Tropical Small Island Agriculture Management (TSIAM)

Journal Homepage: https://ojs3.unpatti.ac.id/index.php/tsiam

Volume 2	Issue 1	June (2022)	DOI: 10.30598/tsiam.2022.2.1.56	Page: 56 - 67

The Establishment of Twelve Island Groups as a Development Directive in Maluku Province

P.P.E. Papilaya[1], I.N.S. Jaya[2], T. Rusolono[2], N. Puspaningsih[2]

[1]Departemen of Forestry. Pattimura University Ambon. Mollucas-Indonesia
[2]Forest Management, IPB University Sorong Bogor. West Java-Indonesia

Email Corresponding Author: patrichp@yahoo.com

A R T I C L E I N F O

Keywords: *Small Island, Typology, Cluster analysis, Cophenetic Distance.*

Received : 22 March 2022
Revised : 13 May 2022
Accepted : 20 June 2022

A B S T R A C T

This research was created to test the existence of the determination of twelve groups of development islands that are by the development of development in the field of forestry. Biophysical, social and economic data were used to test the twelve island clusters' determination. Cluster analysis is used in testing the determination of twelve island clusters. The kaiser-Meyer-Elkin (KME) and multicollinearity tests are used to obtain data quality in cluster analysis. The cluster analysis results using twelve variables free of multicollinearity (VIF) resulted in five island clusters. Cluster analysis using Ward's method with Kophenitic distance (Cophenetic Distance) provides maximum results, where the distance between the furthest groups (between classes) of 76028208246067.70 or 97.15% and the closest distance of fellow groups 2228914432702.33 or 2.85 %.

Introduction

The management of natural resources in the archipelago has challenges because the area has limitations, such as carrying capacity, capacity, accessibility (isolation), environmental pressures, vulnerability to natural disasters, and social pressures.

Therefore, the islands are not a large expanse of land then make management arrangements through the development of typology and approach to the carrying capacity/class of land capabilities and the tendency of spatial changes due to environmental pressures (Polido et al., 2014; Beller et al., 1990)

Establishing twelve island groups in Maluku Province is one of the concrete steps to accelerate equitable development in Maluku Province. The twelve island groups are determined based on geographical proximity, cultural similarity, the potential of natural resources, similarity in the economy and economic orientation [3]. Twelve Island Groups formed in an area of ± 712 480 km2 is a wide range with several islands ± 1412 islands. The establishment of a development island group in hopes of accelerating the development process; on the other hand, the human development index in Maluku Province provides a sad picture, where only Ambon city (80.24) and central Maluku regency (70.6) are in good condition (Badan Pusat Statistik Provinsi Maluku (BPS), 2019)

The question that arises departs from the above considerations, whether the formation of twelve island groups is the right policy to accelerate Maluku Province's development. This typological development aims to improve the efficiency and effectiveness of natural resource management by maximizing productivity and minimizing the impact of damage to environmental impacts, specifically on forest resource management.

The study aimed to establish a region typology with (1) an existing island cluster approach and (2) a "natural classes" approach by establishing a region typology taking into account biophysical, social and economic conditions. In this approach, the cluster analysis approach is used to find the typology of the region regardless of the existing island cluster. Typology of the newly formed island group is the typology of a new island group formed naturally, taking into account biophysical, social and economic factors, especially those related to forest resource management.

Materials and Methods

Time and Place

This study was conducted for 3 months, starting from June to August 2021 located in Maluku Province (Figure 1).

Tools and Materials

The data used in the study came from the Central Bureau of Statistics (BPS) of Maluku Province in 2019 and land cover from the Ministry of Environment and Forestry. The data used consists of biophysical, social and economic from 118 sub-districts. The data is tabulated and detailed into fourteen (14) main variables plus spatial variables, i.e. central clusters (x, y), so that they become sixteen (16) variables.

Statistical static analysis was conducted using IBM SPSS 25 and XLSTAT 365 software. IBM SPSS 25 is used for discrete analysis and tests sample feasibility (The Kaiser-Meyer-Olkin (KMO) and multicollinearity tests).

