

The Application of Water Quality Monitoring Data in a Reservoir Watershed Using AMOS Confirmatory Factor Analyses

Edward Ming-Yang Wu · Chia Cheng Tsai ·
Juey Fu Cheng · Shu Lung Kuo · Wei Ting Lu

Received: 24 September 2013 / Accepted: 18 February 2014
© Springer International Publishing Switzerland 2014

Abstract This study investigates six water quality monitoring stations in the watershed of the Feitsui Reservoir. It uses nine parameters of water quality collected in an interval of two and half years for factor analyses, which was first conducted to determine four types of factors, respectively, those for organic pollution, eutrophication, seasonal influence, and sediment pollution. The analysis results effectively help to determine water quality in the watershed of the reservoir. The authors reutilize analysis of moment structures (AMOS) to acquire further results in order to confirm the goodness of fit of the previous factor analysis model. During the confirmation, we examine the hypothesized orthogonal results as well as utilize oblique rotation to explore the goodness of fit of the reflective indicators of the orthogonal rotation. As shown in the algorithm results, as long as the covariance curve is included in the four factors, no related issues are detected in the goodness of fit of reflective indicators and interior and external quality is reported with excellence. The orthogonal model, thus, stands. Additionally, when the analysis of structural

equation modeling (SEM) is conducted, sample data mismatches the hypotheses of multivariate normality. Therefore, this study adopts the generalized least square (GLS) for an algorithm. Research results of this study have been submitted to the reservoir management authorities in Taiwan for the improvement of statistical application and strategic evaluation of water quality monitoring data in order to strengthen the managerial effectiveness of water quality in watersheds.

Keywords Factor analyses · Orthogonal rotation · Oblique rotation · Analysis of covariance structures · SEM

1 Introduction

Analysis of moment structures (AMOS) implements a general approach to data analysis known as analysis of covariance structure, analysis of linear structural relations, structural equation modeling, or casual modeling. AMOS will analyze data from several populations at once. It will estimate means for exogenous variables, and it will estimate intercepts in a regression equation. The program will also compute full information maximum likelihood estimates in the presence of missing data. Any parameter can be fixed at a known value in advance, and any parameter can be constrained to be equal to any other parameters. AMOS offers a choice of the four estimation criteria discussed [1]: (a) maximum likelihood, (b) unweighted least squares, (c) generalized least squares, and (d) Browne's asymptotically distribution-free criterion.

AMOS estimates the following quantities: (a) model parameters; (b) standardized regression weights; (c) squared multiple correlation for each endogenous variable in the model, indicating the proportion of the variance of that variable that is accounted for by the remaining variables in the model; (d) total effects; and (e) regression weights for regressing unobserved variables on the observed variables (factor score weights) [4].

E. M.-Y. Wu · J. F. Cheng · W. T. Lu
Department of Civil and Ecological Engineering, I-Shou University,
No. 1, Sec. 1, Syecheng Rd., Daishu District,
Kaohsiung City 840, Taiwan, Republic of China

E. M.-Y. Wu
e-mail: edmywu@isu.edu.tw

C. C. Tsai
Department of Marine Environmental Engineering,
National Kaohsiung Marine University,
Kaohsiung City 811, Taiwan, Republic of China

S. L. Kuo (✉)
Kelee Environmental Consultant Corporation, 6F.-2, No. 288-8,
Sinya Road, Kaohsiung City 806, Taiwan, Republic of China
e-mail: singsuey@ms28.hinet.net

C. C. Tsai
International Wave Dynamics Research Center, National Cheng
Kung University, Tainan City 901, Taiwan, Republic of China

Multivariate monitoring methods that consider all available data simultaneously can extract key information about the relationships and combined effects of environmental pollutants. Additionally, when failures occur in environmental quality management systems, univariate monitoring methods are often inadequate to identify causes because the signal-to-noise ratio is very low in each pollutant measurement. But multivariate monitoring can improve the signal-to-noise ratio through averaging, resulting in a more realistic evaluation of the environmental context [13].

In the chemometrics area, multivariate statistical techniques have become one of the most active research areas in modeling and analysis over the last decade [15, 21, 7]. However, to the authors' knowledge, only limited research on the effectiveness of multivariate models for the assessment and management of air pollution has been conducted thus far [16–18]. Additionally, multivariate statistical techniques such as cluster analysis (CA), factor analyses (FA), principal component analysis (PCA), and discriminant analysis (DA) have been widely used as unbiased methods in the analysis of water quality data for drawing meaningful information [26, 19]. The multivariate treatment of data is widely used to characterize and evaluate surface and freshwater quality, and it is useful for evidencing temporal and spatial variations caused by natural and anthropogenic factors linked to seasonality [11, 25].

