

# Piecewise Positivity Preserving Cubic Bezier Interpolation for Estimating Solar Radiation Missing Value in Penang, Malaysia

<sup>1</sup>Azizan Saaban, <sup>1</sup>Mohd Lutfi Zainudin and <sup>2</sup>Mohd Nazari Abu Bakar

<sup>1</sup>School of Science Quantitative, UUMCAS, Universiti Utara Malaysia, 06010 Sintok, Kedah Malaysia

<sup>2</sup>Faculty of Applied Science, Universiti Teknologi MARA, 02600 Arau, Perlis Malaysia

## Article history

Received: 07-09-2016

Revised: 01-12-2016

Accepted: 27-12-2016

## Corresponding Author:

Azizan Saaban

School of Science Quantitative,  
UUMCAS, Universiti Utara  
Malaysia, 06010 Sintok, Kedah  
Malaysia

Tel: +6049286304

Fax: +6049286309

Email: azizan.s@uum.edu.my

**Abstract:** Complete solar radiation data is important in variety of disciplines such as solar radiation development system or solar project which is strongly affected by hourly variation of the solar radiation. However, the available data may contain missing data for several days in the absence of measurement due to several factors such as instrument failure and human errors. In this study, the secondary data of solar radiation at Penang, Malaysia taken from Malaysian Meteorology Department was investigated. In order to fill the gap of the above missing values, the interpolation method based on positivity preserving Bezier curves to ensure that the solar radiation are always positive is proposed. This paper start with the derivation of positivity preserving condition for Bezier curves by imposing lower bound of the Bezier control points. The proposed method is tested using simulated missing value of solar radiation dataset. Result from RMSE and Coefficient of Determination ( $r^2$ ) for different set of missing value from 10, 30 and 50%, show that the proposed method produce a better result compared to existing interpolation method without implementing positivity preserving condition. Finally, the proposed method was used to fill the gap of missing values of solar radiation at Penang, Malaysia.

**Keywords:** Solar Radiation, Bézier Curves, Positivity Preserving, Missing Values

## Introduction

Most of solar radiation data in Malaysia were obtained from a pyranometer instrument. However due to the instrument failure and human errors, missing value of solar radiation reading will occurs and consequently lack of and/or incorrect data can appear (Moradi, 2009). Dataset with reliable values in the dataset is very important for other related analysis.

The proper interpolation methods of the missing data is an interesting problem that will requires a well understanding of the data (Little and Rubin, 2002) and large of focus has been given to this problem in climatology field (Dobesch *et al.*, 2007) but much less in solar radiation data. There exists several interpolation methods which have been used to interpolate solar radiation data such as Inverse Distance Weighting (IDW) (Park, 2013), Kriging (Park, 2013) and splines (Park, 2013; Cruz and Rivera, 2012; Tait and Liley, 2009;

Xia *et al.*, 2000). However all the previous works are focused on interpolation of solar radiation based on other meteorological elements such as temperature, cloudiness, precipitation and sunshine duration instead of only using solar radiation intensity element.

These methods allows the relationship between the meteorology variable and two or three dependent variables based on latitude, longitude and elevation of the measurement stations which involving higher implementation costs. To overcome the drawback and based on the limitation of solar radiation data in Malaysia, Saaban *et al.* (2014) proposed a simple piecewise interpolation method of single variable based on available solar radiation intensity (MJm-2) by using Bézier and Ball curves functions to fill the gap of hourly missing data at Alor Setar Meteorological Station based on simulated missing value of the completed data. They shows that using cubic Bézier curve for missing value estimation at Alor Setar, Kedah have given a promising

results. However the drawback of the method is that occur numbers of solar intensity value estimation are negative and which is meaningless for further analysis.

Thus this paper will propose convincing method for preserving the positive value of solar intensity by imposing sufficient positivity condition for cubic Bézier curve. The motivation of this works is obtained by the previous work of bivariate interpolation method as given in Saaban *et al.* (2009) and Brodlić *et al.* (2005). They focus on visualization and estimation of rainfall data using triangulation and point based interpolation methods respectively.

### Positivity Preserving Condition for Cubic Bézier Curve

The following proposition give the sufficient conditions for positivity-preserving of cubic Bézier polynomial curve.