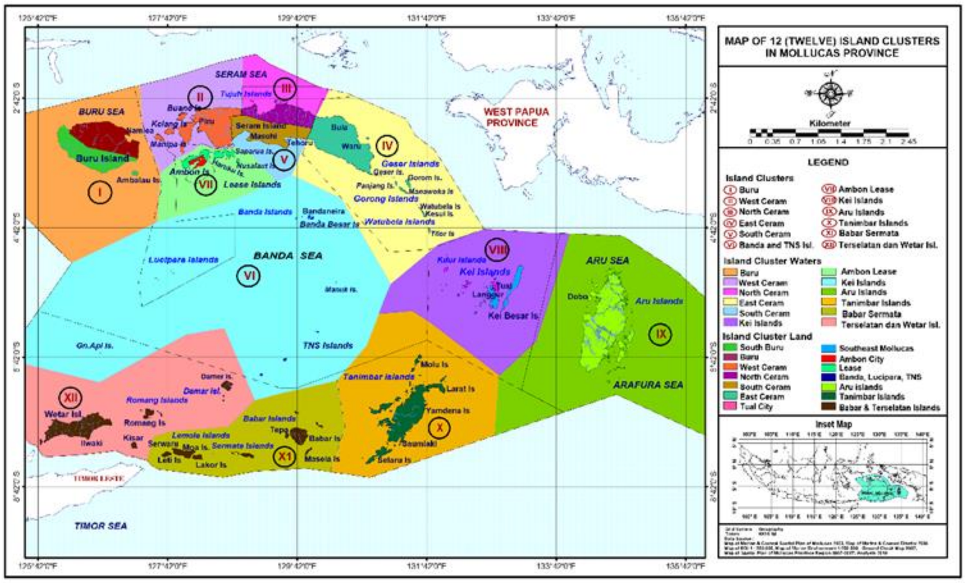

Figure 1. Twelve clusters of islands in Maluku Province

Research Methods

Cluster analysis

Cluster analysis is concerned with separating data into groups whose identities were not previously known. In general, even the correct number of groups in which data should be sorted in cluster analysis is not previously known. In contrast, the degree of similarity and difference between individual observations is used to define groups and establish group membership. Unlike discriminant analysis, it does not procedurally contain rules for establishing membership for future observations, but cluster analysis can give rise to groupings of data that the analysis discriminant may have missed and lead to empirically proper data stratification (Legendre & Legendre, 2012). A good cluster must meet two main requirements, namely 1) high homogeneity (similarity) in one cluster (within-cluster) and 2) high heterogeneity (difference) between one cluster and another (between clusters). The study used distance metrics (Euclidean distance, euclidean squared D, Bhattacharya D, Minskosky D). The method chosen is the one that produces the highest between classes and produces the high Chopenetic correlation (Johnson & Winchern, 2007). This method searches for and groups similar data (similarity) and the other. There are three (3) methods that measure similarity:

1. Measure the correlation between a pair of objects on several variables.
2. Measure the distance between two objects. The most popular measurement is the euclidean distance method.
3. Measure the associations between objects.

Results And Discussion

Descriptive analysis

The descriptive analysis aims to describe essential characteristics, which refer to numbers that provide information on the topics of interest we are studying (Pérez-Vicente & Expósito Ruiz, 2009). The results of the descriptive analysis of the twelve island groups are shown in table 1. The results provide data distribution information from fourteen variables used in cluster analysis. The highest average values are obtained on fishery production variables (x13) and the lowest on sex ratio variables (x7). This data explains the significant difference in the average distribution of each variable.

Table 1. Results of discrete analysis of twelve island clusters

Variable	Mean	Se Mean	Stdev	Min	Max
X_1	0	2 02949	24687	34075	0
X_2	17	15428	1920	3386	17
X_3	0	6191	275	698	0
X_4	0	5440	428	987	0
X_5	0	30209	2229	3870	0
X_6	766	161262	15527	20316	766
X_7	83	116	103	5	83
X_8	0	36823	3970	4774	0
X_9	0	8081	677	1167	0
X_{10}	0	29208	1417	3495	0
X_{11}	0	19803	1344	3329	0
X_{12}	0	11462	994	1519	0
X_{13}	0	78241522	1185352	8846259	0
X_{14}	0	168168	15264	27959	0
X_{15}	126	135	130	2	126
X_{16}	-8	-3	-5	2	-8

