# A STUDY ON CORRELATION AMONG THE MARINE ENVIRONMENTAL DATA BY USING GEOGRAPHICALLY WEIGHTED REGRESSION

Jae-Moon Park[a] and Hong-Joo Yoon[b]

[a] Dept. of Spatial Information Engineering, Pukyong National University

45, Yongso-ro, Nam-gu, Busan;
Tel: +82-51-629-6656; Fax. +82-51-629-6653
E-mail: parkjm56@gmail.com

[b] Dept. of Spatial Information Engineering, Pukyong National University

45, Yongso-ro, Nam-gu, Busan;
Tel: +82-51-629-6656; Fax. +82-51-629-6653
E-mail: yoonhj@pknu.ac.kr

**KEY WORDS**: Spatial statistics, GWR, Marine environment

**Abstract:** In this study, bimonthly *in-situ* data was used to know correlation and change of trend among SST(Sea Surface Temperature), salinity, DO(Dissolved Oxygen) and chlorophyll-a from Aug. 2001 to Dec. 2002 in Gwangyang bay, Korea. GWR(Geographically Weighted Regression) method was used to identify correlation that considered geographical characteristics in the observed stations. Chlorophyll-a set up as the dependent variable. And SST, salinity and DO was set up as the independent variables for applying to GWR

## INTRODUCTION

It was studied that any substance were spatially distributed and had affected each other. Especially, knowing if there's a pattern in your data is useful if you need to gain a better understanding of a geographic phenomenon, monitor condition on the ground, compare patterns, or track changes (Mitchell, 2005). When correlation among factors establish, correlation coefficient calculated by generally statistics method shows characteristics and information of relation between two variable, but spatial autocorrelation explain relation to several observation value included one variable. Therefore it is not same correlation coefficient as Pearson's (Kim, 2003).
The spatial distribution of data should be identified raster data interpolated field data by naked eye or Moran's *I* value. And correlation among several data uses generally Pearson's correlation coefficient. This method does not clearly identify data included geographically factor. Thus the correlation of marine data was identified to use Geographically Weighted Regression (GWR) method reflected geographically factors of field observing station.

## METHODS AND EQUATION

This study used 9 times field observation data bimonthly from Aug. 2001 to Dec. 2002 in Gwangyang bay, Korea(Figure 1). The data is Sea Surface Temperature (SST), salinity, DO (Disolved Oxygen) and chlorophyll-a(Table 1) and used to identify correlation and the trend of change.
And this study was using the GWR 3.0 program computable Geographically Weighted Regression. The dependant variable set up chlorophyll-a and the independent variables set up sea surface temperature(SST), salinity and Disolved Oxygen(DO) influence on survivable environment of chlorophyll-a. It identified to correlation between SST, salinity and DO and Chlorophyll-a.

