

PAPER • OPEN ACCESS

Prediction analysis of shield vertical attitude based on GRU

To cite this article: Liang Li *et al* 2020 *J. Phys.: Conf. Ser.* **1651** 012032

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

You may also like

- [An Achieving Data Exchange Cross-Chain Alliance Protocol](#)
Qianwen Wang, Shen Wang, Pan Zhang et al.
- [Method Study on Establishing of Ship Sewage Pollutants Discharging Inventory Based on AIS](#)
Rongchang Chen, Jing Shi and Zheng Wang
- [Analysis and Experimental Research on Dynamic Characteristics of Horizontal Shaft System of Photoelectric Equipment](#)
Lei Liu, WeiWei Jiang and Bi He



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

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Prediction analysis of shield vertical attitude based on GRU

Li Liang^{1 a}, Hu Changming^{1,2*} and Hou Yajun¹

¹Civil Engineering Construction and Management, Xi'an University of Architecture and Technology, Xi'an, Shaanxi, 710055, China

²Shaanxi Key Lab of Geotechnical and Underground Space Engineering (XAUAT), Xi'an University of Architecture and Technology, Xi'an, Shaanxi, 710055, China

^aliliang@xauat.edu.cn *Corresponding author's e-mail: 13609161448@163.com

Abstract. To solve the active control of the shield vertical attitude, this paper established the GRU model to realize the prediction of the pitch angle, which can reflect the shield vertical attitude. Firstly, the tunneling parameters, geological parameters and geometric parameters were selected as the influencing factors sets, and a series of data pre-processing strategies were established. And then, the GRU model was established, mainly including the division of training data, the optimization of hyperparameters, the determination of the objective function and optimization algorithm. Eventually, the shield interval from Chengdu was used to verify the new model. As a result, compared with the BP model and LSTM model, it could be found that the MGRU model had higher accuracy. This study could provide references for the prediction and help the engineers to realize the active control in the field of the shield attitude.

1. Introduction

With the rapid development of urban rail transit construction in China, the shield has been widely used because of its construction advantages. During the shield construction, untimely or improper shield attitude control will cause problems such as segment misalignment and rupture, ground subsidence. In serious cases, it is even necessary to redesign the tunnel route. Therefore, timely and effective control is one of the key links to ensure construction safety and quality.

To realize shield attitude control, scholars committed to research the prediction of shield attitude. Sugimoto M., et al (2002, 2007)^[1,2] realized the prediction of the shield attitude through the dynamic mechanical model. GAO B. (2011)^[3] established a expert system to predict the shield attitude. LIANG R., et al (2015)^[4] analyzed the evaluation law of the vertical attitude in soft ground based on the statistical analysis. LI Y. (2017)^[5] elaborated on the modeling process of the BP network to complete the forecasting of the shield segment attitude in breccia ground. ZHANG A.(2018)^[6] constructed the forecasting model of the shield vertical attitude in upper-soft and lower-hard ground based on the BP algorithm. CHEN Y.(2018)^[7] combined the XGBoost with the SVR algorithm to construct a forecast model of the shield attitude. Shen X., et al.(2019)^[8] proposed a prediction model of shield attitude based on the calculation of the shield-soil interaction. In order to predict the position and attitude of the shield, ZHOU C.,et al. (2019)^[9] proposed a hybrid deep learning model.

Literatures show that the current research is mainly focused on the modeling of the shallow neural network to complete the prediction of the shield attitude, and little research considers the “lag effect” about the tunneling process between the shield attitude and the influence factors. Based on the research status, this paper commits to constructing the deep neural network, Gate Recurrent Unit (GRU), which has fused the “lag effect” to predict the shield vertical attitude.



2. The GRU algorithm

Recurrent neural network (RNN), deep learning algorithm, shows a better performance in forecasting the time series problem. It has the advantages of considering the present and past situations. But the traditional RNNs can not solve the vanishing and explosion gradient problem. On this basis, improved algorithms such as long - term memory (LSTM) and gate cycle unit (GRU) are proposed.

