Multivariate Analysis of Coastal Water Quality: Impact of Physicochemical Parameters at Beach Outfalls

Ohoud Bushaibah* **, Adel Al-Haddad, Fatemah Dashti**

* Environmental Pollution and Climate Program, Environmental and Life Sciences Research Center, Kuwait Institute for Scientific Research. P.O. Box 24885, Safat, 13109,

Email: *ashaiba@kisr.edu.kw

Orchid ID number: *https://orcid.org/0000-0002-5704-5491

Corresponding Author*: Ohoud Bushaibah.

ABSTRACT:
Maintaining coastal water quality is vital to environmental health, public safety, and ecosystem resilience. This study employs a variety of statistical methods, including T-tests, single-factor ANOVA, regression analysis, and Principal Component Analysis (PCA), to investigate the relationships between critical physicochemical parameters—temperature, pH, dissolved oxygen (DO), and electrical conductivity (EC)—at two beach outfalls, OC3 and OC18.The aim is to discern the water quality characteristics of the two outfalls to identify relationships and trends between these parameters. T-tests and ANOVA revealed significant differences in electrical conductivity ($p = 0.000125$, $F = 20.549$) and dissolved oxygen ($p = 0.008305$, $F = 7.640$) between OC3 and OC18, illustrating that these parameters vary considerably between sites. However, there were no discernible variations in pH or temperature. Regression analysis revealed that at OC3, none of the predictors (pH, EC, or DO) substantially predicted water temperature. At OC18, pH was a significant predictor ($p = 0.0016$), indicating an associated connection, while EC and DO were not. PCA determined that the primary sources of variability were pH, temperature, and electrical conductivity. The first two principal components (PC1 and PC2) accounted for 81.3% of the total variation, with PC1 (45.2%) influenced by temperature and pH and PC2 (36.1%) by electrical conductivity and dissolved oxygen. OC3 was associated with slightly higher temperatures and pH, whereas OC18 showed higher electrical conductivity and dissolved oxygen levels. These findings emphasize the significance of targeted, site-specific water quality control approaches to preserve coastal ecosystems and public health. The application of multidimensional statistical techniques allows for a more comprehensive assessment of water quality and provides valuable insights into the complicated dynamics of beach outfalls.

Keywords:-Water quality, Ecosystem health, Multivariate analysis, Physicochemical Parameters, Principal Component Analysis (PCA)

1) Introduction:

The coastal water quality is a major concern for environmental researchers, legislators, and the general populace since it has far-reaching consequences for ecosystem health, public safety, and economic activity. Coastal zones are dynamic boundaries between land and marine habitats, with intricate interactions among physical, chemical, and biological processes. [1][2][3]. Significant anthropogenic pressures, such as industrial discharges, agricultural practices, urban runoff, and climate change, might deteriorate water quality and jeopardize the sustainability of coastal ecosystems in these locations [4][5]. Effective management of water quality and preservation of marine biodiversity depends on understanding the physicochemical characteristics of coastal waters. Important factors, including temperature, pH, electrical conductivity (EC), and dissolved oxygen (DO), act as markers of water quality and shed light on the underlying mechanisms influencing coastal habitats. [1][6][7]. Temperature regulates metabolic rates and the solubility of gases; pH affects the chemical speciation and bioavailability of nutrients and pollutants; DO is critical for aerobic respiration and the survival of aquatic species; and EC reflects the ionic composition and salinity of the water. [7][8][9].

Notwithstanding the significance of these variables, to distinguish between the water quality profiles of various coastal locations and to clarify the intricate interactions between them, extensive research incorporating a variety of statistical approaches is required. Conventional methods frequently depend on univariate analysis, which may ignore the multidimensionality of data on water quality and the relationships between different components [10] [11]. Multivariate statistical methods, including Principal Component Analysis (PCA), are effective techniques for highlighting the primary drivers of variability in water quality datasets, reducing the dimensionality of the data, and spotting significant patterns [5] [12] [13] [14]. Furthermore, T-tests and ANOVA single-factor analyses offer a strong foundation for detecting values that show notable spatial variability and comprehending the possible environmental causes of these variations. Furthermore, T-tests and ANOVA single-factor analyses are employed to compare the mean values of each parameter between different locations and evaluate the statistical significance of the observed differences. These tests offer a strong foundation for determining which characteristics show notable regional variability as well as for comprehending the possible environmental factors causing these variations. [15] [16] [17]. Regression analysis is an essential statistical technique in ecological research that helps forecast important environmental outcomes and understand the intricate correlations between many physicochemical characteristics. Using this technique, researchers can measure the direction and intensity of correlations between dependent and independent variables, offering crucial insights for practical management and scientific understanding [1] [18].

The quality of coastal waters is a multidimensional issue that requires a comprehensive understanding of the interactions between several physicochemical characteristics. This study adds to this comprehension by employing a suite of statistical techniques involving T-tests, ANOVA single factor, regression analysis, and PCA to investigate the relationships between physiochemical water quality parameters consisting of temperature, pH, DO, and EC at the two beach outfalls, OC3 and OC18, on Kuwait Bay's southern beach. These outfalls provide a perfect case study for evaluating the geographical variability in coastal water quality owing to their disparate environmental contexts and anthropogenic effects [19] [20]. Furthermore, our study pinpoints significant water temperature predictors, providing insight into the complex interactions between water quality indicators and how they affect thermal dynamics. Most aquatic organisms primarily depend on the thermal properties of their immediate surroundings, rendering water temperature a crucial factor in their existence. Furthermore, dissolved oxygen levels decrease with rising temperatures, which presents serious difficulties for fish and other aquatic life forms trying to survive. Temperature fluctuations can also significantly impact microbial metabolic activities and processes, such as gas exchange rates and sedimentation properties [21] [22]. The findings of this study possess significant implications for protecting marine ecosystems and regulating water quality. Additionally, this study supports targeted management strategies by identifying the key variables affecting regional variations in water quality and elucidating their interrelationships.

2) <u>Methods and Methodology</u>:

(a) Study Area and Sampling Sites

This investigation was carried out at two beach outfalls along the coast, OC3 and OC18. According to Al-Yamani et al. (2004), Kuwait Bay is under considerable strain along the shoreline mainly due to excessive contamination generated by anthropological activities [23]. These locations were selected based on distinctions between anthropogenic influence and unique environmental contexts. OC3 is located in the interior of Kuwait Bay, at the coordinates [longitude 47.863044, latitude 29.320015], in a less developed and more stressed area than OC18, which is located at [longitude 47.989117, latitude 29.391525] in an urban area less exposed to environmental pollution in the Gulf's outer part as shown in fig (1). The samples were taken concurrently from both sites at regular intervals throughout 2022 to accurately capture changes in water quality over time. Water samples were collected from both outfalls using a standardized protocol to ensure consistency and reliability [19] [20]. A pre-cleaned Teflon bailer collected samples 0.5 meters below the water's surface at each site. In situ, portable multi-parameter water quality meters were used to measure temperature, pH, dissolved oxygen (DO), and electrical conductivity (EC) [24]. The quality control procedures were followed, and standard solutions were used to calibrate all instruments before each sample session to guarantee the dependability and correctness of the results. Furthermore, the research applied various statistical approaches to investigate the water quality at these beach outfalls and evaluate the correlations between physicochemical indicators.

T-tests, one-factor ANOVA, regression analysis, and PCA were used to detect significant differences, predictive associations, and principal components. The T-tests and ANOVA single-factor analyses were applied to compare the mean values of each parameter between various locations and assess the statistical significance of the observed differences. These analyses provide a solid basis for understanding the potential environmental variables producing these variances and for identifying which features exhibit significant regional variability [15] [16] [17].

Furthermore, principal component analysis (PCA) was used to decrease the dimensionality of the dataset and detect primary components responsible for the variance in water quality measurements. The technique above facilitates the detection of significant patterns and the distinction of water quality profiles between the two outfalls by allowing the visualization of complex datasets in fewer dimensions [25] [26].

