

PAPER • OPEN ACCESS

Application of Improved Multiple Linear Regression Model in the Tracking of Strip Welding Seam

To cite this article: Yuanbo Shi *et al* 2021 *J. Phys.: Conf. Ser.* **2026** 012025

View the [article online](#) for updates and enhancements.

You may also like

- [The Influence of Control Parameters on Precision of Welding Seam Tracking in Manually Control Master-slave Robot Remote Welding System](#)
Hongtang Chen, Haijun Ma, Hongli Jiang et al.
- [The effect of welding speed on mechanical and microstructural properties of 5754 Al \(AlMg₂\) alloy joined by laser welding](#)
Bekir Çevik and Behçet Gülenç
- [Research on Detection Method for Welding Seam Defects in Ultrasonic TOFD Image Based on Mask R-CNN](#)
Yimin Wang, Yongbei Cao, Birun Ye et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Application of Improved Multiple Linear Regression Model in the Tracking of Strip Welding Seam

Yuanbo Shi¹, Chunpeng Lv¹ and Jianhui Wang²

¹ School of Computer and Communication Engineering, Liaoning Petrochemical University, Fushun Liaoning 113001, China

² School of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Email: syb2011@yeah.net

Abstract. In the process of hot dip galvanization. An improved multiple linear regression model is established to predict the welding seam position of strip steel when two continuous strip steels pass through the air knife. Firstly, a multiple linear regression model is established according to the production parameters of hot dip galvanizing process. Secondly, the least square method is used to improve the multiple linear regression model. Finally, in order to verify the effectiveness of the method, the model is simulated by SPSS software, and the experimental result show that the method achieves the purpose of quantification.

Keywords. Welding seam; steel strip; air knife; multiple linear regression; least squares.

1. Introduction

1.1. Hot Dip Galvanizing Process

Strip steel is one of the main products in our life. It can be used as a raw material for many industries, for example, automobile, home appliances, shipbuilding and aeronautics industry [1]. But the strip steel often needs to be coated with a layer of zinc, for preventing the plate surface corrosion.

At present, the most of the world's steel enterprises have to adopt a continuous hot dip galvanizing process for surface of galvanized strips [2].

1.1.1. Galvanizing Technology. There are a lot of galvanizing methods. For example, Electro galvanized which belong to cold galvanizing method. But now, the hot-dip galvanizing method is used widely as a more mature method. Hot dip galvanizing is a kind of galvanizing method using metal corrosion, which is mainly used in the galvanizing production of industrial steel. This technology immerses the steel strip into the liquid zinc at 500°C and the zinc layer is attached to the surface of the steel strip to achieve the purpose of anti-corrosion.

1.1.2. Continuous Annealing Process Technology. Continuous annealing process technology is the most important step of hot-dip galvanizing, the strip steel enters into annealing furnace for heating after pretreatment [3]. In the annealing furnace, the strip steel passes preheating zone, heating zone and soaking zone. The preheating zone is used to heat the strip steel step by step, which prevent the strip steel deformation caused by the rapid rise of strip steel temperature; The heating zone is to heat the strip steel to the specified temperature at high temperature; The soaking zone is the buffer zone between the heating zone and the zinc pot, which makes the strip steel have better shape and more uniform



Content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

temperature before entering the zinc pot, and improves the quality of zinc plating. After heating, the strip steel enters the zinc pot through the furnace nose. The technological process is shown as figure 1 [4].

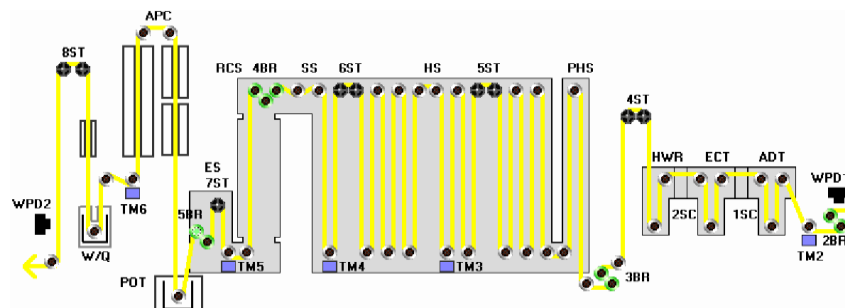


Figure 1. The technological process of continuous annealing.