Gate Recurrent Unit (GRU) proposed by Cho, et al. (2014)^[10] to simplify the network structure of the LSTM algorithm. Compared with LSTM, GRU algorithm is an improved LSTM with better efficiency and accuracy. Therefore, it is used for predicting the shield vertical attitude in this paper.

The GRU network controls the information through the update gate and reset gate, corresponding architecture as shown in figure 1. In this way, it can increase efficiency under the premise of predictive accuracy.

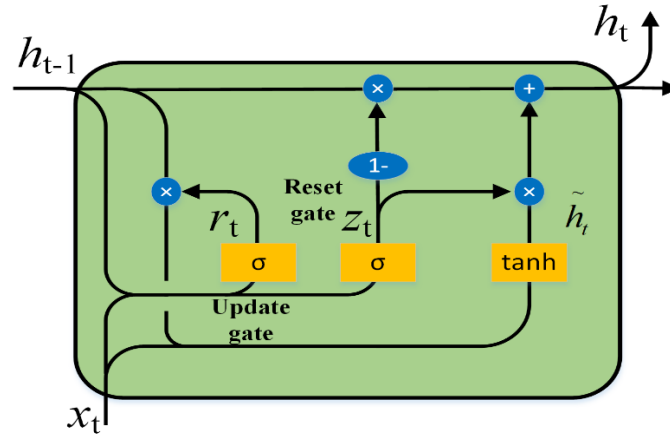


Figure 1. The architecture of the GRU network

Specifically, the update gate is used to obtain long-term dependency, and the reset gate can achieve short-term dependency in the operation of GRU. The principle of the GRU prediction is that set the number of hidden units $=h$, the small-batch input is $X_t \in R^{n \times d}$ (n is the number of samples, and d is the number of input) and the hidden state is $H_{t-1} \in R^{n \times h}$ when the time step is t , thereby:

$$\begin{cases} R_t = \sigma(X_t W_{xr} + H_{t-1} W_{hr} + b_r) \\ Z_t = \sigma(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \\ \hat{H}_t = \tanh(X_t W_{xh} + (R_t \odot H_{t-1}) W_{hh} + b_h) \end{cases} \quad (1)$$

Where W_{xr} , W_{xz} , $W_{xh} \in R^{d \times h}$ and W_{hr} , W_{hz} , $W_{hh} \in R^{h \times h}$ is the weight parameters, b_r , b_z , $b_h \in R^{1 \times h}$ is the bias parameters. When R_t is approximately 1, the foretime state will be transited and saved. Conversely, R_t tend to 0, it will be not kept.

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \hat{H}_t \quad (2)$$

Equation (2) illustrates that the reset gate can combine the hidden state in the past time and the alternative state in the current. When the Z_t is approximately 1, the early state can be transited to the current state.

3. The MGRU predictive model of shield vertical attitude

This study proposed the MGRU predictive model of shield vertical attitude to describe the high-dimensional and non-linear mapping relationship between the shield vertical attitude and it's influencing factors. The corresponding predictive framework has shown in figure 2.

