Analysis of in situ water temperature and satellite-derived sea surface temperature for Lian, Batangas.

Regine C. Robles*[1], Jose Lorenzo M. Javid[1], Rechel G. Arcilla[1] and Wilfredo Y. Licuanan[2]
[1] *Mathematics Department, De La Salle University*
[2] *Biology Department and the Br. Alfred Shields FSC Ocean Research Center, De La Salle University*
**Corresponding Author: roblesreginec@gmail.com*

Abstract: The Philippine coral reefs are currently threatened by increasing sea surface temperatures (SST), a major factor contributing to mass coral bleaching. SST monitoring is done with the use of situ temperature recorders and satellite-derived measurements. Due to its wide spatial coverage, high temporal resolution, and convenience, satellite-derived SSTs are now more frequently used as a substitute for in situ water temperatures. However, few studies have been dedicated to evaluate the accuracy of these substitutes, especially at depths relevant to reef corals. The objective of this paper is to examine the relationship of satellite-derived SST (NOAA) and in situ water temperature (THERM) of the 7 sites in Lian, Batangas monitored by the Br. Alfred Shields Ocean Research (SHORE) Center. Through ordinary least squares (OLS) regression, it was verified that a relationship exists between NOAA and THERM. The resulting regression model, which includes dummy variables for significant sites and station, yielded an adjusted R^2 of 76.79%. This level of explanatory capability is enough to suggest that satellite-derived SSTs may replace the in situ temperatures. Since satellite-derived data is readily available online, it will be more convenient for use in temperature monitoring and reef management.

Key Words: Sea surface temperature; OLS regression

1. INTRODUCTION

Increased sea surface temperature (SST) is one of the major factors contributing to mass coral bleaching. Corals are highly sensitive to slight changes in SST (Winter et al., 2000) and those that exceeds seasonal maxima even by two degrees over a few weeks can induce bleaching (Jokiel and Coles, 1990; Brown 1997)

Coral bleaching is a common stress response by corals which happens when the symbiotic algae called zooxanthellae detach themselves from the coral making the coral's white skeleton visible. If the mild stress decreased in time, the coral could usually recover from bleaching within several weeks up to a few months. However, this leads to coral mortality if the stress remained and/or intensified (Buccheim, 1998). This would then result to a decrease in the live coral cover.

The impacts of the 2010 mass coral bleaching are most evident in Lian, Batangas and Bolinao-Anda reef system, two of the most monitored reefs in the Philippines. In summer of 2010, the average coral cover in Lian was 19 percent. However, there was an observed loss of 6 percent in coral cover even before the mass bleaching in May. Severe ocean warming was also experienced in the Tubbataha reef for the same year with temperatures above 29℃ that lasted for at least 246 days. (Project MIRROR, 2011)

There are two ways on which temperature data can be recorded: in situ recording and satellite measurements. Due to its wide spatial coverage and high temporal resolution, satellite-derived SSTs are now more frequently used as a substitute for in situ water temperatures particularly for reef monitoring.

However, few studies have been dedicated to evaluate the accuracy of these substitutes. In a study conducted by Kim et al. (2010), satellite-derived temperatures were validated using in situ temperatures through regression analysis which resulted in coefficient of determination, R^2, values not lower than 82%. One disadvantage of satellite-derived SSTs is that cloud contamination may occur. That is, a satellite image could not be taken whenever a cloud is blocking the satellite.

In another study, linear regression was used to verify the accuracy of the satellite-derived SSTs with the use of sea surface water temperatures measured by drifters (Notarstefano et al., 2004). The data from drifters, which are thermistor loggers mounted on drifting buoys, are in situ measurements. Thus, they are more reliable and considered to be the true temperature values. The R^2 values from the four satellites used are at least about 96%.

The objective of this study is to obtain an ordinary least squares (OLS) regression model for satellite-derived SST and in situ water temperature.

