SAMPLING DESIGN FOR WATER QUALITY MONITORING IN MARINE RESERVE: A STUDY CASE AT BANDA SEA CONSERVATION PARK

Hedi Indra Januar

Indonesian Research Center for Marine and Fisheries Product Processing and Biotechnology, KS Tubun Petamburan VI Slipi DKI Jakarta 10260 – idjanuar@kkp.go.id

Abstract. *Design of sampling site water quality monitoring at marine reserve is critical, to optimize the effectiveness of periodic assessment. A simple stratified systematic design, that been usually used in most of monitoring analysis, may not maximize the information of spatial data in marine hydrology. The present work applied a multivariate statistical analysis and spatial autocorrelation methods to develop an optimal sampling design for water quality assessment in tropical marine reserve, Banda Sea Conservation Park, Indonesia. Seasonal (west, intermediate, and east monsoon season in Indonesia) and spatial (38 stations) water quality analysis (salinity, dissolved oxygen, pH, and nutrient) in 3 zones of Marine Reserve were conducted. Principal Component Analysis (PCA) showed dissolved oxygen (DO) was the principal variable for the sampling design criterion. Spatial DO Variograms suggested elocation of the sampling stations, to optimize the design of water quality monitoring. Therefore, even the principal variable may vary at other locations, depends on hydrology and other area specific characteristics, the proposed technique could be applied in sampling design concerning water quality* monitoring.

Keywords: Environmental monitoring, marine reserve, samplirng design, water quality

(Diterima: 19-09-2017; Disetujui: 04-12-2017)

1. Introduction

Degradation of water quality has been known to have a major impact on the coral reef diversity. Previous studies have shown that an increase of pollutant materials, for example inorganic nitrogen and other nutrients, may cause a biodiversity shifts and decreases of coral organism cover (De'ath and Fabricius, 2010; Polónia *et al.*, 2015; Duprey *et al.*, 2016). Therefore, water quality monitoring is important as a regular activity for marine conservation authorities. Spatial and temporal changes of water quality can serve as a baseline study for conservation monitoring (Hedley *et al.*, 2016). Periodic monitoring will obtain information on water quality and the spatial or temporal impact of point or non-point souces of terrestrial runoff (Najar and Khan, 2012).

A systematic grid design is generally used in spatial environmental monitoring design. Sampling location is determined by grid boxes of the same size encompassing all of the monitoring area (Crooks *et al.*, 2014). However, previous studies have shown that grid design has some weaknesses. Statistically, systematic grid design needs to optimize, to maximize spatial data information (Wang *et al.* 2009; Huang and Yang, 2011). Moreover, spatial autocorrelation of a water variable may occur between points on adjacent grid sampling (Mattsson *et al.*, 2013). The existence of spatial autocorrelation is the source of error in the statistical analysis of environmental data (Fu and Wei, 2013).

The optimization of the sampling point can be done by pre-survey analysis of optimal distance quantification to avoid the spatial autocorrelation within targeted water variables between sampling points. The spatial autocorrelation distance can be determined by semivariogram analysis (Maas *et al.*, 2010). This research aimed to describe the optimization technique of point sampling design in water quality monitoring at Banda Sea Conservation Park, by calculation of spatial autocorrelation. The results of the analysis are expected to be used as a baseline study for water quality monitoring on coral reef conservation in the region.

2. Methodology

2.1. Study Area

The study was conducted in Banda Sea Conservation Park, Maluku Province. 38 sampling stations were selected with stratified grid boxes of 1 x 2 km^2 (Figure 1). The entire station was grouped by location, which were zone I (Point 1 - 8), zone II (Point 9 - 23), and zone III (point 24-38). Accurate locations from each point were recorded with Garmin ETrex GPS instrumentation 10 and then plotted by Surfer software v.11.