#### Proposition

Consider the cubic Bézier polynomial curve:

$$r(x) = b_0(1-x)^3 + 3b_1(1-x)^2x + 3b_2(1-x)x^2 + b_3x^3, 0 \leq x \leq 1 \quad (1)$$

where,  $b_0, b_1, b_2$  and  $b_3$  are Bézier control points. If  $b_0, b_3 > 0$  and  $b_i \geq -t_0, i \neq 0, i \neq n$  with  $t_0 > 0$  is a lower bound of  $b_i$  and the real root of:

$$3t^4 + 4(b_0 + b_3)t^3 + 6b_0b_3t^2 - b_0^2b_3^2 = 0 \quad (2)$$

then  $r(x) \geq 0, \forall x \in [0, 1]$ .

#### Proof

Let all Bézier ordinates of cubic Bézier polynomial curve of  $b_1$  and  $b_2$  are equal to  $-t$  ( $t > 0$ ). Thus (1) can be written as:

$$r(x) = b_0(1-x)^3 - 3x(1-x)t + b_3x^3 \quad (3)$$

Recall that  $3x(1-x) = 1 - (1-x)^3 - x^3$ , hence (3) becomes:

$$r(x) = (b_0 + t)(1-x)^3 + (b_3 + t)x^3 - t.$$

Assume that  $t$  is fixed. Then, the first and second derivatives of  $r$  with respect to  $x$  are:

$$r'(x) = -3(1-x)^2(b_0 + t) + 3(b_3 + t)x^2 \quad (4)$$

and:

$$r''(x) = 6(1-x)(b_0 + t) + 6(b_3 + t)x \quad (5)$$

Setting  $r'(x) = 0$  gives:

$$\frac{(1-x)^2}{x^2} = \frac{(b_3 + t)}{(b_0 + t)}, x \neq 0 \text{ or } \frac{1-x}{x} = \left(\frac{b_3 + t}{b_0 + t}\right)^{\frac{1}{2}}, x \neq 0.$$

Thus, we get:

$$x = \frac{(b_0 + t)^{\frac{1}{2}}}{(b_0 + t)^{\frac{1}{2}} + (b_3 + t)^{\frac{1}{2}}} \quad (6)$$

It follows that  $b_0 > 0$  and  $b_3 > 0, 0 < x \leq 1$ , implies that  $r''(x) > 0$ , then the minimum value of  $r$  occurs when

$$x = \frac{(b_0 + t)^{\frac{1}{2}}}{(b_0 + t)^{\frac{1}{2}} + (b_3 + t)^{\frac{1}{2}}} \text{ that is:}$$

$$r_{\min}(x) = \frac{(b_0 + t)(b_3 + t)}{\left((b_0 + t)^{\frac{1}{2}} + (b_3 + t)^{\frac{1}{2}}\right)^2} - t \quad (7)$$

Observe that if  $t = 0, r_{\min} > 0$ . Thus, the lower bound of  $b_1$  and  $b_2$  which is  $t_0$  occurs when the minimum value of  $r(x)$  is equal to 0. Now, we shall described how to obtain the value of  $t_0 > 0$  to ensure that  $r(x)$  preserve the positivity for all  $0 \leq x \leq 1$ .

Setting  $r_{\min} = 0$  in (7), we have:

$$\frac{(b_0 + t)(b_3 + t)}{\left(\sqrt{b_0 + t} + \sqrt{b_3 + t}\right)^2} = t$$

or equivalent to:

$$3t^4 + 4(b_0 + b_3)t^3 + 6b_0b_3t^2 - b_0^2b_3^2 = 0,$$

which is (2):

Now, we shall proof that exist  $t_0 \in (0, b_0 + b_3)$  as a real root of (2). Let:

$$f(t) = 3t^4 + 4(b_0 + b_3)t^3 + 6b_0b_3t^2 - b_0^2b_3^2$$

then:

$$f'(t) = 12t^3 + 12(b_0 + b_3)t^2 + 12b_0b_3t.$$

Since  $b_0, b_3 > 0, f$  is an increasing function for  $t$ . Then  $f(0) = -b_0^2b_3^2 < 0$  and:

$$f(b_0 + b_3) = 7(b_0 + b_3)^4 + b_0b_3(6b_0^2 + 6b_3^2 + 11b_0b_3) > 0$$

Therefore exists  $0 < t_0 < b_0 + b_3$ , such that  $f(t_0) = 0$  where  $t_0$  is a real root of (2) in  $(0, b_0 + b_3)$  as shown in Fig. 1.

