Prediction of seawater salinity using truncated spline regression method

by Faisol Faisol

Prediction of Seawater Salinity Using Truncated Spline Regression Method

Faisol[1, a)], Tony Yulianto[1)], Moh. Yaqin[1)], Achmad Basuki[2)], and Muhammad Agus Zainuddin[2)]

[1]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Islam Madura, Pamekasan, 69317, Indonesia.*
[2]*Creative Multimedia Technology Department, Surabaya State Electronics Polytechnic, Surabaya, 60111, Indonesia.*

[a)] *Corresponding author: faisol.munif@gmail.com*

Abstract. Salinity is the level of dissolved salt in water, which is one of the factors that affect salt production, the higher the salinity dissolved in seawater, the better the resulting salt production. The factors that affect seawater salinity include air humidity, wind speed and seawater temperature. In this study, the Spline Truncated Regression method was applied to predict seawater salinity based on the variables that influence it. Data collection was obtained using satellite images obtained from Landsat 8. The data was taken within 1 year from January to December 2019. From the analysis results, it was found that the linear spline model with 1 knot point is the best model with a minimum GCV value.0.3648021178, with an R-Sq of 0.580443194, the MSE value is 0.136237833, and the MAPE is 0.98349512. Based on the MAPE value, the prediction model is said to be very accurate forecasting.

INTRODUCTION

As an archipelagic country, Indonesia has great potential to produce and be self-sufficient in salt. However, so far the amount of existing salt production has not been able to meet domestic salt needs. Salt-making centers are concentrated on the islands of Java and Madura, with an area of 10,231 Ha and 15,347 Ha, respectively. Other salt-making locations are in West Nusa Tenggara, South Sulawesi, and Sumatra with an area of 1,155 Ha, 2,040 Ha, and 1,885 Ha, respectively. Thus, the total area of salt fields in Indonesia is 30,658 hectares, of which around 25,542 hectares are traditionally managed by the people [1].

In general, salt refers to a chemical compound with the name Sodium Chloride or Sodium Chloride (NaCl). Salt is one of the complementary needs for food and a source of electrolytes for the human body. The need for good quality salt is still met through imports, especially iodized salt and industrial salt. One of the largest salt-producing islands in Indonesia is Madura Island [2].

Madura has long been known as the Salt Island which has a salt area of 15,347 hectares located in several districts, namely in Sumenep with a salt area of 10.067 hectares, Pamekasan 3,075 hectares and in Sampang 1,046 hectares. The district in Madura which is the largest salt producer in Indonesia is Pamekasan Regency, which is located in three sub-districts, namely Galis, Pademawu and Tlanakan. The widest people's salt pond area is in Pademawu District reaching 500.7 Ha, while the people's salt pond area in Galis District is 353 Ha and in Tlanakan District only 35 Ha [1].

The quality of salt produced traditionally is generally very low and must be reprocessed to be used as salt for consumption and industrial salt. This low salt quality is caused by several factors, one of which is salinity. Salinity is the amount of dissolved salts in water. Salinity is part of the physical and chemical properties of a water, in addition to temperature, pH, substrate and others. Water salinity describes the salt content in a water. The salt in question is a variety of ions dissolved in water, including table salt (NaCl). In general, salinity is caused by 7 main ions, namely sodium (Na), chloride (Cl), calcium (Ca), magnesium (Mg), potassium (K), sulfate (SO4) and bicarbonate (HCO3).To improve the quality of national salt, it is necessary to predict the salinity value. Salinity is defined as the total weight of salt dissolved in 1 liter of air, usually expressed in units of ‰. Measuring salinity is an important biological factor that must be measured to see life under the sea. Salinity detection is also useful for determining model circulation and pollution in coastal areas for irrigation. In addition, this measurement is to measure the balance between evaporation and ending and can be identified as a priority measurement. By using Landsat-8 satellite imagery, the level of salinity quality can be done using the salinity algorithm. One method to predict seawater salinity is using the Spline Truncated regression method.