 doi: 10.29244/jpsl.8.3.296-300

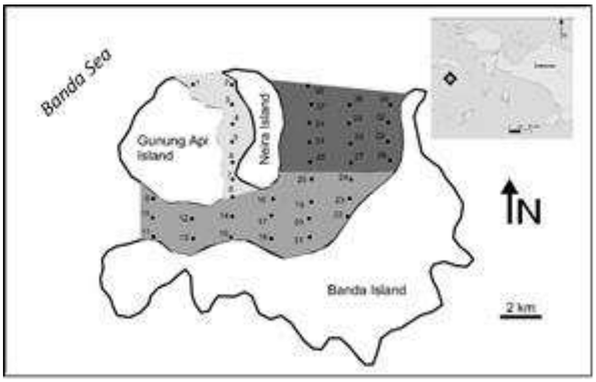


Figure 1. Research station in zones I ( ), zone II ( ), and zone III ( ) at Banda Sea Conservation Park

2.2. Seawater sampling

The sampling activity was conducted in April 2016. Sea water was taken in both tidal conditions with every three replicates for each sample. Seawater samples were collected on the surface (1 m depth) using a Nansen tube. Salinity, pH, DO (dissolved oxygen), and salinity, were analyzed in situ, by HACH HQ40d with DO and pH probes, and Eutech Salt +6 Digital Salt-meter. The measured pH value was pH at the total scale, with a calibrated electrode based on Nemzer and Dickson (2005) method, using tris-buffers in synthetic seawater before the sampling takes place. Moreover, phosphate, nitrate, nitrite and ammonia were also quantified in situ by nutrients colorimetric method by HACH DR-890 portable spectrophotometer.

2.3. Data Analysis

The entire data of water variables were log transformed (1 + x), and then multivariate Principal Component Analysis (PCA) and CDA (Canonical Discriminant Analysis were conducted to identify the main component variable at each point and zone (Januar *et al.*, 2017). The identified variable become the target of autocorrelation distance analysis between sampling stations and the number of station in each zone.

The spatial autocorrelation plot to the lag of distance between stations, according to the formula of Lam (1983) in equation (1).

$$\gamma(h) = \frac{1}{2n}.\sum_{i=1}^{n}\{Z(xi) - Z(xi+h)\}^2 \text{ ....(1)}$$

$\gamma(h)$ is a spatial autocorrelation at the distance h between points, $Z(x_i)$ is the target variable at point x_i, and $Z(x_i+h)$ is the target variable at a point distant h from xi, and n is the number of points the grid that is bypassed at a distance h. The plot between $\gamma(h)$ to h (variogram) was made to determine the autocorrelation distance between sampling stations. Furthermore, the optimal number of sampling stations for each zone was determined according to Kitsiou *et al.* (2001), which is the proportion of standard error in each zone (equation 2). n is the number of optimal sampling points, nT is the sum of all sampling points, and SE is the standard error. Multivariate and geostatistical data analysis of autocorrelation and variogram were performed using PAST Statistical v3.0 (Hammer *et al.*, 2001) and Surfer Software v11.

$n = nT\ x\ (SE\ Zone/SE\ all\ Zones)$ ....(2)