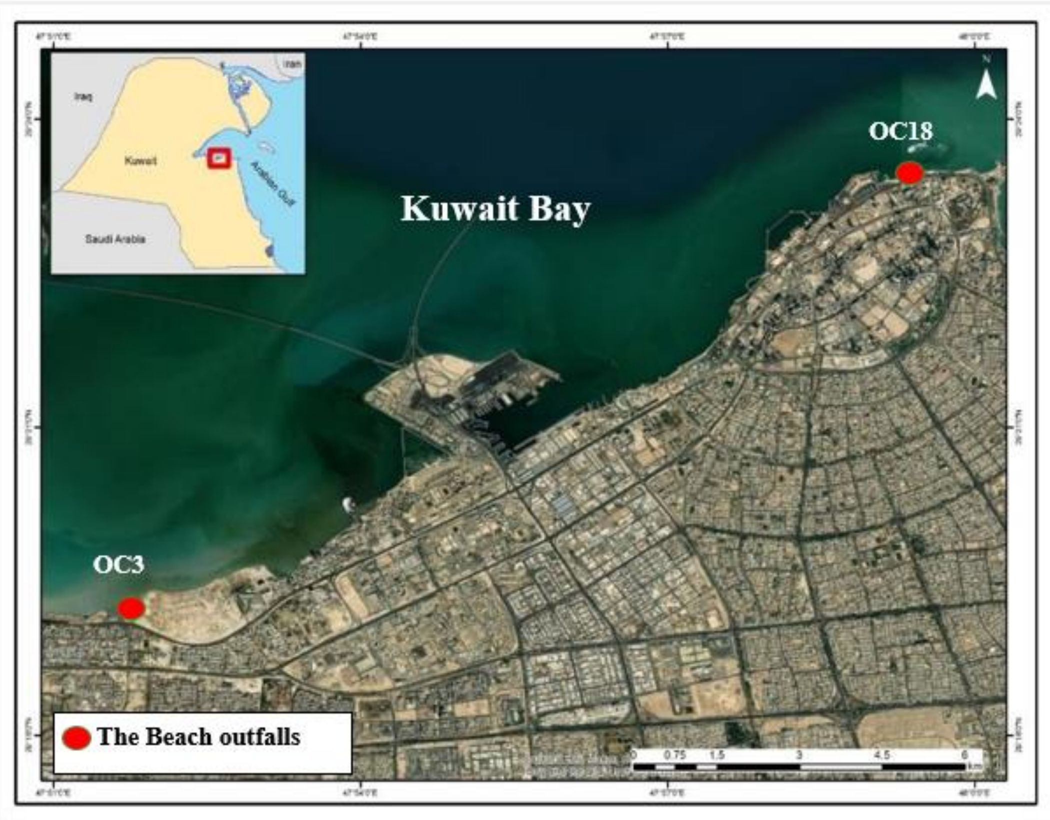


Figure 1 The Two outfalls OC3 and OC18 locations map derived from [19] [20]

The correlations between parameters and the separation of samples from OC3 and OC18 are graphically represented by the PCA biplot, which also highlights the unique environmental features of each site and the primary drivers of variability. Regression analysis was applied to simulate how physicochemical elements affect water temperature, a crucial factor affecting aquatic species' physiological and metabolic functions [1] [27]. The study's regression analysis illustrates how pH and electrical conductivity can predict water temperature, giving valuable tools for anticipating and addressing water quality issues.

<u>3) Results and Discussion</u>

(a)Physicochemical Descriptive Statistics and Trends

The analysis of environmental characteristics Temperature (Temp.), pH, electrical conductivity (EC), and dissolved oxygen (DO) at sample sites OC3 and OC18, as illustrated in Table (1), sheds light on the spatial variability of water quality within the research area these parameters can have a significant impact on the biological and chemical processes occurring inside the seawater, rendering them vital markers of the health of the aquatic environment. At OC3, the average temperature was 28.856°C with a standard error of 0.712, whereas at OC18, the average temperature was somewhat lower at 28.404°C with a more significant standard error of 1.001. Compared to OC18 (15.7°C to 35.8°C), the temperature ranges at OC3 (23°C to 34.5°C) were narrower (Fig.2), suggesting more remarkable significant temperature changes at OC18, which may be related to anthropogenic or localized environmental factors [28] [29]. The solubility of gases, the metabolic rates of aquatic creatures, and the ecosystem dynamics are all significantly impacted by temperature [30] [31]. The pH levels at both locations showed slight fluctuation, with mean values of 7.767 at OC3 and 7.628 at OC18 and standard errors of 0.055 and 0.074, respectively. The pH range was slightly more comprehensive at OC18 (6.91 to 8.13) than at OC3 (7.24 to 8.12), indicating that the water is neither too acidic nor too alkaline, as indicated in (Fig.2). Nonetheless, minute variations in pH may affect the toxicity and solubility of specific substances and minerals [32]. In contrast to OC18, which had a more comprehensive range (2.34 to 10.36 mg/l) and a higher mean DO of 3.837 mg/l with a standard error of 0.410, OC3 had a DO range of 1 to 8.73 mg/l with a mean value of 2.439 mg/l and a standard error of 0.295. Although dissolved oxygen (DO) is essential to the survival of aerobic aquatic organisms and the ecosystem's general health, lower DO levels at OC3 might represent a sign of increased organic pollution or decreased water mixing, which could result in hypoxic conditions [33]. In contradiction to the mean EC at OC18 of 30.470 mS/cm, which was considerably higher, the mean EC at OC3 was 11.431 mS/cm. The standard errors at OC18 were 4.138 and 0.719, respectively, suggesting increased variability. Compared to OC3 (4.9 to 20.39 mS/cm), The EC range at OC18 (7.71 to 73.26 mS/cm) was significantly broader, indicating a higher degree of ionic concentration and perhaps higher pollution levels at OC18 [32]. As shown in Fig (2), a high EC value may indicate dissolved salts and other inorganic elements, which can impact aquatic life and water quality [34]. The statistical analysis demonstrates substantial variations in environmental parameters between the two sites, with higher standard deviations and longer ranges at OC18 for temperature, EC, and DO, indicating increased environmental variability and possible stressors at this location. The high standard error for electrical conductivity (EC) at OC18 is attributed to distinct hydrodynamic and anthropogenic conditions at the sampling sites. OC3, with limited seawater circulation, shows stable but potentially degraded water quality, while OC18, with better circulation, exhibits more significant variability due to mixing and external influences. In addition, untreated wastewater discharges further exacerbate this variability [35] [36]. SE is a vital metric in environmental research for determining the precision of sample mean estimations, with a high SE indicating increased data variability and decreased reliability [35] [36] [37] [38] [7] [39] [40]. Considering these statistical techniques generally assume consistent variance within the data, the high standard error at OC18 shows significant variability in EC measurements, which can complicate the study using additional techniques like t-tests, ANOVA, regression, and PCA. However, low SE values imply narrower confidence ranges for the mean differences, facilitating t-tests and ANOVA, which presuppose normality and homogeneity of variances, to identify significant differences between groups [41] [42]. Cleaning the data might remove substantial variability, which is crucial for understanding the actual characteristics of the outfall. The variability in EC readings can provide valuable insights into the natural fluctuations and anomalies in the system [43]. In addition,

Table 1. The comprehensive summary of the key statistical measures for each parameter at the Two Outfalls OC3 and OC18

Parameter	Statistics	OC3	OC18
Temp. (°C)	**Mean**	28.856	28.404
	Standard Error	0.712	1.001
	Median	29.1	29.3
	Standard Deviation	3.562	5.005
	Range	11.5	20.1
	Minimum	23	15.7
	Maximum	34.5	35.8
pH	**Mean**	7.767	7.628
	Standard Error	0.055	0.074
	Median	7.82	7.75
	Standard Deviation	0.276	0.372
	Range	0.88	1.22
	Minimum	7.24	6.91
	Maximum	8.12	8.13
EC (mS/cm)	**Mean**	11.431	30.47
	Standard Error	0.719	4.138
	Median	10.98	28.73
	Standard Deviation	3.593	20.689
	Range	15.49	65.55
	Minimum	4.9	7.71
	Maximum	20.39	73.26
DO (mg/l)	**Mean**	2.439	3.837
	Standard Error	0.295	0.41
	Median	2.4	3.01
	Standard Deviation	1.477	2.051
	Range	7.73	8.02
	Minimum	1	2.34
	Maximum	8.73	10.36