The functions of reservoir watershed are mainly divided into four types: (i) source of water treatment plant; (ii) water used to maintain sustainable development of a water body; (iii) recreation and leisure including direct contact (swimming) and indirect contact (boating); and (iv) agricultural and other industrial uses. At present, due to improper development and use of land, some reservoir watersheds in Taiwan have encountered the issue of eutrophication. This has a huge impact on water use and flood prevention, as well as shortening the life of reservoirs [12].

This study mainly utilizes AMOS to confirm the goodness of fit of the previous factor analysis model. The basic purpose of AMOS is to perform structural equation modeling (SEM) that confirms models or hypothesized model drawings to be used as analysis of covariance structures and analysis of causal modeling. This study, before the application of AMOS, utilized factor analysis of multivariate statistical analysis to identify common factors for six water quality monitoring stations and selected nine important parameters of water quality: water pH (pH), temperature (Temp), dissolved oxygen (DO), biochemical oxygen demand (BOD), suspended solids (SS), anionic surfactant (MBAS), ammonia nitrogen (NH₃-N), total phosphorous (TP), and chlorophyll (Chloro) to examine the correlation. Additionally, since various results have been identified for the application of factor analysis of multivariate statistical analysis to the analysis of water quality and time series studies at a single watershed in a reservoir [22, 23, 25] and there is lack of the use of AMOS to confirm research results and build goodness of fit of models, the authors are

motivated to research on this topic. This study aims to build a standard set of methods that can be applied by the reservoir authorities in order to dramatically improve the application of statistical analysis of water quality data as well as the enactment of managerial strategies of watershed in reservoirs.

2 Methodology

2.1 The Application of AMOS Software

AMOS is a set of analysis software. In this study, the AMOS 22.0 version of English edition is used and developed by IBM Corporation [2]. Besides, AMOS can share the same database with SPSS software. Along with universality of SPSS and SEM, Linear Structural Relations (LISREL) has been regarded as the same with AMOS.

2.2 Analysis Background of AMOS

Before the AMOS analysis, this study utilized parameters proposed by Wu and Kuo [24] to identify nine parameters of water quality with the SPSS application to six water quality monitoring stations in the watershed of the Feitsui Reservoir in Northern Taiwan. The results of the factor analyses discovered four main factors that influence water quality in the watershed: "organic pollution," "eutrophication," "seasonal," and "sediment pollution" as shown in Table 1. The above factor analyses are conducted with orthogonal rotation. In Table 1, four main factors were highlighted respectively to interpret the water quality parameters, such as the BOD, DO, and NH₃-N in factor 1. The factor loading of the three water quality parameters is 0.771, 0.667, and 0.636, respectively, which reveals the factor loadings are high. It also signifies that there is higher correlation of pollution in watershed area around the reservoir. The factor loading of Chloro is 0.413

Table 1 Matrix of water quality factor loadings for the watershed

Parameters	Factors			
	Organic pollution	Eutrophication	Seasonal	Sediment pollution
BOD	0.771	0.045	0.005	-0.135
DO	0.667	0.157	-0.332	0.356
NH ₃ -N	0.636	0.123	0.321	0.174
Chloro	0.413	0.075	0.311	-0.386
TP	0.057	0.752	0.044	0.002
MBAS	0.147	0.724	0.107	0.070
Temp	0.035	0.156	0.882	-0.028
SS	-0.002	-0.246	0.120	-0.673
pH	0.175	-0.323	0.368	0.653

because the loading value is not as high as others, and it means that there is no significant pollution that occurred by Chloro in this factor, either in behavior or level. This study also utilizes oblique rotation to confirm whether validity coefficient and factor load help to improve the goodness of fit of reflective indicators distributed to orthogonal rotation.

2.3 Study Area

The selected scope is six water monitoring stations in the watershed of the Feitsui Reservoir in Northern Taiwan. It consists of a total area of 303 m² that covers approximately 30 km of Shungxi, Shiding, and Xindian of New Taipei City and Taipei City. The Feitsui Reservoir is the second largest reservoir in Taiwan and is located in Shi Creek. It is the water divide for the six branches that run to the upstream area in Xindian. The purpose for this reservoir is to provide water for household use in Taipei City. The watershed area is the only

protected one in the reservoir. In order to lessen the burden brought by recreation use on the deteriorating water quality, the Feitsui Reservoir Administration, in 2001, banned water recreation activities in the watershed area. Figure 1 shows the geographical location of the Feitsui Reservoir in Northern Taiwan (w1~w6 indicate the monitoring sites of the six water quality monitoring stations).

