© 2018 IJSRSET | Volume 4 | Issue 9 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology

Multivariate Repeated Measurement Model for Analysing Water Quality of Na Thap River During 2004-2014

Felycitia Iradati Yusrina, Khairil A. Notodiputro, Indahwati

Department of Statistics, Bogor Agricultural University, Bogor, West Java, Indonesia

ABSTRACT

The Electricity Generation Authority of Thailand (EGAT), located in Chana, Songkhla province, was built in 2004 and has been sending electricity commercially to the network in 2007. The power plant waste is discharged into the Na Thap River, so there is a great concern that will have an ecological impact on the biodiversity and ecosystems in the Na Thap river waters. This study focuses on the river water quality. The water quality studied was dissolved oxygen (DO), pH and Biochemical Oxygen Demand (BOD) which followed the repeated measurement structure data. Water quality modeling show that the month, zone and their interactions affect the water quality of pH, DO and BOD simultaneously before and after EGAT operation. The variance contribution of between sites smaller than within site. This indicates a variation within sites and no variation between sites.

Keywords : Mixed Model, Multivariate Linear Mixed Model, Na Thap River, Repeated Measurement

I. INTRODUCTION

The rapid growth in electricity demand in Thailand is a major challenge for electricity companies who tried to ensure adequate supplies [1]. The Electricity Generation Authority of Thailand (EGAT) is a state-owned company involved in electricity generation and transmission in the southern region of Thailand. A thermal power plant located in the Chana of Songkhla province was built in 2004 and commercially sent electricity to the network in 2007 [2]. The power plant waste is discharged into the Na Thap River so there is great concern that the waste will have an ecological impact on biodiversity and ecosystems in the Na Thap river waters [3].

Research related to the effect of EGAT that has been done is linear mixed models for analyzing total weights of fish in Na thap river, southern Thailand [4], Polychaeta organism density modeling before and after the power plant operates [5], the use of linear logistic mixed models on Phytoplankton [6], and predictions of fish stock in the Na Thap River, southern Thailand [7]. These studies only focus on organisms and biodiversity, so far there has been no in-depth study that discusses the influence of the construction of power plants in Chana on the quality of the Na Thap River water. Only Chesoh and Lim (2015) have evaluated water quality and say that power plant waste does not significantly affect river water quality but they recommend conducting further studies to monitor river ecology [3].

This study will develop a model to analyze whether the Na river water quality is affected by EGAT construction. Modeling is carried out in a multivariate manner involving water quality response variables those are pH, dissolved oxygen (DO) and Biochemical Oxygen Demand (BOD). Factors affected by response are not only fixed effects but also random effects. For example, the location and site of the survey and it is classified as independent, but random effects. Data is

observed from time to time that has repeated measurement data structure. Between sites assumed to be independent but within site correlation exists between observation time. For this reason, the appropriate model to develop is Multivariate Linear Mixed Model (MLMM) for repeated measurement data.

II. MATERIAL AND METHOD

A. Material

The data used is the data survey from the Electricity Generating Authority of Thailand (EGAT). The survey location is on the Na Thap River in Chana, Songkhla province. The data period used is June 2004 to December 2006 for the period before EGAT operation and January 2007 to May 2014 for the period after EGAT operations. The data used in this study is only water quality data. The water quality variables observed in this study were pH, dissolved oxygen (DO) and Biochemical Oxygen Demand (BOD).

In the survey that was conducted by EGAT, Na Thap River was divided into 3 zones, freshwater, brackish and saline. EGAT is located near a site along the river. The freshwater zone is represented by site numbers 1, 2, 3 and 4. The brackish water zone is represented by numbers 5, 6 and 7, while the saltwater zone is represented by numbers 8, 9 and 10. The EGAT location is around site 3, so the waste from the power plant is disposed around site 3[8].