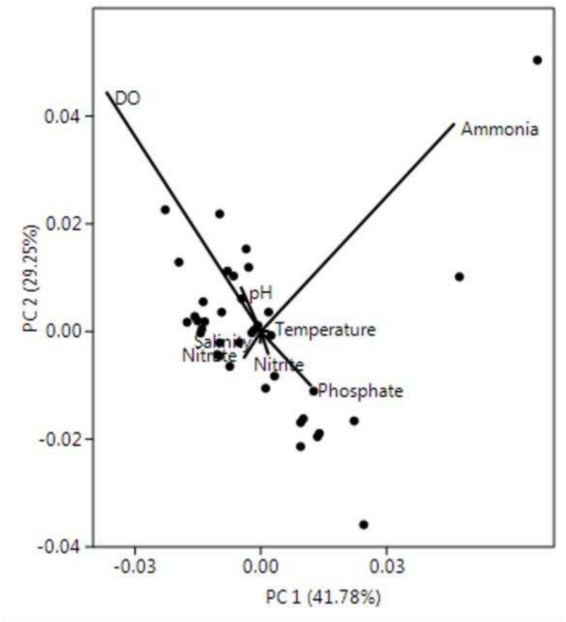
3. Result and Discussion

In general, this study detected that the water variables values were below the ASEAN water quality criteria for aquatic life protection (McPherson *et al.*, 1999), with the exception of the phosphate concentration throughout the study area (Table 1). Moreover, multi-dimensional analysis (Figure 2) showed that DO and ammonia variables were the main factors that distinguish between the overall research stations. Combination between both variables contributed 41.78% to the variation of water characteristics across the whole stations. Meanwhile, phosphate variables contributed as the second main components (29.25%).

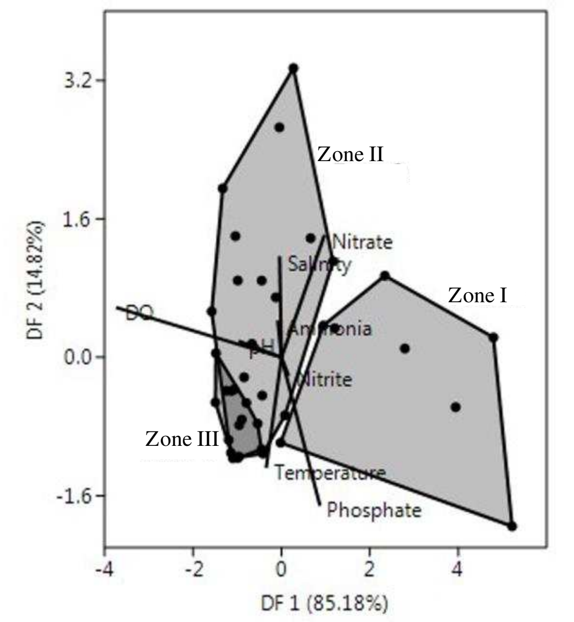
Table 1. Average of water variables in three research zones and ASEAN water quality criteria for aquatic life protection (McPherson et al. 1999)

Variable	Location			water quality criteria
	Zona I	Zona II	Zona III	
Temperature (ºC)	28.74 ± 0.13	28.75 ± 0.43	28.99 ± 0.11	Natural
DO (ppm)	7.01 ± 0.27	7.47 ± 0.28	7.51 ± 0.10	> 4
pH	8.04 ± 0.11	8.17 ± 0.02	8.18 ± 0.01	-
Salinity (ppt)	33.19 ± 0.29	33.35 ± 0.36	33.12 ± 0.07	33 – 34
Phosphate (ppm)	0.09 ± 0.04	0.07 ± 0.02	0.08 ± 0.01	0.045
Nitrite (ppm)	0.001 ± 0.001	0.001 ± 0.001	0.001 ± 0.001	-
Nitrate (ppm)	0.04 ± 0.01	0.04 ± 0.02	0.03 ± 0.02	0.06
Ammonia (ppm)	0.03 ± 0.01	0.03 ± 0.03	0.03 ± 0.02	-

The level of oxygen and ammonia in coastal waters is related to the amount of organic pollutants from domestic areas (Sánchez *et al.*, 2007; Jin *et al.*, 2017). Low DO levels in zone 1 suggests that this zone was the source of pollutants in these waters. High levels of phosphate can be derived from volcanic eruptions (Santana-Casiano *et al.*, 2013). Based on this, it can be seen that the characteristics of waters in the Banda Sea Conservation Area are influenced by two major factors, the domestic anthropogenic and volcanic activities natural run-off.



