ASSESSMENT OF THE SPLINE INTERPOLATION METHOD'S PERFORMANCE IN PREDICTING AND MAPPING THE PHOSPHATE OVER THE COASTAL WATERS OF PULAU TUBA, LANGKAWI, KEDAH

Sharir Aizat Kamaruddin1, Aimie Rifhan Hashim1, Munawwarah Arsyad1, Nur Izzati Mohd Rizal1, Behzad Khairiah Lee1, Noordalila Ramli1, Nurrul Zalaikha Zainal1, Jumiati Mohd Tahir1, Syasya Adhwa Ahmad Afiza1, Nurul Zafirah Hamka1, Eliy Nazira Mat Nazir2

1 Fakulti Sains Gunaan, Universiti Teknologi MARA, Cawangan Perlis, Kampus Arau, 02600, Arau, Perlis, Malaysia
Email: shariraizat@uitm.edu.my, aimierifhan@gmail.com, munawwarahslash71@gmail.com, 2019630718@student.uitm.edu.my, behzadlee.a@gmail.com, noordalilaramli@gmail.com, nurulzainal@gmail.com, jumiatihdir@gmail.com, syasyaafiza97@gmail.com, zafirahhmk@gmail.com

2 Fakulti Pengurusan Perniagaan, Universiti Teknologi MARA, Cawangan Perlis, Kampus Arau, 02600, Arau, Perlis, Malaysia
Email: eliy083@uitm.edu.my

* Corresponding Author

Article Info:

To cite this document:

Abstract:
To date, there has been limited knowledge on the applicability of the spline interpolation method to estimate coastal water quality parameters. This study aims to assess the performance of the regularised and tension spline interpolation techniques for predicting and mapping the phosphate concentration in the surface water of Pulau Tuba, Langkawi, Kedah. Sampling points were set up randomly along the coastal water of Pulau Tuba in November 2018. Samples were obtained using the Niskin Water Sampler at 1 meter below the surface and immediately transported into the marine technology laboratory. Phosphate levels were determined using an Ultraviolet Visible spectrophotometer and the ascorbic acid technique. The Geographic Positioning System (GPS) was used to record the geolocation of each sampling point. The training set (50%) and the testing set (50%) were chosen randomly based on 20 sampling points. The regularized and tension spline interpolation methods were used for this study. Evaluation of prediction of models was carried out using the Root Mean Square Error (RMSE) statistics. The study found that the tension spline interpolation method better predicted phosphate concentration. The Root Mean Square Error (RMSE) was reported at 0.106 and 0.094 for the regularized and tension spline interpolation methods, respectively. Concerned parties can utilize the study's findings as guidelines...
Introduction
Pulau Dayang Bunting and Pulau Tuba are two islands that have gained popularity in the Northern region of Malaysia, which are located near the main island of Pulau Langkawi. Anthropogenic activities can be expected in tourism, agriculture, sea transportation, and residential areas (Kamaruddin et al., 2020). These activities can contribute to decreasing water quality standards and conditions that may directly or indirectly negatively affect living and the non-living things housed in these islands (Kamaruddin et al., 2021b). Introducing excessive water pollutants will negatively impact islanders' way of life and the local economy (Kamaruddin et al., 2018). To date, the baseline information of the phosphate concentration for Pulau Tuba is unknown.

The previous study has shown that the spatial interpolation method can predict the variability of water quality parameters. Kamaruddin et al., (2022a) estimated the concentration of salinity, while Kamaruddin et al., (2021a) estimated the level of water pH using the Inverse Distance Weighted (IDW) interpolation method. A study by Kamaruddin et al., (2022b) managed to map and estimate nitrate concentration using the spline interpolation method recently. Thus, there was a limitation on the knowledge of the spline interpolation method's applicability to map the phosphate concentration.

This study aims to evaluate the applicability of the spline interpolation method to predict the variation of phosphate in the surface water of Pulau Tuba, Langkawi. To achieve this, a comparative assessment was conducted to compare the performance of regularized and tension spline methods.

Literature Review
Phosphate is one of the water parameters that has attracted the attention of environmentalists and ecologists around the globe. In both freshwater and saltwater ecosystems, phosphorus plays a crucial role in controlling the growth of algae and aquatic plants (O’Hare et al., 2018). Phosphorus, at a safe amount, can help plants and soil microorganisms grow typically (Rashid et al., 2016). Phosphorus, on the other hand, can contribute to high levels of eutrophication and toxic algae bloom (Ali & English, 2019). Phosphorus can significantly impact the aquatic ecosystem's balance and water quality, even in small amounts. One approach to estimating phosphate concentration is using the spatial interpolation method that is currently gaining popularity among environmentalists and researchers.