B. Model

The model used is Multivariate linear mixed model (MLMM) which is an extension of linear mixed model (LMM). MLMM has more flexibility in developing models and can accommodate several correlated response variables [9]. In addition, MLMM can be built by adding a random component part with the assumption that each element of the response variable has a linear correlation with the systematic part of the model [10]. pH, DO, and BOD are modeled using MLMM with a fixed effect on time, zone and their interaction. and the random effect of site nested in the zone. The model is as follows [11]

$$\boldsymbol{Y}_{n\,x\,r} = \boldsymbol{X}_{n\,x\,p}\boldsymbol{\beta}_{p\,x\,r} + \boldsymbol{Z}_{n\,x\,q}\boldsymbol{U}_{q\,x\,r} + \boldsymbol{\varepsilon}_{n\,x\,r}$$

with,

Y = Multiple response variable matrix (pH, *DO*, and *BOD*)

$\boldsymbol{X}$ = The combined design matrix for fixed effect (time, zone and their interaction)

$\boldsymbol{\beta}$ = General regression coefficient matrix

$\boldsymbol{Z}$ = The Combined design matrix for site random effect

$\boldsymbol{U}$ = Special coefficient matrix (random effect)

$\boldsymbol{\varepsilon}$ = Random error matrix

r = number of response variables = 3

m = number of fixed effects = 3

s = number of random effects = 1

p = m + 1 = 4

q = s + 1 = 2

The model above is used to model data before and after EGAT operates.

C. Procedure Analysis

Data analysis using R 3.4.4, SAS 9.2 and Minitab 17. The procedures analysis are as follows:

1. Data exploration using graphics to see trends in variable water quality over time.
2. Preparation of repeated measurement modeling data.
   a. Handling outliers with smoothing techniques.
   b. Eliminating seasonal effects on data with deseasonalized.
3. Correlation checks between pH, BOD, and DO variable as the first step of multivariate modeling.
4. pH, BOD, and DO modeling with multivariate linear mixed model before and after EGAT operates.

III. RESULTS AND DISCUSSION

A. Data Exploration

Table 3.1: Descriptive statistics of pH, DO and BOD before and after EGAT operations

Variable	Informations	Average	Variance	Minimum	Median	Maximum	Standar*
pH	Before	7.44	0.25	6.41	7.49	8.55	5-9
	After	7.28	0.72	4.34	7.34	10.3	
DO	Before	6.10	3.65	2.1	6.05	12.5	>2.0
	After	5.18	1.59	0.6	5.3	11.2	
BOD	Before	1.00	0.29	0.15	0.91	3.2	<4.0
	After	1.11	1.32	0	0.8	10.1	

* acceptable range of aquatic animal protection [3]

Data analysis in this study consists of two periods, before and after EGAT operates sending electricity to the network. Table 1 shows that the average pH, DO, and BOD as a whole in the period before and after EGAT operates within an acceptable range of aquatic animal protection [3]. However, the maximum value of pH and BOD in the period after EGAT operates above the standard limit. Variances of pH and BOD after EGAT operates larger than before EGAT operates, indicating there are outliers.

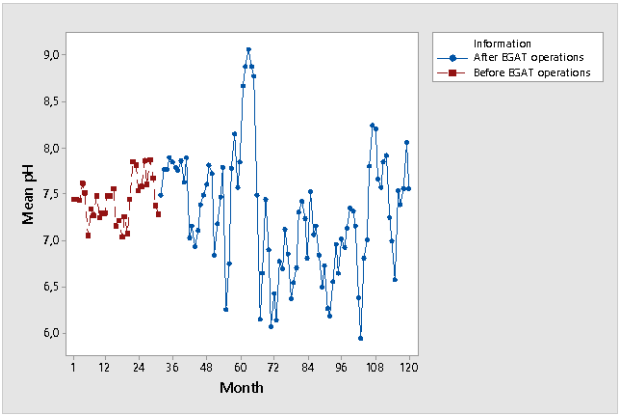

Figure 3.1: Plot the average pH from ten sites in June 2004 to May 2014

Tabel 3.2: Testing simultaneous effect on the quality of Na Thap river water before EGAT operates