(a)



(b)

Figure 2. Graph of Principal Component Analysis (a) and Canonical Discriminant Analysis (b) of water wuality parameters of observed stations Banda Sea Conservation Park

The DA Multivariate showed that the observed zones were distinguished by the degree of oxygen solubility (Figure 2b). Dissolved oxygen level in Zone I was significantly different ($P < 0.05$) compared to Zone II and Zone III. Zone I, especially the western side of Neira Island, are the main port of these islands and the most densed domestic populations within conservation areas. The oxygen levels in zone II and III were shown to be higher, due to their geographical characteristics to the open seas.

Based on a multivariate analysis, oxygen levels are shown as the main variable that distinguish between research points and zones. This underlies the selection of DO as the targeted variable in the monitoring design optimization of sampling stations and the number of station in the observed zone. Oxygen solubility is already known as an important water quality variable, which is linked to various biogeochemical cycles, both biotic and abiotic environments (Harris *et al.*, 2015; Huang *et al.*, 2017).

The variogram plot of DO variable was succesfully detected the optimum spacing stations of each zone (Figure 3). Autocorrelation of DO levels in zone III was detected at a distance of 1.05 km, zone II at 0.72 km, and zone I at 0.44 km (1 nautical degree in equator region approximately 110.567 km). These results showed that the spatial autocorrelation distance of Zone III> Zone II> Zone I. Therefore proportionally, more sampling stations in zone I is needed, compared to Zone III. The calculation of equation (2), showed that the optimal sampling stations for zone I, II, and III was 18, 14, and 6 stations at the distance of 0.5, 1, and 2 km (Table 2). Result of this data was plotted on the map, to obtain the optimal grid sampling stations for water quality monitoring in the Banda Sea Conservation Park (Figure 4).

Table 2. Sum of optimal research station in zone I, zone II, and zone III at Banda Sea Conservation Park

Parameter	Location		
	Zone I	Zone II	Zone III
(1) Ratio of standar error	0.086	0.067	0.029
(2) Number of optimal stations	18	14	6

The variogram application in this study is succesfully detect the spatial autocorrelation, to make a correction of sampling stations design which is based on autocorrelation distance (Dheenan *et al.*, 2016). However, the targeted variable in variogram needs to be selected carefully in order to get representative result. The multivariate analysis techniques, such as PCA and DA, can be performed to identify the main variable that describe the difference of multivariables characteristics in the observation area (Panagopoulos *et al.*, 2016).

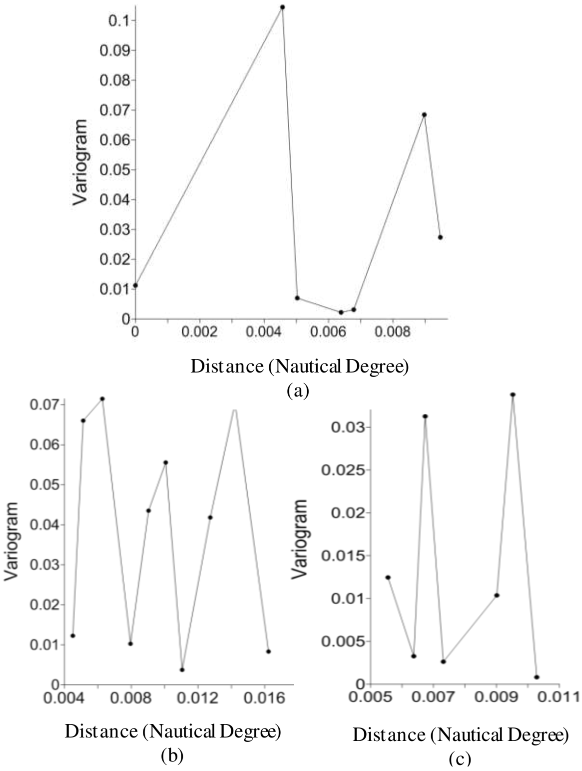


Figure 3. Variogram graph of dissolved oxygen vs distance in zone I (a), zone II (b), and zone III (c) at Banda Sea Conservation park