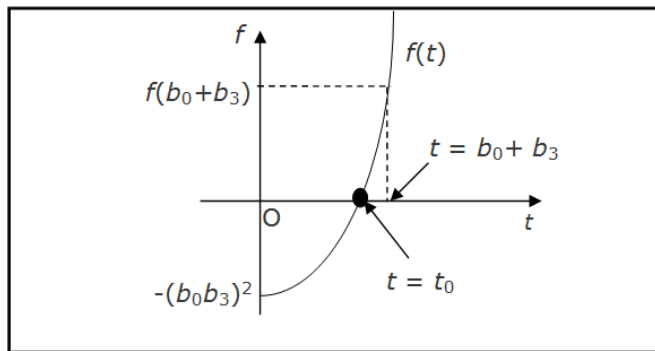


Fig. 1. The unique solution of  $f(t) = 0$  in  $(0, b_0 + b_3)$

### Solar Radiation Dataset

We tested our proposed method using an hourly solar radiation data taken from the Bayan Lepas, Penang Meteorological Station, Malaysia (Latitude  $05^{\circ} 18' 00''N$  and Longitude  $100^{\circ} 16'00'' E$ ) provided by the Malaysian Meteorology Department covering the period from 1 January 2012 to 31 December 2012 equivalent to 8784 data points. Similar to works in Saaban *et al.* (2014), the valid solar radiation intensity measurements are in between 8 am to 7 pm on each day. From the available data, we observe that the most missing values occur during June 6 to July 25, 2012 (19 days) that are need to be filled later. We also observe that the completed data of solar radiation occurs on Jan 10 to May 31, 2012 and was used to validate our proposed method.

### Model Construction and Validation

Let data points be  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  where  $x_i, y_i > 0, i = 1, \dots, 3407$  are the  $n$ -th hourly interval and complete solar radiation intensity at Penang Meteorology Station in Malaysia for the period of Jan 10, 2012 to May 31, 2012 respectively. We want to construct a piecewise interpolant  $y = F_k(x), k = 1, 2, \dots, n-1$  with first derivative continuity that interpolate the data.

To construct our proposed method, we use similar techniques as discussed in Saaban *et al.* (2014) by considering one sub-interval of data point  $A(x_1, y_1)$  and  $B(x_2, y_2)$  consists of the start and end points of hourly interval together with solar radiation intensity  $b_0$  and  $b_3$  respectively which are also the start and end control points of Bézier curve (1). The initial value of intermediate control points of Bézier curve (1),  $b_1$  and  $b_2$ , can be determined by the following relations:

$$b_1 = b_0 + \frac{1}{3}(x_2 - x_1)F'(x_1) \tag{8}$$

$$b_2 = b_3 - \frac{1}{3}(x_2 - x_1)F'(x_2). \tag{9}$$

Estimation of the derivatives at each data point are obtained by using well-known finite difference methods i.e., backward, forward and central difference methods. However, the initial estimate of the above intermediate points may not satisfy the positivity conditions for  $F$ . In view of the Proposition, we need these Bézier coefficients to be greater or equal to  $-t_0$ . If they are not, then the magnitudes of derivatives at the data points need to be reduced so that the condition is satisfied. The modification of these derivatives at points  $A$  and  $B$  is achieved by multiplying each derivative at that points, by a scaling factor  $0 < \alpha, \beta < 1$ .

The scaling factor  $\alpha$  and  $\beta$  can be calculated as follow. If  $b_1 < -t_0$ , then we write (8) as:

$$b_0 + \frac{1}{3}\alpha(x_2 - x_1)F'(x_1) = -t_0$$

and thus, the value of  $\alpha$  is:

$$\alpha = \frac{-(t_0 + b_0)}{(x_2 - x_1)F'(x_1)},$$

otherwise if  $b_1 \geq -t_0$ , we set  $\alpha = 1$ .