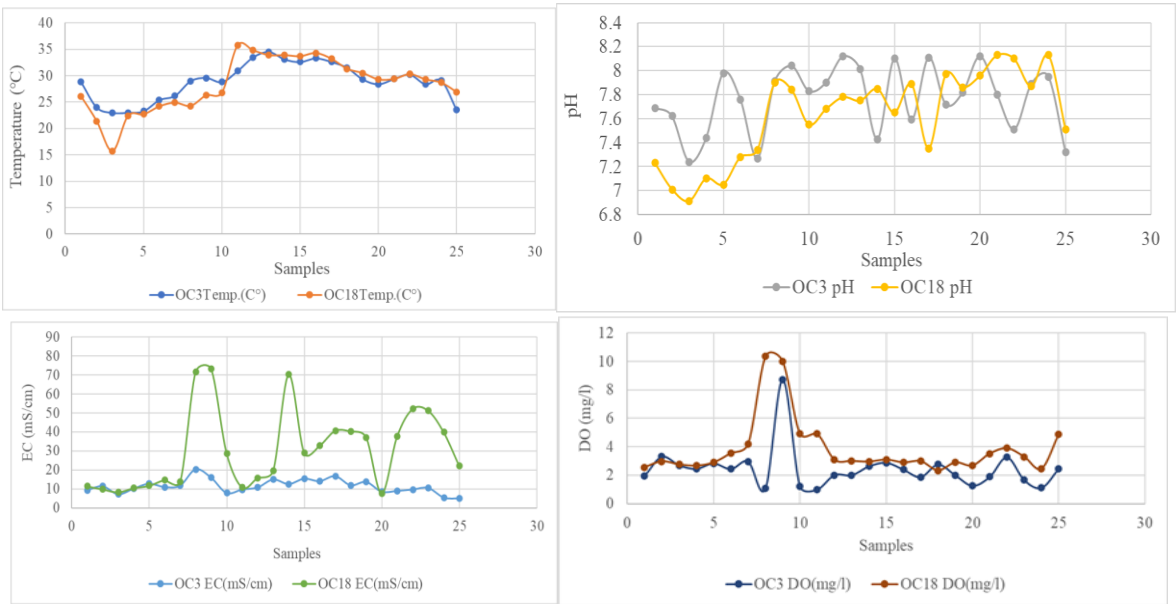


Figure 2 Compares the Temperature, pH, EC, and DO Measurements for OC3 and OC18 Outfalls.

the primary objective of this study is to analyze the raw EC data to understand its natural state and to document all potential variables and causes influencing the EC at OC18 by retaining the raw data. According to a study by Kirchner et al. (2004), this approach facilitates a more comprehensive analysis and comprehension of the data [44]. Low regression analysis standard error (SE) values for the independent variables (DO, temperature, pH) result in more accurate regression coefficient estimations, enhancing the regression model's dependability and simplifying the process of identifying essential predictors [45]. Similarly, PCA is sensitive to the size and variability of the data; low SE values show that the data are well-centered and standardized, which facilitates the extraction of significant principal components and results in more trustworthy interpretations of the data structure [13].

(b) The T-Test comparative analysis

The comparative analysis of physio-chemical parameters between the two sampling sites, OC3 and OC18, using the T-test in Table (2), revealed critical insights into the environmental conditions characterizing these locations. There was statistical indistinguishability between the mean temperature data at OC3 (28.856°C) and OC18 (28.404°C) (p = 0.7148). The two locations appear to have a homogenous thermal regime based on the lack of considerable temperature variation, which their proximity and comparable climatic factors may explain. This temperature homogeneity may also suggest that nearby anthropogenic activities or natural water bodies do not differently impact the thermal parameters of these places. On the other hand, with a p-value of less than 0.0001, there was a significant difference in the mean Electrical Conductivity (EC) values between OC3 (11.43132 mS/cm) and OC18 (30.4696 mS/cm). The noteworthy disparity implies different degrees of potential pollution and ionic content due to rising sea surface temperatures. The increased EC at OC18 may indicate higher salinity, increased mineral runoff, or industrial effluents; thus, more research is needed to pinpoint the precise sources of contamination because it is directly proportional to the number of contaminants and dissolved salts in the water [46] [47]. There was no discernible difference (p = 0.1397) in the pH mean values of OC3 (7.7668) and OC18 (7.6276) between them. The pH values of both locations were in the neutral to slightly alkaline range, which is ideal for most aquatic life forms. Whereas the absence of a significant pH shift suggests that acidification or alkalization processes do not affect these sites differently, continued monitoring is required to spot any changes that could impact natural ecosystems [48]. Nonetheless, constant monitoring is needed to detect future changes that may harm natural ecosystems. When the mean values of OC3 (2.4392 mg/l) and OC18 (3.8368 mg/l) were compared, there was a significant difference in dissolved oxygen levels (p = 0.0083), which may be explained by the influence of increasing temperatures on lowering dissolved oxygen concentrations [46]. The higher DO content at OC18 compared to OC3 would indicate better aeration or lower levels of organic contamination; nonetheless, the regional fluctuation of DO, an essential aspect of aquatic life, might affect aquatic species' distribution and general health [49]. A visual representation of the mean temperature values for the OC3 and OC18 datasets is provided in the box-and-whisker diagram Figure 3. The data does not significantly differ between the two datasets; nevertheless, OC18 has a slightly higher mean temperature than OC3, which could be attributed to the two groups of parameters' varying environmental conditions. Comparably, a statistical analysis of the pH values for OC3 and OC18 considers no discernible variation; nonetheless, the marine environment is more alkaline in OC18 due to its higher mean pH than OC3.

Table 2. The T. Test Analysis Results of the Water Quality Parameters at the Two Outfalls

Parameter	**Mean OC3**	**Mean OC 18**	**T - statistic**	**P-Value**	**Conclusion**
Temp	28.856	28.404	0.367883	0.714765	Not Significant
pH	7.7668	7.6276	1.504242	0.139664	Not Significant
EC	11.43132	30.4696	-4.5331	0.000125	**Highly Significant**
DO	2.4392	3.8368	-2.76412	0.008305	**Significant**

P-values are two-tailed, and the significant result is ($p < 0.05$)

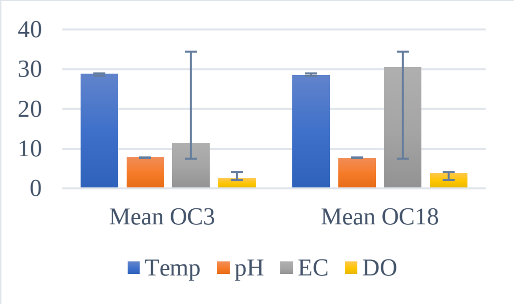


Figure 3. Comparison of Mean Values with Standard Error Bars for Temperature, pH, (EC), and (DO) between OC3 and OC18.

The mean EC varies dramatically between the two sites; OC3 has a value of around 30 mS/cm, while OC18 has a value of about 11 mS/cm. The observed disparities in environmental variables between OC3 and OC18 could be attributed to various factors, including differences in their relative locations, water sources, and anthropogenic activities in the surrounding areas. Moreover, a higher EC at OC3 can indicate more dissolved ions or salts in the water, possibly due to an increase in inadequately treated wastewater effluents in the area [19-20]. There is also considerable variation in the average DO levels; the average for OC3 is approximately 2.4 mg/L, and the average for OC18 is 3.8 mg/L. The reduced DO levels at OC3, which may be the consequence of increased organic matter or nutrient loading in that area, could additionally indicate a higher biological or chemical oxygen requirement [19-20] [28-29] [50]. Significant temporal variability, geographical variability, and ecological complexity in the system are linked to the substantial standard error in ecological complexity (EC) measurements at site OC3. Dissolved ions or salts from wastewater effluents can cause significant regional changes in EC, even in small geographic areas [51-53]. The complex and dynamic nature of the aquatic environment, where several interacting physical, chemical, and biological processes affect the distribution and concentration of dissolved ions, may also be reflected in the significant standard error at OC3 [51-53].