2. METHODOLOGY

2.1 Data

The in situ water temperature (℃) data is from the seven sites monitored by the Br. Alfred Shields FSC Ocean Research (SHORE) Center, the research facility of the College of Science of DLSU in Matuod, Lian, Batangas. The seven sites are Clubhouse, Galvez, Layag-Layag, San Diego, Talim Inner, Talim Outer and Toroso, each with two stations A and B (Fig. 1). Thermistor loggers deployed in each of the 7 sites were used to record temperatures in 15-30 minute intervals from July 2009 to April 2013.

The daily satellite-derived SST data of July 2009 to April 2013 were retrieved from the National Oceanic and Atmospheric Administration- National Operational Model Archive & Distribution System (NOAA-NOMADS) Live Access Server (LAS).

2.2 Data Processing

From the in situ water temperature (THERM) data, the daily mean, median, minimum and maximum temperatures were computed to match with the daily interval of the satellite-derived SST, denoted by the NOAA variable. Dummy variables corresponding to the sites and stations were also created.

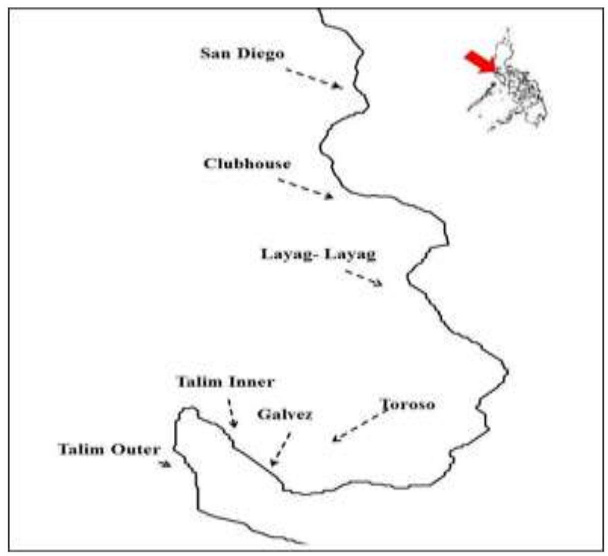

Fig. 1. Map showing the 7 sites monitored by the DLSU SHORE Center

2.3 Analysis

Upon the retrieval of the NOAA data for each site, it was seen that the sites are grouped into two that have similar SSTs. Due to this another dummy variable was created to denote which group the data belongs. Using stepwise selection procedure, the best fit model for predicting THERM with NOAA variable using OLS regression was generated. Shapiro-Wilks and Kolmogorov Smirnov tests for normality, Bartlett's test for homogeneity of variances and Durbin-Watson test for independence were used for residual analysis. All statistical tests were conducted using SAS 9.3 at 5% level of significance.

3. RESULTS AND DISCUSSION

Time series plot of the daily satellite-derived SST data shows some stations having similar SSTs (Fig. 2.). Sites that were found to have similar SSTs were Clubhouse and San Diego, and the sites Galvez, Layag-Layag, Talim Inner, Talim Outer and Toroso. To check if these two groups are significantly different from each other, a paired difference *t*-test was performed. With a p-value of < 0.0001, there is strong evidence to indicate that these two sets differ significantly from each other. Thus, a dummy variable accounting for the group having significantly different average SST was created for the analysis given by

$$\text{GROUP}=\begin{cases}1, & Site = Clubhouse\ or\ San\ Diego \\ 0, & otherwise\end{cases}.$$