Spline Regression is a regression analysis that is able to estimate data that does not have a certain pattern and has a tendency to find its own estimated data from the pattern formed [3]. One of the advantages of the Spline approach is that this model tends to find its own estimate of the data wherever the data pattern moves. This excess occurs because in the Spline there are knot points, which are joint points of integration that indicate changes in data behavior patterns. Truncated is a function that can be interpreted as a cut function [4].

Research using Truncated Spline regression has previously been conducted by D A Widyastuti, A A R Fernandes, H Pramoedyo (2020) titled "Spline estimation method in nonparametric regression using truncated spline approach" which results, based on the best formed regression curve estimation model selected with the criteria of choosing the best model using MSE, the smaller MSE value is the best formed model. From the R2 model we can conclude that the nonparametric model is better than the parametric model. The Coefficient of determination (R2) shows that the variability of response variables can be explained by 87.65% in the model, while the remaining 12.35% is explained by factors not included in the model. And in other studies conducted by Ach. Khozaimi, Yoga Dwitya Pramudita, Eka Mala Sari Rochman, and Aeri Rachmad (2020) titled "Decision Support System for Determining the Quality of Salt in Sumenep Madura-Indonesia" which produces, Provides criteria and parameters about salt can provide the best quality salt. Test functionality that can make all system functions work properly. Therefore, in this study was raised the title "Prediction of Salinity of Sea Water Using Spline Truncated Regression Method" so that with this title is expected to improve the quality of salt in Madura.

LITERATURE REVIEW

Salt

Salt is a white solid in the form of crystals which is a collection of compounds with the main composition of sodium chloride (more than 8%) and other compounds such as magnesium chloride, magnesium sulfate, calcium chloride, etc. Salt has properties or characteristics which mean it is easy to absorb water, bulk density of 0.8 to 0.9 and a melting point at a temperature level of 0.8 to 0.9801°C [2].

Salt produced from the evaporation and crystallization of seawater is known as coarse salt (krosok), this krosok salt has a low quality, namely the average sodium chloride (NaCl) content is only 85%, and contains impurities such as magnesium sulfate (MgSO4). , calcium sulfate (CaSO4), magnesium chloride (MgCl2), potassium chloride (KCl) and soil impurities. This krosok salt cannot be consumed directly by the public or as a raw material or auxiliary material for industrial needs such as the soda industry, oil, textiles and so on because the NaCl content is still below the Indonesian National Standard (SNI). Consumable salt is iodized consumption salt with a minimum sodium chloride (NaCl) content of 94.7% on a dry basis, maximum 7% water, and maximum 0.5% water insoluble part [1].

Salinity

Salinity is defined as the total weight of salt dissolved in 1 liter of air, usually expressed in units of ‰. Measuring salinity is an important biological factor that must be measured to see life under the sea. Salinity detection is also useful for determining model circulation and pollution in coastal areas for irrigation. In addition, this measurement is to measure the balance between evaporation and ending and can be identified as a priority measurement. By using Landsat-8 satellite imagery, the level of salinity quality can be done using the salinity algorithm [9]. The equation is:

$$Salinity = 10^{[-0.141 \times log_{10}(Cp)1,45)}$$

Cp is the attenuation coefficient of light entering the water surface. Cp can be calculated using the following equation:

$$Cp = 10^{[0.71 \times MNDCI^3 + 0.96 \times MNDCI^2 + 1.14 \times MNDCI - 0.25]}$$

MNDCI (Maximum Normal Difference Carbon Index) is calculated using the following formula:

$$MNDCI = \left[\frac{[Rss(555) - max(Rss(412), Rss(443), Rss(490))]}{[Rss(555) + max(Rss(412), Rss(443), Rss(490))]}\right]$$

Rss is the result of digital number processing which has been converted into a reflector value and divided by 3.14. The numbers 412, 443, 490, and 555 are wavelength values.