(c) ANOVA Comparative Analysis

The mean values of temperature, electrical conductivity (EC), pH, and dissolved oxygen (DO) were compared between the two groups, OC3 and OC18, using an ANOVA single-factor analysis (Table 3). The findings shed light on how these factors differed, were comparable across the two groups, and showed no discernible difference between OC3 and OC18's mean temperatures ($F(1,48) = 0.135$, $P = 0.715$). Considering mean temperatures for OC3 and OC18 of 28.856°C and 28.404°C, respectively, indicated that both groups' temperature conditions are statistically comparable. Nonetheless, a noteworthy distinction in electrical conductivity was noted between OC3 and OC18 ($F(1,48) = 20.549$, $P < 0.001$). An appreciable difference in the ionic content or salinity between the two groups was indicated by the mean EC for OC18 (30.470 mS/cm) being significantly higher than that for OC3 (11.431 mS/cm). The pH levels, on the other hand, did not indicate a significant difference between OC3 and OC18 ($F(1,48) = 2.263$, $P = 0.139$), with mean pH values of 7.767 for OC3 and 7.628 for OC18, indicating that the acidity or alkalinity of the environment is equivalent for both groups. According to the ANOVA single-factor analysis results, there was a significant difference ($F(1,48) = 7.640$, $P = 0.008$) in the DO levels between OC3 and OC18. In contrast, the difference in the DO mean value between OC18 (3.837 mg/l) and OC3 (2.439 mg/l) could be attributed to variations in the oxygenation conditions or anthropogenic activity levels between the two groups [28] [29] [50]. Overall, the ANOVA results

Table 3. The ANOVA Single Factor Analysis Results of the Water Quality Parameters at OC3 and OC18

Parameter	Group 1 (OC3)	Group 2 (OC18)	F-Statistic	p-Value	Conclusion
Temperature (°C)	Mean: 28.856 Variance: 12.686	Mean: 28.404 Variance: 25.054	0.135	0.715	Not Significant
pH	Mean: 7.767 Variance: 0.076	Mean: 7.628 Variance: 0.138	2.263	0.139	Not Significant
EC (mS/cm)	Mean: 11.431	Mean: 30.470	20.549	<0.001	**Highly Significant**

	Variance: 12.913	Variance: 428.054			
DO (mg/l)	Mean: 2.439 Variance: 2.183	Mean: 3.837 Variance: 4.208	7.640	0.008	**Significant**

P-values are two-tailed, and the significant results is (p < 0.05)

while temperature and pH levels are similar, there are notable changes between OC3 and OC18 regarding EC and DO levels, which guide future research into the biological and environmental variables in the coastal water surrounding the beach outfalls.

(d) Regression Analysis of OC3 Temperature

The influence of temperature on chemical reactions, physical properties, and biological processes in aquatic habitats makes it a crucial factor in studies on water quality. Temperature considerably impacts the solubility of chemicals, aquatic organisms' metabolic rates, and dissolved oxygen levels. [21] [46] [48] [54]. The approach's interpretive capacity is significantly increased by its highest statistical significance, supported further by temperature's critical role in impacting other water quality indicators and ecosystem health overall. The regression study hypothesizes a significant relationship between temperature (OC3 and OC18) and water quality indices (pH, EC, and DO) [55]. Temperature was assumed to be the dependent variable, and pH, EC (mS/cm), and DO (mg/l) were the independent variables in the multiple linear regression analyses conducted on the OC3 and OC18 datasets. The regression analysis for both OC3 and OC18 shows a substantial link between the temperatures at each site and the temperature at the other location. A comprehensive description of the model's performance is provided by the regression statistics in Table 4, which show the significant insights that the regression analysis for the OC3 dataset produced regarding the relationships between the independent variables OC3 pH, OC3 EC (mS/cm), and OC3 DO (mg/l) and the dependent variable, OC3 Temperature (°C).

The R Square value of 0.3211 suggests that the independent variables in the model can account for about 32.1% of the variance in OC3 Temperature. At the same time, the Multiple R-value of 0.5666 shows a moderately positive correlation between the predicted and actual values of OC3 Temperature. However, the model's explanatory power appears to be somewhat low, as indicated by the Adjusted R Square of 0.2241, which accounts for the number of predictors in the model. Based on the independent variables (OC18 pH, OC18 EC, and OC18 DO), the model's R Square value of 0.5020 suggests that the independent variables may account for roughly 50.2% of the variability in OC18 temperature. The F-statistic of 3.31, with a corresponding p-value of 0.0399, suggests that the regression model is statistically significant at the 5% level in OC3, as shown in Table 5, implying that the independent variables collectively have a substantial impact on the OC3 temperature. However, the F-statistic (7.0566) and significance level (p = 0.0018) indicate that the total regression model is statistically significant. It can be observed that when all predictors are zero, the baseline OC3 Temperature is not significantly different from zero Table (6) since the intercept term is not statistically significant (p = 0.6144). Among the independent variables, OC3 pH has a p-value of 0.0846, indicating that it is marginally significant at the 10% level. OC3 EC (p = 0.1297) and OC3 DO (p = 0.4644) are not statistically significant predictors at conventional significance levels. In OC18, the positive and significant coefficient (9.7590, p = 0.0016) indicates that as pH increases, OC18 temperature tends to increase [55], holding other variables constant. EC does not appear to significantly affect OC18 temperature, as the EC coefficient is very modest (-0.0091) and not statistically significant (p = 0.8738). Furthermore, the coefficient is negative (-0.6374) but not statistically.

Table 4. Regression Statistics of the Water Quality Parameters at OC3 and OC18.

Statistic	OC3	OC18
Multiple R	0.5666	0.7085
R Square	0.3211	0.502
Adjusted R Square	0.2241	0.4309

Standard Error	3.1374	3.7761
Observations	25	25

Table 5. The ANOVA Results of the Regression Model of the Water Quality Parameters at OC3 and OC18.

Source of variation	df	SS	MS	F	Significance F
OC3 Regression	3	97.75	32.58	3.31	0.0399
OC3 Residual	21	206.71	9.84		
OC3 Residual	24	304.46			
OC18 Regression	3	301.8549	100.6183	7.0566	0.0018
OC18 Residual	21	299.4347	14.2588		
OC18 Total	24	601.2896			

Significant (p = 0.1827), implying that DO has no significant effect on OC18 temperature within the measured range. In contrast to the regression analysis results for OC3, these results show that the predictor factors might have different effects on temperature depending on the situation. The data structures and contextual elements differ across the OC3 and OC18 datasets, resulting in variances in predictor significance and degree of effects.

Notably, even while the model sufficiently meets the linear regression assumptions and explains a tiny portion of the variance in OC18 temperature, pH, EC, and DO don’t seem to be significant predictors of the OC18 temperature. The residual plot for OC3 pH (Fig. 7.a) demonstrates residuals dispersed around zero, indicating an adequate fit, albeit some outliers highlight possible constraints in capturing all pH fluctuations. On the other hand, the random scatter around zero in the OC3 EC residual plot (Fig. 7.b) indicates a robust model fit in the absence of any pattern in the residuals. Comparably, the residuals in the OC3 DO plot in Fig. 7.c are centered around zero, suggesting an efficient model fit with a few outliers pointing to possible discrepancies in the DO level prediction [53]. The model's validity is further supported by the line fit plots (Figs. 7.d, e, and f) for pH, EC, and DO, which indicate that the model accurately represents the entire set of data trends. The normal probability plot (Fig. 7.g) verifies that the residuals have an almost normal distribution. On the other hand, a robust model fit with residuals distributed around zero and slight variations that have no discernible impact on the overall fit is demonstrated by the OC18 pH residual plot (Fig. 8.a). The OC18 EC residual plot (Fig. 8.b) shows residuals spread equally about zero with no discernible pattern, demonstrating the model's ability to match the data [56]. The OC18 EC residual plot (Fig. 8.b) reveals residuals evenly distributed around zero, with no discernable trend, indicating the model's ability to fit the data. The model accurately reflects EC behavior, while the residual plot for OC18 DO (mg/l) (Fig. 8.c) reveals a decent match with minimal outliers.

