $14 \text{ m}^3/\text{s}$. Subsequently, the water treated was sent to a disinfection tank (mixed oxidant solution) and contact chamber, which had a volume of $50,000 \text{ m}^3$. Lastly, the water was piped to the megacity through a distribution network. The average water quality characteristics of the different components of the DWSS are shown in Table 1.

Table 1. Water quality parameters in the different components of the DWSS under study.

Parameter	Station		
	Teusacá River	San Rafael Reservoir	WTP
Turbidity (NTU)	44.8	3.83	0.67
Color (CU)	77.4	18.3	3.70
Conductivity (uS/cm)	205	46.3	42.9
pH	7.06	7.06	6.69
Total alkalinity (mg/L CaCO_3)	14.4	15.4	12.2
Chlorides (mg/L Cl^-)	57.1	5.38	4.24
Total iron (mg/L Fe^{+3})	1.85	0.41	0.06
Nitrites (mg/L NO_2)	3.88	0.28	0.01
Nitrates (mg/L NO_3) *	0.75	0.26	0.11
Total hardness (mg/L CaCO_3) *	12.6	17.2	18.9
Dissolved oxygen (mg/L O_2) *	7.12	6.79	N.A.
Sulfates (mg/L SO_4) *	0.48	1.17	5.50
Total coliforms (CFU/100 mL)	66,602	2485	0.00
E. Coli (CFU/100 mL)	19,353	53.6	0.00

Note: sample size per variable = 2922 data (daily information), * = sampled on average once per week, and N.A. = not applicable.

2.2. Data Collection

Three monitoring stations were established for the DWSS: (1) Teusacá river, (2) San Rafael reservoir, and (3) WTP outlet (Figure 1). The water quality parameters considered at each monitoring station were as follows: turbidity, color, conductivity, pH, total alkalinity, chlorides, total iron, nitrites, total coliforms, and E. coli. Water quality information was collected daily and over a period of 8 years. Namely, the sample size for each variable considered was 2922 data. The following water quality parameters were not considered in the time series analysis because they were sampled on average once a week: nitrates, total hardness, dissolved oxygen, and sulfates (Table 1). The collection of water samples at each of the monitoring stations followed the guidelines established by the Standard Methods for the Examination of Water and Wastewater [34]. Water samples from the reservoir were taken at the outlet of the water delivery pipe, inside an incoming hydraulic structure at the WTP. The water came from a pumping station located 42.5 m deep in the reservoir. Water samples were collected in 2 L plastic bottles ($n = 3$), which were previously cleaned and sterilized. The bottles were washed and rinsed several times with water from the sampling point before filling. Samples were taken while the water was flowing at a constant rate. Disposable gloves were also used during sampling. The bottles were filled without leaving air bubbles and closed immediately. Lastly, the bottles were labeled and transported ($5\text{ }^{\circ}\text{C}$) for laboratory analysis [34]. In general terms, the sample collection procedure in the river and at the outlet of the WTP was like that described above. At the latter sampling point there was a valve for taking water samples.

2.3. Data Analysis

Laboratory analysis for each of the water quality parameters considered was performed using the following methods: turbidity (Standard Method 2130), color (Standard Method 2120), conductivity (Standard Method 2510), pH (Standard Method 4500), total alkalinity (Standard Method 2320), chlorides (Standard Method 4500), total iron (Standard Method 3500), nitrites (Standard Method 4500), total coliforms (Standard Method 9222), and E. coli (Standard Method 9223) [34]. All daily time series of water quality parameters were checked for the occurrence of missing data (Table 2). The normal ratio method was used to fill in the missing data [35]. A Kolmogorov–Smirnov test was also applied to evaluate the non-normality of the time series under study ($p\text{-value} > 0.050$) [36]. A principal component analysis (PCA) was also applied to reduce the dimensionality of the set of variables considered. The steps considered during PCA were the following [37]: data standardization, covariance matrix calculation, eigenvector and eigenvalue calculation, principal component selection (95% of total variability), and data transformation. Subsequently, Spearman’s correlation coefficient (r_s) [38] was used to study the association between water quality parameters at each of the monitoring stations. PCA and Spearman’s coefficient were used to detect possible water quality indicator parameters [39] in the DWSS under study. Namely, it was assumed as a hypothesis that the water quality parameters that showed the best significant correlations were those that could be suggested as water quality indicators.

ARIMA and TFARIMA models were developed for each of the parameters identified as possible indicators of water quality in the DWSS. The software used was IBM-SPSS Statistics V.20.0 [40]. The time scales considered for the development of the ARIMA and TFARIMA models were the following: daily, weekly, and monthly. Weekly and monthly time scales were generated from daily information and using 7-day and 30-day moving averages, respectively [41]. This was to use time series with a significant amount of data ($n = 2922$) and to try to develop ARIMA and TFARIMA models with a better statistical fit. During the development of ARIMA and TFARIMA models, the orders p , d , and q were identified using the Expert Modeler tool of the IBM-SPSS Statistics V.20.0 software [40]. This tool followed the stages established by Box–Jenkins for the development of ARIMA and TFARIMA models: identification, parameter estimation (calibration), and verification (validation) [23]. Only additive outliers (occurring in a single observation) were considered for the development of univariate ARIMA models [40]. For the development of TFARIMA models [21], level shift outliers (has a

permanent effect) were also considered in order to be able to develop a statistically adequate model. This is in case it was not possible to develop a TFARIMA model with additive outliers only. In the validation of the ARIMA and TFARIMA models, the following statistics were considered: determination coefficient (R^2 , goodness-of-fit), mean absolute percentage error (MAPE, forecast accuracy), normalized Bayesian information criterion (BIC), and the significance of the Ljung–Box statistic (p -value > 0.050) [42]. Given two ARIMA models, the model with the lower value in the normalized BIC was selected definitively. A lower normalized BIC implied a smaller number of explanatory variables and a better fit of the ARIMA model [43]. The appropriateness of the modeling was assessed by the Ljung–Box statistic. This statistic tests the null hypothesis of no remaining significant autocorrelation in the residuals of the model and provides an indication of whether the model is appropriately specified. A p -value greater than 0.05 indicates that the model is properly specified to describe the correlation information in time series [44].

Table 2. Missing data percentage for each of the water quality parameters according to the monitoring stations considered.

Parameter	Station		
	Teusacá River	San Rafael Reservoir	Teusacá River
	Missing Data (%)	Missing Data (%)	Missing Data (%)
Turbidity	0.14	0.17	0.17
Color	0.55	0.17	0.21
Conductivity	0.21	0.24	0.10
pH	0.21	0.24	0.24
Total alkalinity	0.24	0.31	0.17
Chlorides	4.14	3.15	2.94
Total iron	11.9	12.3	9.55
Nitrites	0.68	2.50	4.07
Total coliforms	13.3	13.4	10.9
E. coli	13.7	14.2	10.9

Note: sample size per variable = 2922 data.

A comparative analysis was made between the components of the DWSS (river, reservoir, and WTP) based on the structure of the ARIMA and TFARIMA models developed. That is, the memory (term ‘ p ’), trend (term ‘ d ’), and variability (term ‘ q ’) reported by the ARIMA models [23], and the numerator, trend, and denominator reported by the TFARIMA models [21] of the time series of that parameter identified as a water quality indicator were analyzed. Lastly, the influence of the rainfall regime on the occurrence of atypical water quality episodes (outliers) in the different components of the water supply system was analyzed.