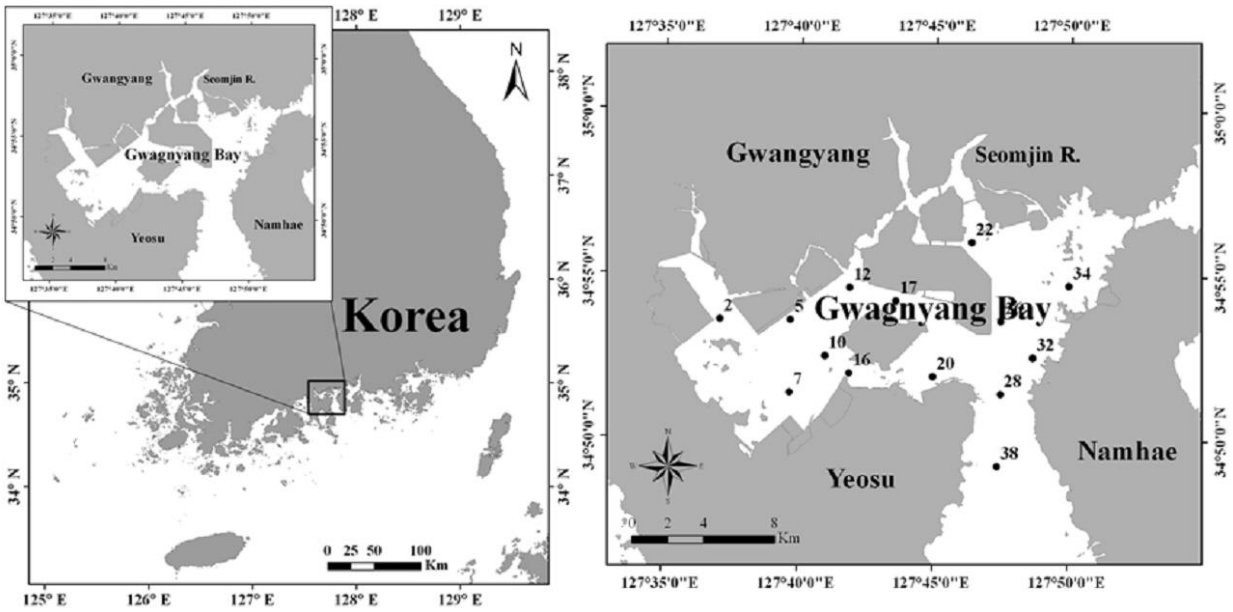


**Figure 1: Studied area and observation stations**

**Table 1: The factors, total number of stations and periods of observation data**.

| Factors | Total number of stations | Periods |
|---|---|---|
| SST | 139 | Aug. 2001 ~ Dec. 2002 (bimonthly) |
| Salinity | 139 | |
| DO | 135 | |
| Chlorophyll-a | 139 | |

## RESULTS

This study used field observation data of SST, salinity, DO, chlorophyll-a in nine times bimonthly from Aug. 2001 to Dec. 2002 in Gwangyang bay, Korea. The weighed option of GWR method was selected to Adaptive kernel type because of irregular location of observation station.

The Fixed kernel type was that using same kernel function of regression analysis at each location, but the Adaptive kernel type was that kernel became smaller in compact distribution and larger in sporadic distribution considering constant size in any location (Jo, 2009).

It was identified that the independent variable seasonally impact of the dependant variable of marine environment factors, calculated GWR seasonally average of factors for making GWR function. The result, the regression coefficients and R-squared were calculated seasonally such as table 2. The impact of independent variable to chlorophyll-a was that SST showed high in summer and low in winter, salinity showed high in winter and low in summer and DO showed high in fall and low in summer. And R-squared was 0.659. The standardized residual was calculated seasonally each observation station to the dependant variables such as figure 2.

**Table 2: Seasonally estimated values of SST, salinity and DO and adjusted R-squared.**

| | Spring | Summer | Fall | Winter |
|---|---|---|---|---|
| SST | 0.174 | 2.364 | 0.588 | -0.923 |
| Salinity | 0.298 | 0.101 | 0.189 | 1.078 |
| DO | 0.315 | -0.721 | 1.827 | 0.061 |
| $R^2_{adj}$ | 0.059 | 0.659 | 0.163 | 0.062 |