Test assumptions

Sample adequacy test

The kaiser-Meyer-olkin (KMO) measure of sampling adequacy is a statistical test that shows the proportion of variances in variables caused by underlying factors. If the KMO value is very high (close to the absolute value,1), then the sample tested is not viable. A high value (close to 1.0) generally indicates that the data should be analyzed with another test, i.e. factor analysis. Bartlett's Sphericity tested the sample correlation matrix hypothesis tested. If the Bartlett sphericity test value is less than 0.05 significance level means the sample tested did not pass the sample feasibility test and needs to be continued with another test, for example, the analytical factor test (IBM, 2019). Results of the measure of sampling adequacy (KMO) test of fourteen variables give a value of 0. 508, which means this test meets the adequacy of the sample. Bartlett's sphericity test results gave a significant value of 0.00, meaning that there was multicollinearity in fourteen variable island groups tested. The multicollinearity test is the next step to demonstrating data's feasibility in analysing twelve island clusters.

Table 2. KMO and Bartlett Test Results

KMO dan an Bartiett's Test		Nilai
kaiser-meyer-olkin measure of sampling adequacy		0.508
bartlett's test of sphericity	chi-square	173.41
siq.		0.00

Table. 3. Multicollinearity Test Results (VIF) Fourteen Variables

Variabel	Tolerance	VIF
Forest cover (Bf)	0.729	1.371
Cluster (Bf)	0.614	1.628
Area of food crops (Bf)	0.082	12.143
Horticultural plant area (Bf)	0.443	2.257
Area of plantation plants (Bf)	0.693	1.443
Population (S)	0.060	16.660
Sex ratio (S)	0.867	1.154
School-age population (S)	0.064	15.730
Number of RTH catches (S)	0.727	1.375
Food crop production (E)	0.090	11.161
Horticultural planting production (E)	0.519	1.926
Plantation crop production (E)	0.539	1.855
Fishing production (E)	0.586	1.706
Potential of livestock (E)	0.568	1.760
Central_X (Bf)	0.687	1.456
Central_Y(Bf)	0.638	1.566

Multicoleniarity test

TA multicollinearity test is performed to ascertain whether there is an interrelationship or colinearity between free variables in a regression model. An interrelational is a linear or substantial relationship between one variable and another in a regression model. Interrelationation can be seen with the value of the correlation coefficient between the free variable, the variance inflation factor (VIF) and tolerance values, the eigenvalue value and condition index, as well as the traditional value of beta coefficient error or partial regression coefficient (Rockwell, 1975; Ghozali, 2018). Table 3 shows that a VIF value that exceeds 10 indicates that the variable is experiencing a period of multicolerianity. Four variables have a value above 10, namely the area of food crop plants (biophysical), the number of residents (social), the number of the school-age population (social) and the production of food crops (E). These four variables are issued for further testing and not used for subsequent analysis for several reasons:

1. Partial regression coefficients are not precisely measured. Therefore, the standard error value is significant,
2. Small changes in data from a sample to sample will cause drastic changes in the partial regression coefficient value,
3. Changes in one variable can cause significant changes in the value of the partial regression coefficient of other variables and
4. The confidence interval value is extensive, so it would be complicated to reject the null hypothesis in a study if there is multicollinearity (Hidayah et al., 2016).

Cluster analysis

Island group typology

Discrete data for cluster analysis of twelve island cluster typologies is displayed in Table 4. These results show considerable variation between the variables forming twelve island groups.

Results of cluster classification of five island cluster typologies

The results of cluster analysis using twelve variables free of multicoleniarity (VIF) resulted in five island clusters. Cluster analysis using Ward's method with Cophenetic Distance provides maximum results, where the distance between the furthest groups (between-classes) is 76028208246067.70 or 97.15% and the closest distance between groups 2228914432702.33 or 2.85%, Table 5. The results of this classification give the meaning that this clustering process provides excellent results in the formation of island cluster typology, where the farthest distance (heterogeneity) of each cluster is classified by 97.11% and the closest distance (homogeneity) of the cluster is only 2.89%.