3. Results and Discussion

3.1. Water Quality Indicators

The PCA results identified three, four, and three principal components for the water quality parameters at the three monitoring stations of the river, reservoir, and WTP, respectively (Figure 2). These principal components explained 57.3%, 53.2%, and 56.4% of the total variability of the data, respectively. The results suggested that this explained variance close to 55% could be related to the stochastic nature and occurrence of outliers in the time series considered [37]. At all three monitoring stations, the association of turbidity, color, and iron in the same principal component was observed. The findings suggested that in order of importance the best water quality indicator in the different DWSS components were the following: turbidity $>$ color $>$ total iron. Thus, we consider turbidity as the main water quality indicator in the three monitoring stations of the DWSS under study. These findings were consistent with those reported by [45] in a pilot plant that treated water from the city of São Paulo (Brazil). Moreover, turbidity was frequently used as a water quality criterion in water sources and treatment processes [46]. In this study, the best observed

association of turbidity with the other water quality parameters at the monitoring stations was with color ($r_s > 0.826$, Table 3). The authors of [47] reported similar results between these two water quality parameters. There are studies that reported other parameters as possible indicators of water quality in water supply systems. For example, under certain conditions, pH was reported as an indicator parameter of water quality [48]. Although pH was an important indicator of water quality in many cases, there are situations where it may not be sufficient or did not provide a complete representation of water quality due to interferences in the measurement (e.g., presence of organic substances, minerals, or detergents) [49].

Table 3. Spearman’s correlation coefficients between the main water quality parameters (river, reservoir, and WTP).

	Turbidity	Color	Conductivity	Total Alkalinity	Chlorides	Total Iron
River						
Turbidity	1.000					
Color	0.929 *	1.000				
Conductivity	−0.777 *	−0.741 *	1.000			
Total alkalinity	−0.770 *	−0.741 *	0.834 *	1.000		
Chlorides	−0.720 *	−0.693 *	0.804 *	0.773 *	1.000	
Total iron	0.698 *	0.693 *	−0.553 *	−0.522 *	−0.509 *	1.000
Reservoir						
Turbidity	1.000					
Color	0.827 *	1.000				
Conductivity	−0.028	−0.065	1.000			
Total alkalinity	−0.036	0.006	0.302 *	1.000		
Chlorides	−0.053	−0.110	0.455 *	−0.146	1.000	
Total iron	0.588 *	0.670 *	−0.022	0.041	−0.015	1.000
WTP						
Turbidity	1.000					
Color	0.523 *	1.000				
Conductivity	−0.072	−0.006	1.000			
Total alkalinity	−0.039	0.116	0.257 *	1.000		
Chlorides	−0.166	−0.023	0.740 *	−0.184	1.000	
Total iron	0.290 *	0.281 *	−0.054	−0.044	−0.021	1.000

Note: * = significant correlations, p -value < 0.010.

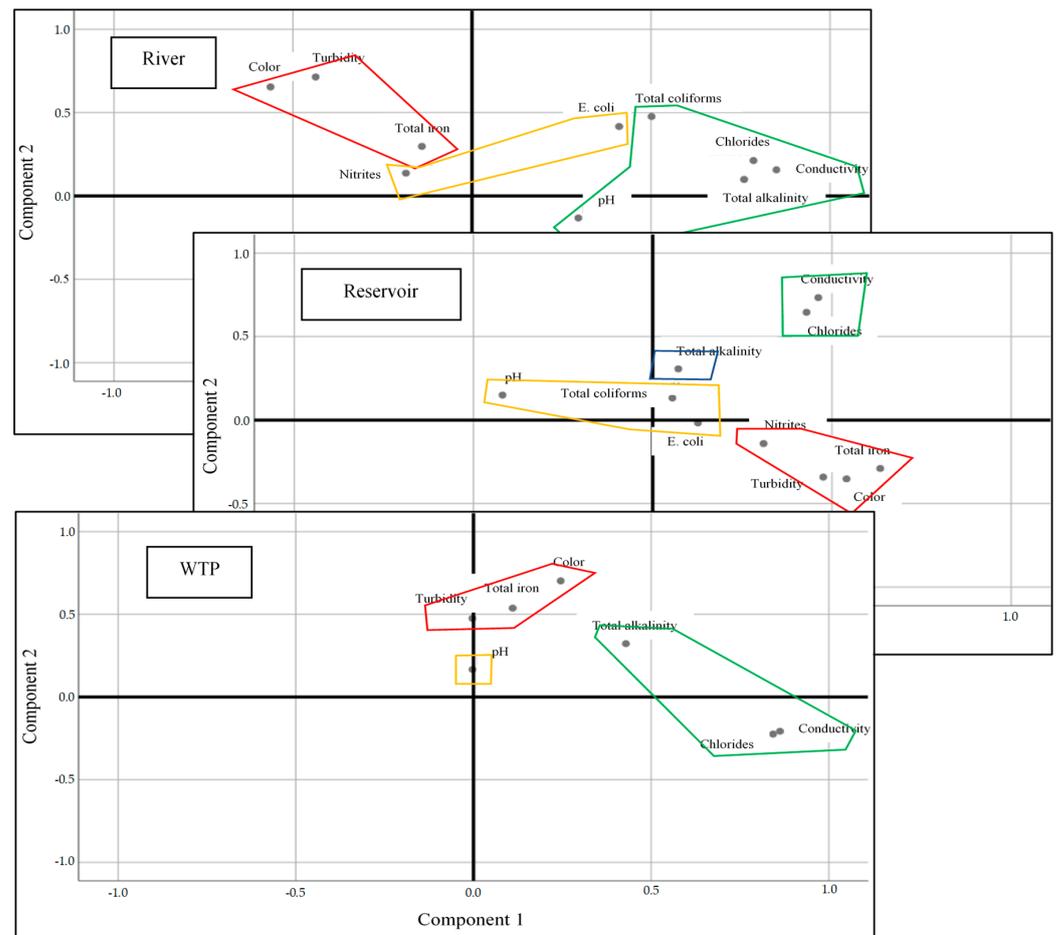


Figure 2. PCA biplot for water quality parameters at monitoring stations (river, reservoir, and WTP).

3.2. ARIMA Models

ARIMA models were developed for the water quality indicator parameter (turbidity) in the three components of the DWSS under study (river, reservoir, and WTP). In this study, no model exhibited stationarity. In relation to the river, the results showed an ARIMA model (0,1,6) under a daily time scale. The ARIMA model developed included additive outliers ($n = 34$, Table 4). The results suggested that river turbidity, at the reservoir inlet, was not influenced by preceding daily observations ($AR = 0$ days). ARIMA modeling also hinted at a decreasing trend ($I = 1$ day, average = -5.36 UNT/year) and variability ($MA = 6$ days) of turbidity in the river during the study period. This variability in daily turbidity possibly supported the need to consider these additive outliers (unexpected value for a single observation) to develop an adequate ARIMA model ($R^2 = 0.760$). Otherwise, the model would have had a worse fit ($R^2 = 0.206$). These outliers, mainly high turbidity in the river (>118 UNT), were possibly associated with periods of increased magnitude and frequency of rainfall at the study site (Figure 3). An analysis with Spearman's correlation coefficient showed a very weak significant correlation ($r_s = 0.108$, p -value < 0.050) between observed daily turbidity in the river and daily rainfall at the study site. The authors of [50] also hinted at this trend of increasing turbidity in rivers during rainy periods. Indeed, turbidity in the river comparatively tended to decrease during dry weather periods.

Table 4. Univariate ARIMA models for turbidity at the different monitoring stations (river, reservoir, and WTP).

