

Irregular GIS Curve Fitting based High Speed Railway Earthquake Influence Range Calculation Model

Zhaobing Hu¹, Zhiming Liu¹ and Lianbao Yang²

¹China Academy of Railway Sciences, Signal & Communication Research Institute, 100081 Beijing, China

²China Academy of Railway Sciences, Railway Big Data Research and Application Innovation Center, 100081 Beijing, China

Abstract. In this paper, to guarantee that the train can take measures to reduce the damage caused by the earthquake, it propose an irregular GIS curve fitting based high-speed railway earthquake influence range calculation model. Firstly, this model eliminates the abnormal points, calculates feature points and finds demarcation points of the high-speed railway GIS curve to get the processed point collection in Mercator coordinate. Secondly, though using the processed point collection, this model applies least square polynomial segmentation fitting method to implement complex high-speed GIS curve fitting. Thirdly, calculate the earthquake influence rang on high-seed railway line, according to the scope of the earthquake equation and the high-speed railway GIS curve fitted equation. Finally, the paper selects the Beijing South to Dezhou East high-speed railway section which is part of Beijing-Shanghai line as a case study, which proves that the model can calculate the earthquake influence scope on the railway line offering decision support for train operation to ensure safety.

1 Introduction

By the end of 2016, China's total mileage of high-speed railway was more than 20,000 kilometers [1, 2], accounted for over 60% of the world's total mileage of high speed railway, which ranked first in the world. China is located in the earthquake area which is vulnerable to earthquake disasters [3]. When the earthquake occurs, within its influence scope, high-speed railway line will vibrate and distort as the ground, leading to the trains jumping up and down, which will cause a devastating disaster [4, 5].

At present, the discretization of high-speed railway lines' GIS curve mainly through the direct method and the indirect method for the data acquisition [6]. Both ways use a finite number of coordinate points to characterize the railway line but without a mathematical equation to represent. The main process of curve discretization of high-speed railway lines is easy to produce abnormal points and lose feature points, which are caused by the system error or random selection of high-speed railway GIS curve points, leading to the coordinate points deviate from the correct position gravely. How to eliminate abnormal points, calculate feature points and on this basis to achieve the high-speed GIS curve fitting is an important problem urgent to tackle [7, 8].

The high-speed railway lines are mainly composed of line segment and circular arc. In view of commonly used methods of irregular curve fitting [9, 10] and combined with the characteristics of the high-speed railway lines,

this paper puts forward an irregular GIS curve fitting based high-speed railway earthquake influence range calculation model. The model mainly includes two parts. The first part is to use the least square polynomial fitting method [11] to get the mathematical model of GIS curve including eliminate abnormal points, calculate feature points, find demarcation points. The second part is based on the mathematical model in combination with influence scope of earthquake equation to get the earthquake influence range of the high-speed railway line. As shown in the Fig 1.

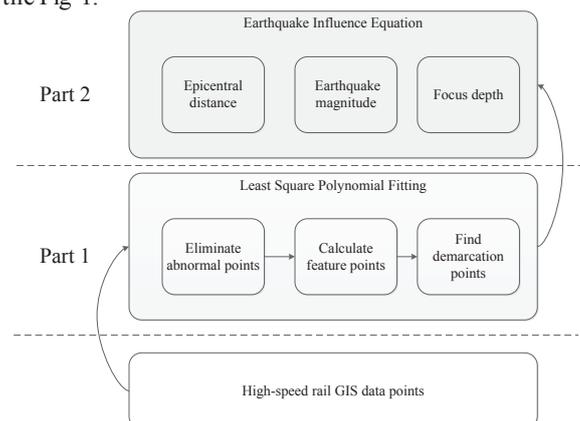


Figure 1. Irregular high-speed GIS curve fitting based on radius of high speed railway earthquake calculation mode

Finally, this paper uses Beijing South to Dezhou East high speed railway section which is part of the Beijing-Shanghai high-speed railway line as a case to implement the proposed model. The R-square and Adjusted R-square

are all above 0.99, shows the excellent goodness of fit by the least square polynomial fitting method proposed above. Moreover though simulate the earthquake happens and the curve fitted, this paper validates the correctness of the high-speed railway earthquake influence range calculation model.