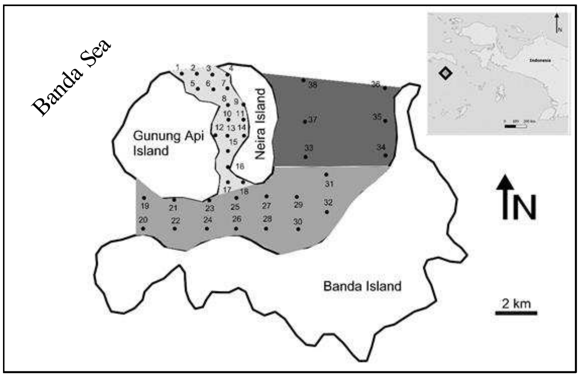


Figure 4. Optimal water quality monitoring stations in zones I ( ), zone II ( ), and zone III ( ) at Banda Sea Conservation Park

4. Conclusion

Systematic grid design is not optimal to maximize the spatial data information in water quality monitoring. Combination of multivariate and spatial variogram statistic is a potential technique to optimize the sampling design of water quality monitoring. Therefore, even the principal variable may vary at other locations, depends on hydrology and other area specific characteristics, the proposed technique could be applied in sampling design concerning water quality monitoring.

References

[1] Crooks, S. M., A. L. Kay, H. N. Davies, V. A. Bell, 2014. From catchment to national scale rainfall-runoff modelling: demonstration of a hydrological modelling framework. Hydrology, 1(1), pp. 63-88.

[2] Dheenan, P. S., D. K. Jha, A. K. Das, N. V. Vinithkumar, M. P. Devi, R. Kirubagaran, 2016. Geographic information systems and multivariate analysis to evaluate fecal bacterial pollution in coastal waters of Andaman, India. Environmental Pollution, 214, pp. 45-53.

[3] De'ath, G., K. Fabricius, 2010. Water quality as a regional driver of coral biodiversity and macroalgae on the Great Barrier Reef. Ecological Applications, 20(3), pp. 840-850.

[4] Duprey, N. N., M. Yasuhara, D. M. Baker, 2016. Reefs of tomorrow: eutrophication reduces coral biodiversity in an urbanized seascape. Global change biology, 22(11), pp. 3550-3565.

[5] Fu, S., C. Y. Wei, 2013. Multivariate and spatial analysis of heavy metal sources and variations in a large old antimony mine, China. Journal of soils and sediments, 13(1), pp. 106-116.

[6] Hammer, O., D. A. T. Harper, P. D. Ryan, 2001. Past: Palaeontological statistics software package for education and data analysis. Palaeontological Electronic, 4(1), 92pp.

[7] Harris, L.A., C. L. S. Hodgkins, M. C. Day, D. Austin, J. M. Testa, W. Boynton, L. Van Der Tak, N. W. Chen, 2015. Optimizing recovery of eutrophic estuaries: impact of destratification and re-aeration on nutrient and dissolved oxygen dynamics. Ecological Engineering, 75, pp. 470-483.

[8] Hedley, J. D., C. M. Roelfsema, I. Chollett, A. R. Harborne, S. F. Heron, S. Weeks, W. J. Skirving, A. E. Strong, C. M. Eakin, T. R. Christensen, V. Ticzon, 2016. Remote sensing of coral reefs for monitoring and management: a review. Remote Sensing, 8(2), pp. 118-168.

[9] Huang, Q., C. Yang, 2011. Optimizing grid computing configuration and scheduling for geospatial analysis: An example with interpolating DEM. Computers & Geosciences, 37(2), pp. 165-176.

[10] Huang, J., H. Yin, S. C. Chapra, Q. Zhou, 2017. Modelling Dissolved Oxygen Depression in an Urban River in China. Water, 9(7), Article number 520, DOI: 10.3390/w9070520.

[11] Januar, H. I., N. P. Zamani, D. Soedarma, E. Chasanah, A. D. Wright, 2017. Tropical coral reef coral patterns in Indonesian shallow water areas close to underwater volcanic vents at Minahasa Seashore, and Mahengetang and Gunung Api Islands. Marine Ecology, 38, e12415, DOI: 10.1111/maec. 12415.