2.4 Factor Analyses

FA is a useful tool for extracting latent information, such as relationships between variables that are directly observable [10]. The original data matrix is decomposed into the product of a matrix of factor loadings and a matrix of factor scores plus a residual matrix. In general, by applying the eigenvalue-one criterion, the number of extracted factors is less than the number of measured features. So, the dimensionality of the original data space can be decreased by means of FA. After



Fig. 1 Geographical location of the Feitsui Reservoir watershed

rotation of the factor loading matrix (e.g., by varimax rotation), the factors can often be interpreted as origins or common sources. Factor analyses were performed on a correlation matrix of rearranged data in order to explain the structure of the underlying data set. The correlation coefficient matrix measures how well the variance of each constituent can be explained by the relationship with each of the others [14]. Then, the variance/covariance and correlation coefficients of the variables were computed. The model for the relevant factor analyses is shown below:

$$\begin{aligned}
 x_1 &= \mu_1 + \ell_{11}f_1 + \ell_{12}f_2 + \dots + \ell_{1q}f_q + e_1 \\
 x_2 &= \mu_2 + \ell_{21}f_1 + \ell_{22}f_2 + \dots + \ell_{2q}f_q + e_2 : \\
 x_i &= \mu_i + \ell_{i1}f_1 + \ell_{i2}f_2 + \dots + \ell_{iq}f_q + e_i : \\
 x_p &= \mu_p + \ell_{p1}f_1 + \ell_{p2}f_2 + \dots + \ell_{pq}f_q + e_p
 \end{aligned} \tag{1}$$

Table 2 SEM evaluation indices and standard for the goodness of fit of model

Statistical test	Standard or threshold value of goodness of fit in AMOS test
Absolute goodness of fit index	
X^2	$p > 0.05$
GFI	>0.90
AGFI	>0.90
RMR	<0.05
RMSEA	<0.05 (good goodness of fit), <0.08 (fair goodness)
NCP	90 % confidence interval includes 0, the smaller the better
ECVI	Theoretic ECVI value shall be smaller than the independent ECVI value
Incremental goodness of fit index	
NFI	>0.90
RFI	>0.90
IFI	>0.90
TLI	>0.90
CFI	>0.90
Parsimony goodness of fit index	
PGFI	>0.50
PNFI	>0.50
NC	1 < NC < 3 indicates the existence of parsimony goodness of fit index. NC > 5 indicates that the model is required to be modified; theoretical AIC value shall be smaller than the independent AIC value
AIC	AIC value
CAIC	The theoretical CAIC value shall be smaller than the independent CAIC value

GFI goodness of fit index, *AGFI* adjusted goodness-of-fit index, *RMR* root mean squared residual, *RMSEA* root mean square error of approximation, *ECVI* expected cross-validation index, *RMSEA* root mean square error of approximation, *NCP* non-centrality parameter, *AIC* Akaike information criterion, *CAIC* consistent Akaike information criterion, *PGFI* parsimony goodness of fit index, *PNFI* parsimony normed fit index, *NC* non-centrality, *RFI* relative fit index, *IFI* incremental index of fit, *TLI* Tucker–Lewis index, *CFI* comparative fit index

Table 3 Results of factor analyses and the variance explained

Components	Initial eigenvalues	% of total variance	Cumulative variance %
1	2.016	22.440	22.240
2	1.275	14.171	36.571
3	1.151	12.794	49.365
4	1.016	11.291	60.656
5	0.883	9.807	70.464
6	0.795	8.830	79.293
7	0.768	8.532	87.826
8	0.621	6.903	94.729
9	0.471	5.271	100.0000

where f_1, \dots, f_q are common factors contained in each variable x_i ; e_i is a special factor contained only in the i th variable (x_i); and ℓ_{ij} is the loading of i th factor to the j th common factor (f_j).

2.5 Confirmation of AMOS Software Program

This study applies factor analyses to identify four factors related to organic pollution, eutrophication, seasonal, and sediment pollution; these four factors are analyzed with AMOS modeling software. AMOS is mainly used to analyze p values and chi-squares. We apply AMOS to analyze the goodness of fit, reliability, and validity of four factors under factor analyses. AMOS simulates the goodness of fit of factor analysis model from the perspective of confirmation. Therefore, a p value higher than 0.05 is defined as good [2, 5] and is different from a statistical p value defined as lower than 0.05; it has its own principle variance and explanation. In this study, p value and chi-square are used to confirm the goodness of fit, but the model also requires that it should be confirmed with diverse indices for self tests in order to

Table 4 Evaluation of normal distribution (factor analyses)

Variable	Min	Max	Skew	C.R.	Kurtosis	C.R.
Temp	0.000	1.000	-0.710	-4.859	0.767	2.625
SS	0.000	1.000	5.921	40.520	41.255	141.163
pH	0.000	1.000	-0.200	-1.366	0.448	1.532
Chloro	0.000	1.000	2.266	15.508	5.349	18.302
MBAS	0.000	1.000	8.423	57.645	97.548	333.785
TP	0.000	1.000	2.714	18.571	10.194	34.881
NH3N	0.000	1.000	2.407	16.474	7.884	26.978
DO	0.000	1.000	-0.386	-2.639	0.336	1.150
BOD	0.000	1.000	2.938	20.105	15.642	53.521
Multivariate					191.214	113.897

C.R. critical ratio

understand the fairness of the internal quality of the model. The evaluation standard is shown in Table 2 [9].