Time Scale	ARIMA Model (p,d,q)	Transformation/ Outliers	R ²	MAPE (%)	BIC	Ljung-Box Q (18) GL	p-Value (Q)
River							
Daily	(0,1,10)	Natural logarithm	0.206	83.308	9.599	19.339 14	0.152
	(4,1,2)	Natural logarithm	0.208	83.065	9.610	11.113 12	0.519
	(0,1,6) *	Natural logarithm/34	0.760	64.471	8.324	21.252 14	0.095
Weekly	(0,1,8)	Natural logarithm	0.868	14.704	6.110	36.743 15	0.001
	(1,1,7) *	Square root	0.901	16.904	5.838	18.029 10	0.054
	(1,1,7)	Natural logarithm/47	0.872	13.280	6.226	9.053 13	0.769
Monthly	(1,1,1)	Natural logarithm	0.974	3.988	3.554	11.602 16	0.771
	(2,1,1)	Natural logarithm	0.974	3.970	3.561	6.533 15	0.969
	(2,1,2) *	Natural logarithm/68	0.990	3.448	2.825	18.283 14	0.194
Reservoir							
Daily	(0,1,5)	Natural logarithm	0.640	38.448	3.729	17.477 15	0.291
	(1,1,5)	Natural logarithm	0.640	38.428	3.739	12.177 12	0.432
	(0,1,5) *	Natural logarithm/36	0.889	31.833	1.712	17.200 15	0.307
Weekly	(0,1,7)	Natural logarithm	0.289	6.923	0.208	26.497 11	0.005
	(1,1,7)	Square root	0.902	7.264	0.002	16.946 10	0.076
	(0,1,9) *	Natural logarithm/36	0.921	6.039	−0.092	14.759 9	0.098
Monthly	(0,2,1)	Natural logarithm	0.975	2.520	−2.275	23.547 17	0.132
	(1,2,1)	Natural logarithm	0.975	2.490	−2.275	2.995 16	0.143
	(1,1,1) *	Natural logarithm/74	0.996	1.706	−3.919	25.781 16	0.057
WTP							
Daily	(0,1,11)	Natural logarithm	0.150	20.696	2.748	31.458 14	0.005
	(1,0,7)	Square root	0.236	35.029	2.506	8.053 10	0.624
	(1,1,9) *	No transformation/13	0.988	18.517	−1.629	15.251 8	0.054
Weekly	(0,1,7)	Natural logarithm	0.822	4.118	−0.450	199.126 13	0.000
	(0,1,10)	Square root	0.929	6.053	−1.351	13.687 8	0.090
	(3,1,8) *	Natural logarithm/26	0.982	3.129	−2.683	12.769 12	0.386
Monthly	(0,1,6)	Natural logarithm	0.969	1.329	−3.394	9.263 17	0.932
	(1,1,6)	Natural logarithm	0.969	1.312	−3.387	0.503 11	1.000
	(2,1,14) *	Natural logarithm/34	0.998	0.960	−6.185	1.914 2	0.384

Note: * = selected ARIMA models.

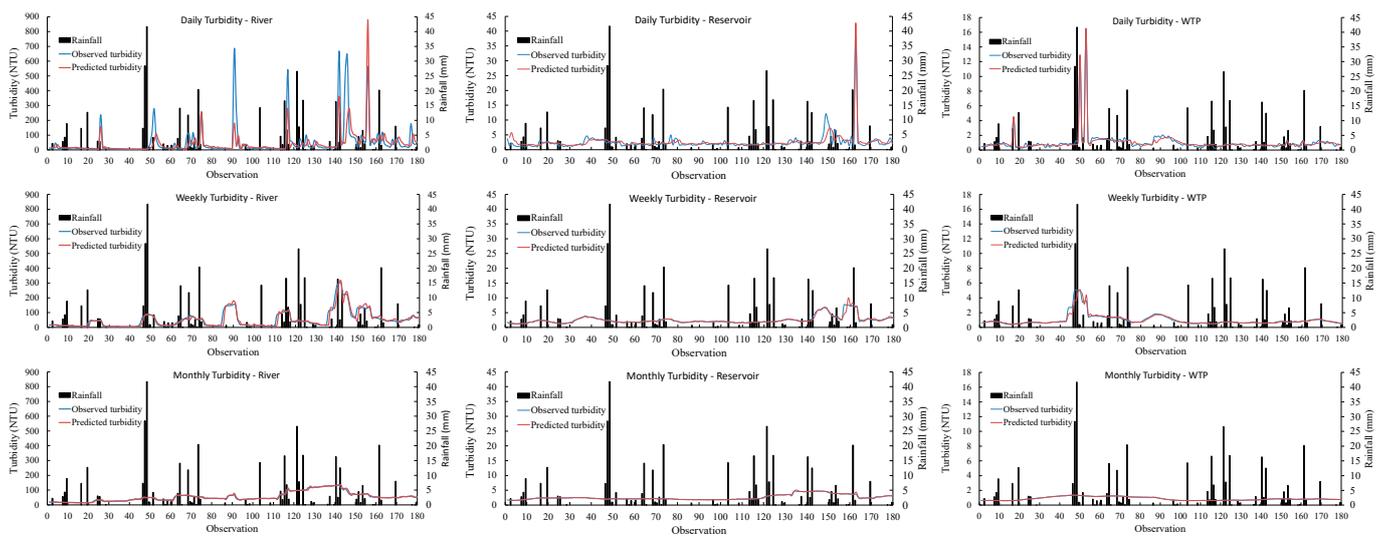


Figure 3. Observed and predicted (ARIMA) values of turbidity in the different DWSS components (river, reservoir, and WTP), and observed daily rainfall. Daily, weekly, and monthly time scales (selected period: 1–180, out of 2922 observations).

Daily, the findings showed an ARIMA model (0,1,5) for turbidity in the reservoir ($R^2 = 0.889$). This model considered 36 additive outliers (Table 4). The model developed suggested that turbidity within the reservoir was not influenced by preceding daily observations (AR = 0 days). This trend was possibly associated with the large dilution capacity of the reservoir from the significant water volume stored (average = 75 million m^3). The authors of [51] also reported the turbidity buffering capacity of reservoirs due to large volumes of stored water. Daily modeling also suggested a decreasing trend (I = 1 day, average = -0.070 UNT/year) and variability (MA = 5 days) of turbidity inside the reservoir during the study period. This daily variability of turbidity possibly supported the need to consider these additive outliers to develop an adequate ARIMA model. Otherwise, the model fit would have been worse ($R^2 = 0.640$). These outliers, mainly high turbidity (>37.6 UNT), were comparatively associated with periods of increased rainfall (magnitude and frequency, Figure 3). During dry periods turbidity tended to decrease comparatively in the reservoir.

In relation to daily turbidity at the WTP outlet, the results showed an ARIMA model (1,1,9). This model included 13 additive outliers ($R^2 = 0.988$, Table 4). The results suggested that daily turbidity at the WTP outlet was influenced by the immediately preceding observation (AR = 1 day). Namely, the turbidity at the WTP outlet was probably influenced by the water treatment operations performed on the immediately preceding day. The authors of [52] also reported that water quality in WTPs was significantly influenced by treatment activities performed on previous days. ARIMA analysis suggested a decreasing trend (I = 1 day, average = -0.13 UNT/year) and variability (MA = 9) in turbidity at the WTP outlet during the study period. This variability in turbidity possibly supported the need to consider these outliers in the development of the ARIMA model. Otherwise, the model would have had a worse fit ($R^2 = 0.236$). These outliers, mainly high turbidity (>4.50 UNT), were possibly associated with the turbidity of the raw water, which came from the reservoir. Comparatively, these additive outliers of high turbidity coincided with rainy periods at the study site (Figure 3). The authors of [53] also reported this trend of increased water turbidity due to a decline in raw water quality at the WTP intakes.

On a weekly basis, the results showed an ARIMA model (1,1,7) for turbidity in the river ($R^2 = 0.901$). This model did not include additive outliers (Table 4). The model suggested that weekly river turbidity was influenced by what occurred in the immediately preceding week (AR = 1). That is, the effects of river turbidity were possibly only significant during this time window. The weekly ARIMA model also suggested a decreasing trend (I = 1

week, average = -5.36 UNT/year) and variability (MA = 7) of turbidity in the river during the study period. Because the weekly turbidity time series was preprocessed with a 7-day moving average, then the effective variability of turbidity in the river corresponded to two weeks. This preprocessing of the turbidity time series (smoothing) possibly also influenced the non-detection of outliers by the developed weekly ARIMA model, as compared to the ARIMA model obtained under a daily time scale. In addition, the findings showed in the reservoir a weekly ARIMA model (0,1,9) for turbidity. This model considered 36 additive outliers. The model suggested that turbidity within the reservoir was not influenced by preceding weekly observations (AR = 0). This trend was possibly related to the large buffering capacity of the reservoir to absorb the pollutant load [54]. The weekly model also suggested a decreasing trend (I = 1, average = -0.070 UNT/year) and variability (MA = 9) of turbidity in the reservoir. This weekly variability of turbidity possibly supported the need to consider these additive outliers ($n = 36$) to obtain an adequate ARIMA model. These high turbidity outliers (>7.70 UNT) were comparatively related to the occurrence of rainy periods at the study site (Figure 3).