2 Irregular high-speed lines GIS curve fitting

Irregular GIS curve fitting of high-speed railway line is to select the appropriate type of fitting method to the coordinate points. Curve fitting model consists of linear and nonlinear fitting and the polynomial fitting belongs to the linear fitting. Exponential function and trigonometric function belong to nonlinear fitting. In base of weather the fitting method can generate parameter equation or not, curve fitting, can also be divided into the parameter fitting and nonparametric fitting. The parameter fitting is mainly based on the least square fitting. Nonparametric fitting mainly use interpolation method such as cubic spline interpolation which can pass all data points and smoothness is better, but it is difficult to get the parameter equation. High-speed railway line is mainly made up of line segments and the circular arc with the large radius. Meanwhile the polynomial fitting can adapt to the characteristics of high-speed railway line. Therefore, in order to get parameters equation model, and guarantee the good fitting effect, this article adopts the method of least square polynomial fitting.

For the discretization of high-speed railway lines GIS curve, the discretization coordinate points may have abnormal points and redundancy points, which serious influence the goodness of curve fitting. Thus the first step of curve fitting is to eliminate abnormal points and reduce the redundancy points. To the complex high-speed lines curve also need to extract the boundary point. This chapter will introduce each part of the process in detail.

2.1 Eliminate abnormal points

Abnormal points refer to the points of GIS curve deviates from the position gravely. The abnormal point elimination is according to the distance ratio as shown in Fig 2.

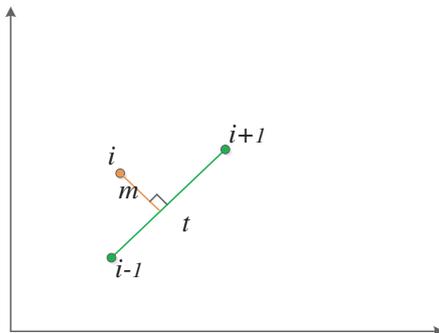


Figure 2. The distance ratio to eliminate abnormal points

Among them, m represents the point i to point $i-1$ and point $i+1$ straight line distance, t represents the distance between point $i-1$ and point $i+1$. The method to calculate the distance is shown in (1)-(3).

$$m = \sqrt{\frac{((x_i - x_{i-1})(y_{i+1} - y_{i-1}) - (y_i - y_{i-1})(x_{i+1} - x_{i-1}))^2}{(x_{i+1} - x_{i-1})^2 + (y_{i+1} - y_{i-1})^2}} \quad (1)$$

$$t = \sqrt{(x_{i+1} - x_{i-1})^2 + (y_{i+1} - y_{i-1})^2} \quad (2)$$

$$l = \frac{m}{t} \quad (3)$$

Generally l value of GIS curve is small, but the l value of abnormal point will increase significantly.

2.2 Calculate feature points

Feature points are generally recognized as the major points of the curve, which can describe the curve effectively. The method of selecting feature points mainly consists of angle test method and a polygon approximation method. This paper adopts an improved feature extraction method of vector curves, which is more simple and accurate than Douglas-Peucker algorithm [12].

Suppose the points of the high-speed railway GIS line are $A_0(x_0, y_0), A_1(x_1, y_1), \dots, A_m(x_m, y_m)$. The slope of $A_{i-1}A_i$ is as follows.

$$k_{i-1,i} = (y_i - y_{i-1}) / (x_i - x_{i-1}) \quad (4)$$

The slope of A_iA_{i+1} is as follows.

$$k_{i,i+1} = (y_{i+1} - y_i) / (x_{i+1} - x_i) \quad (5)$$

Set the threshold ϵ be a small value, if the

$$|k_{i,i+1} - k_{i-1,i}| \leq \epsilon \quad (6)$$

Which means that the point i and point $i+1$ slope is very similar, the point A_i is abandoned then. Otherwise the point A_i is selected as the feature point.

The selection of feature points is mainly evaluated by the data compression and fidelity, which is measured by compression ratio and integral square error commonly. The compression ratio is calculated as (7).

$$CR = \frac{n_1}{n_2} \quad (7)$$

n_1 is the number of original data points and n_2 is the resulting feature points. ISE is the integral error caused by the approximation curve, calculated by (8).

$$ISE = \sum_{i=2}^{n_2-1} m^2 \quad (8)$$

2.3 Find demarcation points

Demarcation points represent the key position points of the curve, such as turning point. The demarcation points are mainly to select the largest slope point among the feature points.