1.2. Effect of Welding Seam on the Air Knife

In the process of galvanizing, the position of the air knife is located on the export of the zinc pot. The main function is to scrape off the liquid zinc on the surface of strip steel, which ensure the uniform and smooth surface of strip steel, and control the thickness of zinc layer [5, 6].

1.2.1. The Work Principle of Air Knife. The air knife body is composed of a pair of cylinder blocks, which are respectively located on both sides of the strip steel, the structure is shown as figure 2. The reason why the air knife can scrape off the liquid zinc on the surface of the strip steel and control the thickness of the zinc layer accurately is that the air knife passes the gas through the gap of the cylinder block to form a knife shaped airflow, which makes the liquid zinc smooth.

1.2.2. Effect of Weld Passing Through Air Knife. A pair of movable baffles are designed to fix the position of the strip steel in order to prevent the strip steel from skewing and bumping when the air knife is working. When the welding seam passes through the baffle, it is easy to scrape off the baffle because the weld is crescent shaped. If the baffle falls into the zinc pot, the loss is great. The process is shown as figure 3.

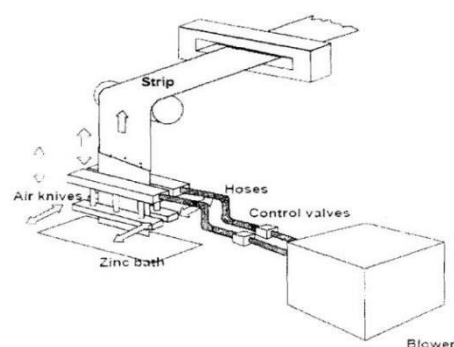


Figure 2. The structure of the air knife.

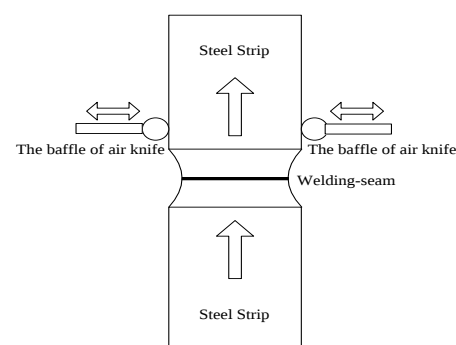


Figure 3. The process of the welding seam passes through the baffle of air knife.

2. The Traditional Welding Seam Detection

Therefore, the welding seam tracking process is very important for hot-dip galvanizing. But, because of the high temperature in the vicinity of the zinc pot, it cannot install detection equipment, how to accurately detect welding seam reach the air knife, it is unable to realize. So, the traditional welding seam detection adopt early prediction or roller speed calculation method to control the air knife.

2.1. Pre Prediction Method

2.1.1. The Method of Pre Predicted. The pre prediction method does not consider the elongation of the strip after heating in the annealing furnace, calculates the position of the welding seam according to the original length of the strip steel, opens the baffle in advance, and closes the air knife when the weld completely passes through the air knife.

2.1.2. The Disadvantage of Pre Predicted. Although air knife baffle can't be scratched by welding seams for prediction in advance, the tailgate opening will cause the strip steel shaking, so the effect of air knife scraping the zinc layer is not good. Therefore, by using the method of pre prediction, the yield of high-quality strip is low and the loss of strip steel is large. In general, for iron and steel enterprises, air knife with welding seams should be opened in ± 5 meters, so that the rate of strip steel waste is less.

2.2. Calculation with Roller Speed

2.2.1. The Method of Calculation with Roller Speed. Due to the waste of strip steel uncoiling in advance, so far, most enterprises choose to use roll speed to calculate. References [7, 8] used the method is that relies on moving forward of many rollers rotating in the process of moving strip steel. At the entrance of the annealing furnace, there is a velocity measuring roller 3BR. At the exit of the furnace, there is a velocity measuring roller 5BR. So according to roller speed differential of two points, the elongation of strip steel can be calculated.

2.2.2. The Disadvantage of Calculation with Roller Speed Differential. The disadvantage of calculation with roller speed differential is relatively simple [9, 10]. Because of the roller speed is affected by working speed, the roller slip often occurs between strip steel and the rollers. Once the roller slips, the speed of the roller will produce mistakes, which makes the calculation results inaccurate.

3. Multiple Linear Regression

Multiple linear regression is a statistical method. It is often used in big data samples to predict the relationship between dependent variables and multiple independent variables [11].