As with OC3, the line fit plots (d, e, and f) for OC18 in Fig. 8 show that the model accurately represents the trends in the data. The normal probability plot (Fig. 8.g) confirms the residuals' approximate normal distribution, which lends credence to the suitability of the model. Both OC3 and OC18 residual plots exhibit residuals centered around zero, indicating good model fits for pH, EC, and DO. However, OC18 demonstrates slightly fewer outliers and more consistent residual distribution than OC3. The average probability plots for both datasets confirm that the residuals follow a normal distribution, validating the models' appropriateness, and the models for both OC3 and OC18 effectively capture the trends in the data, with OC18 showing a marginally better fit. The residuals' behavior indicates that the models are robust, with random distribution around zero and minimal significant outliers. These findings highlight the complexity of environmental temperature regulation and underscore the need for comprehensive models that account for a broader spectrum of influencing factors [46] [54] [48] [55].

Table 6. The Coefficients of the Regression Model of the Water Quality Parameters at OC3 and OC18 and Their Corresponding Statistical Significance.

Predictor	Coefficient	Standard Error	t-Stat	P-value	Lower 95%	Upper 95%

Intercept	-9.8089	19.1787	-0.5114	0.6144	-49.6932	30.0755
OC3 pH	4.6162	2.5505	1.8099	0.0846	-0.6879	9.9202
OC3 EC (mS/cm)	0.3182	0.2017	1.5773	0.1297	-0.1013	0.7378
OC3 DO (mg/l)	-0.3384	0.4542	-0.7452	0.4644	-1.2829	0.606
Intercept	-43.3097	19.888	-2.1777	0.041	-84.6691	-1.9504
OC18 pH	9.759	2.6855	3.634	0.0016	4.1742	15.3438
OC18 EC (mS/cm)	-0.0091	0.0568	-0.1608	0.8738	-0.1273	0.109
OC18 DO (mg/l)	-0.6374	0.4626	-1.3781	0.1827	-1.5994	0.3245

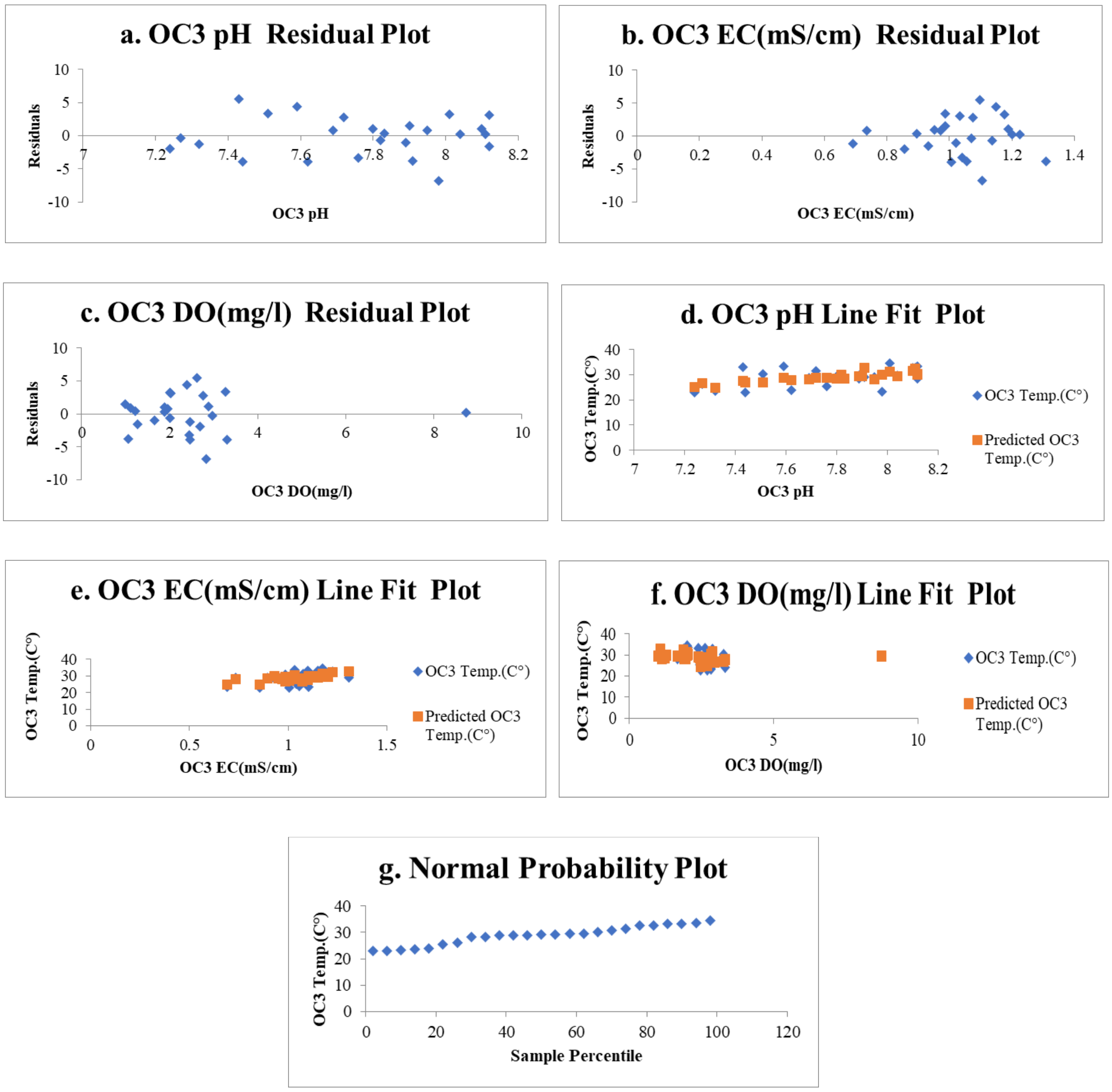
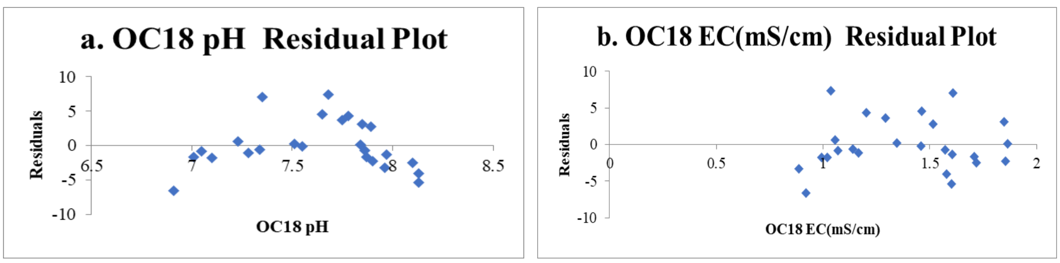


Figure 7. a. OC3 pH Residual Plot. **b**. OC3 EC Residual Plot. **c**. OC3 DO Residual Plot. **d.** OC3. pH Line Fit Plot. **e.** OC3 EC Line Fit Plot. **f**. OC3 DO (mg/l) Line Fit Plot. **g**. OC3 Normal Probability Plot