Before the evaluation of the goodness of fit, the model needs to be examined with an offending estimate. In SEM analysis, the “maximum and minimum” standard deviation estimates are often not defined [2, 8, 20].

2.6 AMOS Modification Indices

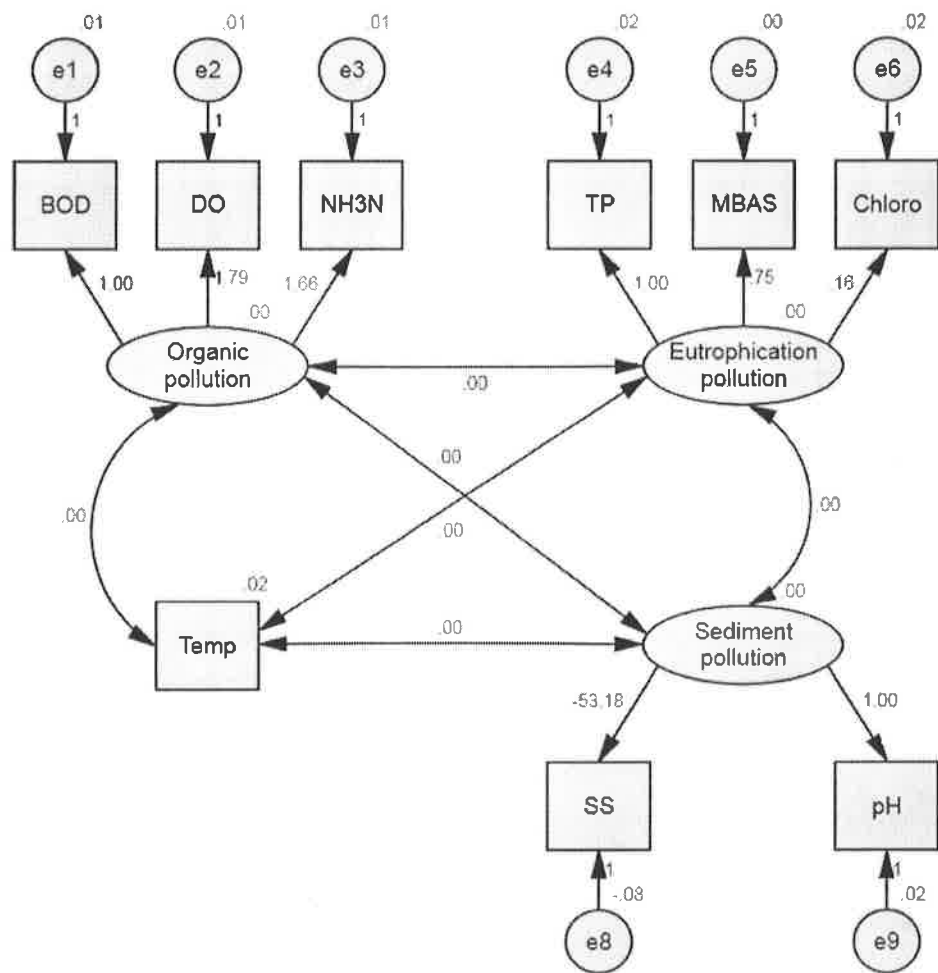
Modification indices are important clues used to examine sequence errors, but the value of modification indices needs to be discussed for the modification requirement; as proposed

by scholars [3, 8, 27], modification is needed when modification indices are higher than 3.84.

2.7 AMOS Specification Search

Specification search is the last modification tool used in this research. The purpose of the model search function of AMOS is to conduct various possibility analyses for the selected covariance curve in the hypothesized model in order to identify the optimal goodness of fit with the highest *p* value with the connected combination of covariance curves among the four factors [3, 8].