In relation to weekly turbidity at the WTP outlet, the findings showed an ARIMA model (3,1,8). This model included 26 additive outliers ($R^2 = 0.982$, Table 4). The results suggested that turbidity at the WTP outlet was influenced by what occurred during the three immediately preceding observations (AR = 3). Because the weekly turbidity time series was preprocessed (7-day moving average), the effective time window for the AR term corresponded to 1.43 weeks. ARIMA analysis also suggested a decreasing trend (I = 1, average = -0.13 UNT/year) and variability (MA = 8) in turbidity at the WTP outlet. This variability possibly made it necessary to consider additive outliers in the development of the weekly ARIMA model. Comparatively, these additive outliers of high turbidity (>2.16 UNT) occurred during rainy periods at the study site, possibly causing outliers of high turbidity to also be observed at the reservoir (>7.70 UNT). During dry weather periods, turbidity tended to decrease at the WTP outlet (Figure 3).

On a monthly basis, the findings showed an ARIMA model (2,1,2) for turbidity in the river ($R^2 = 0.990$). The selected model included 68 additive outliers (Table 4). The results suggested that monthly in-stream turbidity was influenced by what occurred during the two immediately preceding observations (AR = 2). Because the monthly turbidity time series was preprocessed (30-day moving average), then the effective time window for the AR term corresponded to 1.07 months. The model also hinted at a decreasing trend (I = 1, average = -5.36 UNT/year) and low variability (MA = 2) in river turbidity. Though, this variability possibly made it necessary to consider outliers in the development of the monthly ARIMA model. For example, this model also detected additive outliers of high (199 UNT) and low turbidity (3.84 UNT) during periods of rainy and dry weather at the study site, respectively. In addition, the results showed a monthly ARIMA model (1,1,1) for turbidity at the reservoir. The model included 74 additive outliers. The findings suggested that turbidity in the reservoir was influenced by what occurred in the immediately preceding monthly observation (AR = 1). The ARIMA model also suggested a slight decreasing trend (I = 1, average = -0.070 UNT/year) and low variability (MA = 1) monthly in reservoir turbidity. This model also detected additive outliers of high (12.7 UNT) and low turbidity (0.83 UNT) during periods of rainy and dry weather at the study site, respectively.

In relation to monthly turbidity at the WTP outlet, the results showed an ARIMA model (2,1,14). This model included 34 additive outliers ($R^2 = 0.998$, Table 4). The results suggested that turbidity at the WTP outlet was influenced by what occurred during the two immediately preceding observations (AR = 2). The effective time sale for this influence was then 1.07 months. The model also suggested a decreasing trend (I = 1, average = 0.13 UNT/year) and high variability (MA = 14) in turbidity at the WTP outlet. This variability in turbidity possibly supported the need to consider outliers in the development of the ARIMA model. For example, this model detected additive outliers of high (1.35 UNT) and low turbidity (0.25 UNT) during periods of rainy and dry weather at the study site, respectively. In effect, this underran the range of monthly variation in turbidity at the WTP outlet.

3.3. ARIMA and TFARIMA Comparative Analysis

Daily, the results showed ARIMA models (0,1,6) and (0,1,5) for turbidity in the river and reservoir, respectively (Table 4). It was observed that the models in these two DWSS components were similar with respect to the AR term. The models did not hint at the influence of past events on the behavior of river and reservoir turbidity ($AR = 0$). This trend was possibly associated with the non-occurrence of significant daily events influencing future river and reservoir turbidity behavior (e.g., runoff and downpours). This trend did not also suggest a significant influence of river discharge on the turbidity behavior in the reservoir. The latter was probably associated with the large dilution capacity (buffering of pollutant loads) of the reservoir from the significant volume of stored water. The authors of [55] also reported the turbidity buffering capacity of reservoirs due to significant volumes of stored water. In addition, the results showed a daily TFARIMA model (0,1,5) for reservoir turbidity (dependent variable) from the influence of river turbidity (independent variable) ($R^2 = 0.902$). This model showed that there was no influence of past river turbidity events on reservoir turbidity (numerator/delay = 0, Table 5). Namely, the findings suggested that turbidity in the reservoir responded instantaneously to the behavior of turbidity in the river (Figure 4). In the context of this study, this trend would not be entirely logical because the river contributed only 5% of the water volume stored in the reservoir. The remaining 95% of the water volume was provided by a transfer from the Chuza reservoir. This instantaneous trend was possibly related to the prevailing climatic conditions at the study site. In other words, the climatic conditions (rainfall regime) at the study site were possibly more dominant (instantaneous effect) on river and reservoir turbidity behavior compared to the dominance of river turbidity on reservoir turbidity behavior [56,57].

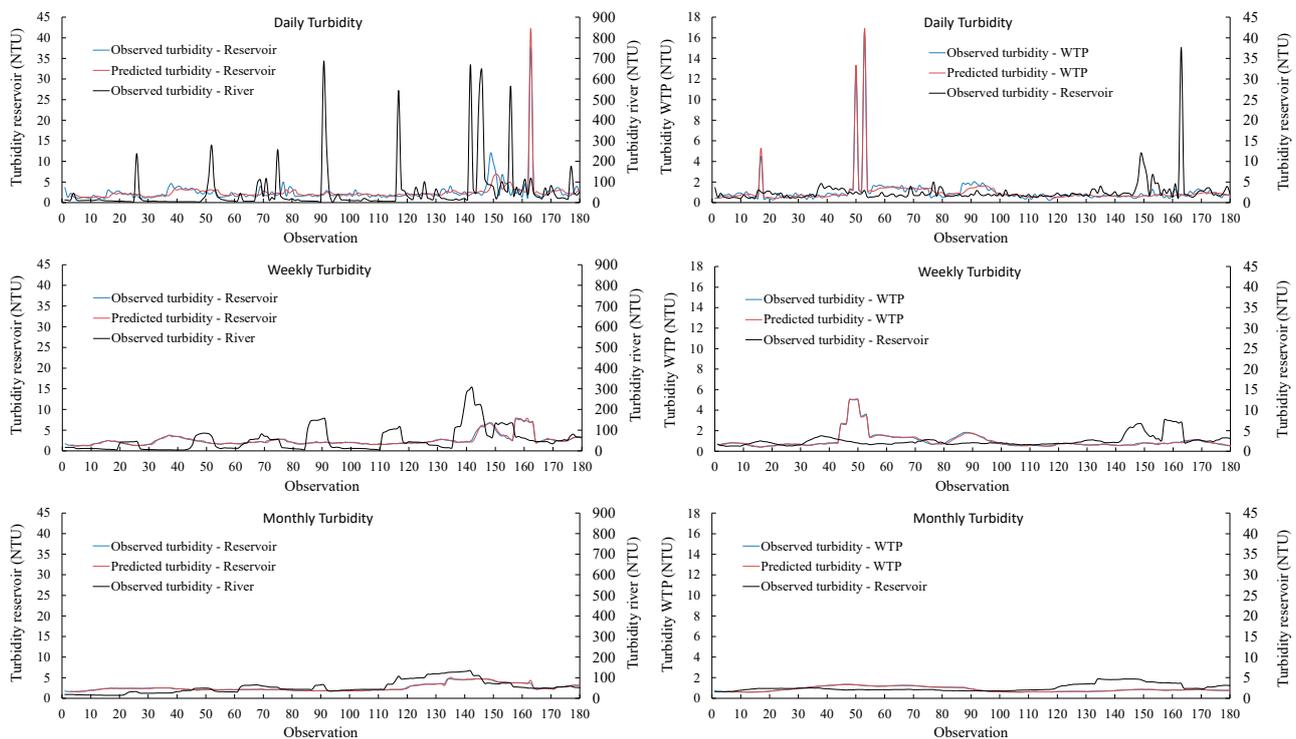


Figure 4. Observed and predicted values (TFARIMA) of turbidity in the different DWSS components (river versus reservoir and reservoir versus WTP). Daily, weekly, and monthly time scales (selected period: 1–180, out of 2922 observations).