[12] Jin, Q., J. Lu, J. Wu, Y. Luo, 2017. Simultaneous removal of organic carbon and nitrogen pollutants in the Yangtze estuarine sediment: The role of heterotrophic nitrifiers. Estuarine, Coastal and Shelf Science, 191, pp. 150-156.

[13] Kitsiou, D., G. Tsirtsis, M. Karydis, 2001. Developing an optimal sampling design. A case study in a coastal marine ecosystem. Environmental Monitoring and Assessment, 71(1), pp. 1-12.

[14] Lam, N. S. N., 1983. Spatial interpolation methods: a review. The American Cartographer, 10(2), pp. 129-150.

[15] Maas, S., R. Scheifler, M. Benslama, N. Crini, E. Lucot, Z. Brahmia, S. Benyacoub, P. Giraudoux, 2010. Spatial distribution of heavy metal concentrations in urban, suburban and agricultural soils in a Mediterranean city of Algeria. Environmental pollution, 158(6), pp. 2294-2301.

[16] Mattsson, B. J., E. F. Zipkin, B. Gardner, P. J. Blank, J. R. Sauer, J. A. Royle, 2013. Explaining local-scale species distributions: relative contributions of spatial autocorrelation and landscape heterogeneity for an avian assemblage. PLoS One, 8(2), p.e55097, DOI: doi.org/10.1371/journal.pone. 0055097.

[17] McPherson, C. A., P. M. Chapman, G. A. Vigers, K. S. Ong, 1999. ASEAN Marine Water Quality Criteria: Contextual Framework, Principles, Methodology and Criteria for 18 Parameters. ASEAN Marine Environmental Quality Criteria - Working Group (AMEQCWG), ASEAN-Canada Cooperative

Programme on Marine Science - Phase II (CPMS-II). EVS Environment Consultants, North Vancouver and Department of Fisheries, Malaysia. 568pp.

[18] Najar, I. A., A. B. Khan, 2012. Assessment of water quality and identification of pollution sources of three lakes in Kashmir, India, using multivariate analysis. Environmental Earth Sciences, 66(8), pp.2367-2378.

[19] Nemzer, B. V., A. G. Dickson, 2005. The stability and reproducibility of Tris buffers in synthetic seawater. Marine Chemistry, 96(3), pp. 237–242.

[20] Panagopoulos, G. P., D. Angelopoulou, E. E. Tzirtzilakis, P. Giannoulopoulos, 2016. The contribution of cluster and discriminant analysis to the classification of complex aquifer systems. Environmental monitoring and assessment, 188(10), article number 591, DOI: https://doi.org/10.1007/s10661-016-5590-y.

[21] Polónia, A. R. M., D. F. R. Cleary, N. J. de Voogd, W. Renema, B. W. Hoeksema, A. Martins, N. C. M. Gomes, 2015. Habitat and water quality variables as predictors of community composition in an Indonesian coral reef: a multi-taxon study in the Spermonde Archipelago. Science of the total environment, 537, pp. 139-151.

[22] Santana-Casiano, J. M., M. González-Dávila, E. Fraile-Nuez, D. De Armas, A. G. González, J. F. Domínguez-Yanes, J. Escánez, 2013. The natural ocean acidification and fertilization event caused by the submarine eruption of El Hierro. Scientific reports, 3, 1-8.

[23] Sánchez, E., M. F. Colmenarejo, J. Vicente, A. Rubio, M. G. García, L. Travieso, R. Borja, 2007. Use of the water quality index and dissolved oxygen deficit as simple indicators of watersheds pollution. Ecological Indicators, 7(2), pp. 315-328.

[24] Wang, Y., X. Zhang, C. Huang, 2009. Spatial variability of soil total nitrogen and soil total phosphorus under different land uses in a small watershed on the Loess Plateau, China. Geoderma, 150(1), pp. 141-149.