Spatial interpolation techniques can be used for monitoring phosphorus levels and understanding its sources in a study area. Spline interpolation is a deterministic method for predicting pollutant variability (Wu, 2016). Its application in predicting phosphate variation is restricted, particularly in coastal ecosystems and mangrove habitats. The use of spline
interpolation methods to explain the pattern of water pollution has been demonstrated to be both cost and time effective.

Methodology
This section will focus on the description of the study area, sampling activities and design, development of the model, explanations of the mathematical functions of spline (spline equation, tension spline method, and regularized spline method), statistical analysis, and map accuracy assessment.

Description of the Study Area
Pulau Tuba is one of Pulau Langkawi's largest islands. Tourism, agriculture, marine transportation, and residential settlements near coastal areas are all examples of anthropogenic activity that can be expected. Mangrove ecosystems can be found stretched along the strait and the inner part of Pulau Tuba and Pulau Dayang Bunting. Figure 1 shows the sampling area for this study.

![Figure 1: Study and Sampling Area of Pulau Tuba, Kedah.](image)

Sampling Activities and Design
Sampling activities were conducted during the high tide of November 2018. 20 Sampling points were randomly chosen and geolocated using the Global Positioning System (GPS). Samples were collected using the Niskin Water Sampler at 1 meter below the surface depth.
Samples were collected in bottles, labeled from SP01 to SP20, preserved with crushed ice, and immediately transferred to the Marine Technology Laboratory of the Faculty of Applied Sciences, Universiti Teknologi MARA, Cawangan Perlis, Kampus Arau for determination of the concentration of phosphate. Phosphate was determined using the Ultraviolet-visible (UV-VIS) spectroscopy using the ascorbic acid method.

**Development of the Model**
The model was developed and manipulated using ArcMap and ArcCatalog applications available in ArcGIS software. Half of the total sampling points were randomly chosen for the calibration of models, while the leftovers were used to validate the model. The value of 0.1 is selected for the weight parameter as the matter is the default. The number of points is reduced to 10 as ten sampling points were used only to calibrate the model (the default number of points set by the software is 12). The output interpolated surface raster was in the form of a floating-point raster.

**Mathematical Functions for Spline Are Explained through Equation (1) to Equation (5)**

**Splines Equation**

\[ S(x, y) = T(x, y) + \sum_{j=1}^{N} \lambda_j R(r_j) \] 

(1)

Where:
- \( J = 1, 2, \ldots, N \) 
- \( N \) = Number of points 
- \( \lambda_j \) = Coefficients found by the solution of a system of linear equations. 
- \( r_j \) = Distance from the point \((x, y)\) to the \( j^{th} \) point.

**Tension Spline Method**

\[ T(x, y) = a_1 \] 

(2)

Where:
- \( a_1 \) = Coefficients found by the solution of a system of linear equations.

And

\[ R(r) = \frac{1}{2\pi \varphi^2} [1n \left( \frac{r \varphi}{2} \right) + c + K_c(r \varphi)] \] 

(3)

Where:
- \( \varphi \) = Weight parameter 
- \( K_c \) = Modified Bessel function 
- \( C \) = Constant equal to 0.0577215
**Regularized Spline Method**

\[ T(x, y) = a_1 + a_2x + a_3y \] (4)

Where:
\[ a_1 \] = coefficients found by the solution of a system of a linear equation.

And,
\[ R(r) = \frac{1}{2\pi} \left\{ r^2 \left[ \ln \left( \frac{r}{2\pi} \right) + c - 1 \right] + \tau^2 \left[ K_\circ \left( \frac{r}{\tau} \right) + c + 1 \ln \left( \frac{r}{2\pi} \right) \right] \} \] (5)

Where:
\[ r \] = Distance between the point and the sample,
\[ \tau^2 \] = Weight parameter
and \[ K_\circ \] = Modified Bessel function.

**Statistical Analyses**

Descriptive statistics, the Pearson product-moment correlation coefficient analyses (1-tailed), linear regression assessment, and paired sample T-test were conducted using the Statistical Package for the Social Sciences (SPSS) Version 21. The alpha for all the analyses was set at 0.05. Evaluation of the model was carried out using the Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and Mean Absolute Error (MSE) statistics tabulated in an Excel spreadsheet.