Table 5. TFARIMA models of turbidity between the different DWSS components (river versus reservoir and reservoir versus WTP).

Time Scale	ARIMA Model (p,d,q)	Transformation	R ²	MAPE (%)	Outliers	BIC	Ljung-Box Q (18) GL	p-Value (Q)
River/Reservoir								
Daily	(0,1,5)	Natural logarithm	0.080	36.636	No	3.725	18.955 15	0.216
	(0,1,5) *	Natural logarithm	0.902	30.306	36 (AD)	1.593	19.071 15	0.211
Numerator/Delay: 0, Difference: 1, Denominator/Delay: 1								
Weekly	(0,1,7)	Natural logarithm	0.889	7.096	No	0.138	26.073 11	0.006
	(0,1,9)	Natural logarithm	0.988	5.308	49 (AD and LS)	-1.948	21.069 9	0.012
Numerator/Delay: 0, Difference: 1, Denominator/Delay: 1								
Monthly	(1,1,1)	Natural logarithm	0.976	2.534	No	-2.324	8.043 16	0.948
	(1,1,13) *	Natural logarithm	0.994	2.264	44 (AD)	-3.515	17.108 15	0.312
Numerator/Delay: 0, Difference: 1, Denominator/Delay: 1								
Reservoir/WTP								
Daily	(0,1,9)	Natural logarithm	0.013	20.580	No	2.750	32.755 15	0.005
	(0,1,9) *	Natural logarithm	0.997	18.353	17 (AD and LS)	-2.883	19.246 15	0.203
Numerator/Delay: 0, Difference: 1, Denominator/Delay: N/A								
Weekly	(0,1,7)	Natural logarithm	0.867	5.510	No	-0.772	189.890 16	0.000
	(0,1,10)	None	0.999	4.882	12 (AD and LS)	-6.180	57.717 11	0.000
Numerator/Delay: N/A, Difference: N/A, Denominator/Delay: N/A								
Monthly	(0,2,18)	Natural logarithm	0.995	1.136	No	-8.828	44.387 14	0.000
	(0,2,9) *	None	0.998	1.076	20 (AD and LS)	-9.602	18.684 14	0.177
Numerator/Delay: 0, Difference: 2, Denominator/Delay: 2								

Note: * = selected TFARIMA models, AD = additive outliers, and LS = level shift outliers.

Additionally, the TFARIMA model showed a decreasing trend in the daily turbidity of the river (difference = 1) to explain the behavior of turbidity in the reservoir. Thus, the results hinted at a decreasing trend in river and reservoir turbidity during the study period. The findings also showed that it was necessary to consider 36 additive outliers to develop an adequate TFARIMA model ($R^2 = 0.902$, Table 5). Otherwise, the model would have had a worse fit ($R^2 = 0.080$). These additive outliers only occurred during 1.23% of the study period and corresponded mainly to periods of high turbidity, which was probably related to an increase in rainfall at the study site (Figure 3). The TFARIMA model also suggested a short influence of the variability of turbidity in the river on the observed turbidity in the reservoir (denominator/delay = 1). This is possibly because the river contributed only 5% of the water volume stored in the reservoir. As mentioned, the climatic conditions (rainfall regime) of the study site were possibly more dominant in the behavior of river and reservoir turbidity compared to the effects transmitted from the river to the reservoir.

On a weekly basis, the results showed ARIMA models (1,1,7) and (0,1,9) for turbidity in the river and reservoir, respectively (Table 4). The results suggested a different behavior in the AR term. Initially, it was hinted that in the river past events had more influence on turbidity behavior (AR = 1) compared to turbidity behavior in the reservoir (AR = 0).

Although the difference in the magnitude of the AR term was only one unit. The results also showed an alternative ARIMA model (1,1,7) for turbidity in the reservoir. Thus, the findings also suggested that the turbidity behavior in the river and reservoir was influenced by what happened in the immediately preceding week. This trend was conditioned by the preprocessing of the weekly time series (7-day moving average). Namely, turbidity in the river and reservoir was influenced by the daily average behavior observed during the immediately preceding week. This weekly trend suggested a short-term influence. In addition, the results showed a TFARIMA model (0,1,9) for reservoir turbidity from river turbidity ($R^2 = 0.988$). Although the model initially suggested that there was no influence of river turbidity on reservoir turbidity (numerator/delay = 0, Table 5), this model could not be considered because it did not meet the significance criterion in the Ljung–Box statistic (p -value > 0.050). In other words, in this study it was not possible to adequately develop a TFARIMA model for reservoir turbidity from river turbidity (Figure 4). This could also suggest the non-existence of an influence of river turbidity on reservoir turbidity under a weekly time scale. Lastly, by additionally considering level shift outlier, it was also not possible to develop an adequate TFARIMA model (p -value < 0.050).

On a monthly basis, the results showed ARIMA models (2,1,2) and (1,1,1) for turbidity in the river and reservoir, respectively (Table 4). The results showed different behavior in the AR term. Thus, the findings suggested that the turbidity behavior in the river tended to be more influenced by past events (AR = 2) compared to turbidity in the reservoir (AR = 1). However, the difference in the magnitude of the AR term was only one unit. This shorter time window in the reservoir was possibly associated with its higher dilution capacity compared to the river. The results also showed an alternative ARIMA model (1,1,1) for turbidity in the river. In general terms, the turbidity of the river and reservoir were influenced by the daily average behavior observed during the immediately preceding month. In addition, the results showed a TFARIMA model (1,1,13) for turbidity in the reservoir from turbidity in the river ($R^2 = 0.994$). This model implied that there was no monthly influence of past river turbidity events on reservoir turbidity (numerator/delay = 0, Table 5). This trend was like that observed under a daily time scale. Namely, turbidity in the reservoir was possibly more conditioned by the climatic regime (rainfall) of the study site than by turbidity contributed from the river.

Additionally, TFARIMA model showed a decreasing trend in the monthly turbidity of the river (difference = 1) to explain the behavior of the reservoir turbidity. Thus, the results suggested a decreasing trend in river and reservoir turbidity throughout the study period. The findings also showed that it was necessary to consider 44 additive outliers to develop a more adequate monthly TFARIMA model ($R^2 = 0.994$, Table 5), although without considering these additive outliers, the alternative TFARIMA model also had an adequate fit ($R^2 = 0.976$, Table 5). The results hinted that these additive outliers in river turbidity possibly did not have a significant influence on the monthly behavior of turbidity in the reservoir. Indeed, preprocessing of the monthly turbidity time series (30-day moving average) led to a smoothing of the observed turbidity values. These outliers occurred during only 1.51% of the study period (unexpected value for a single observation) and corresponded mainly to periods of high turbidity, which was probably related to periods of increased rainfall at the study site. Lastly, TFARIMA model consistently suggested a low monthly influence of in-stream turbidity variability on reservoir turbidity (denominator/delay = 1). As mentioned, climatic conditions (rainfall regime) of the study site were possibly the most dominant in the turbidity behavior in the river and reservoir. Apparently, the influence of rainfall on monthly river turbidity was greater (AR = 2) compared to the influence of rainfall on monthly reservoir turbidity (AR = 1).

Daily, the results showed ARIMA models (0,1,5) and (1,1,9) for turbidity in the reservoir and WTP, respectively (Table 4). Different behavior was observed in the AR term of the developed models. The results suggested that the turbidity behavior in the WTP was influenced by what happened the day before (AR = 1). Namely, the treatment operations conducted at the WTP possibly had a time window of one day. Conversely, the reservoir

was not influenced by past events ($AR = 0$), possibly due to its great capacity to dilute pollutants [51]. Therefore, the findings suggested that the turbidity behavior at the WTP outlet was more conditioned by the water treatment operations conducted the day before than by the turbidity behavior in the reservoir (raw water). In addition, the results showed a daily TFARIMA model (0,1,9) for WTP turbidity (dependent variable) from reservoir turbidity (independent variable) ($R^2 = 0.997$). The model suggested that there was no influence of past reservoir turbidity events on turbidity at the WTP outlet (numerator/delay = 0, Table 5). Due to the large variability of daily turbidity at the WTP, it was not possible to develop an adequate TFARIMA model without considering outliers (Table 5). It was necessary to consider both additive and level shift outliers ($n = 17$) to develop an adequate model (Figure 4). These outliers, primarily high turbidity at the WTP outlet, occurred only 0.58% of the time during the study period. The findings also showed that the variability of turbidity in the reservoir had no influence on the variability of turbidity at the WTP outlet (denominator/delay: N/A). Indeed, this variability in turbidity was possibly conditioned by the treatment operations conducted inside the WTP [52]. Lastly, TFARIMA model suggested a decreasing trend in the turbidity of the reservoir and WTP during the study period (difference = 1).