Fig. 2 Results of orthogonal rotation (factor analyses)



Orthogonal Rotation Factor Analysis
 Unstandardized estimates
 Chi-Square = 93.003
 df = 28 P-value = .000

Table 5 Covariance of modification indices

			M.I.	Par change
NH ₃ -N	↔	Temp	10.011	0.003
DO	↔	Temp	12.642	-0.004
BOD	↔	Chloro	5.540	0.002

3 Results and Discussion

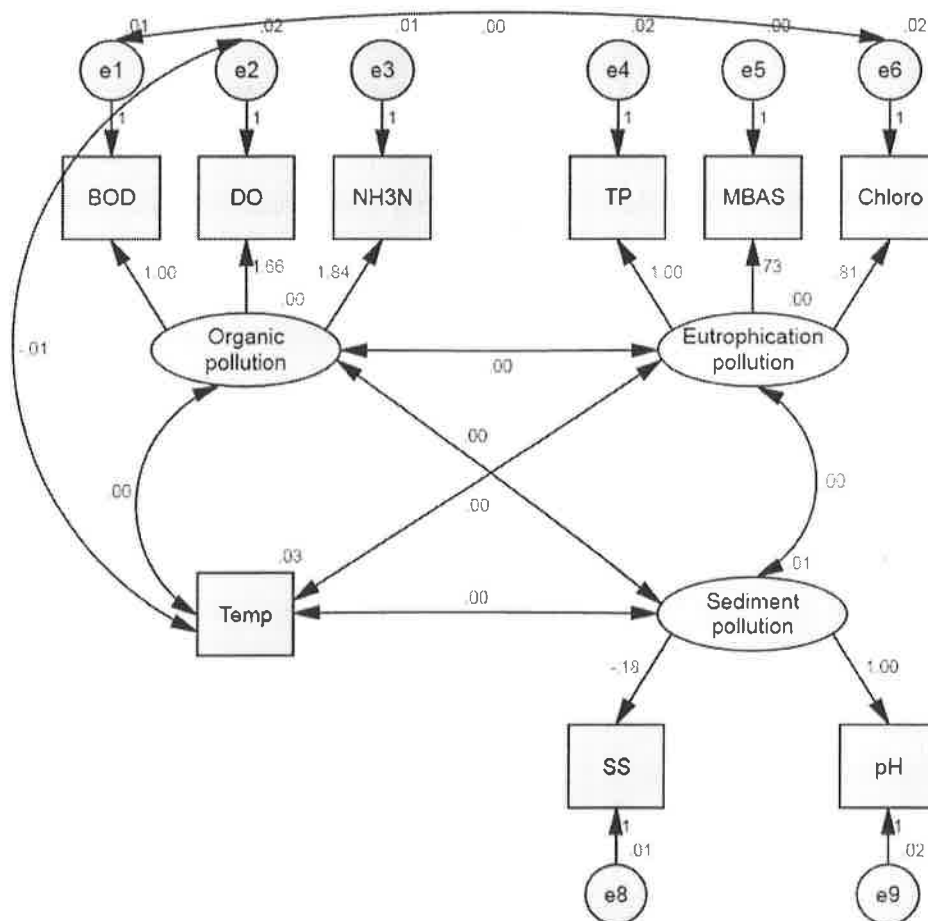
3.1 Application and Analysis of Factor Analyses

Before AMOS is used in this study, Wu et al. [25] utilized varimax rotation in factor analyses for orthogonal rotation to interpret number characteristics. Since a number value lower than 1.00 cannot be shown directly when SPSS 17.0 calculates

a characteristic value, researchers select a value higher than 1.00 as the latent factor for the convenience of future studies. In practical analysis, a characteristic value smaller than 1.00 but close to 1.00 shall also be manipulated for the purpose of analysis. Therefore, according to the analysis results, there are four factors with a characteristic value higher than 1.00 as shown in Table 3, and they are selected as the main factors used to explore the influence on water quality in the watershed.

In this study, the above numbers with characteristic value higher than 1 decide the main factors; we then select the parameters of each factor with the component matrix acquired through orthogonal rotation. Table 1, provided in the "Methodology" section, shows the matrix, after orthogonal rotation, that explains the characteristics of the four factors. It can also be utilized to describe the main factors that influence water quality in the watershed area as well as the variance between the four factors.

Fig. 3 Modification indices of oblique rotation



ObliqueRotation FactorAnalysis
 Unstandardized estimates
 Chi-Square = 18.729
 df = 20 P-value = .539

3.2 Selection of AMOS

Table 4 shows the kurtosis of the nine parameters of water quality as a non-normal distribution because they all are not close to the number value of 3. Normal distribution refers to a number value close to 3, and the value of the whole model reaches to 191.214, a very steep distribution. Generalized least square (GLS) shall be used to replace maximum likelihood in the algorithm.

3.3 Orthogonal Rotation Analysis

In this study, at the beginning of the factor analyses, assume that there is no correlation between the four factors, and orthogonal rotation should be utilized for multivariate statistics. As shown in Fig. 2, there is a connected covariance line,

but this study hypothesizes the line as 0.00001. With a low correlation, it can be regarded as a zero correlation.