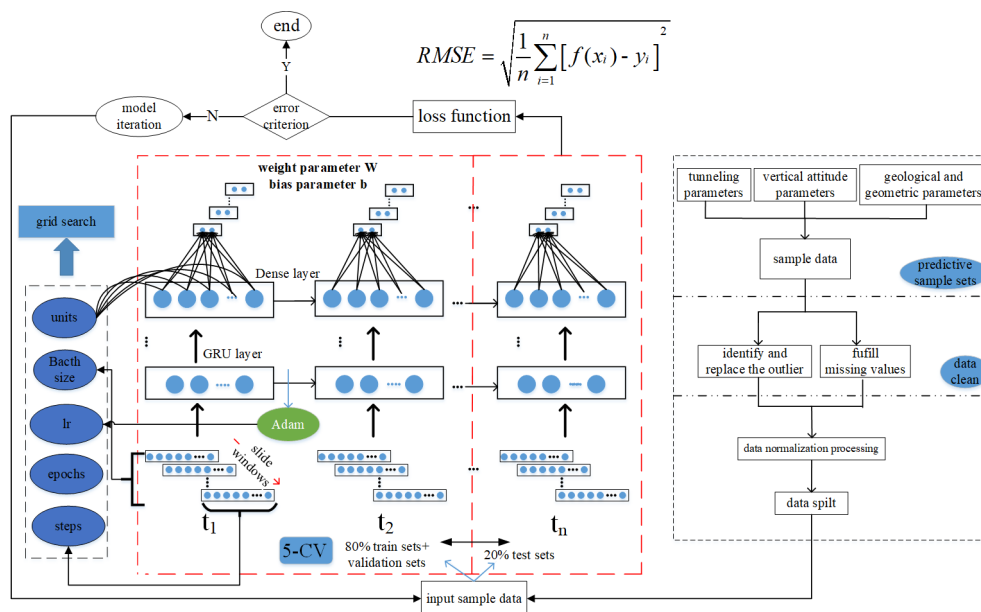


Figure 2. the GRU predictive framework of shield vertical attitude

The specific steps of the GRU predictive model are as follow:

- 1) Determine the input and output variables. In this study, the tunneling parameters, geological parameters and the geometric parameters are considered as the input variables. And then, the pitch angle is the index to reflect the shield attitude.
- 2) Data pre-processing. In order to ensure the predictive quality, a series of data pre-processing strategies are implemented, such as identify and replace the outlier in the box plot, fulfill the missing values in linear interpolation, normalization processing in maximum-minimum method, data split and so forth.
- 3) Transform data formation. Considering the “lag effect” of the shield attitude, the model is conducted to establish the relationship between past and current influencing factors and the future shield attitude. The input sample data is collected by use of the slider window, and the data is converted to the formation of (samples, width window, the dimension of the input variables).
- 4) Construct the predictive model. The GRU predictive model can be constructed under the Keras based on the above mentioned steps.
- 5) Optimize hyperparameters. The paper optimizes the hyperparameters (steps, batch_size, units of hidden layers N_i (is the i th hidden layer), epoch, lr) based on the Grid search method after setting the range of the values of hyperparameters and the various steps.
- 6) Train model. The grid search method is used to train the model using the training set and the validation set, and the Adam algorithm is used as the optimization algorithm.
- 7) If the MSE (objective function) $<$ the error criterion or the biggest epochs, the training process of the model is finished, otherwise continued from step 6).
- 8) The test sets can be used to test the GRU predictive model. If the predictive results can not fulfill the accuracy requirement, the new train should be performed from step 5).

4. The case study

To verify the proposed prediction model, the shield interval from Chengdu in China was used to elaborate on the prediction of the shield vertical attitude.

4.1. Project overview

The study implements the case application based on the Chengdu shield interval in China. In this project, the shield diameter is about 6m, and the buried depth is about 16-18m. The propulsion system of the shield is four-partition. The shield is mainly across five types of ground, including slightly dense pebble,

medium dense pebble, dense pebble, strongly weathered mudstone and moderately weathered mudstone. The main physical and mechanical parameters are shown in table 1.

Table 1. The main physical and mechanical parameters

ground types	density (g/cm^3)	Deformation modulus (MPa)	Permeability coefficient (m/d)	cohesion (KPa)	Internal friction angle ($^\circ$)
slightly dense pebble	2.10	30	22	0	30
medium dense pebble	2.20	35	22	0	35
dense pebble	2.30	40	22	0	40
strongly weathered mudstone	2.10	23.2	0.44	65	30
moderately weathered mudstone	2.32	23	0.44	420	35.60

4.2. Constructing influence factors sets

The influencing factors sets are constructed from three aspects, the tunneling parameters, the geometric parameters and the geological parameters. The corresponding variables and data sources are shown in table 2.