On a weekly basis, the results showed ARIMA models (0,1,9) and (3,1,8) for turbidity in the reservoir and WTP, respectively (Table 4). The results showed different behavior in the AR term. The findings showed that the turbidity behavior at the WTP was influenced by what occurred during the three immediately preceding observations ($AR = 3$). Due to the preprocessing of these time series (7-day moving average), the effective time window was 1.29 weeks. The results suggested that the treatment operations executed at the WTP were influenced by the daily average behavior observed during the immediately preceding 1.29 weeks. In contrast, in the reservoir there was no evidence of the influence of past events on the turbidity behavior ($AR = 0$). The results also showed a weekly TFARIMA model (0,1,10) for the WTP turbidity based on the reservoir turbidity. This model could not be considered because it did not meet the significance criterion in the Ljung–Box statistic (p -value > 0.050), despite having considered additive and level shift outliers (Figure 4). In other words, it was not possible to adequately develop a transfer function model for turbidity at the WTP outlet from the reservoir turbidity. This could also suggest the non-existence of a weekly influence of reservoir turbidity on turbidity at the WTP outlet. In this study, the turbidity behavior was possibly more conditioned by the treatment operations inside the WTP than by the quality of the raw water coming from the reservoir.

On a monthly basis, the results showed ARIMA models (1,1,1) and (2,1,14) for turbidity in the reservoir and at the WTP outlet, respectively (Table 4). Differences were observed in the AR term. The findings suggested that the turbidity behavior at the WTP tended to be more influenced by past events ($AR = 2$) compared to turbidity at the reservoir ($AR = 1$). Possibly, this shorter time window at the reservoir was associated with its greater pollutant dilution capacity compared to the WTP [51]. The findings also showed that the behavior of turbidity at the WTP (treatment operations) was influenced by what occurred during the two immediately preceding observations ($AR = 2$). Due to the preprocessing of these time series, the effective time window at the WTP was 1.03 months. At the reservoir, the time window of past events was then similar, one month ($AR = 1$). In general terms, the turbidity of the reservoir and WTP were influenced by the daily average behavior observed during the immediately preceding month. In addition, the results showed a monthly TFARIMA model (0,2,9) for WTP turbidity from reservoir turbidity ($R^2 = 0.998$). On this occasion, it was additionally necessary to consider level shift outliers (they have a permanent effect), and not exclusively additive outliers (unexpected value for a single observation) to develop a satisfactory TFARIMA model. These mainly high turbidity outliers only occurred during 0.68% of the study period (Figure 4). This model also suggested that there was no influence of past turbidity events in the reservoir on turbidity at the WTP outlet (numerator/delay = 0, Table 5). The TFARIMA model showed a decreasing trend in monthly turbidity in the reservoir (difference = 2) to explain the behavior of turbidity at the WTP outlet. Thus, the results suggested a decreasing trend in turbidity at the WTP outlet and in the

reservoir during the entire study period. Lastly, TFARIMA model suggested a short monthly influence (1.03 effective months) of the variability of turbidity in the reservoir on turbidity at the WTP outlet (denominator/delay = 2).

4. Conclusions

The findings of this study on the use of ARIMA models to analyze the temporal behavior of water quality parameters in the initial components of a DWSS allow us to visualize the following conclusions.

The best water quality indicator parameters are as follows: turbidity > color > total iron. Depending on the DWSS component, the order of importance for these indicator parameters is as follows: river > reservoir > WTP.

Increasing the time window of the ARIMA analysis (daily, weekly, and monthly) suggests an increase in the magnitude of the AR term in the univariate models. This trend is evident as follows: WTP > river > reservoir. This trend suggests that the turbidity behavior in the WTP is more influenced by past observations compared to the turbidity behavior in the river and reservoir, respectively. Indeed, the results hint that the time window of water quality analysis in DWSSs should be differentiated according to the component under study.

The DWSS component that shows the greatest variability in turbidity according to the ARIMA analysis is the WTP, followed by the river and reservoir, respectively. This trend suggests the influence of the following factors: Water treatment operations performed at the WTP, observed rainfall regime, and pollutant dilution capacity of the reservoir. In addition, by increasing the time window of the ARIMA analysis (daily, weekly, and monthly) an increase in the detection of additive outliers is observed. Thus, the smoothing of the data series (moving average) as the time window of the ARIMA analysis increases possibly leads to a greater sensitivity of the model for outlier detection.

TFARIMA models suggest that there is no significant influence of past river turbidity events on turbidity in the reservoir, and of reservoir turbidity on turbidity at the WTP outlet. This is for all time scales considered in this study. In addition, a decreasing trend in turbidity is observed in all DWSS components during the study period.

As the TFARIMA analysis time window increases (daily, weekly, and monthly), the models tend to have a better fit. Between river and reservoir turbidity, it is possible to develop models fitted only with additive outliers. Between reservoir and WTP turbidity, it is necessary to additionally consider level shift outliers. Hence, the results suggest that turbidity outlier events between the river and reservoir occur mainly in a single observation, and between the reservoir and WTP also have a permanent effect over time. In effect, this possibly conditions the water treatment operations inside the WTP.

The findings suggest that the AR term of the models is useful for studying the transfer of effects between DWSS components, and the MA term is useful for studying the influence of external factors on water quality in each DWSS component. For example, the latter term is useful for studying the performance of the WTP. Lastly, this study is relevant to visualize the usefulness of ARIMA analysis in the operation of DWSSs, for decision making by research institutes, and agencies monitoring and controlling the water quality supplied in megacities.