2.4 The least square polynomial fitting

The least square polynomial fitting is mainly according to the given set of data points in the space of polynomial:

$$\phi = span\{1, x, x^2, \dots, x^n\} \quad (9)$$

To confirm $c_0, c_1, c_2, \dots, c_n$, get the math model:

$$\varphi(x) = \sum_{k=0}^n c_k x^k, \forall c_k \in R \quad (10)$$

Guarantee it the minimum square error:

$$\sum_{i=0}^m (y_i - \varphi(x_i))^2 = \min \quad (11)$$

Let

$$Y = (y_0, y_1, \dots, y_m) \quad (12)$$

$$C = (c_0, c_1, \dots, c_{n+1}) \quad (13)$$

And

$$A = \begin{bmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_m & \dots & x_m^n \end{bmatrix}_{(m+1) \times (n+1)} \quad (14)$$

So (11) can transfer to:

$$\|Y - AC\|_2^2 = \min \quad (15)$$

After get the (15) partial derivatives and let the gradient extreme value to be 0 then it changes to:

$$A^T AC = A^T Y \quad (16)$$

Then by matrix operation:

$$C = (A^T A)^{-1} A^T Y \quad (17)$$

Above all, the equation of high-speed railway line under the least square polynomial method is obtained.

3 High speed railway earthquake influence range calculation model

At the time of the earthquake happens, it produces *P* wave and the *S* wave, which is also known as the

longitudinal wave and transverse wave. The *P* wave's vibration direction is parallel to the direction of earthquake and its transmission speed is very fast. However the *S* wave's vibration direction is vertical to the direction of earthquake and its transmission speed is very slow. *P* wave is more destructive than the *S* wave. High-speed railway earthquake monitoring and early warning system is based on the *P* wave and *S* wave velocity difference to realize earthquake early warning. The principle of high-speed railway earthquake monitoring and early warning is shown in Fig 3.

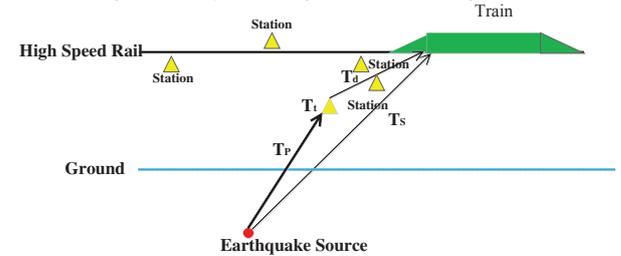


Figure 3. Principle of high-speed railway earthquake monitoring and early warning

T_s, T_p, T_d represents *S* wave, *P* wave and electromagnetic wave transmission time respectively, T_t refers to the stations identify seismic waves and start to send electromagnetic waves time. In order to make the earthquake early warning be efficient, the data processing time and the sum of T_s, T_p, T_t must be as small as possible. When T_s is greater than the sum of T_s, T_p, T_t the early warning and effective and the greater the better.

When the earthquake happens, it can be represented an equation of circle, with epicenter as the center and the impact scope as radius. In combine with the fitted curve function, the influence scope of earthquake on high-speed railway line can be transformed into the problem of solving the equation as shown in (18).

$$\begin{cases} \varphi(x) = \sum_{k=0}^n c_k x^k \\ E(x) = (x - x_e)^2 + (y - y_e)^2 = R^2 \\ x_{\min} \leq x \leq x_{\max} \\ y_{\min} \leq y \leq y_{\max} \end{cases} \quad (18)$$

$\varphi(x)$ is the fitted high-speed railway GIS, line equation $E(x)$ is the earthquake influence range equation. R is the epicentral distance, which can be calculated according to the parameters such as magnitude and focal depth etc. (x_{\min}, y_{\min}) is the minimum coordinate value of high-speed railway GIS line and (x_{\max}, y_{\max}) is the maximum. When calculate the intersection points, then the earthquake specific influence range on the line should be accurately calculated.

4 Case study

To validate the proposed model, this paper selects Beijing South to Dezhou East high-speed railway section which is part of the Beijing-Shanghai high-speed railway line as a test case. At the same time, the railway line is through the earthquake zone as shown in Fig 4.



Figure 4. Principle of high-speed railway earthquake monitoring and early warning

After converting it to the ArcGIS Mercator plane coordinate, there has 5788 sampling points and the gis curve line in is shown as the Fig 5.