Salt Farmer

Salt farmers are business actors in the area who make krosok salt from seawater raw materials with a certain process. The processing of salt pond land in Pamekasan Regency is generally divided into two, namely processing by PT Garam and private salt companies and management by individuals known as people's salt fields. The salt pond area managed by the company is 980.0 Ha (52.44%) while the remaining 888.70 Ha (47.56%) is the people's salt pond. People's salt ponds are located in three sub-districts, namely Galis, Pademawu and Tlanakan. The area of the largest people's salt ponds is in Pademawu District reaching 500,7 Ha, while Galis Subdistrict has 353 people's salt ponds and Tlanakan Subdistrict is an area with the narrowest people's salt pond area of 35 Ha. In order to increase the amount of salt production.

Truncated Spline Regression

Truncated Spline Regression is an approach towards fitting the data while taking into account the smoothness of the curve. Spline is a segmented polynomial model. The segmented nature will provide good flexibility. This property allows the spline Regression model to adapt effectively to the characteristics of the data. In the longitudinal data there are $i = 1,2,\dots,m$ subject and observations in each subject, then the Spline Truncated function is of the order of knots $j = 1,2,\dots,n_i pK = \{K_1, K_2, \dots, K_r\}$ [4] for longitudinal data can be given by the following equation.

$$f(x_i) = \sum_{s=0}^{p-1} \beta_{is} x_i^s + \sum_{s=1}^{r} \beta_{1(p+s-1)} (x_i - K_{is})_+^{p-1}$$

With truncated function of equation $(x_i - K_{is})_+^{p-1}$ obtained

$$(x_i - K_{is})_+^{p-1} = \begin{cases} (x_i - K_{is})^{p-1} & ; x_i - K_{is} \geq 0 \\ 0 & ; x_i - K_{is} < 0 \end{cases}$$

From the equation above obtained nonparametric regression model Spline Truncated for longitudinal data of the p-order in the following equation:

$$y_i = \sum_{s=0}^{p-1} \beta_{is} x_i^s + \sum_{s=1}^{r} \beta_{i(p+s-1)} (x_i - K_{is})_+^{p-1} + \varepsilon_i$$

In conjunction with the estimation of smooth curves $f(x_{ij})$ in spline nonparametric regression for longitudinal data with knot points $K = \{K_1, K_2, \dots, K_r\}$ then the estimate for β_i is in the following equation:

$$\hat{\beta}_{iK} = (X_{iK}^T X_{iK})^{-1} X_{iK}^T y_i$$

And the estimation function is $f(X_{ij})$

$$\hat{f}_K(X_i) = X_{iK} \hat{\beta}_{iK} = X_{iK} (X_{iK}^T X_{iK})^{-1} X_{iK}^T y_i = A_{iK} y_i$$

Criteria for Selection of the Best Estimate

Indicators of the goodness of the model from nonparametric regression can be seen from the following measures:

Mean Square Error (MSE)

The method that is often used to evaluate forecasting results is the Mean Square Error (MSE) method. By using MSE, the error that shows how big the estimation results are with the results to be estimated, in the forecasting phase, using MSE as a measure of accuracy can also cause problems. This measure does not facilitate comparisons between different time series and for different time intervals, because MSE applies absolute measures. In addition, the interpretation is not intuitive even for specialists, because this measure involves squaring a series of values [5]. The equation for calculating the MSE is given, namely:

$$MSE[K] = \frac{1}{n}\mathbf{y}^T(\mathbf{I} - \mathbf{A}[K])^T(\mathbf{I} - \mathbf{A}[K])\mathbf{y}$$

Generalized Cross Validation (GCV)

Generalized Cross Validation (GCV) is one of several criteria to determine the optimal value of k, the value of k is the value of the oscillation parameter. GCV is often used in nonparametric regression because it has asymptotic optimal properties when compared to other model validation methods [6]. GCV is determined by the following equation:

$$GCV(K) = \frac{MSE\ (K)}{(n^{-1}trace(I - \boldsymbol{A}[K]))^2}$$

with:

K	: oscillation parameters
n	: amount of data
I	: identity matrix
$A[K]$	: hat matrix

To obtain the oscillation parameters, it can be seen from the minimum GCV value.