Author Contributions: Conceptualization, C.A.Z.-M.; Methodology, C.A.Z.-M. and H.A.R.-Q.; Software, C.A.Z.-M. and H.A.R.-Q.; Validation, C.A.Z.-M.; Formal analysis, C.A.Z.-M.; Investigation, C.A.Z.-M., H.A.R.-Q., and C.F.U.-B.; Resources, C.A.Z.-M., H.A.R.-Q., and C.F.U.-B.; Data curation, C.A.Z.-M., H.A.R.-Q., and C.F.U.-B.; Writing—original draft, C.A.Z.-M.; Writing—review & editing, C.A.Z.-M., H.A.R.-Q., and C.F.U.-B.; Visualization, C.A.Z.-M. and H.A.R.-Q.; Supervision, C.A.Z.-M.; Project administration, C.A.Z.-M.; Funding acquisition, C.A.Z.-M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors wish to acknowledge the logistical support provided by the Empresa de Acueducto y Alcantarillado de Bogotá—E.A.A.B. (Colombia). We also thank the participating institutions (Universidad Distrital Francisco José de Caldas and Universidad Militar Nueva Granada, Colombia) for the support granted to researchers. In the case of the author Carlos Felipe Urazán-Bonells, it is mentioned that it is a product of his academic work as a professor at the Universidad Militar Nueva Granada.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, T.; Yue, W.; Wu, T.; Zhang, X.; Xia, C. Human Well-Being Related Analysis on Urban Carrying Capacity: An Empirical Study in Chinese Mega-Cities. *J. Urban Aff.* **2023**, *45*, 1–17. [[CrossRef](#)]
2. Abu-Zeid, M.A. Water and Sustainable Development: The Vision for World Water, Life and the environment. This Paper Is Based on a Keynote Address Made at the International Conference on Water and Sustainable Development, Paris, March 19, 1998. *Water Policy* **1998**, *1*, 9–19. [[CrossRef](#)]
3. Cosgrove, W.J.; Loucks, D.P. Water Management: Current and Future Challenges and Research Directions. *Water Resour. Res.* **2015**, *51*, 4823–4839. [[CrossRef](#)]
4. Tzanakakis, V.A.; Paranychianakis, N.V.; Angelakis, A.N. Water Supply and Water Scarcity. *Water* **2020**, *12*, 2347. [[CrossRef](#)]
5. Senbore, S.; Oke, S.A. Urban Development Impact on Climate Variability and Surface Water Quality in Part of Mangaung Metropolis of South Africa. *Dev. S. Afr.* **2023**, *40*, 293–312. [[CrossRef](#)]
6. Ramsay, L.; Petersen, M.M.; Hansen, B.; Schullehner, J.; van der Wens, P.; Voutchkova, D.; Kristiansen, S.M. Drinking Water Criteria for Arsenic in High-Income, Low-Dose Countries: The Effect of Legislation on Public Health. *Environ. Sci. Technol.* **2021**, *55*, 3483–3493. [[CrossRef](#)]
7. Wang, H.-J.; Peng, C.-W.; Han, X.; Wang, Y.; Zhang, J.; Liu, J.-L.; Zhou, M.-X.; Tang, F.; Liu, A.-L. Toxicological Characteristics of Drinking Water in Two Large-Scale Municipal Water Supply Systems of a Metropolitan City in Central China. *Environ. Sci. Pollut. Res.* **2023**, *30*, 64058–64066. [[CrossRef](#)] [[PubMed](#)]
8. Li, P.; Wu, J. Drinking Water Quality and Public Health. *Expo. Health* **2019**, *11*, 73–79. [[CrossRef](#)]
9. Pérez-Vidal, A.; Escobar-Rivera, J.C.; Torres-Lozada, P. Development and Implementation of a Water-Safety Plan for Drinking-Water Supply System of Cali, Colombia. *Int. J. Hyg. Environ. Health* **2020**, *224*, 113422. [[CrossRef](#)]
10. Benítez, J.S.; Rodríguez, C.M.; Casas, A.F. Disinfection Byproducts (DBPs) in Drinking Water Supply Systems: A Systematic Review. *Phys. Chem. Earth Parts A/B/C* **2021**, *123*, 102987. [[CrossRef](#)]
11. Wu, S.; Sumari, N.S.; Dong, T.; Xu, G.; Liu, Y. Characterizing Urban Expansion Combining Concentric-Ring and Grid-Based Analysis for Latin American Cities. *Land* **2021**, *10*, 444. [[CrossRef](#)]
12. Feola, G.; Suzunaga, J.; Soler, J.; Goodman, M.K. Ordinary Land Grabbing in Peri-Urban Spaces: Land Conflicts and Governance in a Small Colombian City. *Geoforum* **2019**, *105*, 145–157. [[CrossRef](#)]
13. Rodríguez-Jeangros, N.; Camacho, L.A.; Rodríguez, J.P.; McCray, J.E. Integrated Urban Water Resources Model to Improve Water Quality Management in Data-Limited Cities with Application to Bogotá, Colombia. *J. Sustain. Water Built Environ.* **2018**, *4*, 04017019. [[CrossRef](#)]
14. Rodríguez, J.P.; McIntyre, N.; Díaz-Granados, M.; Quijano, J.P.; Maksimović, Č. Monitoring and Modelling to Support Wastewater System Management in Developing Mega-Cities. *Sci. Total Environ.* **2013**, *445–446*, 79–93. [[CrossRef](#)]
15. Salamanca-Cano, A.K.; Durán-Díaz, P. Stakeholder Engagement around Water Governance: 30 Years of Decision-Making in the Bogotá River Basin. *Urban Sci.* **2023**, *7*, 81. [[CrossRef](#)]
16. Poveda, G.; Álvarez, D.M.; Rueda, Ó.A. Hydro-Climatic Variability over the Andes of Colombia Associated with ENSO: A Review of Climatic Processes and Their Impact on One of the Earth’s Most Important Biodiversity Hotspots. *Clim. Dyn.* **2011**, *36*, 2233–2249. [[CrossRef](#)]
17. Restrepo, J.D.; Kettner, A.J.; Syvitski, J.P.M. Recent Deforestation Causes Rapid Increase in River Sediment Load in the Colombian Andes. *Anthropocene* **2015**, *10*, 13–28. [[CrossRef](#)]
18. Pintilie, L.; Torres, C.M.; Teodosiu, C.; Castells, F. Urban Wastewater Reclamation for Industrial Reuse: An LCA Case Study. *J. Clean. Prod.* **2016**, *139*, 1–14. [[CrossRef](#)]
19. Savun-Hekimoğlu, B.; Erbay, B.; Hekimoğlu, M.; Burak, S. Evaluation of Water Supply Alternatives for Istanbul Using Forecasting and Multi-Criteria Decision Making Methods. *J. Clean. Prod.* **2021**, *287*, 125080. [[CrossRef](#)]
20. Zhu, M.; Wang, J.; Yang, X.; Zhang, Y.; Zhang, L.; Ren, H.; Wu, B.; Ye, L. A Review of the Application of Machine Learning in Water Quality Evaluation. *Eco-Environ. Health* **2022**, *1*, 107–116. [[CrossRef](#)]
21. Barrientos-Torres, D.; Martínez-Ríos, E.A.; Navarro-Tuch, S.A.; Pablos-Hach, J.L.; Bustamante-Bello, R. Water Flow Modeling and Forecast in a Water Branch of Mexico City through ARIMA and Transfer Function Models for Anomaly Detection. *Water* **2023**, *15*, 2792. [[CrossRef](#)]
22. Veerendra, G.T.N.; Kumaravel, B.; Rao, P.K.R.; Dey, S.; Manoj, A.V.P. Forecasting Models for Surface Water Quality Using Predictive Analytics. *Environ. Dev. Sustain.* **2023**. [[CrossRef](#)]