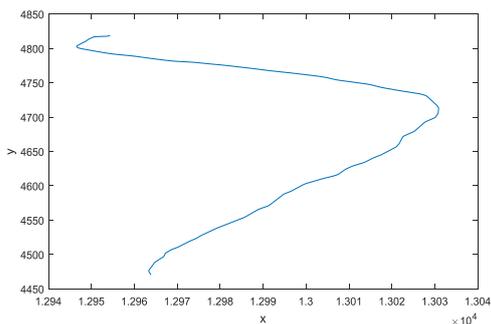


Figure 5. Beijing South-Dezhou East high-speed railway gis curve in mercator coordinate

4.1 Eliminate abnormal points

According to the 2.1's method, calculate the l value of each point and get the histogram figure as shown in Fig 6.

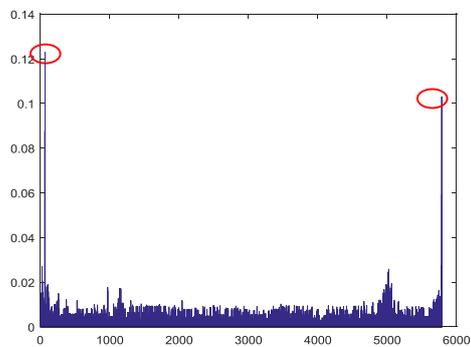


Figure 6. The l value of Beijing South-Dezhou East high-speed railway gis curve

There are two points' l value are bigger than others in Fig 6, which are abnormal points that need to eliminate. The Mercator coordinate of the abnormal points are (12952.26988,12952.26988) and (12963.85045, 12963.85045).

4.2 Calculate feature points

According to the 2.2's method, calculate the absolute slope value of each point as shown in the Fig 7.

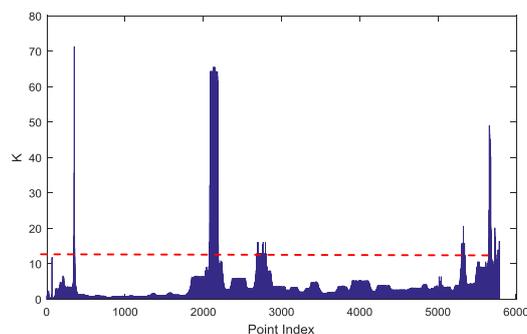


Figure 7. The absolute slope value of Beijing South-Dezhou East high-speed railway gis curve

The absolute slope values of the coordinate points are mainly lower than 10, so choose the threshold $l > 10$ and extract a total of 391 feature points. Moreover, the feature points can represent the original 5676 points with less distortion as shown in Fig 8.

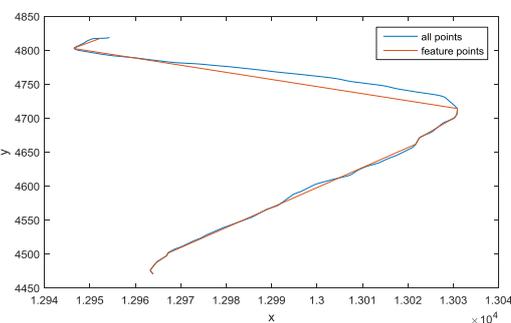


Figure 8. The comparison of Beijing South-Dezhou East high-speed railway gis curve by feature points and all points

4.3 Find demarcation points

According to the 2.3's method, choose the threshold $l > 60$ and extract the demarcation points as shown in Fig 9.

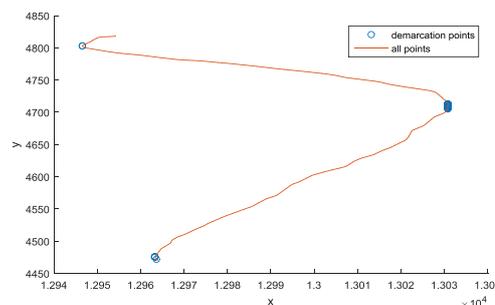


Figure 9. The demarcation points of Beijing South-Dezhou East high-speed railway gis curve

4.4 The least square polynomial fitting

For the processed Mercator coordinate points set of Beijing South-Dezhou East high-speed railway GIS curve, adopt the least square polynomial fitting, which cannot obtain good fitting effect. The fitted figure is shown in Fig 10 (a) and the fitted polynomial equation is (19).

$$\varphi(x) = 44.59x^7 + 4.409x^6 - 195.5x^5 + 42.72x^4 + 201.2x^3 - 67.7x^2 + 18.92x + 4638 \quad (19)$$

By analyzing the parameters of fitting effect, such as the SSE, R square, Adjusted R square and RMSE, which are shown in Table 1. The R square and Adjusted R square both are about 0.46, which are far away from 1. So it's necessary to use the demarcation points calculated above to do piecewise fitting. The fitted figure result is shown in Fig 10 (b-d) and the fitted polynomial equation is (20). The R-square is nearly 1 and SSE is very small, which are shown in Table 1.