3.1. The Model of Multiple Linear Regression

In general, the multiple linear regression model is as follows:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \varepsilon_i, \quad i = (1, 2, 3 \cdots, n) \quad (1)$$

where $y_i (i = 1, 2 \cdots, n)$ are predicted values of the independent variables; $x_{ip} (i = 1, 2 \cdots, n)$ are known values of independent variables; P are the number of independent variables; $\beta_{ip} (i = 1, 2 \cdots, n)$ are the regression coefficient; β_0 are intercept; ε_i are the random value.

In order to write convenient, we use the matrix form as:

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \\ \beta_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{21} & \cdots & \cdots & x_{p1} \\ 1 & x_{22} & \cdots & \cdots & x_{p2} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & x_{2n} & \cdots & \cdots & x_{np} \end{pmatrix}_{p \times n}, \quad \varepsilon_i = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_i \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

The multiple linear regression model can be expressed as:

$$Y = X\beta + \varepsilon \quad (2)$$

$$\varepsilon \sim N(0, \sigma^2 I_n) \quad (3)$$

$$E(Y) = X\beta \quad (4)$$

3.2. The Multivariate Linear Regression Model Is Improved by Least Square Method

In order to minimize the actual value and the measured value, the least square method is used to solve the optimal value problem. The least squares method is used to search the $\beta_0, \beta_1, \dots, \beta_p$ estimate $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p$ to minimize the sum of squares of the deviations as follows:

$$Q(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_p) = \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_p x_{ip})^2 \quad (5)$$

The partial derivative can be obtained by the following equation:

$$\begin{cases} \frac{\partial Q}{\partial \hat{\beta}_0} \Big|_{\hat{\beta}_0 = \hat{\beta}_0} = -2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_p x_{ip}) = 0 \\ \frac{\partial Q}{\partial \hat{\beta}_1} \Big|_{\hat{\beta}_1 = \hat{\beta}_1} = -2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_p x_{ip}) x_{i1} = 0 \\ \dots \dots \dots \\ \frac{\partial Q}{\partial \hat{\beta}_p} \Big|_{\hat{\beta}_p = \hat{\beta}_p} = -2 \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_p x_{ip}) x_{ip} = 0 \end{cases} \quad (6)$$

It converts to the normal equations as follows:

$$\begin{aligned} \sum (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}) &= \sum y_i \\ \sum (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}) x_{i1} &= \sum y_i x_{i1} \\ \dots \dots \dots \\ \sum (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_p x_{ip}) x_{ip} &= \sum y_i x_{ip} \end{aligned} \quad (7)$$

The above equations is expressed by matrix as follows:

$$X^T X \hat{\beta} = X^T Y \quad (8)$$

Using the least square method, $\hat{\beta}$ can be expressed as follows:

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (9)$$

Thus, the regression coefficient of multiple linear regression is calculated, and the linear regression equation is written as follows:

$$l_{ij} = \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) \quad (i, j = 2, 3, \dots, p) \quad (10)$$

$$l_{iy} = \sum_{k=1}^n (x_{ki} - \bar{x}_i)(y_k - \bar{y}) \quad (i = 2, 3, \dots, p) \quad (11)$$

$$l_{yy} = \sum_{k=1}^n (y_k - \bar{y})^2 \quad (12)$$

$$L = (l_{ij})_{p \times p}, L_{xy} = (l_{iy})_{p \times 1}$$

$$\bar{x} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{pmatrix} \quad (13)$$

where

$$\bar{x}_i = \frac{1}{n} \sum_{k=1}^n x_{ki}, i = 2, 3, \dots, p \quad (14)$$

$$\bar{y} = \frac{1}{n} \sum_{k=1}^n y_k \quad (15)$$

So, we can get:

$$l_{ij} = \sum_{k=1}^n x_{ki} x_{kj} - n \bar{x}_i \bar{x}_j \quad (16)$$

$$L = C^T C - n \bar{x} \bar{x}^T \quad (17)$$

$$x^T x = \begin{pmatrix} n & E^T C \\ C^T E & C^T C \end{pmatrix} = \begin{pmatrix} n & n \bar{x}^T \\ n \bar{x} & L + n \bar{x} \bar{x}^T \end{pmatrix} \quad (18)$$

$$l_{iy} = \sum_{k=1}^n x_{ki} y_i - n \bar{x}_i \bar{y} \quad (19)$$

$$x^T Y = \begin{pmatrix} E^T Y \\ C^T Y \end{pmatrix} = \begin{pmatrix} n \bar{y} \\ L_{xy} + n \bar{y} \cdot \bar{x} \end{pmatrix} \quad (20)$$

where $C = (x_{ij})_{n \times p}$, $E = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}_{n \times 1}$.