Coefficient of Determination R^2

The coefficient of determination is a measure of the contribution of the predictor variables to the response variable. Given the formula for the coefficient of determination as follows:

$$R^2 = \frac{(\hat{y} - \bar{y})^T(\hat{y} - \bar{y})}{(y - \bar{y})^T(y - \bar{y})}$$

with $\bar{y}$ is a vector that contains the average response. A good model can be measured from a large R^2 value [7].

Mean Absolute Percentage Error (MAPE)

According to Pakaja (2012), the Mean Absolute Percentage Error (MAPE) is calculated using errorabsolute value in each period divided by the actual observed value for that period. Then, average the absolute percentage error. MAPE is an error measurement that calculates the size of the percentage deviation between actual data and forecasted or predicted data [8].

Another researcher explained that the MAPE value provides an indication of how much the average absolute forecast error is compared to the true value. The formula for calculating MAPE is as follows:

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

RESEARCH METHODS

Several stages of research are described that will be used to achieve research objectives

Descriptive statistics

The description of the data is useful for knowing the general description of the data about the predictor and response variables used. The general description discussed is the maximum, minimum, mean, standard deviation measures.

In this study there is one response variable and two predictor variables. The response variable used is Salinity (Y) and the predictor variables used are humidity (X_1), wind speed (X_2) and sea water temperature (X_3). The data can be shown in a table form as following:

Table 1 Descriptive Statistics of Seawater Salinity Data

Data	Minimum	Maximum	Mean (Average)	Standard Deviation
Y	29.3110	31.75882	30.80594	0.574608
X_1	69.3	91.5	77.89091	6.757494
X_2	4.2	19.5	12.56364	4.215435
X_3	22.2	26.4	23.86818	1.187389

Based on Table 1, it can be seen that seawater salinity data taken from January 2019 to December 2019 obtained minimum salinity data (Y) of 29.3110, maximum data of 31.75882, with an average of 30.80594 and a standard deviation of 0.574608. The minimum data on air humidity (X_1) is 69.3, the maximum data is 91.5, with an average of 77.89091 and a standard deviation of 6.757494. Then the minimum data wind speed (X_2) is 4.2, the maximum data is 19.5, with an average of 12.56364 and a standard deviation of 4.215435. And the minimum data for sea water temperature (X_3) is 22.2, the maximum data is 26.4, with an average of 23.86818 and a standard deviation of 1.187389.

Scatterplot

Before being analyzed using spline regression analysis. First, the pattern of the relationship between each predictor variable and the response variable was investigated by using a scatter plot. The plot is carried out with the aim of knowing the early detection of data patterns. The scatter plot made is done one by one between each predictor variable and the response variable so that 3 scatter plots are formed. The results of the scatter plot can be seen in Figure 1.