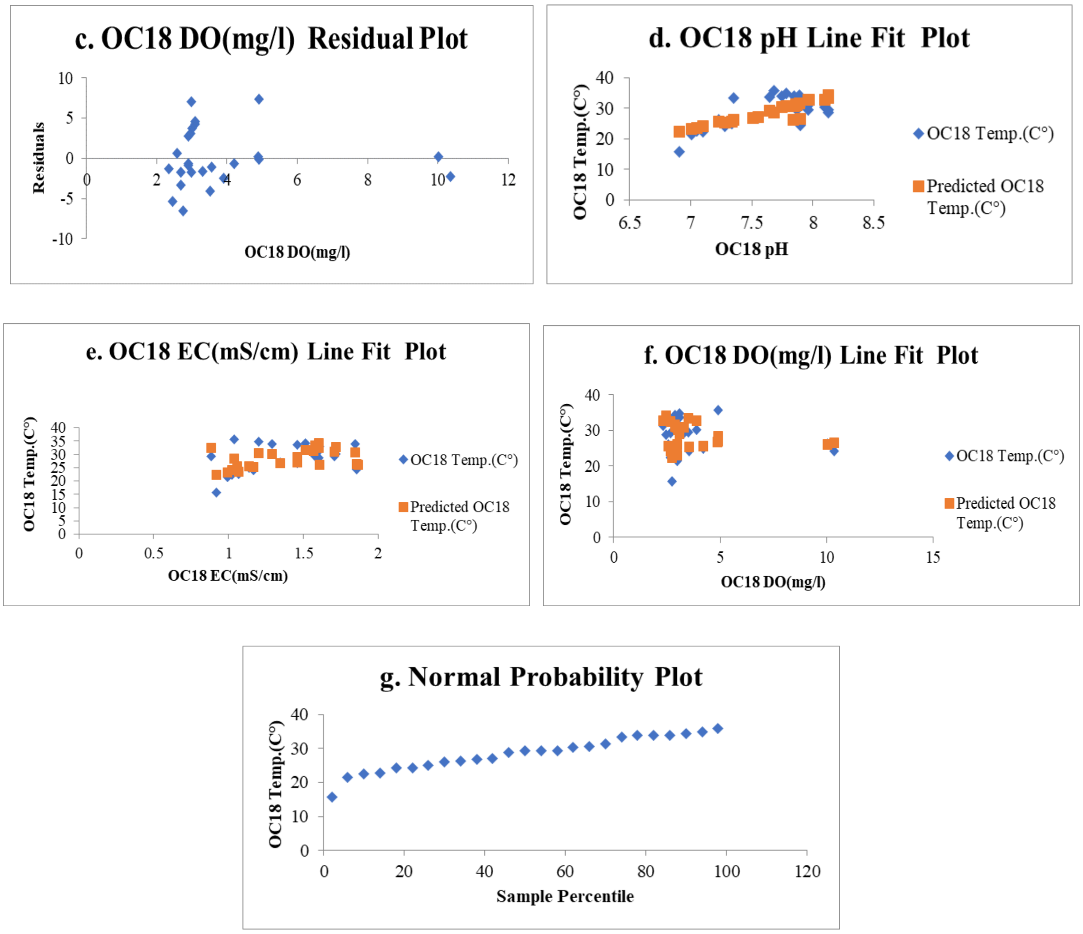


Figure 8. a. OC18 pH Residual Plot. **b.** OC18 EC Residual Plot. **c.** OC18 DO Residual Plot. **d.** OC3 pH Line Fit Plot**. e.** OC18 EC Line Fit Plot. **f.** OC18 DO Line Fit Plot. **g.** OC18 Normal Probability Plot

(e) Principal Component Analysis (PCA)

Principal component analysis (PCA) was employed to identify the principal components that account for the variance in the observed water quality measures and minimize the dataset's dimensionality. The key elements were ascertained by computing the eigenvalues and eigenvectors using standardized data (z-scores), which were used to verify consistency across variables. Additionally, a biplot was created to show the correlations between the parameters and the distinctions between the samples that came from OC3 and OC18The eigenvalues associated with each principal component are displayed in the scree plot (Fig. 9). As shown in Table 7, the first principal component (PC1) explains 45.2% of the total variance. In comparison, the second principal component (PC2) accounts for an additional 36.1%. Combining the first two significant components yields an 81.3% variance in the dataset, indicating that these two components are sufficient to represent the majority of the data [50]. The eigenvalues of the following two primary components (PC3 and PC4) fall considerably, indicating that they have minimal influence on explaining the variation. This is demonstrated by the "elbow" in the scree plot, which appears after the second component and reinforces the decision to focus on the first two principal components for further research. The PCA biplot (Fig. 10) provides a two-dimensional representation of the data regarding the first two principal components (PC1 and PC2). The biplot shows unique clustering for the two outfalls, OC3 and OC18, thereby rendering it more straightforward to distinguish between them. The principal component 1 (PC1) axis, which accounts for 45.2% of the variance, appears to capture the primary distinctions between the data from the two outfalls; on the other hand, the second principal component (PC2), which accounts for 36.1% of the variation and suggests additional underlying variability, further separates the samples. The biplot (Fig. 10) additionally demonstrated how the initial variables (TEMP, pH, EC, and DO) were loaded onto the principal components. Variables like EC and DO, which have larger vector lengths, had more impact determining the main components. The potent loading vectors and the distinct distinction between the clusters of OC3 and OC18, as shown in the biplot, further suggest that the environmental characteristics assessed help differentiate between the two outfalls. The scree plot and biplot show that the top two principal components contain the majority of the variance in the dataset, efficiently distinguishing between samples from OC3 and OC18. The efficacy of PCA in decreasing data dimensionality while keeping crucial information is highlighted by this analysis, which offers insightful information about the critical environmental parameters causing the observed variability [25] [26] [13]. PC1 and PC2 explain 81.3052% of the overall variance, illustrated in Table 7, demonstrating that these two components are essential in capturing the dataset's variability. The eigenvalues indicate the major components' variance, whereas the covariance matrix's diagonal represents the original variance. These findings highlight the significance of PCA in extracting valuable insights from complicated environmental datasets and guiding future monitoring efforts for assessing beach outfall conditions.

Table 7. Eigenvalues and Variance:

Parameter	**PC_1**	**PC_2**	**PC_3**	**PC_4**
Eigenvalue	1.8082	1.4441	0.3908	0.357
% of Variance	45.2039	36.1013	9.7695	8.9253
Cumulative (%)	45.2039	81.3052	91.0747	100

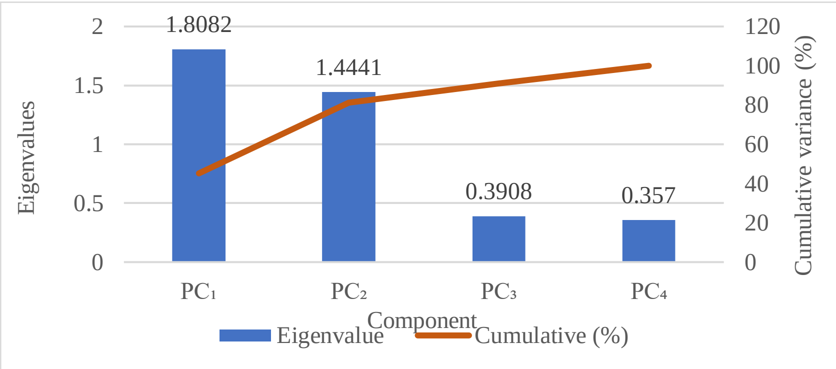


Figure 9. The Scree Plot of the PCA Analysis.