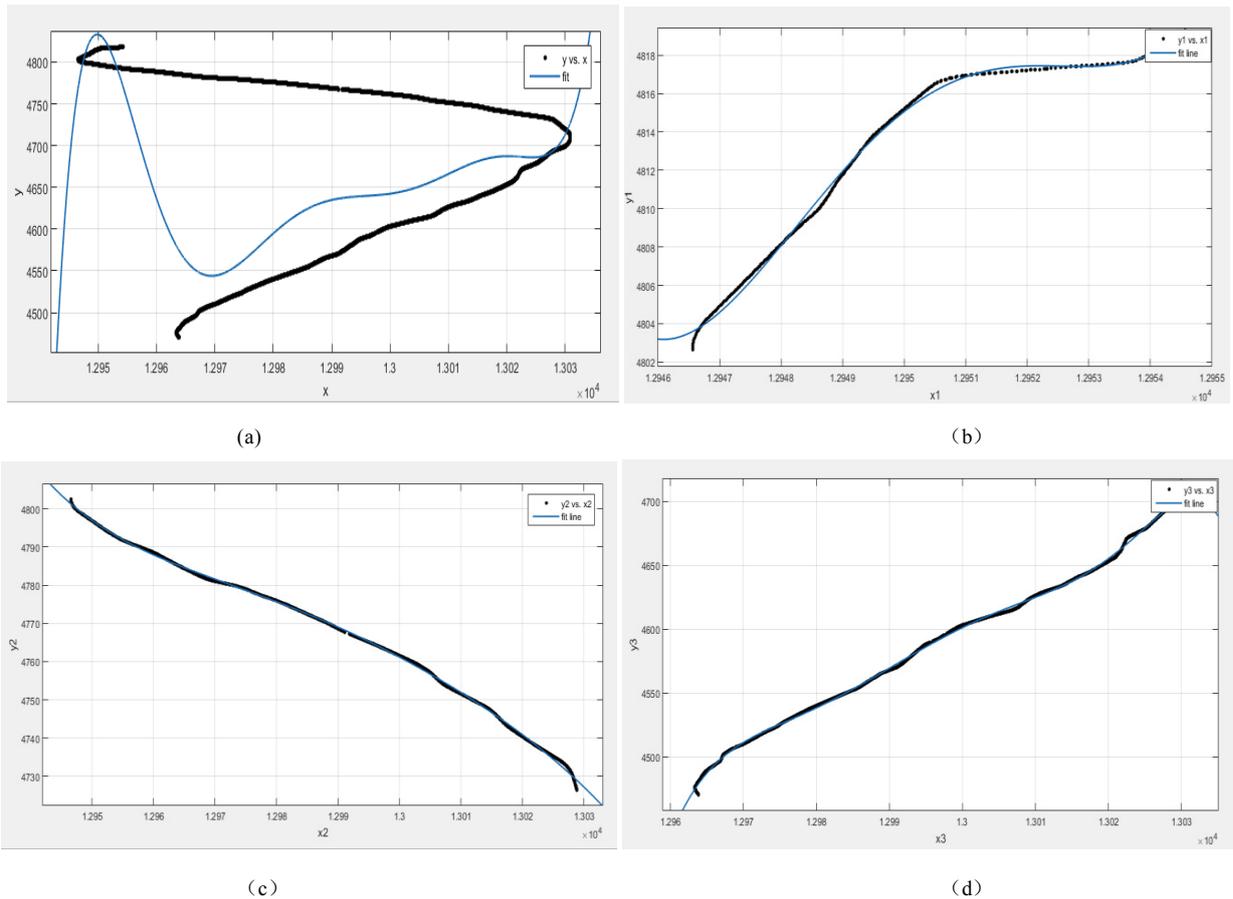


Figure 10. Beijing South-Dezhou East high-speed railway gis curve fitting results