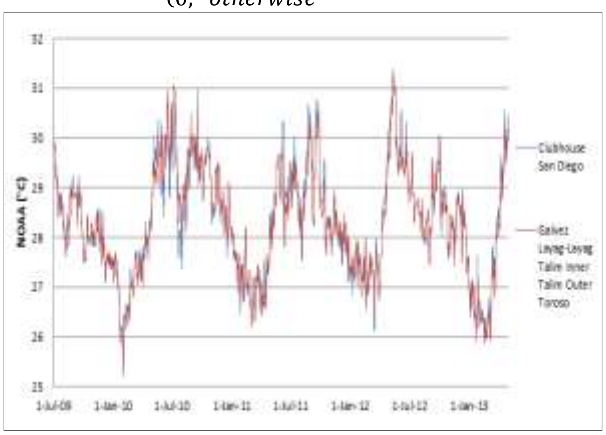

Fig. 2. Time Series plot of NOAA variable

From the in situ water temperature (THERM) data collected, the daily mean, median, minimum and maximum temperatures were computed to match with the daily interval of the satellite-derived SST (NOAA). Table 1 shows that daily mean THERM, denoted by MEANTHERM, have the highest correlation with NOAA ($r = 0.87391$). This indicates a very strong direct linear relationship between NOAA and MEANTHERM (Fig. 3). Hence, MEANTHERM was used as the dependent variable for this analysis.

Table 1. Pearson Correlation Coefficients, n=8546

	NOAA
MEANTHERM	0.87391
MEDIANTHERM	0.87238
MINTHERM	0.86683
MAXTHERM	0.86280

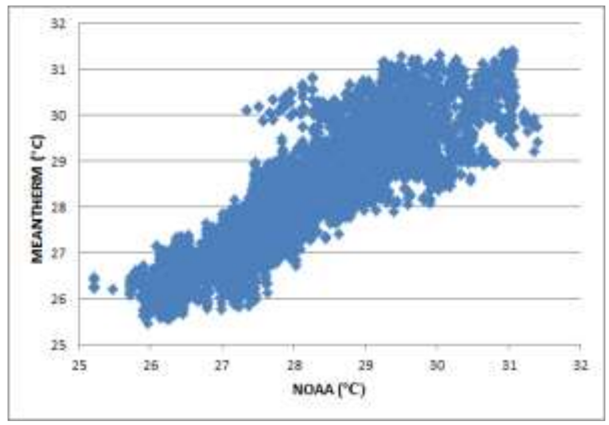

Fig. 3. Scatterplot of NOAA vs. MEANTHERM

The fitted model (MODEL 1) with NOAA and GROUP dummy variable presented an adjusted $R^2 = 0.7658$ and a Root MSE (RMSE) equal to 0.59135. The model is able to explain 76.58% of the variation in MEANTHERM. A p-value < 0.0001 strongly suggests that the fitted model is significant. All variables are significant at $\alpha = 0.05$ (Table 2). There is an expected increase of 0.9692℃ in MEANTHERM for every 1℃ increase in NOAA temperature, holding all other variables constant. MEANTHERM temperatures from sites Clubhouse and San Diego are, on the average, lower by 0.1260℃ compared to all other sites.

Table 2. Parameter estimates for Model 1, n=8546

Variable	Parameter Estimates	Standard Error	Pr>|t|	VIF
Intercept	0.99995	0.1641	<.0001	0
NOAA	0.9692	0.0058	<.0001	1.0013
GROUP	-0.126	0.0143	<.0001	1.0013

Using the Kolmogorov-Smirnov test, it was found that, with a p-value < 0.01, the error terms are not normally distributed. The Durbin-Watson test, which displayed a p-value < 0.0001, showed strong evidence to reject the null hypothesis that there is no serial correlation between error terms. With a p-value < 0.0001 for the Bartlett's test for homogeneity of error variances, there is a strong evidence of heteroscedasticity. Based on Table 2, none of the variables have a Variance Inflation Factor (VIF) value greater than 10. Therefore, non-multicollinearity of variables is satisfied. Also, none of the Cook's D values exceeded the threshold of 0.04328 indicating that there are no influential points in the data.

The set of dummy variables accounting for the seven sites and two stations were also used to examine their effect on MEANTHERM. The reference points for the site and station dummy variables were Toroso and station B, respectively.

The fitted model (Model 2) showed an adjusted $R^2 = 0.7679$ and an $RMSE = 0.58874$. The model is able to explain 76.79% of the variation in MEANTHERM. A p-value < 0.0001 strongly suggests that the model is significant. All variables are significant at $\alpha = 0.05$ (Table 3). For every 1℃ increase in NOAA temperature, there is an expected increase of 0.9672℃ in MEANTHERM, holding all other variables constant.