Table 2. Variables and data sources of the influence factors sets

The types of variables	The names of variables	unit	data sources
input variables	the total thrust (X_1)	t	PLC
	the torque of cutter head (X_2)	t*m	
	penetration (X_3)	mm/r	
	synchronous grouting volume (X_4)	m^3	
	the difference between the upper and lower cylinder (X_5)	mm	
	the buried depth (X_6)	m	Geological section drawing and survey report
	compound ratio (X_7)	-	Geological section drawing
	the deformation modulus (X_8)	kPa	Geological section drawing and survey report
	internal friction angle (X_9)	$^\circ$	
	cohesion (X_{10})	MPa	
Output variable	the pitch angle (Y)	mm/m	Automatic guidance system

4.3. Data Collection, pre-processing and split

The construction data are collected based on table 2 from ring 170 to 240 to develop the prediction of the vertical attitude model, and the gathered frequency is 5 minutes. To improve the accuracy and efficiency of the model, the box plot method is used to dispose the outlier. And then, the missing values are fulfilled based on linear interpolation. Eventually, the maximum-minimum method is applied to increase the speed of the model convergence. In total, there are 500 groups of sample data, and the descriptive statistic is shown in table 3.

Table 3. The descriptive statistic of sample data

variables	unit	maximum	minimum	average	standard deviation
the total thrust (X_1)	t	1115.00	665.00	893.66	96.71
the torque of cutter head (X_2)	t*m	295.00	185.00	232.04	23.91
penetration (X_3)	mm/r	9.33	5.83	7.77	0.83
synchronous grouting volume (X_4)	m^3	6.90	5.55	6.34	0.28

the difference between the upper and lower cylinder(X_5)	mm	31.00	-87.00	-29.99	27.66
buried depth(X_6)	m	16.08	14.33	15.20	0.52
compound ratio (X_7)	-	0.71	0.58	0.63	0.04
deformation modulus (X_8)	KPa	29.87	26.38	28.32	0.92
internal friction angle (X_9)	°	36.35	33.17	35.15	0.62
cohesion(X_{10})	MPa	229.69	168.06	197.30	18.37
the pitch angle (Y)	mm/m	1.83	-2.00	-0.26	0.81

The sample data split into 80% training and validation sets and the 20% test sets. At the same time, the data is divided by use of slide windows. Training and validation sets are partitioned in a five-fold cross-validation method.

4.4. Results

The computer configuration and software environment used in the model development is Intel (R) Core (TM) i5-7200U CPU, 16.0GB memory. The system is Windows 10 (64-bit), the programming language version is Python 3.6.5, the integrated development environment is spyder 3.2.3 version in the Anaconda package, and the model is conducted based on the Keras package which puts the Tensorflow as the backend.

The predictive model is implemented based on the GRU predictive framework of shield vertical attitude (Figure 2). The correlation coefficient (R^2) and root mean square error ($RMSE$) are the evaluated indexes of the models. The $RMSE$ gradually decreases and tends to stabilize when the epoch is 200 (figure 3).

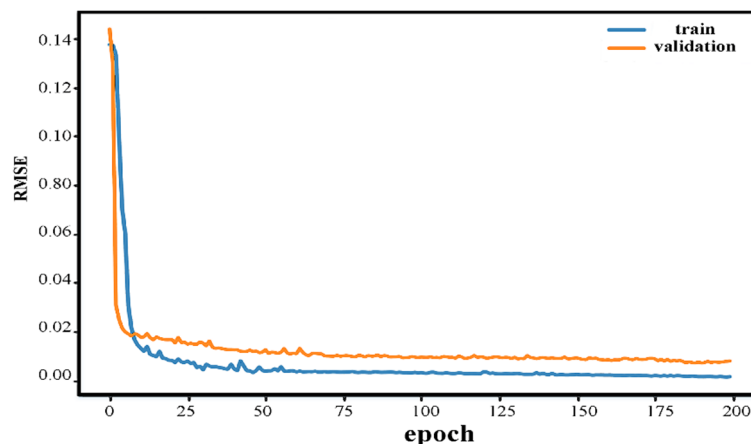


Figure 3. the train results of the GRU predictive model

To verify the performance of the proposed GRU predictive model, the test data sets are used to evaluate. And then, the BP and LSTM algorithms are conducted to compare with the GRU model based on the evaluated indexes. Corresponding results are shown in table 4.