**Map Accuracy Assessment**

The assessment was conducted to determine the correct occurrences predicted by the model or method. Classification of ranges was analyzed and set up before the calculation. Classification of phosphate concentration and the number of classes were done manually using ArcMap applications.

\[ \text{Overall Accuracy (OA)} = \frac{\text{Number of Correct Occurrences}}{\text{Total Occurrences}} \times 100 \] (6)

**Result and Discussion**

Spatial models developed using regularized and tension spline were successfully developed for phosphate. The spatial model developed using the regularized method had the range of phosphate at 0.063mg/L to 1.090mg/L, while for spatial model predicted using the tension method had the range of phosphate at 0.021mg/L to 0.393mg/L. Figure 2 shows the spatial model produced by both models.
The attribute table for both models was open and transferred to excel spreadsheets for further investigation. The total number of observations (N) is reported at 10. Table 1 shows the measured and predicted values for all sampling points.

<table>
<thead>
<tr>
<th>Sampling Points</th>
<th>Measured values (mg/L)</th>
<th>Regularized Model (mg/L)</th>
<th>Tension Model (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP02</td>
<td>0.020</td>
<td>0.221</td>
<td>0.248</td>
</tr>
<tr>
<td>SP04</td>
<td>0.326</td>
<td>0.422</td>
<td>0.371</td>
</tr>
<tr>
<td>SP06</td>
<td>0.205</td>
<td>0.106</td>
<td>0.129</td>
</tr>
<tr>
<td>SP08</td>
<td>0.172</td>
<td>0.178</td>
<td>0.170</td>
</tr>
<tr>
<td>SP10</td>
<td>0.295</td>
<td>0.444</td>
<td>0.363</td>
</tr>
<tr>
<td>SP12</td>
<td>0.065</td>
<td>0.110</td>
<td>0.160</td>
</tr>
<tr>
<td>SP14</td>
<td>0.157</td>
<td>0.293</td>
<td>0.227</td>
</tr>
<tr>
<td>SP16</td>
<td>0.263</td>
<td>0.177</td>
<td>0.174</td>
</tr>
<tr>
<td>SP18</td>
<td>0.148</td>
<td>0.175</td>
<td>0.174</td>
</tr>
<tr>
<td>SP20</td>
<td>0.132</td>
<td>0.097</td>
<td>0.107</td>
</tr>
</tbody>
</table>

Descriptive statistics were used to summarize the data distribution for both measured and predicted values generated by both models. The skewness values were reported at -.018, .978, and .987 for the measured values, regularized model, and the tension model, respectively. Moreover, the kurtosis values were reported at -.620, -.284, -.120 for the measured values, regularized model, and the tension model, correspondingly. Standard Errors for skewness and
Kurtosis were calculated at 0.687 and 1.334 for all models, including the measured phosphate values. Table 2 shows the data for the descriptive analysis.

### Table 2: Descriptive Statistics Between the Measured and Predicted Values Generated by Both Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Range Statistic</th>
<th>Minimum Statistic</th>
<th>Maximum Statistic</th>
<th>Mean Statistic</th>
<th>Std. Error</th>
<th>Std. Deviation Statistic</th>
<th>Variance Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured</td>
<td>.31</td>
<td>.02</td>
<td>.33</td>
<td>.1783</td>
<td>.03069</td>
<td>.09705</td>
<td>.009</td>
</tr>
<tr>
<td>Regularized model</td>
<td>.35</td>
<td>.10</td>
<td>.44</td>
<td>.2223</td>
<td>.03974</td>
<td>.12566</td>
<td>.016</td>
</tr>
<tr>
<td>Tension model</td>
<td>.26</td>
<td>.11</td>
<td>.37</td>
<td>.2123</td>
<td>.02884</td>
<td>.09121</td>
<td>.008</td>
</tr>
</tbody>
</table>

A Pearson product-moment correlation coefficient was computed to assess the correlation between the observed and predicted value of phosphate, based on regularized and tension spline methods at 1-tailed correlation analysis (alpha = .05). For estimation made by regularized spline method, there was a positive correlation between the 2 variables, \( r = .614, n = 10, p = .03 \). On the other hand, for the prediction made by the tension spline method, there was a positive correlation between the 2 variables, \( r = .525, n = 10, p = .06 \). Overall, there was a strong, positive correlation between the observed and predicted value of phosphate for both methods. The analysis indicates that only the prediction made by the regularized model had a significant value \( p < .05 \). Table 3 shows the correlation analysis between the measured and predicted values generated by both models.