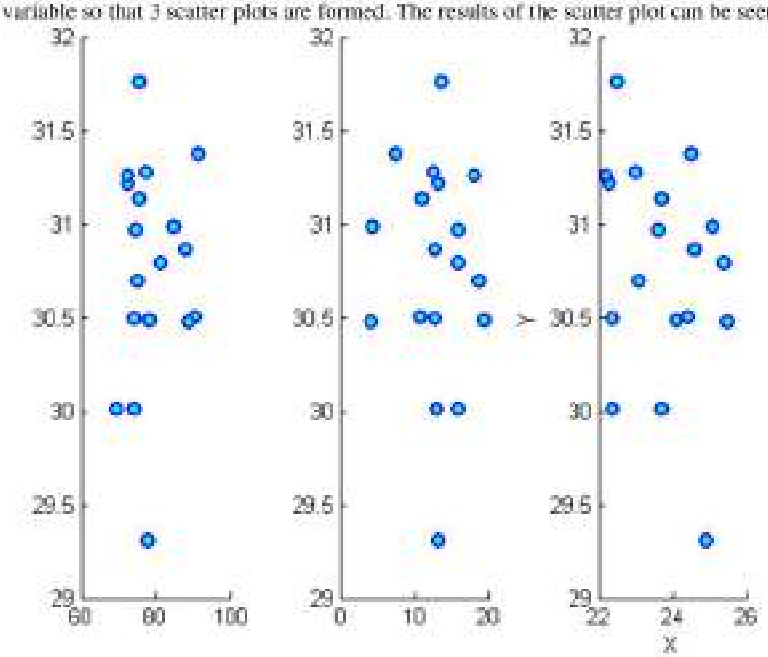

Figure 1 Plot between each predictor variable X_1, X_2, and X_3 with **Y** response variable

In Figure 1 it can be seen that from the 3 existing scatter plots, it shows that the plots formed do not have the same pattern between one scatter plot and another. So it can be said that the pattern of the relationship between each predictor variable and the response variable has a nonparametric pattern which should not be analyzed using parametric regression, but it is better to use nonparametric regression analysis.

RESEARCH RESULT

Optimal Knot Point Selection

In truncated spline regression, it is very important to determine the optimal knot point. If the optimal knot point is obtained, the best spline truncated model will be obtained. Optimal knot point selection is done by looking at the value of Generalized Cross Validation (GCV). Then the experiment was carried out until the minimum GCV value was obtained. There are several steps taken before getting the best spline model based on the minimum GCV, namely by comparing the GCV value between the linear spline model with 1 knot point to 100 knot points. When the addition of knot points results in a larger GCV, then the addition of knot points is not optimal.

Linear Spline Model with 1 Knot Point

The best model in the spline starts with the selection of a minimum GCV value so as to provide optimum knot points. Spline models with optimum knot points are the best spline models. However, the model goodness criteria by looking at the coefficient of determination values R^2, MSE and MAPE are still considered. Here is a linear spline model in general with one knot formed.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_{1.1})_+ + \hat{\beta}_3 x_2 + \hat{\beta}_4 (x_2 - k_{2.1})_+ + \hat{\beta}_5 x_3 + \hat{\beta}_6 (x_3 - k_{3.1})_+$$

With experiments in forming linear spline models with 1 knot obtained the optimal knot point from the first model obtained the optimal knot point based on the minimum GCV, with a GCV value is 0.3648021178. Knot points and estimated parameters are presented in Table 2.

Table 2 Knot Points and Estimated Linear Spline Regression Parameters with 1 Knot

i-th Knot	Variable	Parameter Estimation
1st Model		
		$\hat{\beta}_0 = 80.79465$
	x_1	$\hat{\beta}_1 = 0.2197413$
$K_1 = 81.47237$		$\hat{\beta}_2 = -0.2728400$
	x_2	$\hat{\beta}_3 = -0.39902985$
$K_2 = 12.69868$		$\hat{\beta}_4 = 0.43338245$
	x_3	$\hat{\beta}_5 = -2.901381e + 00$
$K_3 = 22.38119$		$\hat{\beta}_6 = 2.127684e + 00$
2nd Model		
i-th Knot	Variable	Parameter Estimation
		$\hat{\beta}_0 = 28.70435$
	x_1	$\hat{\beta}_1 = 0.2042626$
$K_1 = 75.12671$		$\hat{\beta}_2 = -0.1663454$
	x_2	$\hat{\beta}_3 = -0.36036855$
$K_2 = 13.86244$		$\hat{\beta}_4 = 0.1096818$
	x_3	$\hat{\beta}_5 = -0.5260388$
$K_3 = 25.42822$		$\hat{\beta}_6 = -5.278851e + 00$