Table 4. Results of statistically described analysis of typological clusters of twelve island clusters

Variable	Minimum	Maximum	Mean	Std. deviation
X_2	17	15428	1920	3386
X_4	0	5440	428	987
X_5	0	30209	2229	3870
X_7	102	116	102	10
X_9	0	8081	677	1167
X_{11}	0	19803	1344	3329
X_{12}	0	11462	994	1519
X_{13}	0	78241522	1185352	8846259
X_{14}	0	168168	15264	27959
X_{15}	126	135	130	2
X_{16}	-8	-3	-5	2

Table 5. Results of Classification of Cluster Analysis 5 Island Cluster Typology

Similarity	**Absolute**	**Percent**
Within-Class	2228914432702.33	2.89%
Between-Classes	76028208246067.70	97.11%
Total	78257122678770.00	100.00%

Table 6 describes the centroid of the five typologies of island clusters resulting from clustering twelve island clusters. Clustering results produce the 11 best variables. The furthest centroid distance cluster is in variable x13 (plantation crop production), while the nearest one is in variable x16 (centroid y). This applies to cluster groups 3 and 4. The variable with the closest distance is in variable x16 with centroid group 4.

Table 6. Variable centroid class of five island group typology

Cluster	X2	X4	X5	X7	X9	X11	X12	X13	X14	X15	X16
1	3055	69	1927	103	544	412	869	4845	17577	129	-4
2	1094	2155	1689	106	1170	6549	980	1852	38690	128	-4
3	777	97	6249	107	2126	53	5731	67325656	11612	129	-5
4	97	79	18530	104	3869	118	992	1059752	13883	129	-3
5	1495	37	1224	100	261	65	915	15077	3504	131	-7

Table 7. Cluster constituent variable centroid distribution value

No.	X2	X4	X5	X7	X9	X11	X12	X13	X14	X15	X16
1	3055	69	1927	103	544	412	869	4853	17577	129	-4
2	1094	2155	1689	106	1170	6549	980	1860	38690	128	-4
3	777	97	6249	107	2125	53	5731	67325663	11611	129	-5
4	97	79	18530	103	3869	118	992	1059760	13883	129	-3
5	1495	37	1224	100	261	65	915	15085	3504	131	-7

Table 7 informs the value of variable distribution on the typology formation of five island groups. These results can be explained as follows:

a. Cluster 1. This cluster has the largest group area with low horticulture plant area, medium category and a low sex ratio compared to the number of households captured.

This cluster also has a medium to high horticultural crop production capacity and low category fishing capacity but a high potential livestock capacity. It also has a relatable location close to the centre of the cluster.

b. Cluster 2. This cluster has the largest number two group area and the largest horticulture plant area, medium category, and a high sex ratio group comparison followed by a small number of household catches. The cluster also has the most significant horticultural crop production capacity, but its plantation capacity and low fishing production capacity. However, it has the most significant potential livestock capacity in all clusters. It also has a relatable location close to the centre of the cluster.

c. Cluster 3. This cluster has a medium category group area equal to the area of horticulture plants, the area of high category plantation crops, and the highest sex ratio compared with the second largest number of fishing households. This cluster also has a low horticultural crop production capacity but has the highest crop production capacity and catch fisheries and has great livestock potential. It also has a relatable location close to the centre of the cluster.

d. Cluster 4. Cluster 4. This cluster has the smallest group area with the largest area of horticulture plants classified as moderate but has the highest area of plantation crops; the sex ratio group comparison is low but has the most significant number of fishing households. This cluster also has a low horticultural crop production capacity but a low plantation crop production capacity, the second-largest fishing potential, and moderate livestock potential. It also has a relatable location close to the centre of the cluster.

e. Cluster 5. This cluster has the second largest group area but has the lowest horticulture crop area, but has a high area of plantation crops; the sex ratio group comparison is shallow but has the most significant number of households. This cluster also has a low production capacity of plantation horticulture crops, which then has the same fishing potential as the potential of medium-sized livestock. It also has a relatable location close to the centre of the cluster.