Similarly, if  $b_2 < -t_0$ , then we write (9) as:

$$b_2 - \frac{1}{3}\beta(x_2 - x_1)F'(x_2) = -t_0$$

and thus, the value of  $\beta$  is:

$$\beta = \frac{(t_0 + b_2)}{(x_2 - x_1)F'(x_2)},$$

otherwise if  $b_2 \geq -t_0$ , we set  $\beta = 1$  (Saaban *et al.*, 2006) for further details of the scheme.

In order to validate the accuracy of the proposed methods, we simulate three different types of missing values consisting of 10, 30 and 50% missing based on training dataset as stated earlier. The result based on statistical indicator RMSE and  $r^2$  coefficients of all missing value patterns for the new proposed method with positivity preserving and without positivity preserving conditions is shown in Table 1. The result show that based on the RMSE and  $r^2$  coefficients, the proposed method with positivity preserving perform more better compared to the previous method without implementing positivity conditions.

Based on the above validation, the new proposed cubic Bézier interpolation method is used for estimating solar radiation intensity at Penang, Malaysia as discuss in the following section.

### Implementation for Overall Estimation of Solar Radiation Data in Penang

We implement our proposed method in estimation of solar radiation as the following steps. First we fill the

gap of missing values by, firstly interpolated the missing value at time 7 AM, 1 PM and 7 PM each day of the year. Then, by using this interpolated values, we interpolate solar radiation at hourly missing data for each day to produce a complete hourly solar data set of the year in Penang. The missing value occurs in the period of Jan 1 to Jan 9, 2012 and Jun 1 to December 31, 2012 will be estimated and the completed solar radiation data set in Penang was obtained. Figure 2 shows an example of missing value occurs at 12 PM to 5 PM on Jan 10, 2012 and Fig. 3 shows the estimated value of solar radiation using our proposed method. Uncompleted solar radiation set for the year of 2012 at Penang, Malaysia is illustrated in Fig. 4 while Fig. 5 shows the completed data set after implemented our proposed method and may be can be used for further analysis with regard to the application of the solar radiation in related areas. Can be seen that because of the long period of missing value of solar radiation at Penang during the given date, the estimation values pattern is lower compare to the pattern of true solar radiation measurement throughout of the year.

Table 1. RMSE and  $r^2$  statistical indicator based on missing value percentages for all methods

Percentage of missing value	Error indicator	Method	
		Cubic Bézier (without positivity preserving)	Cubic Bézier (with positivity preserving)
10%	RMSE	0.100120151008629	0.0996666589140811
	$r^2$	0.920750595469546	0.921109553961819
30%	RMSE	0.139331522163505	0.138739874097182
	$r^2$	0.88257449292969	0.883073120757165
50%	RMSE	0.226000234875972	0.197408881098049
	$r^2$	0.819328681622124	0.842185461324591