### Table 3: Correlation Between the Measured and Predicted Values Generated by Both Models

<table>
<thead>
<tr>
<th></th>
<th>Regularized Model</th>
<th>Tension Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measured</td>
<td>Pearson Correlation: .614*</td>
<td>.525</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.030</td>
<td>.060</td>
</tr>
<tr>
<td>Sum of Squares and</td>
<td>.067</td>
<td>.042</td>
</tr>
<tr>
<td>Cross-products</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Covariance</td>
<td>.007</td>
<td>.005</td>
</tr>
<tr>
<td>N</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 Level (1-tailed).

A paired-samples t-test was conducted to compare the value of phosphate in measured and predicted conditions. For Pair 1 (Measured vs. predicted made by regularized model), there was an insignificant difference in the value of phosphate for measured (\( M = .1783, SD = .09705 \)) and predicted (\( M = .2223, SD = .12566 \)) conditions; \( t (9) = -1.375, p = .202 \). In contrast, For Pair 2 (Measured vs. predicted made by tension model), there was an insignificant difference in the value of phosphate for measured (\( M = .1783, SD = .09705 \)) and predicted (\( M = .2123, SD = .09121 \)) conditions; \( t (9) = -1.170, p = .272 \). Table 4 shows the result of paired samples test between the measured and predicted values generated by both models.
A simple linear regression analysis was used to predict the relationship between the observed and predicted values of phosphate, based on the regularized and tension spline methods at a 95% confidence level. The prediction value of phosphate estimated by the regularized spline method did not explain a significant amount of variance in the observed value of phosphate, $F(1,8) = 4.834$, $p = .059$, $R^2 = .377$, $R^2_{\text{adjusted}} = .299$. Similarly, the prediction value of phosphate estimated by the tension spline method did not explain a significant amount of variance in the observed value of phosphate, $F(1,8) = 3.043$, $p = .119$, $R^2 = .276$, $R^2_{\text{adjusted}} = .185$. A simple scatter plot with a regression line between the predicted and observed phosphate reading produced by both methods is shown in Figure 3.

![Figure 3: Simple Scatter Plots with a Regression Line Between the Predicted and Observed Phosphate Reading Produced by Both Methods](image_url)

Evaluation of the model based on regularized and tension spline methods was conducted using the MAE, MAPE, RMSE, and MSE. The MAPE values are high due to a large amount of prediction error at Sampling points SP02 and SP12. MAPE is very sensitive to an outlier or
extreme data. As a result, many researchers used RMSE as the best evaluation tool to evaluate the model’s accuracy. Based on the error analysis, the study found that the tension spline method produced fewer errors when compared to the regularized method. Results indicate that the spline interpolation method can map and estimate the concentration of water contaminants (Kamaruddin et al., 2019). Table 5 shows the result of the error analysis for both models.

**Table 5: The Result of the MAE, MAPE, RMSE, and MSE Analysis for Both Models.**

<table>
<thead>
<tr>
<th></th>
<th>Regularized Method</th>
<th>Tension Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>0.088</td>
<td>0.072</td>
</tr>
<tr>
<td>MAPE</td>
<td>136.96%</td>
<td>147.69%</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.106</td>
<td>0.094</td>
</tr>
<tr>
<td>MSE</td>
<td>0.011</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Next, the transformation of the spatial model to a map was carried out. Maps elements such as title, scale, north arrow, legend, and scale text were inserted into the map. Map accuracy assessment was carried out to investigate the performance of the map to predict the variation of phosphate concentration. The range for legend is fixed at < 0.2, 0.2-0.39, >0.39. The map accuracy was determined at 60%. Overestimation and underestimation by spline interpolation methods reduced the accuracy of the model. The same situation has been found by Kamaruddin et al., (2020). Table 6 shows the correct and incorrect occurrences tabulated for the accuracy of the map developed using the tension method.