Spline Linier Model with 2 Knot Point

As in the 1 knot linear spline model, the determination of the best model in a linear spline with 2 knots is obtained by conducting experiments. Here is a linear spline model with 2 knots formed.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - k_{1.2})_+ + \hat{\beta}_3 (x_1 - k_{2.2})_+ + \hat{\beta}_4 x_2 + \hat{\beta}_5 (x_2 - k_{3.2})_+ + \hat{\beta}_6 (x_2 - k_{4.2})_+ + \hat{\beta}_7 x_3 + \hat{\beta}_8 (x_3 - k_{5.2})_+ + \hat{\beta}_9 (x_3 - k_{6.2})_+$$

From the experiment, a combination of knots with a minimum GCV of 0.730764734 was selected. The knot points and estimated parameters formed are presented in Table 3.

Table 3 Knot Points and Estimated Linear Spline Regression Parameters with 1 Knot

i-th Knot	Variable	Parameter Estimation
1st Model		
		$\beta_0 = 23.92162$
	x_1	$\beta_1 = 0.2128783$
$K_1 = 81.42848$		$\beta_2 = -0.5633605$
$K_2 = 83.08286$		$\beta_3 = -0.3811240$
	x_2	$\beta_4 = 2.464998e + 00$
$K_3 = 7.585429e + 00$		$\beta_5 = 1.531224e + 00$
$K_4 = 5.395012e + 00$		$\beta_6 = -4.002728e + 00$
	x_3	$\beta_7 = -0.8209125$
$K_5 = 22.89770$		$\beta_8 = 0.1210547$
$K_6 = 23.99162$		$\beta_9 = -0.1322199$
2nd Model		
i-th Knot	Variable	Parameter Estimation
		$\beta_0 = 28.98945$
	x_1	$\beta_1 = 0.1222731$
$K_1 = 90.27161$		$\beta_2 = 8.369654e + 00$
$K_2 = 90.00538$		$\beta_3 = -9.807809e + 00$
	x_2	$\beta_4 = 1.284114e + 00$
$K_3 = 11.12028$		$\beta_5 = 4.783943e + 00$
$K_4 = 9.645453e + 00$		$\beta_6 = -6.016156e + 00$
	x_3	$\beta_7 = -0.5610341$
$K_5 = 23.47799$		$\beta_8 = 0.3163153$
$K_6 = 23.04882$		$\beta_9 = -0.3513394$

Best Truncated Spline Models

From the best model of each knot, from 1 to [illegible]0 knots is selected the best model based on minimum GCV. The GCV, R^2, MSE, and MAPE values of both models can be seen in Table 4.

Table 4 GCV, R^2, MSE, and MAPE Values of both Spline Models

Knot (K)	GCV Value	R Square Value	MSE Value	MAPE Value
1st Knot	**0.36480212**	**0.580443194**	**0.136237833**	**0.98349512**

Best Fashion Interpretation

From the results of the selection of optimum knot points obtained, the best model formed is linear spline regression with 1 knot point that is

$$\hat{y} = 80.7947 + 0.21974. X_1$$

Based on the results of the parameter test it was obtained that x_1 had a significant effect on y. The coefficient β_2 positively indicates that there is a positive relationship between x_1 to y.

If the x_1 increases by one unit, then y will also increase by 0.2197413. This means that if the humidity increases, it is likely to cause the addition of seawater salinity by 0.2197413.

Based on the above comparison results can be predicted based on the model using seawater salinity data. The predicted results are presented in Table 5 below:

Table 5 Data Testing Prediction Results

Y Data	Y Prediction	Error
30.57217	15.49176	15.08041
31.09285	15.99717	15.09569
31.74977	15.31597	16.4338
30.66296	17.27167	13.39129

Based on the best model obtained, the prediction result for 4 period as in Table 5, have a little difference from the original data, this is indicated by a relatively small error value.

CONCLUSION

The conclusions obtained based on the previous discussion are in the best model in truncated spline regression to predict the salinity of sea water are as follows:

$$\hat{y} = 80.7947 + 0.21974.X_1$$

The model is the best model with 1 knot point that produces GCV values of 0.36480212, MAPE values of 0.98349512, MSE 0.136237833, and R-Square 0.580443194. Based on the MAPE value, the prediction model is said to be very accurate forecasting.