Sources	$v1$	$v2$	F- value	P-value
Intercept	3	1862	18596.70	<.0001
Month	264	1862	6.42	<.0001
Zone	6	1862	61.16	<.0001
Month* Zone	528	1862	1.68	<.0001

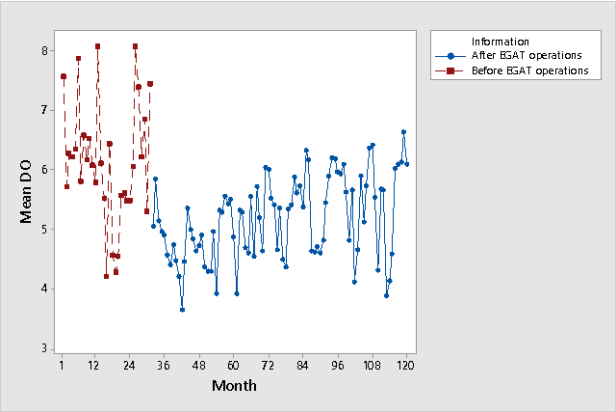

Figure 3.2: Plot the average DO from ten sites in June 2004 to May 2014

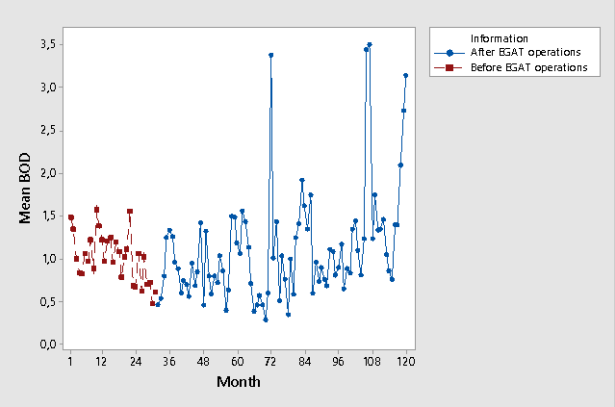

Figure 3.3: Plot the average BOD value from ten sites from June 2004 to May 2014

B. Multivariate Linear Mixed Model

BOD values at several sites differ greatly from the majority of other observations. This shows that there are outliers, so handling is needed for it does not cause bias in data analysis, damaging data reliability and eliminating data accuracy. The method of treating outliers in this study is the smoothing technique. Observation responses of pH, DO, and BOD were obtained by repeated measurements based on time per month so that there was a seasonal effect

in the data. Seasonal effects are eliminated by deseasonalized techniques.

Figure 2 shows that the average pH of ten sites before EGAT operates had relatively stable fluctuations compared to after EGAT operates. As for the average dissolved oxygen (DO) of the ten sites after EGAT operates, the fluctuation is quite stable but decreased compared to before EGAT operates (Figure 3). Decreased DO values indicate polluted water. Despite the decline, the average of the ten sites DO still conform to standards. Figure 4 shows that the average BOD for ten sites before EGAT operates had relatively stable fluctuations, whereas after EGAT operates there are a number of average BOD values that are higher than the average in other months but the value is still at the standard.

Tabel 3.3: Testing simultaneous effect on the quality of Na Thap river water after EGAT operates

Sources	$v1$	$v2$	F- value	P-value
Intercept	3	554	6421.99	<.0001
Month	90	554	7.33	<.0001
Zone	6	554	59.14	<.0001
Month* Zone	180	554	1.69	<.0001