23. Ristow, D.C.M.; Henning, E.; Kalbusch, A.; Petersen, C.E. Models for Forecasting Water Demand Using Time Series Analysis: A Case Study in Southern Brazil. *J. Water Sanit. Hyg. Dev.* **2021**, *11*, 231–240. [[CrossRef](#)]
24. Kaur, J.; Parmar, K.S.; Singh, S. Autoregressive Models in Environmental Forecasting Time Series: A Theoretical and Application Review. *Environ. Sci. Pollut. Res.* **2023**, *30*, 19617–19641. [[CrossRef](#)]
25. Katimon, A.; Shahid, S.; Mohsenipour, M. Modeling Water Quality and Hydrological Variables Using ARIMA: A Case Study of Johor River, Malaysia. *Sustain. Water Resour. Manag.* **2018**, *4*, 991–998. [[CrossRef](#)]
26. Wang, J.; Zhang, L.; Zhang, W.; Wang, X. Reliable Model of Reservoir Water Quality Prediction Based on Improved ARIMA Method. *Environ. Eng. Sci.* **2019**, *36*, 1041–1048. [[CrossRef](#)]
27. Elevation, S.; Uzgören, N.; Bingöl, D.; Elevation, B. Drinking Water Quality Control: Control Charts for Turbidity and pH. *J. Water Sanit. Hyg. Dev.* **2016**, *6*, 511–518. [[CrossRef](#)]
28. Sentas, A.; Psilovikos, A.; Karamoutsou, L.; Charizopoulos, N. Monitoring, Modeling, and Assessment of Water Quality and Quantity in River Pinios, Using ARIMA Models. *Desalination Water Treat.* **2018**, *133*, 336–347. [[CrossRef](#)]
29. Azhar, S.C.; Aris, A.Z.; Yusoff, M.K.; Ramli, M.F.; Juahir, H. Classification of River Water Quality Using Multivariate Analysis. *Procedia Environ. Sci.* **2015**, *30*, 79–84. [[CrossRef](#)]
30. Issa, H.M.; Alrwai, R.A. Long-Term Drinking Water Quality Assessment Using Index and Multivariate Statistical Analysis for Three Water Treatment Plants of Erbil City, Iraq. *UKH J. Sci. Eng.* **2018**, *2*, 39–48. [[CrossRef](#)]
31. Arain, M.B.; Ullah, I.; Niaz, A.; Shah, N.; Shah, A.; Hussain, Z.; Tariq, M.; Afridi, H.I.; Baig, J.A.; Kazi, T.G. Evaluation of Water Quality Parameters in Drinking Water of District Bannu, Pakistan: Multivariate Study. *Sustain. Water Qual. Ecol.* **2014**, *3–4*, 114–123. [[CrossRef](#)]
32. Maiolo, M.; Pantusa, D. Multivariate Analysis of Water Quality Data for Drinking Water Supply Systems. *Water* **2021**, *13*, 1766. [[CrossRef](#)]
33. Skandalos, N.; Wang, M.; Kapsalis, V.; D’Agostino, D.; Parker, D.; Bhuvad, S.S.; Udayraj; Peng, J.; Karamanis, D. Building PV Integration According to Regional Climate Conditions: BIPV Regional Adaptability Extending Köppen-Geiger Climate Classification against Urban and Climate-Related Temperature Increases. *Renew. Sustain. Energy Rev.* **2022**, *169*, 112950. [[CrossRef](#)]
34. Baird, R.B.; Eaton, A.D.; Rice, E.W. (Eds.) *Standard Methods for the Examination of Water and Wastewater*, 23rd ed.; American Water Works Association: Washington, DC, USA, 2017; ISBN 978-0-87553-287-5.
35. Burhanuddin, S.N.Z.A.; Deni, S.M.; Ramli, N.M. Imputation of Missing Rainfall Data Using Revised Normal Ratio Method. *Adv. Sci. Lett.* **2017**, *23*, 10981–10985. [[CrossRef](#)]
36. de Gois, G.; de Oliveira-Júnior, J.F.; da Silva Junior, C.A.; Sobral, B.S.; de Bodas Terassi, P.M.; Junior, A.H.S.L. Statistical Normality and Homogeneity of a 71-Year Rainfall Dataset for the State of Rio de Janeiro—Brazil. *Theor. Appl. Clim.* **2020**, *141*, 1573–1591. [[CrossRef](#)]
37. Ivošev, G.; Burton, L.; Bonner, R. Dimensionality Reduction and Visualization in Principal Component Analysis. *Anal. Chem.* **2008**, *80*, 4933–4944. [[CrossRef](#)]
38. Amiri, B.J.; Nakane, K. Modeling the Linkage Between River Water Quality and Landscape Metrics in the Chugoku District of Japan. *Water Resour. Manag.* **2009**, *23*, 931–956. [[CrossRef](#)]
39. Beamonte Córdoba, E.; Casino Martínez, A.; Veres Ferrer, E. Water Quality Indicators: Comparison of a Probabilistic Index and a General Quality Index. The Case of the Confederación Hidrográfica Del Júcar (Spain). *Ecol. Indic.* **2010**, *10*, 1049–1054. [[CrossRef](#)]
40. Morgan, G.A.; Barrett, K.C.; Leech, N.L.; Gloeckner, G.W. *IBM SPSS for Introductory Statistics: Use and Interpretation*, 6th ed.; Routledge: London, UK, 2019; ISBN 978-1-00-000491-5.
41. Dimri, T.; Ahmad, S.; Sharif, M. Time Series Analysis of Climate Variables Using Seasonal ARIMA Approach. *J. Earth Syst. Sci.* **2020**, *129*, 149. [[CrossRef](#)]
42. Viccione, G.; Guarnaccia, C.; Mancini, S.; Quartieri, J. On the Use of ARIMA Models for Short-Term Water Tank Levels Forecasting. *Water Supply* **2019**, *20*, 787–799. [[CrossRef](#)]
43. Mahla, S.K.; Parmar, K.S.; Singh, J.; Dhir, A.; Sandhu, S.S.; Chauhan, B.S. Trend and Time Series Analysis by ARIMA Model to Predict the Emissions and Performance Characteristics of Biogas Fueled Compression Ignition Engine. *Energy Sources Part A Recovery Util. Environ. Eff.* **2023**, *45*, 4293–4304. [[CrossRef](#)]
44. Ljung, G.M.; Box, G.E.P. On a Measure of Lack of Fit in Time Series Models. *Biometrika* **1978**, *65*, 297–303. [[CrossRef](#)]
45. Cruz, D.; Pimentel, M.; Russo, A.; Cabral, W. Charge Neutralization Mechanism Efficiency in Water with High Color Turbidity Ratio Using Aluminium Sulfate and Flocculation Index. *Water* **2020**, *12*, 572. [[CrossRef](#)]
46. Stevenson, M.; Bravo, C. Advanced Turbidity Prediction for Operational Water Supply Planning. *Decis. Support Syst.* **2019**, *119*, 72–84. [[CrossRef](#)]
47. García-Ávila, F.; Avilés-Añazco, A.; Sánchez-Cordero, E.; Valdiviezo-González, L.; Ordoñez, M.D.T. The Challenge of Improving the Efficiency of Drinking Water Treatment Systems in Rural Areas Facing Changes in the Raw Water Quality. *S. Afr. J. Chem. Eng.* **2021**, *37*, 141–149. [[CrossRef](#)]
48. Saalidong, B.M.; Aram, S.A.; Otu, S.; Lartey, P.O. Examining the Dynamics of the Relationship between Water pH and Other Water Quality Parameters in Ground and Surface Water Systems. *PLoS ONE* **2022**, *17*, e0262117. [[CrossRef](#)]
49. Verma, A.K.; Singh, T.N. Prediction of Water Quality from Simple Field Parameters. *Environ. Earth Sci.* **2013**, *69*, 821–829. [[CrossRef](#)]
50. Girardi, R.; Pinheiro, A.; Garbossa, L.H.P.; Torres, É. Water Quality Change of Rivers during Rainy Events in a Watershed with Different Land Uses in Southern Brazil. *RBRH* **2016**, *21*, 514–524. [[CrossRef](#)]
51. Wang, K.; Gelda, R.K.; Mukundan, R.; Steinschneider, S. Inter-Model Comparison of Turbidity-Discharge Rating Curves and the Implications for Reservoir Operations Management. *JAWRA J. Am. Water Resour. Assoc.* **2021**, *57*, 430–448. [[CrossRef](#)]

52. Besmer, M.D.; Hammes, F. Short-Term Microbial Dynamics in a Drinking Water Plant Treating Groundwater with Occasional High Microbial Loads. *Water Res.* **2016**, *107*, 11–18. [[CrossRef](#)] [[PubMed](#)]
53. Price, J.I.; Heberling, M.T. The Effects of Source Water Quality on Drinking Water Treatment Costs: A Review and Synthesis of Empirical Literature. *Ecol. Econ.* **2018**, *151*, 195–209. [[CrossRef](#)]
54. Li, D.; Bu, S.; Li, Q.; Chen, S.; Zhen, Z.; Fu, C. Water Environment Capacity Estimation and Pollutant Reduction of Yanghe Reservoir Basin in Hebei Province, China, via 0-D Water Quality Model. *Environ. Earth Sci.* **2021**, *80*, 508. [[CrossRef](#)]
55. Grochowska, J. Assessment of Water Buffer Capacity of Two Morphometrically Different, Degraded, Urban Lakes. *Water* **2020**, *12*, 1512. [[CrossRef](#)]
56. Huynh, T.T.; Kim, J.; Kim, W.; Hur, J.; Ho, Q.N.; Bi, Q.; Bui, T.V.; Kim, J.J.; Lee, S.D.; Choi, Y.Y.; et al. Dynamics of Suspended Particulate Matter in an Impounded River Under Dry and Wet Weather Conditions. *Water Resour. Res.* **2023**, *59*, e2022WR033629. [[CrossRef](#)]
57. Zeb, H.; Yaqub, A.; Ajab, H.; Zeb, I.; Khan, I. Effect of Climate Change and Human Activities on Surface and Ground Water Quality in Major Cities of Pakistan. *Water* **2023**, *15*, 2693. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.