Table 1. Goodness of fit by direct fitting and piecewise fitting

Fit Method	SSE	R-square	Adjusted R-square	RMSE
Direct fitting	34510000	0.4634	0.4627	77.28
Piecewise fitting	21.38	0.9973	0.9973	0.2486
	406	0.9994	0.9994	0.513
	10380	0.9994	0.9994	1.7

$$\varphi(x) = \begin{cases} 0.579x^4 - 0.6665x^3 - 3.087x^2 + 6.151x + 4814, & 12946.5614196 \leq x \leq 12954.3304068 \\ -0.3694x^5 + 0.4256x^4 - 0.6263x^3 - 4.045x^2 - 18.09x + 4769, & 12946.5614196 \leq x \leq 13028.842108 \\ -0.853x^8 - 1.867x^7 - 2.29x^6 + 15.41x^5 + 15.15x^4 - 22.67x^3 - 16.14x^2 + 70.55x + 4586, & 12963.8816191999 \leq x \leq 13028.842108 \end{cases} \quad (20)$$

4.5 Earthquake influence range calculation

This paper simulates a earthquake that happens with the magnitude of 5.9 and the epicentral distance of 63 KM.

By using the proposed model, the influence range of the earthquake on the high-speed railway line was obtained and as shown in Fig 11.

5 Conclusion

The model proposed by this paper still need to be improved to reduce the earthquake damage to the high-speed railway. The future work of this paper is to simplify the equation of the high-speed railway line and improve the efficiency of the calculation speed to quickly get the influence scope of earthquake on the high-speed railway line.

Also, trying to use the big data and cloud computing technology to improve the efficiency of earthquake early warning computing is also urgent to study.



Figure 11 Influence range of the earthquake on the Beijing-Shanghai East high-speed railway gis curve

Acknowledgement

The authors would like to thank the support of National Natural Science Foundation of China under Grant no. U1534202.

References

1. K. Takagi, "Development of high-speed railways in China". Japan Railway & Transport Review, (2011)
2. A. Ryder, "High speed rail", Journal of Transport Geography, **22**, pp. 303-305,(2012)
3. L. Huang, W.L. Bao, K. Chen, "The comparison analysis of Chinese public perception of earthquakes on different time scales", Nat. Hazards, **73** (2),pp. 613–62,(2014)
4. Z. Jin, S. Pei, X. Li, H. Liu, S. Qiang, "Effect of vertical ground motion on earthquake-induced derailment of railway vehicles over simply-supported bridges", Journal of Sound and Vibration, vol. **383**, pp. 277-294,(2016)
5. J. Koseki, M. Koda, S. Matsuo, H. Takasaki, T. Fujiwara, "Damage to railway earth structures and foundations caused by the 2011 off the Pacific Coast of Tohoku Earthquake", Soils and Foundations, vol. **52**, pp. 872-889 ,(2012)
6. P. Bo, G. Luo, K. Wang, "A graph-based method for fitting planar B-spline curves with intersections", Journal of Computational Design and Engineering, vol. **3**(1), pp. 14-23,(2016)
7. J. Wang, L. Feng, "Curve-fitting models for fossil fuel production forecasting: Key influence factors, Journal of Natural Gas Science and Engineering", vol. **32**, pp. 138-149,(2016)
8. Diogo E.V. Andrade, Rubens R. Fernandes, Tainan G.M. Santos, Emanuel Vitor Ceccon, Ana C.B. da Cruz, Admilson T. Franco, Cezar O.R. Negrão, "Curve-fitting equation for prediction of the start-up stress overshoot of an oil-based drilling fluid", Journal of Petroleum Science and Engineering, vol. **146**, pp. 902-908, (2016)
9. S. Biswas, D. Ghoshal, R. Hazra, "A new algorithm of image segmentation using curve fitting based higher order polynomial smoothing", Optik - International Journal for Light and Electron Optics, vol. **127**, pp. 8916-8925, (2016)
10. M. Marji, P. Siy, "A new algorithm for dominant points detection and polygonization of digital curves", Pattern Recognition, vol. **36**, pp. 2239-2251, (2003)
11. A. Björck, "Numerical Methods for Least Squares Problems", SIAM, Philadelphia, (1996)
12. K. Ebisch, "A correction to the Douglas–Peucker line generalization algorithm", Computers & Geosciences, vol. **28**, pp. 995-997, (2002)