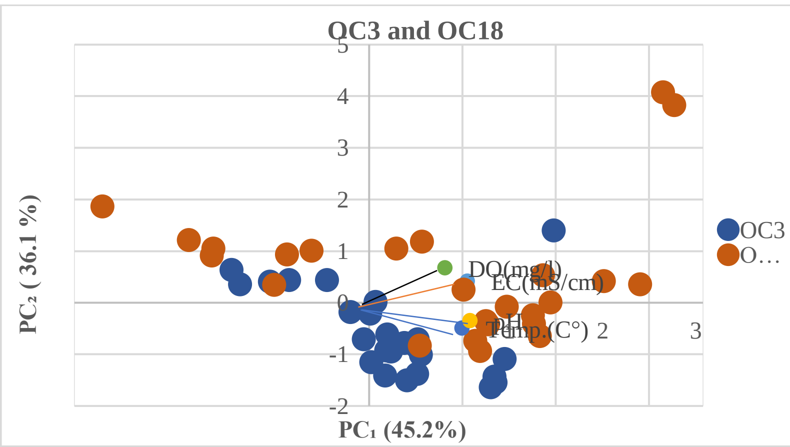


Figure 10: PCA Biplot showing the first two principal components and the clustering of the Water Quality Parameters at OC3 and OC18.

4) Conclusion

The significance of multivariate methods in comprehending the intricate dynamics of coastal water quality has contributed to shedding light on the complex interactions among the significant physicochemical factors affecting the water quality at coastal beach outfalls. A sophisticated understanding of the dynamics of the coastal environment was established by utilizing an intensive multivariate approach that included descriptive statistics, t-tests, ANOVA, regression analysis, and principal component analysis (PCA). The knowledge acquired can guide wise management and policy decisions that safeguard and maintain these essential coastal ecosystems in light of mounting anthropogenic pressures and changing environmental conditions. The primary contribution of this research is additionally contributing to the scientific literature on coastal water quality evaluation. Additionally, it provides a solid framework for future studies protecting these vital natural resources. The data analysis in this study indicated significant regional variability in parameters such as temperature, pH, EC, and DO at the OC3 and OC18 outfall locations. Significant variances in the mean values of these variables, as demonstrated by t-tests and ANOVA, highlight the coastal environment's heterogeneity and the impact of localized influences. The regression analysis revealed the interconnectedness of physicochemical processes, with water temperature emerging as a significant predictor of DO levels at each site, emphasizing the vital role of thermal conditions in coastal water quality. This research finding emphasizes the sensitivity of coastal ecosystems to heat changes and their consequences for other crucial water quality indices. The PCA biplot effectively visualizes multivariate correlations, revealing temperature, pH, and electrical conductivity as the key drivers of variability in the coastal water quality dataset. The vectors representing these factors revealed intricate interactions, providing information on their correlations and potential synergistic or antagonistic effects. While this study offers beneficial insights, it is crucial to acknowledge several limitations, such as the limited time scope and spatial coverage, which are constrained to two distinct outfall sites. Future research should address these limitations by including long-term monitoring to capture seasonal fluctuations and expanding the study to include more outfall locations and larger coastal areas for better spatial resolution. It will also be vital to look into the possible effects of climate change on the quality of coastal water, especially concerning rising temperatures and altered precipitation patterns.

6) Acknowledgement

The authors thank the Kuwait Institute for Scientific Research for providing the financial support that made this research possible and to the Water Research Center (WRC) for the invaluable guidance and unwavering support provided throughout this study. Gratitude is also extended to the research team members for their insightful engagement and remarkable contributions.

9) Conflict of interest

The authors declare no conflicts of interest related to this research.

References:

Boyd, C. E. Water quality: an introduction. Springer Nature, (2019).

1. Bhateria, Rachna, and Disha Jain. "Water quality assessment of lake water: a review." Sustainable water resources management 2 (2016): 161-173.
2. Zhu, Mengyuan, Jiawei Wang, Xiao Yang, Yu Zhang, Linyu Zhang, Hongqiang Ren, Bing Wu, and Lin Ye. "A review of the application of machine learning in water quality evaluation." Eco-Environment & Health 1, no. 2 (2022): 107-116.
3. Newton, A., Icely, J., Cristina, S., Perillo, G. M., Turner, R. E., Ashan, D., ... & Kuenzer, C. (2020). Anthropogenic, direct pressures on coastal wetlands. Frontiers in Ecology and Evolution, 8, 144.
4. Newton, Alice, John Icely, Sonia Cristina, Gerardo ME Perillo, R. Eugene Turner, Dewan Ashan, Simon Cragg et al. "Anthropogenic, direct pressures on coastal wetlands." Frontiers in Ecology and Evolution 8 (2020): 144.
5. Newton, Alice, John Icely, Sonia Cristina, Gerardo ME Perillo, R. Eugene Turner, Dewan Ashan, Simon Cragg et al. "Anthropogenic,

direct pressures on coastal wetlands." Frontiers in Ecology and Evolution 8 (2020): 144.
6. Chapman, Deborah V. Water quality assessments: a guide to the use of biota, sediments and water in environmental monitoring. CRC Press, 2021.
7. Tanjung, Rosye Hefmi Rechnelty, and Baigo Hamuna. "Assessment of water quality and pollution index in coastal waters of Mimika, Indonesia." Journal of Ecological Engineering 20, no. 2 (2019).
8. Gorde, S. P., and M. V. Jadhav. "Assessment of water quality parameters: a review." J Eng Res Appl 3, no. 6 (2013): 2029-2035.
9. Bierman, Paul, Megan Lewis, Bertram Ostendorf, and Jason Tanner. "A review of methods for analysing spatial and temporal patterns in coastal water quality." Ecological Indicators 11, no. 1 (2011): 103-114.
10. Zhang, Xuan, Qishan Wang, Yanfang Liu, Jing Wu, and Miao Yu. "Application of multivariate statistical techniques in the assessment of water quality in the Southwest New Territories and Kowloon, Hong Kong." Environmental monitoring and assessment 173 (2011): 17-27.
11. Jolliffe, Ian T. Principal component analysis for special types of data. Springer New York, 2002.
12. Zhao, Y., X. H. Xia, Z. F. Yang, and F. Wang. "Assessment of water quality in Baiyangdian Lake using multivariate statistical techniques." Procedia Environmental Sciences 13 (2012): 1213-1226.
13. Fonseca, Dangelys, Marcus Paulo de Oliveira Correa, Ramona Rodrigues Santos, Ronald Palandi Cardoso, José Salvador da Motta Reis, and Nilo Antonio de Souza Sampaio. "Effect of pollution on physical and chemical water data: a multivariate statistical analysis." Revista de Gestão e Secretariado 14, no. 5 (2023): 7353-7366.
14. Weissgerber, Tracey L., Oscar Garcia-Valencia, Vesna D. Garovic, Natasa M. Milic, and Stacey J. Winham. "Why we need to report more than'Data were Analyzed by t-tests or ANOVA'." Elife 7 (2018): e36163.
15. Quirk, Thomas J., and Thomas J. Quirk. "One-way analysis of variance (ANOVA)." Excel 2007 for Educational and Psychological Statistics: A Guide to Solving Practical Problems (2012): 163-179.
16. Rohlf, F. James, and Robert R. Sokal. Statistical tables. Macmillan, 1995.
17. Montgomery, Douglas C., Elizabeth A. Peck, and G. Geoffrey Vining. Introduction to linear regression analysis. John Wiley & Sons, 2021.
18. Bushaibah, O., Adel Al-haddad, Mishari Khajah, Farshad Dashti and E. Ibrahim. “Preliminary Assessment of Wastewater Quality near Emergency Outfalls in Kuwait Bay.” International Journal of Environmental Science and Development (2023): n. pag.
19. Ohoud Bushaibah, Adel Al-Haddad, Fatemah Dashti. 2024. “Wastewater Effluent from Outfalls in Kuwait Bay: A Threat to Recreational Water Quality”. Migration Letters 21 (7). London, UK:84-103.
20. Verma, A. K., and T. N. Singh. "Prediction of water quality from simple field parameters." Environmental earth sciences 69 (2013): 821-829.
21. Tchobanoglus, George, Franklin Burton, and H. David Stensel. "Wastewater engineering: treatment and reuse." American Water Works Association. Journal 95, no. 5 (2003): 201.
22. Al-Yamani, Faiza Yousef, James Bishop, Essa Ramadhan, M. Al-Husaini, and A. N. Al-Ghadban. "Oceanographic atlas of Kuwait's waters." (2004).
23. Baird, R., E. Rice, and A. Eaton. "Standard methods for the examination of water and wastewaters." Water Environment Federation, Chair Eugene W. Rice, American Public Health Association Andrew D. Eaton, American Water Works Association 1 (2017): 71-90.
24. Abdi, Hervé, and Lynne J. Williams. "Principal component analysis." Wiley interdisciplinary reviews: computational statistics 2, no. 4 (2010): 433-459.
25. Greenacre, Michael, Patrick JF Groenen, Trevor Hastie, Alfonso Iodice d’Enza, Angelos Markos, and Elena Tuzhilina. "Principal component analysis." Nature Reviews Methods Primers 2, no. 1 (2022): 100.
26. Chatterjee, Samprit, and Jeffrey S. Simonoff. Handbook of regression analysis. John Wiley & Sons, 2013.
27. Devlin, Michelle J., Mark Breckels, Carolyn A. Graves, Jon Barry, Elisa Capuzzo,