The results of cluster analysis using fourteen variables produce 5 typologies with excellent grouping, Table 7 and 8. These results show that the grouping produced by this clustering process can be used for the next analysis process. These results also provide "validation" for what has been done before, using discriminant analysis. The use of discriminant analysis and cluster analysis in answering environmental problems has also been reported in several journals. (Hajigholizadeh & Melesse, 2017) used cluster analysis (CA) and discriminant analysis (DA) to assess water quality and evaluate its spatial and temporal variation in South Florida. To this end, the 15-year (2000–2014) data set of 12 variable water quality variables includes 16 monitoring stations. (Shrestha & Kazama, 2007) also used multivariate statistical techniques, such as cluster analysis (CA), major component analysis (PCA), factor analysis (FA) and discriminant analysis (DA), applied to the evaluation of temporal/spatial variation and interpretation of large complex water quality data sets from fuji watersheds. In line with that, some researchers have also used multivariate statistics together to solve environmental problems (Singh et al., 2004; Azhar et al., 2015)

Table 8. Results of Cluster Analysis with fourteen variables

Kluster	1	2	3	4	5
Object	81	24	2	2	9
Within-class variance	4245646175	742343475390	528431310	238313021310644	2602435174
Minimum distance to centroid	9270.929	166066.168	16254.712	10915883.412	21333.935
Average Distance to centroid	36439.833	345605.113	16254.712	10915883.412	42932.726
Maximum distance to centroid	472301.645	4034491.677	16254.712	10915883.412	97655.299

Average distance centroid typology of five island groups

Table 9 shows the distribution of centroid distances between the typologies of five groups. The closest distance centroids of the first cluster are with typology clusters 2 and 5, while the furthest distance is with typology clusters 3 and 4. The closest distance centroid second cluster is with typology clusters 1 and 5, while the furthest distance is with typology clusters 3 and 4. The third cluster only has a distance approximately as far away as the cluster camp. The closest distance centroid fifth cluster is with typology clusters 1 and 2, while the furthest distance is with typology clusters 3 and 4. This result is the same as the results of the discriminant analysis, which is shown in Figures 2 and 3 (Papilaya et al., 2021).

Table. 9. Distance between centroid five typologies of island clusters

cluster	1	2	3	4	5
1	0	22384	67320811	1055053	17489
2	22384	0	67323809	1058350	38219
3	67320811	67323809	0	66265905	67310579
4	1055053	1058350	66265905	0	1044877
5	17489	38219	67310579	1044877	0

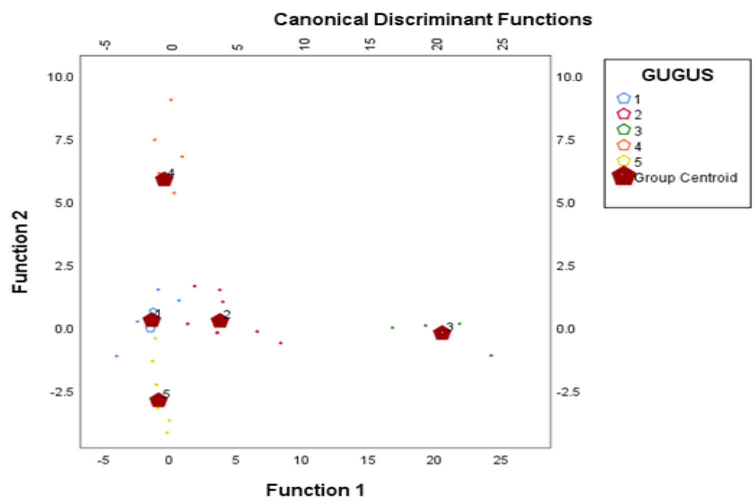

Figure 2. Distribution of centroid 5 typology of island clusters

Conclusion

The typology that can be built using the cluster approach is five classes using 14 variables. These classes have a Cophenetic correlation of 0.69 with a pretty good variance decomposition result, namely within-class 2.89% and between class 97.11%. The distance to the nearest centroid is at the centroid of clusters 1, 2 and 5, while the centroids of clusters 3 and 4 are the farthest from the other three clusters.