Table 4. The comparison results of the three models

The predictive model	$RMSE$	R^2
GRU	0.083	0.998
LSTM	0.114	0.995
BPNN	0.167	0.990

5. Conclusions

Based on the comparison results of the three models, it can be seen that the GRU model of the shield vertical attitude has a better performance than other models. Therefore, the proposed model in this study, the GRU predictive model, can achieve an accurate prediction of the shield vertical attitude.

Appendices

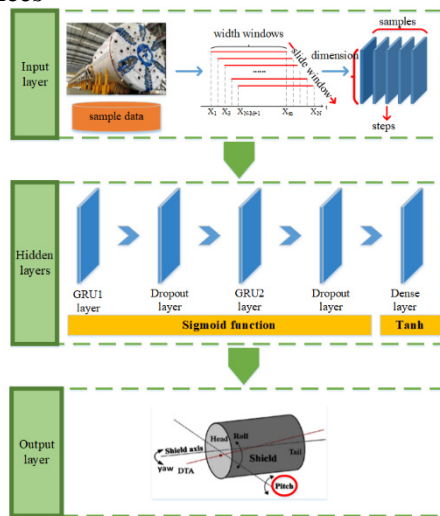


Figure 4. the initial network structure of the GRU model

Input layer (cell=10)	input	(none,5,10)
GRU layer (Activation function=Sigmoid, cell=128)	input	(none,5,10)
	output	(none,5,128)
GRU layer (Activation function=Sigmoid, cell=128)	input	(none,5,128)
	output	(none,128)
Dense layer (Activation function=Sigmoid, cell=256)	input	(none,128)
	output	(none,256)
Output layer (Activation function=Tanh, cell=1)	input	(none,128)
	output	(none,1)

Figure 5. the final network structure of the GRU model

Acknowledgments

Authors wishing to acknowledge the data support of the Chengdu Project Department of China Railway First Bureau for this study.

References

- [1] SRAMOON A., SUGIMOTO M., KAYUKAWA K..(2002) Theoretical Model of Shield Behavior During Excavation. II: Application[J]. Journal of Geotechnical and Geoenvironmental Engineering, 128:156-165.
- [2] SUGIMOTO M., SRAMOON A., KONISHI S., et al. (2007) Simulation of Shield Tunneling Behavior Along a Curved Alignment in a Multilayered Ground[J]. Journal of Geotechnical and Geoenvironmental Engineering, 133:684-694.
- [3] GAO B.(2011) The study on attitude and orbit automatic control system of EPB shield machine[D]. Shenyang: Northeastern University.
- [4] LIANG R., XIA T., LIN C., et al. (2015) Control of the shield's vertical attitude during shield tunnelling in soft soils[J]. Modern Tunneling Technology, 52: 152-157, 183.
- [5] LI Y., WANG X., LI Q., et al.(2017)A method for optimizing the operating track parameters of shield machine in breccia formation based on neural network [P]. China Patent: CN108710940A.
- [6] ZHANG A..(2018) Study of fuzzy control on tunneling parameters of shield attitude in upper-soft and lower-hard ground[J]. Journal of Railway Science and Engineering, 15: 2920-2927.
- [7] CHEN Y. (2018) Analys is and Visualization of Shield Construction Data:A Study and Implementation [D]. Xi'an: Xidian University.
- [8] SHEN X., YUAN D., JIN D..(2019) Influence of Shield Attitude Change on Shield-Soil Interaction[J]. Applied Sciences,9:1-19.
- [9] ZHOU C., DING L., ZHOU Y., et al. (2019) Hybrid Support Vector Machine Optimization Model for Prediction of Energy Consumption of Cutter Head Drives in Shield Tunneling[J]. Journal of Computing in Civil Engineering, 33:1-17.
- [10] CHO K., MERRIENBOER M., DZMIRY D., et al. (2014) On the Properties of Neural Machine Translation: Encoder-Decoder Approaches[J]. Computer Science,9:103-111.