Francisco P. Huerta, Fahad Al Ajmi, Mona M. Al-Hussain, William JF LeQuesne, and Brett P. Lyons. "Seasonal and temporal drivers influencing phytoplankton community in Kuwait marine waters: documenting a changing landscape in the Gulf." Frontiers in Marine Science 6 (2019): 141.
28. De Rességuier, Laure, Séverine Mary, Renan Le Roux, Théo Petitjean, Hervé Quénol, and Cornelis Van Leeuwen. "Temperature variability at local scale in the Bordeaux area. Relations with environmental factors and impact on vine phenology." Frontiers in plant science 11 (2020): 515.
29. Battino, Rubin, and H. Lawrence Clever. "The solubility of gases in liquids." Chemical Reviews 66, no. 4 (1966): 395-463.
30. Dalpadado, Padmini, Randi B. Ingvaldsen, Leif Christian Stige, Bjarte Bogstad, Tor Knutsen, Geir Ottersen, and Bjørnar Ellertsen. "Climate effects on Barents Sea ecosystem dynamics." ICES Journal of Marine Science 69, no. 7 (2012): 1303-1316.
31. Stumm, Werner, and James J. Morgan. Aquatic chemistry: chemical equilibria and rates in natural waters. John Wiley & Sons, 2013.
32. Hutchison, Vivian B., Thomas E. Burley, Kyle W. Blasch, Paul E. Exter, Gregory L. Gunther, Aaron J. Shipman, Courtney M. Kelley, and Cheryl A. Morris. US Geological Survey data strategy 2023–33. No. 1517. US Geological Survey, 2024.
33. Rhoades, J. D. "Salinity: Electrical conductivity and total dissolved solids." Methods of soil analysis: Part 3 Chemical methods 5 (1996): 417-435.
34. Fournier, Matthieu, Nicolas Massei, Michel Bakalowicz, Ludivine Dussart-Baptista, Joël Rodet, and Jean-Paul Dupont. "Using turbidity dynamics and geochemical variability as a tool for understanding the behavior and vulnerability of a karst aquifer." Hydrogeology Journal 15 (2007): 689-704.
35. Tilburg, Charles E., Linda M. Jordan, Amy E. Carlson, Stephan I. Zeeman, and Philip O. Yund. "The effects of precipitation, river discharge, land use and coastal circulation on water quality in coastal Maine." Royal Society open science 2, no. 7 (2015): 140429.
36. Lee, Dong Kyu, Junyong In, and Sangseok Lee. "Standard deviation and standard error of the mean." Korean journal of anesthesiology 68, no. 3 (2015): 220-223.
37. Emerson, Steven, and John Hedges. Chemical oceanography and the marine carbon cycle. Cambridge University Press, 2008.
38. Thompson, J., C. E. Pelc, and T. E. Jordan. "Water quality sampling methods may bias evaluations of watershed management practices." Science of the Total Environment 765 (2021): 142739.
39. Singh, Ajay S., and Micah B. Masuku. "Sampling techniques & determination of sample size in applied statistics research: An overview." International Journal of economics, commerce and management 2, no. 11 (2014): 1-22.
40. Cohen, Jacob. Statistical power analysis for the behavioral sciences. routledge, 2013.
41. Field, Andy. "Discovering statistics using IBM SPSS statistics." (2013): 115-21.
42. Kannel, Prakash Raj, Seockheon Lee, Young-Soo Lee, Sushil Raj Kanel, and Siddhi Pratap Khan. "Application of water quality indices and dissolved oxygen as indicators for river water classification and urban impact assessment." Environmental monitoring and assessment 132 (2007): 93-110.
43. Kirchner, James W., Xiahong Feng, Colin Neal, and Alice J. Robson. "The fine structure of water-quality dynamics: The (high-frequency) wave of the future." Hydrological processes 18, no. 7 (2004): 1353-1359.
44. Kutner, Michael H., Christopher J. Nachtsheim, John Neter, and William Li. Applied linear statistical models. McGraw-hill, 2005.
45. Mugwanya, Muziri, Mahmood AO Dawood, Fahad Kimera, and Hani Sewilam. "Anthropogenic temperature fluctuations and their effect on aquaculture: A comprehensive review." Aquaculture and Fisheries 7, no. 3 (2022): 223-243.
46. Frost, Matthew, John M. Baxter, Paul J. Buckley, Martyn Cox, Stephen R. Dye, and Narumon Withers Harvey. "Impacts of climate change on fish, fisheries and aquaculture." Aquatic Conservation: Marine and Freshwater Ecosystems 22, no. 3 (2012): 331-336.

47. Duarte, Carlos M., Iris E. Hendriks, Tommy S. Moore, Ylva S. Olsen, Alexandra Steckbauer, Laura Ramajo, Jacob Carstensen, Julie A. Trotter, and Malcolm McCulloch. "Is ocean acidification an open-ocean syndrome? Understanding anthropogenic impacts on seawater pH." Estuaries and Coasts 36 (2013): 221-236.
48. Brown, Cheryl A., and James H. Power. "Historic and recent patterns of dissolved oxygen in the Yaquina Estuary (Oregon, USA): importance of anthropogenic activities and oceanic conditions." Estuarine, Coastal and Shelf Science 92, no. 3 (2011): 446-455.
49. Wu, Mei-Lin, You-Shao Wang, Cui-Ci Sun, Haili Wang, Jun-De Dong, and Shu-Hua Han. "Identification of anthropogenic effects and seasonality on water quality in Daya Bay, South China Sea." Journal of Environmental Management 90, no. 10 (2009): 3082-3090.
50. Ahmed, A., N. Akter, S. Hasan, and M. Ataullah. "Spatio-temporal variations of water quality and phytoplankton diversity of the different rivers flowing within Sundarban mangrove wetland ecosystem of Bangladesh." Journal of Biodiversity Conservation and Bioresource Management 5, no. 1 (2019): 61-76.
51. Fondriest Environmental, Inc. “Conductivity, Salinity and Total Dissolved Solids.” Fundamentals of Environmental Measurements. 3 Mar 2014. Web. <
52. https://www.fondriest.com/environmental-measurements/parameters/water-quality/conductivity-salinity-tds/ >.Tyler, Robert H., Tim P. Boyer, Takuto Minami, Melissa M. Zweng, and James R. Reagan. "Electrical conductivity of the global ocean." Earth, Planets and Space 69 (2017): 1-10.
53. Howarth, Robert W. "Coastal nitrogen pollution: a review of sources and trends globally and regionally." Harmful algae 8, no. 1 (2008): 14-20.
54. Manasrah, Riyad, Mohammed Raheed, and Mohammed I. Badran. "Relationships between water temperature, nutrients and dissolved oxygen in the northern Gulf of Aqaba, Red Sea." Oceanologia 48, no. 2 (2006).
55. Harrell Jr, Frank E., Kerry L. Lee, and Daniel B. Mark. "Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors." Statistics in medicine 15, no. 4 (1996): 361-387.