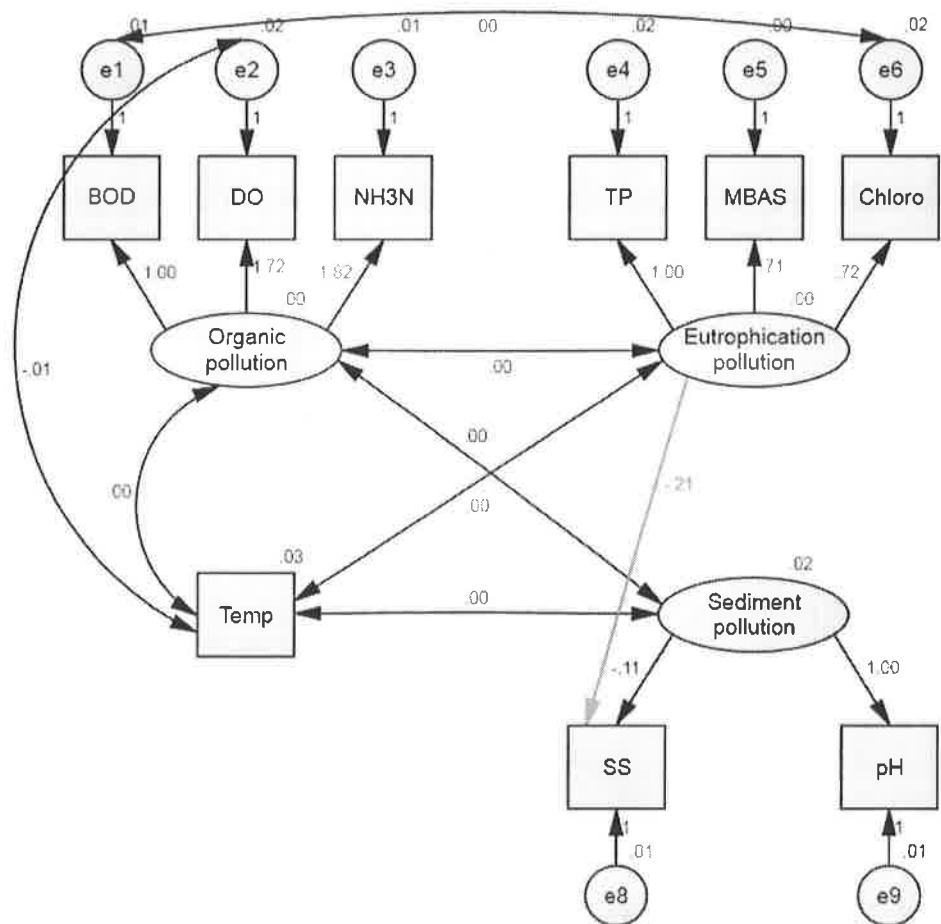
According to Fig. 2, p value only reaches to 0.000 and chi-square is 93.003 implying poor initial goodness of fit of the hypothesized model and a failure to explain the estimate.

3.4 Analysis of Oblique Rotation

This study uses oblique rotation to analyze and modify the original orthogonal rotation to identify the correlation of the four factors because of the biological, chemical, and physical relationship of the nine parameters of water quality.

In terms of observation of factor load, reflective indicators of the three items in organic pollution improve their factor load after oblique rotation to ensure the goodness of fit of the hypothesized model for the factor of organic pollution.

Fig. 4 Completion of oblique rotation model



ObliqueRotation FactorAnalysis
 Unstandardized estimates
 Chi-Square = 17.643
 df = 20 P-value = .611

Reflective indicators of the three items in eutrophication dramatically improve their factor load after oblique rotation, and it is worth noting that originally, in orthogonal rotation, Chloro has a factor load of 0.05, but after oblique rotation, it increases to 0.28 indicating there is no hypothesis error of the three indicators of the eutrophication factor. The pH value and SS of sediment pollution also maintain a negative correlation indicating the goodness of fit of sediment pollution in the hypothesized model.

From the analysis results of oblique rotation, there is a covariance connection between the four factors, and among them, the correlation coefficient between organic pollution and eutrophication is the highest with a value of 0.52 indicating a correlation of organic pollution to the other three factors as well as stronger correlation with eutrophication. The explanation is excessive algae growth due to eutrophication, and the death of a high number of algae produces excessive BOD. Likely, the amount of ammonia nitrogen also increases.

3.5 Application of AMOS Modification Indices

This study selects the best results of *p* value, chi-square, and interior quality acquired via oblique rotation between the two types of rotation for the use of modification indices and model search. Sequence modification focuses first on modification indices because there is a need to examine whether any parameters (covariance) need to be redefined due to its significant influence on the continuous model search.

Due to smaller *p* value as well as the failure to complete the iteration estimate procedure after several modifications, we still cannot acquire the results and do not choose to use orthogonal rotation.

Table 5 provides three modification indices higher than 3.84 and the covariance relationship between ammonia nitrogen, DO, and temperature. Moreover, it indicates that BOD and Chloro need to be redefined.

