Advanced Statistical Techniques and Tools for Water Quality Measurement

An Academic presentation by
Dr. Nancy Agnes, Head, Technical Operations, Statswork Group www.statswork.com
Email: info@statswork.com

TODAY'S DISCUSSION

Outline

- Introduction
- Advanced Statistical Techniques
- Future Scope

Statswork
Pioneer Statistical Consulting since 2001

Introduction

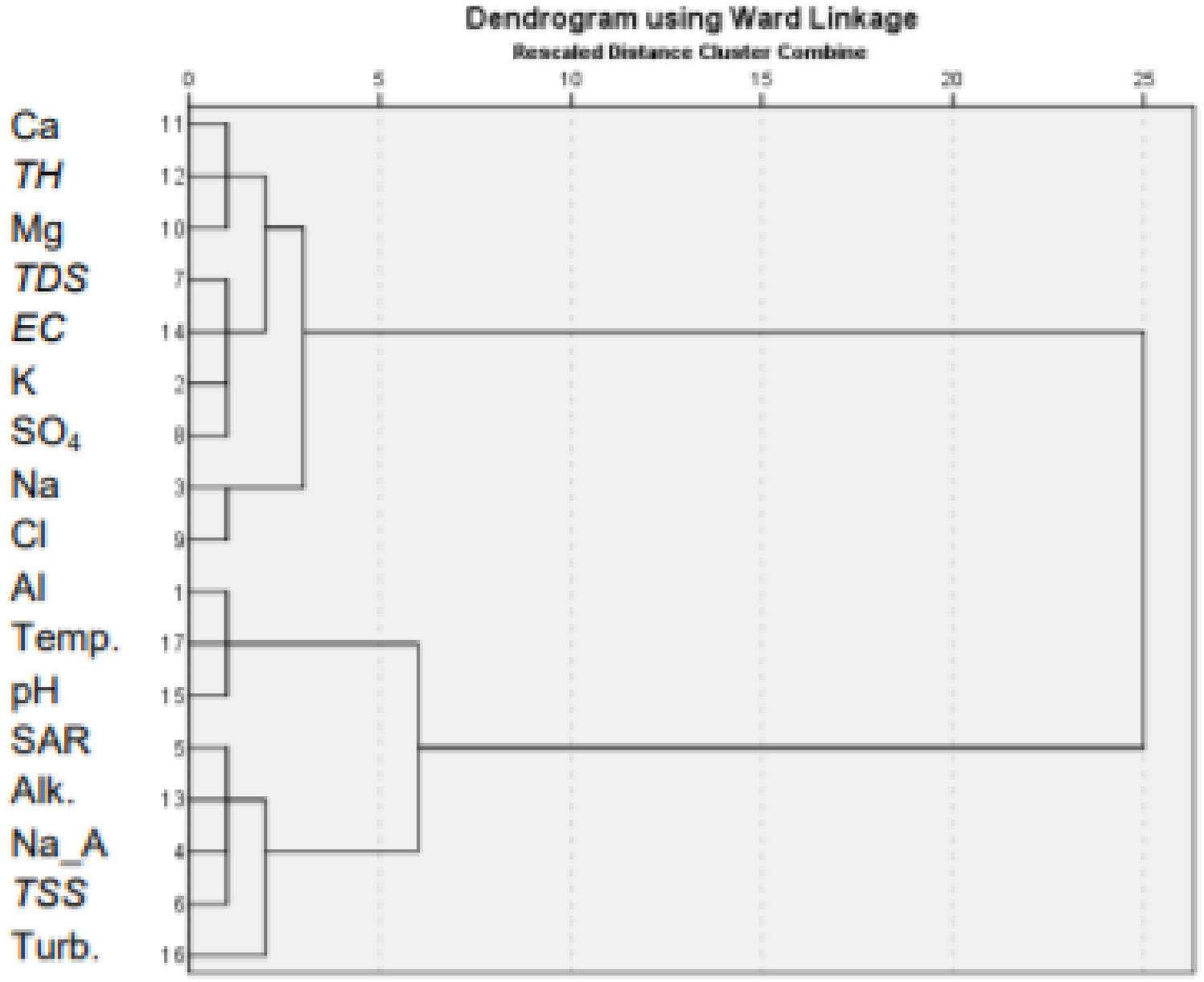
- A water quality data typically involves a large number of measurements.
- These data serve as a foundation for plant operation, modelling the process, treatment planning, and **economic assessments**.
- Since having several measurements in the data, it may affect the prediction and accuracy.
- Thus, popular statistical techniques such as principal component analysis, factor analysis, cluster analysis, etc., can be used to group the important variables for the study.

Statswork
Pioneer Statistical Consulting since 2001

Advanced Statistical Techniques

- Principle component analysis or factor analysis is basically a dimension reduction technique and it converts the original variable into latent variable.
- The association between the **variables** can be identified using Kaiser-Meger-Olkin statistic.
- Similar to factor analysis, cluster analysis also group the data or variables into homogenous clusters and after getting the clusters one can identify the correlation between the homogeneous clusters.
- The dendrogram for cluster analysis is shown in figure

Statswork
Pioneer Statistical Consulting since 2001



Statswork
Pioneer Statistical Consulting since 2001

Figure 1: Dendrogram (Source: Saad and Hamdam (2020))

- Ustaoğlu et al (2020) studied the impact of water quality due to agriculture and pollution in a tropical river and adopted principle component analysis, correlation, and clustering technique and identified that the considered river data in Turkey have good quality of water and the risk factors are not affecting the public.