Table 3.2 and Table 3.3 present the results of the testing on the fixed effects of pH, DO, and BOD water quality simultaneously before and after EGAT operates. All fixed effects in the model have a p-value less than 5% so it can be concluded that the month, zone and their interactions have an affect on the water quality of pH, DO, and BOD simultaneously in both periods. Although the factors that affect pH, DO, and BOD before and after EGAT operates simultaneously are similiar, but as shown in Table 3.1, in terms of the characteristics of pH, DO, and BOD there are some changes before and after EGAT operates. The pH variances before and after EGAT operates are quite differently, which are 0.253 and 0.728. The variance after EGAT operates higher than before EGAT operates, this indicates that pH stability is disrupted after EGAT operates. The average DO decreases after EGAT operates, whereas oxygen dissolved in water is needed to support the life of aquatic organisms [11]. After EGAT operates, the average BOD increased slightly from 1.00 mg/L to 1.11 mg/L. The BOD value of river water indicate the number of organic pollutants in river water [12].

Tabel 3.4: Testing partial effect on the quality of Na Thap river water in the period before EGAT operations

Effect	Variable	Month	Zone	Estimates	SE	DF	P-value
Month	DO	Agst-04		12.99	0.58	554	0.026
Month	DO	Sep-04		11.93	0.58	554	0.041
Month	DO	Agst-05		25.86	0.58	554	<.000
Month	DO	Sept-05		38.58	0.58	554	<.000
Month* Zone	DO	Agst-04	1	46.66	0.80	554	<.000
Month* Zone	DO	Agst-04	2	26.73	0.87	554	0.002
Month* Zone	DO	Sep-04	2	31.49	0.87	554	0.000
Month* Zone	DO	Okt-04	1	25.49	0.80	554	0.001
Month* Zone	DO	Okt-04	2	27.74	0.87	554	0.001
Month* Zone	DO	Okt-05	2	-24.41	0.82	554	0.003
Month* Zone	DO	Ags-06	2	-22.13	0.82	554	0.007

Partial testing of the fixed effects in Table 3.4 and Table 3.5 is a follow-up test of the simultaneous tests (Table 3.2 and Table 3.3) to determine the parameters that influence the response with a significant level of 5% before and after EGAT operates. The partial tests presented in Table 5 and Table 6 are only variables that have a significant effect in August-October because the rainfall in the three months is the lowest in Thailand (Singhrattna et al. 2005). Low rainfall allows EGAT waste to be discharged into the river to further affect water quality so that partial testing focuses on the three months. In the period before EGAT operations, DO levels in August and September in 2004 and 2005 proved to be significantly different from December 2006. Month and zone interactions were also significantly different in August 2004 in

freshwater zones and brackish water zones compared to saltwater zones. Another real interaction was September 2004 with the brackish water zone significantly different from the saltwater zone, and October 2004 with the freshwater and brackish water zones compared to saltwater and October 2005 with the brackish water zone compared to saltwater at the 5% level.

Partial testing after EGAT operates generating DO and pH in August 2008 was significantly different from May 2014 (Table 6). Interpretation of interactions between zones and month is the same as before EGAT operates.

Tabel 3.5: Testing partial effect on the quality of Na Thap river water in the period after EGAT operations

Effect	Variable	Month	Zone	Penduga	SE	DF	P-value
Month	DO	Agst-08		-14.37	0.51	1862	0.005
Month *Zone	DO	Agst-07	1	-22.80	0.64	1862	0.001
Month *Zone	DO	Sep-07	1	-29.70	0.68	1862	<.001
Month *Zone	DO	Sep-07	2	-23.69	0.73	1862	0.001
Month *Zone	DO	Okt-07	1	-14.89	0.68	1862	0.029
Month *Zone	DO	Okt-08	2	20.37	0.73	1862	0.005
Month *Zone	DO	Sep-09	2	-15.03	0.73	1862	0.040
Month *Zone	DO	Agst-10	1	-16.55	0.68	1862	0.015
Month *Zone	DO	Agst-10	2	-13.86	0.68	1862	0.043
Month *Zone	DO	Agst-11	1	-24.89	0.68	1862	0.001
Month *Zone	DO	Agst-11	2	-15.34	0.73	1862	0.036
Month *Zone	DO	Okt-11	1	-18.32	0.68	1862	0.007
Month *Zone	pH	Agst-08	1	-17.30	0.68	1862	0.011

Random effect of the site which is nested within the zone causing the diversity that occurs in the site or between sites. Before EGAT operated the estimated variance at the site is 0.5140 while the estimated variance between sites is 0.1636. As after EGAT, the estimated variance at the site is 0.4018 while the estimated variance between sites is 0.1324. The estimated site variance value that is greater than the estimated variance between sites shows the site's diversity contribution on the model is small.