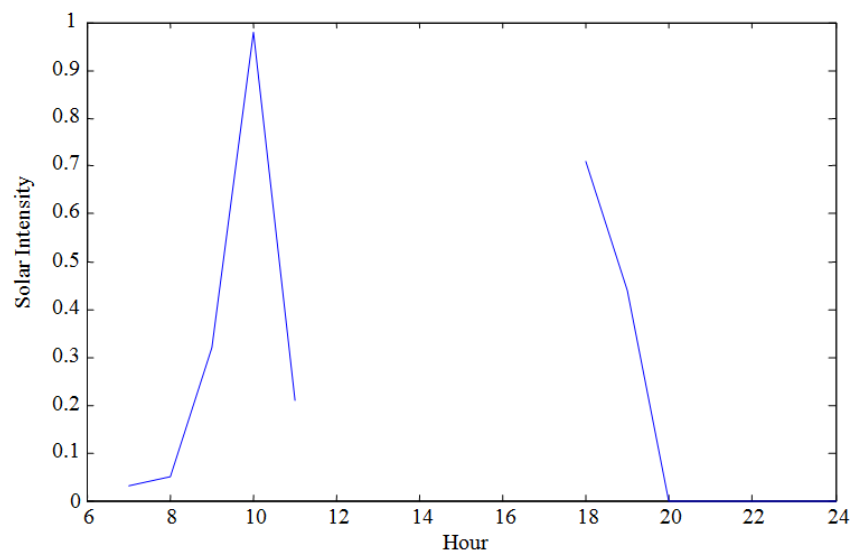


Fig. 2. Missing value on Jan, 10 2012 h 12 to 17 of the day

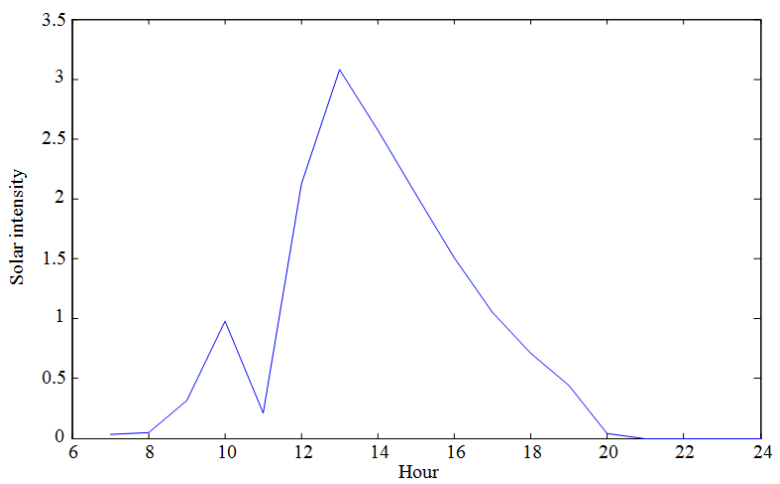


Fig. 3. Estimate value on Jan, 10 2012 h 12 to 17 of the day

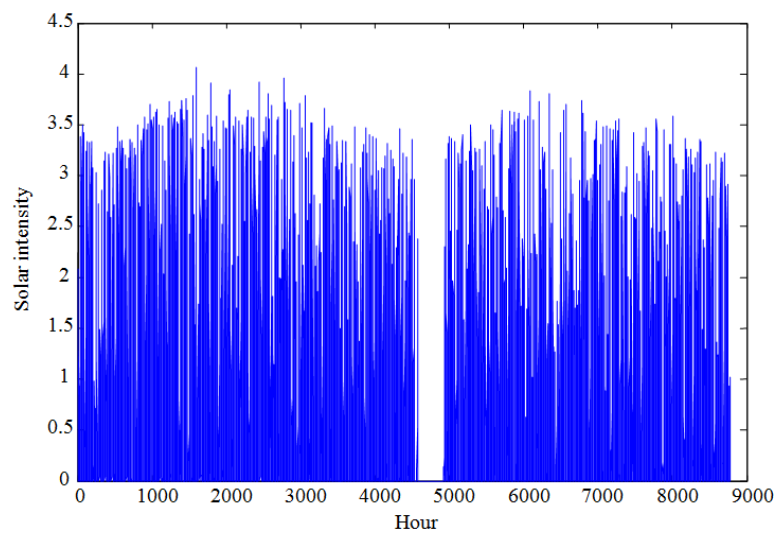


Fig. 4. Solar radiation pattern with missing value on 2012 at Penang, Malaysia

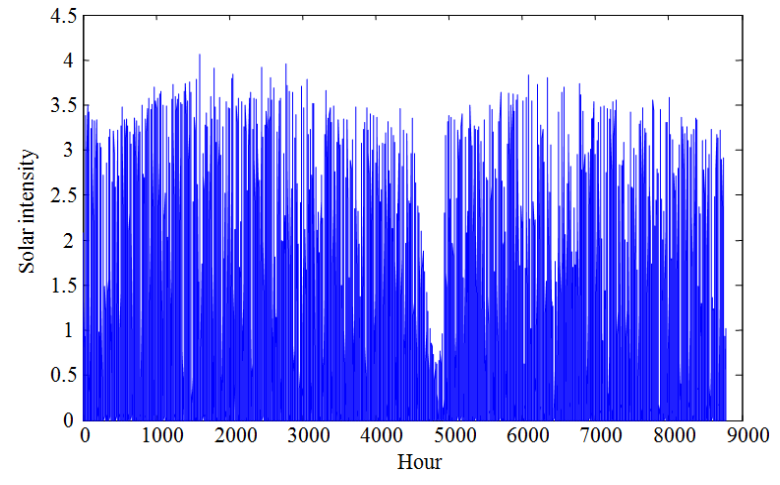


Fig. 5. Completed estimation of solar radiation pattern on 2012 at Penang, Malaysia using the proposed method