- The element concentrations in the water are measured at different season and the **descriptive** summaries are as shown in the following table.

Contd...

Statswork
Pioneer Statistical Consulting since 2001

Metal	Spring	Summer	Autumn	Winter	Mean±SD	Min	Max
Al	112.63	70.67	91	58	83.07± 49	28	195
Mn	19.75	8.47	11.76	6.56	11.63±10.37	5.51	43.31
Fe	104.8	71.33	142.33	56.67	93.53±66.77	33	278
Cu	0.92	0.96	3.88	3.65	2.35±2.23	0.26	7.72
Zn	11.43	12	15.31	7.69	11.60±9.16	1.34	32.03

Statswork
Pioneer Statistical Consulting since 2001

Contd...

- Another **statistical technique** to analyse the laboratory water quality data is the analysis of variance or simply, ANOVA.

- Analysis of variance is basically identifies the mean variability between the variables.

- It uses the F-statistic and based on the test statistic value, one can decide whether the proposed research hypothesis is true or false.

- Saad and Hamdan (2020) considered a study on plant water quality and evaluated the water quality in different areas by collecting the sample water from different rivers.

Contd...

Statswork
Pioneer Statistical Consulting since 2001

- The water quality are grouped into three factors namely, biological factor, chemical factor, and physical factor.
- They adopted the factor analysis to identify the hidden relationship among the data.
- In addition, **cluster analysis** has been considered to group the data into three clusters.
- This study can also be analysed with correlation analysis or ANOVA technique to identify the relationship.

Contd...

Statswork
Pioneer Statistical Consulting since 2001

- Further, Kelly et al (2020) presented a brief review of water quality based on sanitary inspection.

- The relationship between the sanitary inspection and the water quality has been evaluated using binary **logistic** regression and correlation index.

- The following table shows the literature review and various statistical approaches.

Contd...

Statswork
Pioneer Statistical Consulting since 2001

Study	Country	Type of water source	Water quality indicatorb	Statistical modelc
Lloyd and Suyati,1989	Indonesia	PW, CW, OW, RWH, PS, BH, SW	TTC	NS
Lloyd and Bartram,1991	Java	BH, CW, OW	TTC	Linear associations, SHI
Howard et al., 2003	Uganda	PS	TTC, FS	Logistic regression, OR
Haruna et al., 2005	Uganda	PS	TC, TTC, FS	Pearson product-moment correlation coefficients
Godfrey et al., 2006	Mozambique	BH, OW	TTC, Enterococci	Logistic regression
Magrath, 2006	Sierra Leone	CW, BH, OW, PS	TTC	NS
Cronin et al., 2006	Mozambique	BH, CW, OW, SW	TTC	Linear associations
Luby et al., 2008	Bangladesh	BH	TC, TTC, EC	OR
Vaccari et al., 2010	Thailand	CW, OW	TC, EC, TTC	Linear associations
Aldana, 2010	Nicaragua	BH, CW, PS, PW, RWH	TTC, FS	Mantel-Haenszel statistical test
Parker et al., 2010	Uganda	BH, PS, CW, OW, SW, RWH	TTC	Kolmogorov–Smirnov two sample test, Kruskal–Wallis test, Spearman's rank
Barthiban and Lloyd, 2011	Maldives	OW	TTC	Linear associations, SHI
Bacci and Chapman, 2011	Ireland	BH	TTC	NS
Barthiban et al., 2012	Maldives	OW	TTC	Linear associations, SHI
Mushi et al., 2012	Tanzania	CW, OW	TC, EC, CP,SFB	Spearman's rank correlation analysis
Akoachere et al., 2013	Cameroon	CW, OW	TC, Vibrio, Staphylococcus	Pearson's Chi-square test
Sorlini et al., 2013	Chad;Cameroon	BH, OW, PW, SW	EC, Enterococci, Salmonellae	NS
Usha et al., 2014	India	CW	EC	Fischer's exact test, OR
Engström et al., 2015 S	South Sudan	BH, CW	TTC	OR, Chi-square tests
Okotto-Okotto et al., 2015	Kenya	CW, OW	TTC	Interval regression
Gerges et al., 2016	Haiti	BH, CW, OW	EC	Logistic regression
Dey et al., 2017	Bangladesh	BH	EC, TC, TTC	Multiple logistic regression
Ercumen et al., 2017	Bangladesh	BH	EC	Linear associations
Misati et al., 2017	Kenya	BH, RWH, OW, CW, NS, SW, PW	TTC	Wilcoxon rank sum test
Snoad et al., 2017I	India	BH, OW, PW, US	TTC	Logistic regression

Statswork
Pioneer Statistical Consulting since 2001

Future Scope

- There are numerous statistical procedures or techniques to analyse the water quality data.

- Since the water scarcity is increasing due to lack of rainfall in many regions, finding the water quality for the recycled water, **research** related to turning the hard water to soft water, etc. are considered as future scope of research.

Statswork
Pioneer Statistical Consulting since 2001

Contact Us

UNITED KINGDOM

+44-1143520021

INDIA

+91-4448137070

EMAIL

info@statswork.com

Statswork
Pioneer Statistical Consulting since 2001