Table 3. Parameter estimates for Model 2, n=8546

Variable	Parameter Estimates	Standard Error	Pr>|t|	VIF
Intercept	0.8954	0.1644	<.0001	0
NOAA	0.9672	0.0058	<.0001	1.0022
Chouse	-0.0982	0.0201	<.0001	1.3107
Galv	0.0588	0.0200	0.0033	1.3154
Talin	0.1031	0.0202	<.0001	1.2014
Talout	-0.0773	0.0201	0.0001	1.3117
Station	0.1507	0.0199	<.0001	1.2100

Having p-values < 0.01, < 0.0001, and < 0.0001, the assumptions for normality, independence and homogeneity of variance were violated, respectively. Non-multicollinearity is satisfied for all variables based on the VIF values in Table 3. None of the Cook's D values exceeded the threshold of 0.04329 indicating that there are no influential points in the data.

The results of the two models are consistent with the previous findings by Notarstefano, et al. (2004) and Kim et al. (2010).

Models 1 and 2 displayed an adjusted R^2 equal to 0.7658 and 0.7679, respectively. If parsimony is a priority in the selection of the best model, then Model 1 is better. A difference of 0.21% is not that substantial since both models are able to explain about 77% of the variability in MEANTHERM. However, model 2 uses dummy variables for the significant sites which provide more information on the variable MEANTHERM. Also, RMSE is lower than that of model 1. Therefore, model 2 is preferred for MEANTHERM prediction.

4. CONCLUSION

It was found through OLS Regression that a relationship exists between THERM, the in situ water temperature and NOAA, the satellite-derived SST. NOAA, together with the dummy variables for sites and stations (Model 2), can explain 76.79% of the variability in MEANTHERM. This level of explanatory capability is enough to suggest that the satellite-derived SSTs may take the place of in situ temperatures. Data collection will be easier since satellite-derived SST data is readily available online through NOAA-NOMADS.

Since the OLS regression model for MEANTHERM and NOAA failed to satisfy residual assumptions, it is advised that future researchers use a different model such as a Time Series model that could be of better fit to the data. Also, collection of more in situ data points is recommended to better validate the satellite-derived SSTs.

Lastly, due to the accessibility of satellite data, efforts that should have been allotted for carrying out in situ measurements is recommended to be shifted to other data collection such as a more frequent recording of coral cover. By doing so, more information can be obtained regarding the current status of the coral cover which researchers will find valuable.

5. REFERENCES

Brown, B. E. (1997). Coral bleaching: Causes and consequences. Coral Reefs 16: S129-138

Buchheim, J. (1998). Coral Reef Bleaching. Retrieved from http://www.marinebiology.org/coralbleaching.hth.

Jokiel, P.L. and Coles, S.L. (1990) Response of Hawaiian and other Indo-Pacific corals to elevated sea surface temperature. Coral Reefs 8: 155-162

Kim, E. J., Kang, S. K., Jang, S., Lee, J. K., Kim, Y. H., Kang, H., Kwon, Y. Y., & Seung, Y. H. (2010). Satellite-Derived SST Validation based on In-Situ Data during Summer in the East China Sea and Western North Pacific. Ocean Sci. J. 45(3):159-170.

NOAA-NOMADS (2013). Optimum Interpolation SST (AVHRR). [Data file]. Retrieved from http://nomads.ncdc.noaa.gov/las/getUI.do

Notarstefano, G., Mauri, E., & Poulain, P. (2004). Comparison analysis between bulk sea water temperature, as measured by drifters, and sea surface temperature derived by satellite measurements. Rel-I-22 OGA-11, OGS, Trieste, Italy.

Project MIRROR, (2011). Impact of mass coral bleaching and crown of thorns sea stars on selected reefs in the Philippines. De La Salle University Br. Alfred Shields FSC Marine Station. 24 pp.

Winter, A., Appeldoorn, R., Bruckner, A., Williams, E. & Goenaga, C. (1998). Sea Surface Temperature and Coral Reef Bleaching off La Parguera, Puerto Rico (Northeastern Carribean Sea). Coral Reefs 17: 377-382.