## Conclusion

We have shown how a cubic Bézier curve can be constrained in order to preserve the positivity of solar radiation data points. This is achieved by imposing a lower bound on two inner points of cubic Bézier ordinates at hourly interval of data points. Our proposed method has been applied in training completed solar radiation dataset from Jan, 10 to May, 31 2012 at Penang obtained from the Malaysian Meteorological Department. We have also shown that the proposed method is outperform the previous method. We successfully implemented our proposed method to fill the gap of solar radiation missing value on 2012 at Penang, Malaysia.

## Acknowledgment

This research was financially supported by grants Exploratory Research Grant Scheme (ERGS) Specification Code Output (S/O) Numbers 12824 from Institutions of Higher Education, Department of Higher Education. Authors on this paper would like to express a very thankful to University Utara Malaysia (UUM) for their emphasized and pursued them about this potential project in Malaysia, particularly, in development energy field.

## Author's Contributions

**Azizan Saaban:** Participated in all experiments, coordinated the data-analysis and contributed to the writing of the manuscript and proofreading.

**Mohd Lutfi Zainudin:** Participated in all experiments, coordinated the data-analysis and contributed to the writing of the manuscript.

**Mohd Nazari Abu Bakar:** coordinated the data-analysis and contributed to the writing of the manuscript and proofreading.

## Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

## References

Brodlie, K.W., M.R. Asim and K. Unsworth, 2005. Constrained visualization using the shepard interpolation family. *Comput. Graph. Forum*, 24: 809-820. DOI: 10.1111/j.1467-8659.2005.00903.x

Cruz, L.I.F. and E.I.O. Rivera, 2012. Biharmonic spline interpolation for solar radiation mapping using puerto rico as a case of study. *Proceedings of the 38th IEEE Photovoltaic Specialists Conference*, Jun. 3-8, IEEE Xplore Press, pp: 2913-2915.  
DOI: 10.1109/PVSC.2012.6318196

Dobesch, H., P. Dumolard and I. Dyras, 2007. *Spatial Interpolation for Climate Data: The Use of GIS in Climatology and Meteorology*. 1st Edn., Wiley, London, ISBN-10: 1905209703, pp: 284.

Little, R.J.A. and D.B. Rubin, 2002. *Statistical Analysis with Missing Data*. 2nd Edn., Wiley, New York, ISBN-10: 0471183865, pp: 408.

Moradi, I., 2009. Quality control of global solar radiation using sunshine duration hours. *Energy*, 34: 1-6.  
DOI: 10.1016/j.energy.2008.09.006

Park, J.K., 2013. Comparison of spatial interpolation methods for estimating solar radiation in South Korea. *Global J. Technol.*

Saaban, A., A.R.M. Piah and A.A. Majid, 2006. Positivity-preserving scattered data interpolating surface using  $C^1$  piecewise cubic triangular patches. *Proceedings of the International Conference on Computer Graphics, Imaging and Visualisation*, Jul. 26-28, IEEE Xplore Press, Los Alamitos, California, pp: 490-495.  
DOI: 10.1109/CGIV.2006.71

Saaban, A., A.R.M. Piah and A.A. Majid, 2009. Visualization of rainfall data distribution using quantic Bézier triangular patches. *Bull. Malays. Math. Sci. Soc.*, 32: 137-150.

Saaban, A., M.L. Zainudin and M.N.A. Bakar, 2014. On piecewise interpolation techniques for estimating solar radiation missing values in Kedah. *AIP Conf. Proc.*, 1635: 217-221. DOI: 10.1063/1.4903586

Tait, A. and B. Liley, 2009. Interpolation of daily solar radiation for New Zealand using a satellite data derived cloud cover surface. *Weather Climate*, 29: 70-88.

Xia, Y., W. Winterhalter and P. Fabian, 2000. Interpolation of daily global solar radiation with thin plate smoothing splines. *Theoretical Applied Climatol.*, 66: 109-115.  
DOI: 10.1007/s007040070036