**Table 6: Incorrect Occurrences Were Found at Four Sampling Points.**

<table>
<thead>
<tr>
<th>Sampling Points</th>
<th>Measured values</th>
<th>Tension Model</th>
<th>Correct occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>SP02</td>
<td>&lt; 0.2</td>
<td>0.2-0.4</td>
<td>x</td>
</tr>
<tr>
<td>SP04</td>
<td>0.2-0.4</td>
<td>0.2-0.4</td>
<td>/</td>
</tr>
<tr>
<td>SP06</td>
<td>0.2-0.4</td>
<td>&lt; 0.2</td>
<td>x</td>
</tr>
<tr>
<td>SP08</td>
<td>&lt; 0.2</td>
<td>&lt; 0.2</td>
<td>/</td>
</tr>
<tr>
<td>SP10</td>
<td>0.2-0.4</td>
<td>0.2-0.4</td>
<td>/</td>
</tr>
<tr>
<td>SP12</td>
<td>&lt; 0.2</td>
<td>&lt; 0.2</td>
<td>/</td>
</tr>
<tr>
<td>SP14</td>
<td>&lt; 0.2</td>
<td>0.2-0.4</td>
<td>x</td>
</tr>
<tr>
<td>SP16</td>
<td>0.2-0.4</td>
<td>&lt; 0.2</td>
<td>x</td>
</tr>
<tr>
<td>SP18</td>
<td>&lt; 0.2</td>
<td>&lt; 0.2</td>
<td>/</td>
</tr>
<tr>
<td>SP20</td>
<td>&lt; 0.2</td>
<td>&lt; 0.2</td>
<td>/</td>
</tr>
</tbody>
</table>

Increased phosphate concentration can be observed in the inner part of Pulau Dayang Bunting, while the coast of Pulau tuba experiences decreasing phosphate concentration. The increment of phosphate could be due to anthropogenic activities where a cluster of villages can be found here. Furthermore, less concentration of phosphate along the coastal waters of Pulau Tuba was observed could be because of low human activities. Studies have shown that there have been many connections between phosphorus levels and anthropogenic activities (Malone & Newton, 2020). However, the phosphate concentration is still below the limit of 1.0 mg/L.

Poor agricultural practices (Lwin et al., 2017), runoff from cities and lawns (Yang & Toor, 2018), leaking septic systems (Macintosh et al., 2011), and sewage treatment plant outputs (Li et al., 2021) can all contribute to high phosphorus levels. Increased development of algae and
big aquatic plants (Mustafa & Hayder, 2021), as well as decreasing levels of dissolved oxygen (Osaka et al., 2021), can occur from too much phosphorus, a process known as eutrophication (Isiuku & Enyoh, 2020). Algal blooms caused by high phosphorus levels can produce algal toxins (Moll, 2021) that are detrimental to humans (Sidabutar et al., 2020) and animal health (Nishikawa et al., 2009). Awareness of the effects and impact of deteriorating water quality standards on the marine environment can be made by effectively educating youngsters and the public, and the impact of climate change on the water quality should also be acknowledged (Kamaruddin et al., 2022c). Figure 4 shows the distribution of phosphate over the coastal water of Pulau Tuba, Kedah.

Figure 4: Map of Phosphate Developed Using Tension Spline Method

Conclusions
In conclusion, the tension spline method is superior to regularize spline method in predicting the variability of surface water phosphate. The tension spline method produced less prediction error than the regularized spline method. The research concludes that spline can effectively estimate water quality parameters.
This study has two critical flaws that could be addressed in future studies. The study was conducted during a short period, and fluctuations in phosphate levels are expected and influenced by the monsoon season, tidal activities, and the time of sampling activities. Because every parameter has a considerable link with other water parameters, water quality study is complicated. Second, time constraints contribute to the low sampling frequency and point count. Due to a lack of baseline data for water phosphate, this research was hampered by the lack of previous investigations in this field. However, this constraint has been thoroughly examined and supported by literature and data from other researchers. Because of the research's limitations, new research in the study area is now possible. According to the study, sampling operations should be carried out over a more extended period to construct a robust model. The impacts of the tidal activity, seasonal change, and sample period should also be considered.

Nonetheless, the findings of this study can be used by government and non-government organizations as baseline data, allowing them to properly monitor and sustain the coastal and marine water of Pulau Tuba, Langkawi. Because phosphate significantly impacts fish growth and development, the developed map can also be used by entrepreneurs when deciding on aquaculture or mariculture sites.

Acknowledgments
The authors gratefully acknowledge the generous assistance and support from the academic and non-academic staff for their contribution to this research and publication, especially to UiTM Marine Research Station (MARES); the Research Initiative Group (RIG): Ocean Research, Conservation and Advance (ORCA), Integrative Natural Product Research; and the Faculty of Applied Sciences, Universiti Teknologi MARA, Perlis Branch, Arau Campus, 02600 Arau, Peris, Malaysia.

References


of Universiti Teknologi MARA, Perlis branch. *Jurnal Intelek*, 17(1), 236. doi:10.24191/ji.v17i1.16125