Using matrix inverse formula, we can get:

$$(x^T x)^{-1} = \begin{pmatrix} \frac{1}{n} + \bar{x}^T L^{-1} \bar{x} & -\bar{x}^T L^{-1} \\ -L^{-1} \bar{x} & L^{-1} \end{pmatrix} \quad (21)$$

$$\hat{\beta} = (x^T x)^{-1} x^T y = \begin{pmatrix} \bar{y} - \bar{x}^T L^{-1} L_{xy} \\ L^{-1} L_{xy} \end{pmatrix} \quad (22)$$

So, the $\beta_0, \beta_1, \dots, \beta_n$ least squares estimate are:

$$\begin{cases} \hat{\beta}_0 = \bar{y} - \sum_{i=1}^p \hat{\beta}_i \bar{x}_{\cdot i} \\ \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \dots \\ \hat{\beta}_p \end{pmatrix} = L^{-1} \begin{pmatrix} l_{1y} \\ l_{2y} \\ \dots \\ l_{py} \end{pmatrix} = L^{-1} L_{xy} \end{cases} \quad (23)$$

3.3. Application of Improved Multivariate Linear Regression on the Welding Seam Tracking

The model is applied to practical engineering, will receive that

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon \quad i = (1, 2, 3, \dots, n) \quad (24)$$

where y is the deviation of welding seam tracking; β_0 is the data of measured deviation; $\beta_i x_i$ is influence factor for the steel strip. The speed of selected roller are speed of 2BR, 3BR, 4BR, 5BR and 6BR roller.

4. Simulation

4.1. Data Selection

The simulation data chose the information of 150 rollers of the steel strip, which included steel strip volume number, prediction error, measurement error and rollers speed. The actual measurement error is shown as figure 4.

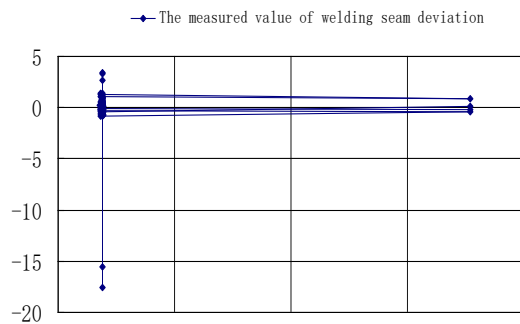


Figure 4. The actual measurement error of welding seam deviation.

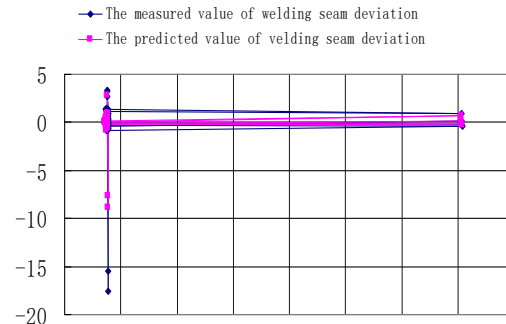


Figure 5. The actual measurement error of welding seam deviation comparing with the predicted value.

The comparison between the prediction results and measured results is shown as figure 5. By comparison, the error range of the prediction results which obtained by the improved multiple linear regression model is smaller than that calculated by the roller speed.

4.2. SPSS Simulation

It used the SPSS to simulate the improved multiple linear regression model. The $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ value of the Eq24 were obtained through the simulation computation. Simulation results are shown in table 1.

Table 1. The result of SPSS simulation.

Model		Unstandardized coefficients		Standardized coefficients		t	Sig.
		B	standard error	version			
1	constant	.679	.280			2.428	.016
	B0	1.210	.027	.956		44.128	.000
	2BR	.099	.140	.838		.706	.481
	3BR	-.143	.192	-1.222		-.748	.455
	4BR	.305	.208	2.617		1.467	.144
	5BR	-.408	.172	-3.509		-2.370	.019
	6BR	.142	.130	1.226		1.091	.277
2	constant	.700	.278			2.521	.013
	B0	1.209	.027	.955		44.221	.000
	3BR	-.045	.132	-.387		-.344	.731
	4BR	.310	.208	2.656		1.492	.138
	5BR	-.410	.172	-3.526		-2.386	.018
	6BR	.140	.130	1.208		1.077	.283
3	constant	.694	.276			2.511	.013
	B0	1.211	.027	.956		45.057	.000
	4BR	.257	.139	2.201		1.850	.066
	5BR	-.399	.168	-3.435		-2.370	.019
	6BR	.137	.129	1.184		1.060	.291
4	constant	.689	.276			2.495	.014
	B0	1.211	.027	.956		45.027	.000
	4BR	.288	.136	2.468		2.121	.036
	5BR	-.293	.135	-2.518		-2.164	.032