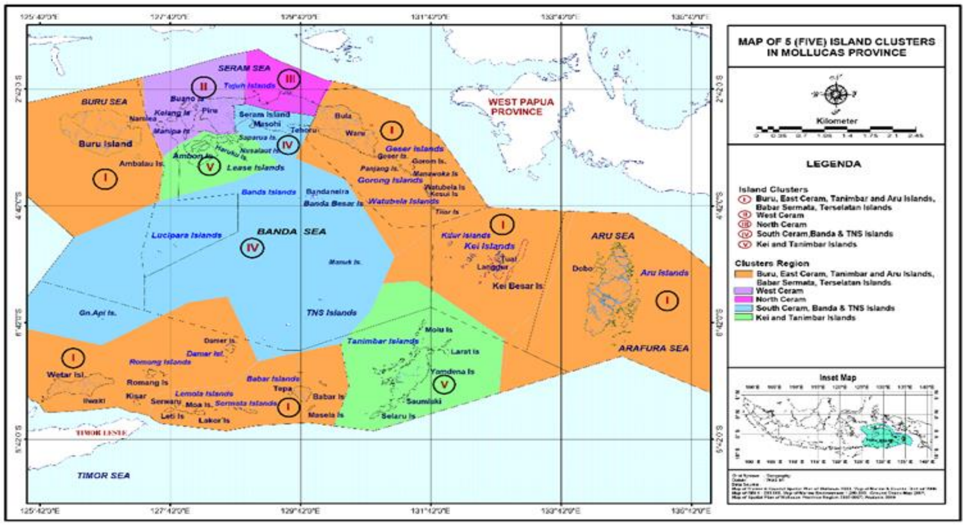

Figure 3. Five Island clusters

REFERENCES

Azhar, S. C., Aris, A. Z., Yusoff, M. K., Ramli, M. F., & Juahir, H. (2015). Classification of River Water Quality Using Multivariate Analysis. *Procedia Environmental Sciences*. https://doi.org/10.1016/j.proenv.2015.10.014

Badan Pusat Statistik Provinsi Maluku (BPS). (2019). *Indeks Pembangunan Manusia (IPM) menurut Kabupaten / Kota di Provinsi Maluku tahun 2018*. https://maluku.bps.go.id/statictable/2019/05/07/324/indeks-pembangunan-manusia-ipm-menurut-kabupaten-kota-di-provinsi-maluku-2017-2018.html

Beller, W., D'Ayala, P., & Hein, P. (1990). Overview: Sustainable development and environmental management of small islands. *Sustainable Development and Environmental Management of Small Islands. Serie: Man and the Biosphere*, *5*, 3–13.

Ghozali, I. (2018). *Aplikasi Analisis Multivariate dengan Program IBM SPSS 25 Edisi 9*. Badan Penerbit Universitas Diponegoro.

Hajigholizadeh, M., & Melesse, A. M. (2017). Assortment and spatiotemporal analysis of surface water quality using cluster and discriminant analyses. *Catena*. https://doi.org/10.1016/j.catena.2016.12.018

Hidayah, Z., Rosyid, D. M., & Armono, H. D. (2016). Planning for Sustainable Small Island Management: Case Study of Gili Timur Island East Java Province Indonesia. *Procedia - Social and Behavioral Sciences*, *227*(November 2015), 785–790. https://doi.org/10.1016/j.sbspro.2016.06.146

IBM. (2019). *KMO and Bartlett's Test*. IBM Knowledge Center.

Johnson, R. A., & Winchern, D. W. (2007). *Applied Multivariate Statistics*. Pearson Education Inc.

Legendre, P., & Legendre, L. (2012). Cluster analysis. In *Developments in Environmental Modelling*. https://doi.org/10.1016/B978-0-444-53868-0.50008-3

Papilaya, P., Suhendang, E., Jaya, I. N. S., & Rusolono, T. (2021). Quantitative approach for reclassification of the spatial cluster of archipelagos in Maluku Province for the basis of forest development. *Telkomnika (Telecommunication Computing Electronics and Control)*, *19*(5). https://doi.org/10.12928/TELKOMNIKA.v19i5.17041

Pérez-Vicente, S., & Expósito Ruiz, M. (2009). Descriptive statistics. *Allergologia et Immunopathologia*. https://doi.org/10.1016/j.aller.2009.10.005

Polido, A., João, E., & Ramos, T. B. (2014). Sustainability approaches and strategic environmental assessment in small islands: An integrative review. In *Ocean and Coastal Management*. https://doi.org/10.1016/j.ocecoaman.2014.05.005

Rockwell, R. C. (1975). Assessment of Multicollinearity: The Haitovsky Test of the Determinant. *Sociological Methods & Research*. https://doi.org/10.1177/004912417500300304

Shrestha, S., & Kazama, F. (2007). Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environmental Modelling and Software*. https://doi.org/10.1016/j.envsoft.2006.02.001

Singh, K. P., Malik, A., Mohan, D., & Sinha, S. (2004). Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) - A case study. *Water Research*. https://doi.org/10.1016/j.watres.2004.06.011