ACKNOWLEDGMENTS

Thank you to the Ministry of Education, Culture, Research, and Technology for funding this research, so that this research can be completed smoothly.

REFERENCES

1. A. khozaimi, Y. D. Pramudita, E. M. S. Rochman dan A. Rachmad, "Decision Support System for Determining the Quality of Salt in sumenep Madura-Indonesia," Journal of Phyics: Conference Series, vol. 1477, no. 052057, pp. 1742-6596, 2020.
2. Fatmawati dan M. Kurdi. "Beyond Salt Industries and Environment in Sumenep: Effective Partnership For People Welfare," IOP Conf. Series: Earth and Environmental Science, vol. 469, no. 012085, pp. 1755-1315, 2020.
3. D. A. Widyastuti, A. A. R. Fernandes dan H. Pramoedyo, "Spline Estimation Method in Nonparametric Regression Using Truncated Spline Approach," Journal of Physics: Conference Series, vol. 1872, no. 012027, pp. 1742-6596, 2021.
4. M. Ramli, V. Ratnasari dan I. N. Budiantara, "Estimation of Matrix Variance-Covariense on Nonparametric Regresion Spline Truncated for Longitudinal Data," Journal of Physics: Conference series, vol. 1562, no. 012014, pp. 1742-6596, 2020.
5. W. Wang dan Y. Lu, "Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model," IOP Conf. Series: Materials Science and Engineering, vol. 324, no. 012049, pp. 1757-899X, 2018.
6. M. Maharani dan D. R. S. Saputro, "Generalized Cross Validation (GCV) in Smoothing Spline Nonparmetric Regression Models," Journal of Physics: Conference Series, vol. 1808, no. 012053, pp. 1742-6596, 2021.
7. A. F. Yoantika dan Susiswo, "Comparing the Principal Regression Analysis Method with Ridge Regression Analysis in Overcoming Multicollinearity on Human Development Index (HDI) Data in Regency/City ofEast Java in 2018," Journal of Physics: Conference Series, vol. 1872, no. 012024, pp. 1742-6596, 2021.
8. U. Khair, H. Fahmi, S. A. Hakim dan R. Rahim, "Forecasting Error Calculation with Mean Absolute Deviation and Mean Absolute Percentage Error," IOP Conf. Series: Journal of Physics: Conf. Series, vol. 930, no. 012002, pp. 1742-6596, 2017.
9. Prabiantissa, C. N., Basuki, A., & Sesulihatien, W. T. (2019). ''Observation of Fish Dissemination Pattern on Madura Coastal Using Segmentation of Satellite Images''. EMITTER International Journal of Engineering Technology, 7(1), 343-365.

Prediction of seawater salinity using truncated spline regression method

ORIGINALITY REPORT

18%	15%	12%	0%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	emitter.pens.ac.id Internet Source	6%
2	ojs3.unpatti.ac.id Internet Source	3%
3	journal.trunojoyo.ac.id Internet Source	2%
4	journal.walisongo.ac.id Internet Source	2%
5	journal.uim.ac.id Internet Source	2%
6	Ach. Khozaimi, Yoga Dwitya Pramudita, Eka Mala Sari Rochman, Aeri Rachmad. "Decision Support System for Determining the Quality of Salt in Sumenep Madura-Indonesia", Journal of Physics: Conference Series, 2020 Publication	2%
7	Wiwit Octavia Nia Putri, Atwal Arifin. "The Effect of Liquidity, Leverage, Institutional Ownership, and Sales Growth on Financial	2%

Distress on Property and Real Estate Companies Listed an The IDX 2016-2019",
Majalah Ilmiah Bijak, 2021
Publication

Exclude quotes	Off	Exclude matches	< 2%
Exclude bibliography	Off		