According to Fig. 3, the model after the second modification has met the requirement for *p* value >0.5, and there is a positive correlation between BOD and Chloro. The explanation is that Chloro means algae and more algae are required for higher BOD.

When factor load exceeds 0.5, the increasing number of water quality indicators reaches standards such as DO of 0.559, BOD of 0.661, MBAS of 0.554, pH of 0.634, and the rest of the water quality indicators are also close to the standard factor load of 0.5. When the factor load is higher, the reliability coefficient also increases.

3.6 Specification Search of AMOS

This study selects all covariance curves to execute the first specification search and a total of 256 models of

Table 6 Evaluation indices of second specification search (oblique rotation)

Statistical test	Goodness of fit standard or threshold value of AMOS program
Absolute goodness of fit index	
χ^2	0.611
GFI	0.986
AGFI	0.968
RMR	0.001
RMSEA	0.000 (<0.05 good goodness of fit)
NCP	0.000
ECVI	Theoretical value 0.242<independent value 0.552
Incremental goodness of fit index	
NFI	0.871
RFI	0.767
IFI	1.020
TLI	1.042
CFI	1.000
Parsimony goodness of fit index	
PGFI	0.438
PNFI	0.484
AIC	Theoretical value 67.643<independent value 154.529
CAIC	Theoretical value 183.602<independent value 196.274

GFI goodness of fit index, *AGFI* adjusted goodness-of-fit index, *RMR* root mean squared residual, *RMSEA* root mean square error of approximation, *ECVI* expected cross-validation index, *RMSEA* root mean square error of approximation, *NCP* non-centrality parameter, *AIC* Akaike information criterion, *CAIC* consistent Akaike information criterion, *PGFI* parsimony goodness of fit index, *PNFI* parsimony normed fit index, *NC* non-centrality, *RFI* relative fit index, *IFI* incremental index of fit, *TLI* Tucker–Lewis index, *CFI* comparative fit index

covariance curves that need to be analyzed. For the results of the first specification search, this study uses Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC) to explain the optimal model of the hypothesized one.

AMOS, during the specification search, uses AIC_0 , BCC_0 , and BIC_0 as the judgment basis; when the three values are close to 0, they present the optimal hypothesized model [6]. Hence, this study conducts two specification searches and selects Fig. 4 as the best hypothesized model. At the same time, the second modification of the specification search in this study maintains good external quality as shown in Table 6.

4 Conclusion

This study uses multivariate statistical analysis to identify the four factors of organic pollution, eutrophication, seasonal, and sediment pollution that mainly influence water quality in the watershed area.

This study uses the AMOS computing program to confirm factor analyses; the initial hypothesized orthogonal rotation model is reported with p value and chi-square of 0.000 and 93.003 completely unfitted. But after oblique rotation, it is confirmed that with the inclusion of a covariance curve, the initial orthogonal rotation is not a complete error estimate. After several explorative factor analyses, this study discovers the optimal model diagram with a p value and chi-square of 0.611 and 17.643 reaching the standards and confirming the normal distribution of reflective indicators of the four factors in the hypothesized model of orthogonal rotation. The good external quality has the biological, chemical, and physical relationship to support the orthogonal rotation model. This also explains the biological, chemical, and physical influences in nature on the nine parameters of water quality.

When multivariate analyses are applied, factor analyses, cluster analyses, and discriminant analyses shall also be conducted at the same time to obtain optimal results of statistical analyses. Due to the length limit, this study only presents factor analyses in stage 1 and confirmatory factor analyses in stage 2 to establish the complete set of methods.

In the future, heavy metals that affect water quality shall be included to improve water quality analyses and managerial completeness.

This study aims to establish a set of standard methods for the reference of the reservoir authorities to dramatically improve the application and statistical analyses of water quality monitoring data as well as the enactment of water quality managerial strategies in the watershed areas of reservoirs.

Acknowledgments The authors sincerely appreciate financial supports for the project (NSC 100-2221-E214-011) from Taiwan's National Science Council for the work.

The authors also deeply appreciate the anonymous reviewers for their insightful comments and suggestions.