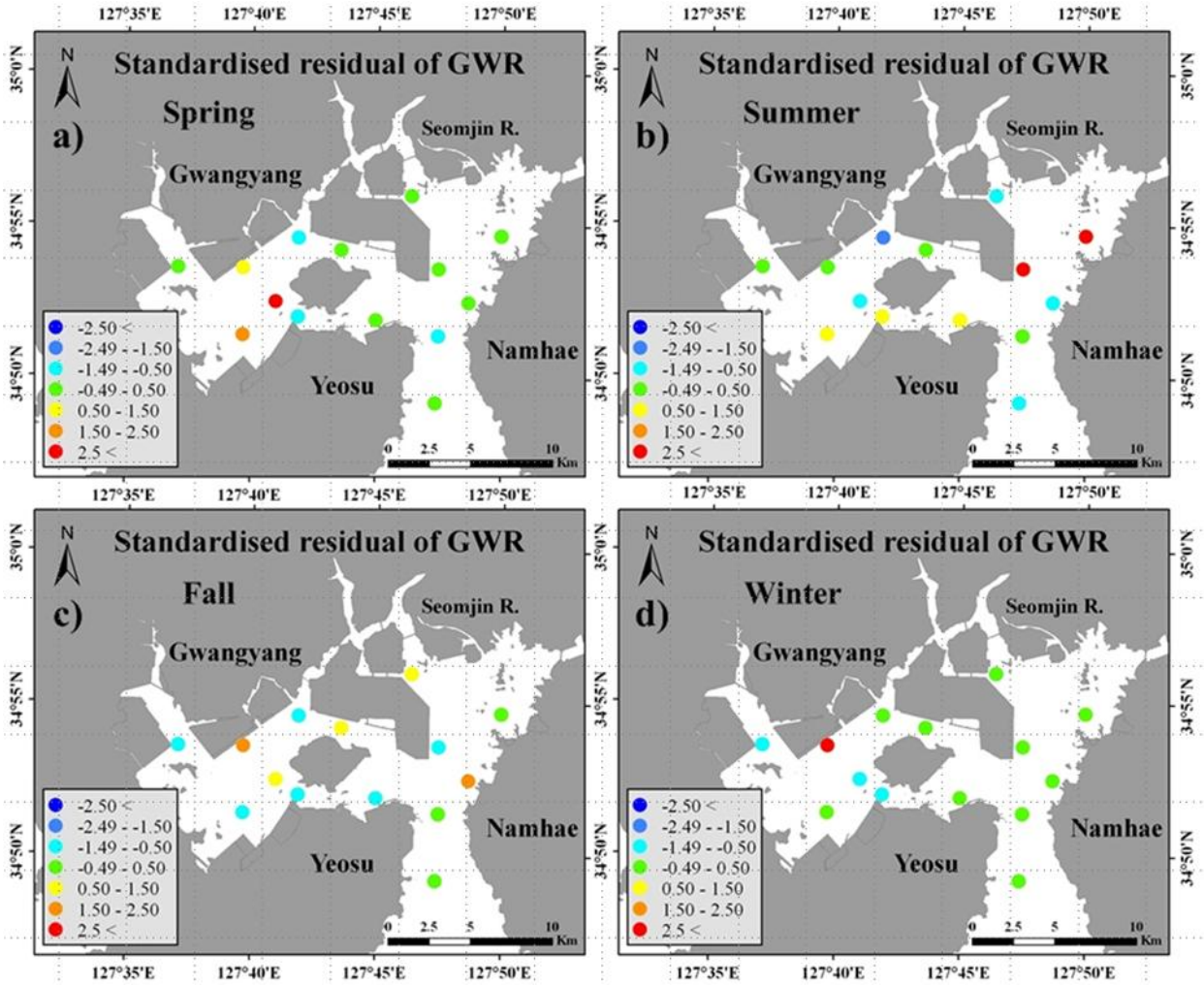


**Figure 2. Seasonally standardized residual using GWR analysis (a)spring, b)summer, c)fall and d)winter)**

## CONCLUSIONS & RECOMMENDATIONS

The Geographically Weighted Regression method was applied to identify influence among SST, salinity, DO and chlorophyll-a of factors composing marine environment. The results, the high R-square showed in condition of high SST, low salinity, low DO in summer.

## REFERENCES:

**References from Journals:**
Jo, D.J., 2009. A Spatial Analysis of Sociodemographic Correlates of Health Related Quality of Life, 32(3), pp.1-20.

**References from Books:**
Michell, A., 2005. The ESRI Guide to GIS Analysis(Volume 2: Spatial Measurements & Statistics). ESRI Press, pp.72.

**References from Other Literature:**
Kim, G.G., 2003. The search of spatial autocorrelation and the application of spatial regression, The Korea Association of Public Administration, pp.983-1001.