IV. CONCLUSION

Water quality modeling results simultaneously show that Month, zone and their interactions affect the water quality of pH, DO, and BOD before and after EGAT operates. Although the influencing factors were the same in both periods, but in terms of pH diversity after EGAT operates is higher than before EGAT operates which indicates pH stability is disrupted after EGAT operates. The characteristics of the DO decreased, while the BOD increased slightly on average after EGAT operates. This shows that EGAT development affects the quality of the Na Thap river water. The contribution of diversity of sites on multivariate linear mixed models is not too large, more diversity is caused by errors.

V. ACKNOWLEDGEMENTS

The data used in this paper has been provided by Dr. Apriadee and Dr. Sarawuth of Prince of Songkla University, signed as a joint research with the Department of Statistics, Bogor Agriculture University, Indonesia.

VI. REFERENCES

[1]. Chesoh S. Community Perception. Satisfaction and Participation towards Power Plant Development in Southernmost of Thailand. Journal of Sustainable Development. Vol 3, No 2, 84-88, 2010.

[2]. Chesoh S and Lim A. Investigation of Aquatic Environment and Social Aspects of Thermal Power Operation Plant in Southern of Thailand. Asian Social Science. Vol 10, No 16, 168-175, 2014.

[3]. Chesoh S and Lim A. Monitoring of Thermal Power Plant Operations on Aquatic Ecosystems in Tropical Estuarine River of Thailand 2008-

2013. E-proceedings of the 36th IAHR World Congress,The Hague, The Netherlands, 2015.

[4]. Angraini Y, Notodiputro KA and Sadik K. Linear Mixed Models for Analyzing Total Weights of Fish in Na Thap River, Southern Thailand. Proceedings of the 2nd ISI Regional Statistics Conference, Indonesia, 2017.

[5]. Lueangthuwapranit C and Saheem N. Polychaeta Organism Density in Na Thap Estuary. International Journal of Biology, Vol 3, No 4 , 30-35, 2011.

[6]. Arisanti R, Notodiputro KA, Sadik K and Lim A. Bias Reduction in Estimating Variance Components of Phytoplankton Existence at Na Thap River Based on Logistics Linear Mixed Models. IOP Conf , 2017.

[7]. Saheem N, Lim A and Chesoh S. Predicting standing crops using lagged fingerling density of freshwater fish in the Na Thap River of southern Thailand, J. Sci. Technol, Vol 36, No 1, 13-19, 2014.

[8]. Saheem N. Statistical Modeling of Aquatic Animal Abundance in the Na Thap River. Prince of Songkla University, 2015.

[9]. Singh V, Rana RK and Singhal R. Analysis of repeated measurement data in the clinical trials. Journal of Ayurveda and Integrative Medicine, Vol 4, No 2,77-81, 2013.

[10]. Latra IN, Linuwih S, Purhadi and Suhartono. Estimation for multivariate linear mixed models. International Journal of Basic & Applied Sciences ,Vol 10, No 06, 2010.

[11]. Siahaan R, Indrawan A, Soedharma D and Prasetyo LB. Kualitas air sungai Cisadane, Jawa Barat – Banten. Jurnal Ilmiah Sains, Vol 11, No2, 268-272,2011.

[12]. Novotny V and Olem H. "Water Quality: prevention, Identification, and Management of Diffuse Pollution". New York (US):Van Nostrand Reinhold, 1994.