So, the multiple linear regression model in the strip steel welding seam tracking is:

$$y = 0.956 + 0.838x_1 - 1.222x_2 + 2.617x_3 - 3.509x_4 + 1.226x_5$$

The data correlation simulation figure is shown as figure 6, it shows that the data curve tends to basically linear. So, data correlation is very well.

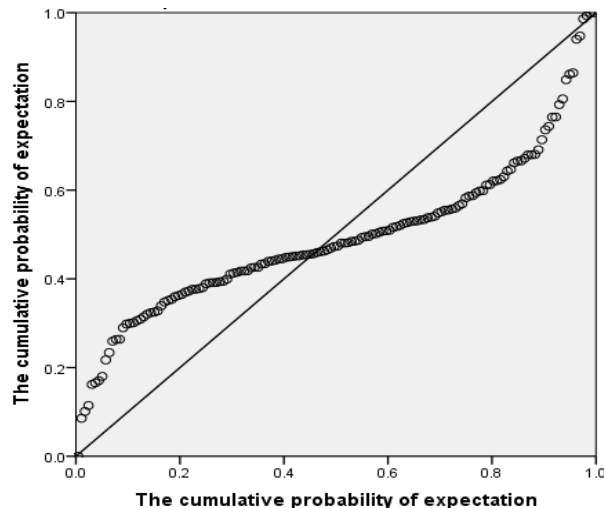


Figure 6. The Standardized regression residuals of the standard P-P figure.

5. Conclusion

In this paper, the least square method is used to improve the multiple linear regression model, and the model is applied to solve the problem of seam tracking in hot-dip galvanizing process. After the model is established, SPSS software is used for simulation calculation, and the relevant parameters are obtained.

Finally, the prediction results of the model are compared with those of the traditional prediction methods. The experimental results show that the model has better prediction effect than the traditional prediction method. Therefore, this method has a certain engineering application value.

References

- [1] Wang Y 2018 Study on influence factors of zinc layer thickness via response surface method, Taguchi method and genetic algorithm *J. Industrial Engineering & Management* **07** (01) 1-10.
- [2] Bao C R, Kang Y L and Wang L 2016 Numerical simulation of effect of air knife process on thickness of zinc layer *Journal of Iron and Steel Research* **20** 16.
- [3] Ma W L 2013 Welding seam tracking system of the hot- dip galvanizing line in Bao Steel *Metallurgical Industry Automation* **37** (1) 53-56.
- [4] Ushakov S N, D'yakonov A A, Gorbunov A V, et al. 2012 Production of high-strength microalloyed steel strip on continuous lines *Steel in Translation* **42** (2) 183-185.
- [5] Gi Y H and Kyoon C M 2010 Development of novel air-knife system to prevent check-mark stain on galvanized strip surface *ISIJ International* **50** (5) 752-759.
- [6] Yuan J and Shu J 2008 Application of modified multivariate nonlinear regression analysis model and neural network model to water quality prediction *Water Resources Protection* **24** (3) 46-48.
- [7] Park S H, Lee I B and Lee E 2020 Effects of thermal shields on temperature of a hot steel strip in rolling process *ISIJ International* **60** (8) 1737-1742.
- [8] Hore S and Das S K 2019 Computer simulation of microstructure evolution in strip-cast silicon steel with cube and fibre texture *Philosophical Magazine* **99** (22) 1-17.
- [9] Zhang C X 2009 A brief analysis of air knife control system *Technology Wind* 167.

- [10] Cheng J M, et al. 2009 Numerical simulation of the air knife in hot-dip galvanizing *Heavy Machinery* (2) 37-40.
- [11] Odinikuku W E, Atadious D and Onwuamaeze I P 2020 Optimization of mild steel welding process parameters using multivariate linear regression *Journal of Engineering Research and Reports* 43-50.