References

1. Arbuckle, J. L. (1994). Computer announcement: AMOS: analysis of moment structures. *Psychometrika*, *59*, 135–141.
2. Arbuckle, J. L., & Wothke, W. (1999). IBM AMOS 4.0 user's guide, Chicago, IL, USA.
3. Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Academy of Marketing Science*, *16*, 76–85.
4. Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
5. Bollen, K. A. (1989). A new incremental fit index general structural equation models. *Sociological Methods & Research*, *17*, 303–310.
6. Burnham, K. P., & Anderson, D. R. (1998). *Model select and inference: a practical information-theoretic approach*. New York: Springer-Verlag.
7. Charlton, A. J., Robb, P., Donarski, A. J., & Godward, J. (2008). Non-targeted detection of chemical contamination in carbonated soft drinks using NMR spectroscopy, variable selection and chemometrics. *Analytica Chimica Acta*, *618*, 196–205.
8. Chen, S. Y. (2005). *Multivariate analysis* (4th ed.). Taipei: Huatai Publisher. in Chinese.
9. Chen, Y. S., Huang, T. W., & Huang, F. M. G. M. (2006). *Basic principles of structural equation modeling*. Kaohsiung: Li-Wen Publisher. in Chinese.
10. Einax, J. W., Truckenbrodt, D., & Kampe, O. (1998). River pollution data interpreted by means of chemometric methods. *Microchemical Journal*, *58*, 315–321.
11. Helena, B., Pardo, R., Vega, M., Barrado, E., Ferná'ndez, J. M., & Ferná'ndez, L. (2000). Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga river, Spain) by principal component analysis. *Water Research*, *34*, 807–815.
12. Hsieh, P. H., Kuo, J. T., Wu, E. M.-Y., Ciou, S. K., & Liu, W. C. (2010). Optimal best management practice placement strategies for nonpoint source pollution management in the Fei-Tsui reservoir watershed. *Environmental Engineering Science*, *27*, 441–447.
13. Lin, C., Wu, E. M.-Y., Lee, C. N., & Kuo, S. L. (2011). Applying multivariate statistical factor analyses on selecting an optimal method for recycling of food wastes. *Environmental Engineering Science*, *28*, 349–356.
14. Liu, C. W., Lin, K. H., & Kuo, Y. M. (2003). Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Science of the Total Environment*, *313*, 77–85.
15. MacGregor, J. F., & Kourti, T. (1995). Statistical process control of multivariate process. *Control Engineering Practice*, *3*, 403–410.
16. Nunnari, G., Dorling, S., Schlink, U., Cawley, G., Foxall, R., & Chatterton, T. (2004). Modelling SO₂ concentration at a point with statistical approaches. *Environmental Modelling and Software*, *19*, 887–897.
17. Pires, J. C. M., Sousa, S. I. V., Pereira, M. C., Alvim-Ferraz, M. C. M., & Martins, F. G. (2008). Management of air quality monitoring using principal component and cluster analysis-part I: SO₂ and PM₁₀. *Atmospheric Environment*, *42*, 1249–1255.
18. Pires, J. C. M., Sousa, S. I. V., Pereira, M. C., Alvim-Ferraz, M. C. M., & Martins, F. G. (2008). Management of air quality monitoring using principal component and cluster analysis—part II: CO, NO₂ and O₃. *Atmospheric Environment*, *42*, 1261–1268.
19. Simeonov, V., Stratis, J. A., Samara, C., Zachariadis, G., Voutsas, D., Anthemidis, A., et al. (2003). Assessment of the surface water quality in Northern Greece. *Water Research*, *37*, 4119–4125.
20. Tajik, S., Ayoubi, S. S., & Nourbakhsh, F. S. (2012). Prediction of soil enzymes activity by digital terrain analysis: comparing artificial neural network and multiple linear regression models. *Environmental Engineering Science*, *29*, 798–808.
21. Theodora, K. (2006). The process analytical technology initiative and multivariate process analysis, monitoring and control. *Analytical and Bioanalytical Chemistry*, *384*, 1043–1551.
22. Ticiavilca, A., & McKee, M. (2011). Multivariate Bayesian regression approach to forecast releases from a system of multiple reservoirs. *Water Resources Management*, *25*, 523–532.
23. Varol, M., Gökot, B., Bekleyen, A., & Şen, B. (2012). Water quality assessment and apportionment of pollution sources of Tigris river (Turkey) using multivariate statistical techniques—a case study. *River Research and Applications*, *28*, 1428–1436.
24. Wu, E. M.-Y., & Kuo, S. L. (2012). Applying a multivariate statistical analysis model to evaluate the water quality of a watershed. *Water Environment Research*, *84*, 2075–2085.
25. Wu, E. M.-Y., Kuo, S. L., & Liu, W. C. (2012). Generalized autoregressive conditional heteroskedastic model for water quality analyses and time series investigation in reservoir watersheds. *Environmental Engineering Science*, *29*, 227–237.
26. Wunderlin, D. A., Díaz, M. P., Ame, M. V., Pesce, S. F., Hued, A. C., & Bistoni, M. A. (2001). Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality—a case study: Suquia river basin (Cordoba-Argentina). *Water Research*, *35*, 2881–2882.
27. Yang, Y. H., Zhou, F., Guo, H. C., Sheng, H., Liu, H., Dao, X., et al. (2010). Analysis of spatial and temporal water pollution patterns in Lake Dianchi using multivariate statistical methods. *Environmental Monitoring and Assessment*, *170*, 407–416.