2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Long and Short-Term Memory Mechanism Hybrid Model for Speed Trajectory Prediction of Heavy Haul Trains*

Kexuan Xu, Qiang Liu, *Senior Member, IEEE*

***Abstract*—The trajectory prediction of heavy heavy trains (HHTs) is crucial for ensuring the safe and automatic operation. It is inevitable to design a train model for predicting train operation trajectory of the HHTs. However, large capacity and the existence of unmodeled dynamics in the operation process caused by air resistance, working condition switching, and external environment make it difficult to establish accurate speed trajectory prediction models (STPMs) using traditional mechanism-driven methods. Most recent research considers using data-driven model to learn information from data, but they have no information about the physical characteristic of the STMP model. This makes it difficult to accurately describe the relationship between the control force and the running speed during the train operation. To overcome these issues, this study combines the mechanism-driven model with the deep learning model to establish a new long and short-term memory mechanism hybrid (LSTMMH) model. Specifically, the mechanism model describes the change of control force, while the LSTM captures the unmodeled dynamics of long time series of running process. The effectiveness of the proposed method is demonstrated while the performance is compared with the traditional LSTM and mechanism models using the real data.**

I. INTRODUCTION

It is an important task to establish an accurate speed trajectory prediction model (STPM) to predict train trajectory and guide drivers for operation safety and stability of the heavy haul trains (HHTs). This model predict the subsequent speed trajectory according to the existing running state characterized by speed, mileage, and control force. However, large capacity of the HHTs leads to uncertainties of its force. Also there are significant unmodeled dynamics in complex track conditions including ramp, curve rate, speed limit, etc. The research on STPM modeling has attracted much attention in the academia and industry.

Some scholars have made efforts to establish STPMs for HHTs. Overall, the existing work can be classified into three categories: mechanism-driven models, data-driven models, and mechanism-data hybrid models. Some of these studies are devoted to establishing mechanism-driven models [1],[2], but there are significant unmodeled dynamics caused by track conditions, external environment and internal wear of equipment. Recent work has focused on the time-varying, nonlinear, and time-delay of control force transmission during train running [3]. For example, taking into account the train motion dynamics containing nonlinearity and parameter uncertainty, the variable with time delay is integrated into the analysis of velocity delay by the Lyapunov-Krasovskii function [4]. Literatures [5-7] use state-space models to establish the dynamic model of HHTs. The linearizing nonlinear forces, however, cannot be described by the mechanism-driven model. In addition, the dynamics of train force transmission are omitted in the existing work. Although the mechanism model is established based on physical characteristics, ignoring the unmodeled dynamics may greatly decrease the STPM accuracy in the complex environment. Moreover, it requires iterative verification by drivers with extensive experience in field experiments [8].

Data-driven models learn from a plurality of historical data and have a wide range of applications in solving practical engineering problems. The STPM data come from speed sensors installed on the locomotive and carriage, which record a large number of historical data during the round-trip running of the HHTs. Aiming at the prediction of train delay, a wavelet neural network model is established [9]. An intelligent prediction and feature recognition method for the large area joint train delay is proposed. Reference [10] takes into account the energy consumption of high-speed trains and proposes the back-propagation model. Although these data-driven models achieve higher accuracy than mechanism-driven models, the structure relies entirely on training data. The physical characteristics of the control force during running for the HHTs cannot be effectively extracted. To capture the time dynamics of the trajectory, reference [11] proposes a recurrent neural network (RNN)-based STPM. The STPM is even more difficult for the HHTs than the ones for high-speed railways and urban railways with dispersed power and short bodies.

References [12-14] propose data-driven models, while the key parameters cannot be used for online strategies. The mechanism model is used to construct the physical characteristics. However, the above methods are to identify the mechanism-driven models, while and data-driven model alternately apply traditional neural networks to improve the prediction accuracy. Reference [15] proposes a gas leakage diagnosis strategy based on a two states mechanism model, and the unscented kalman filter method is developed to diagnose. Reference [16] proposes a mechanism model but that would lead to large errors. Therefore, a data-driven error compensation model is beneficial to speed trajectory prediction. More practial aspects should be taken into account while establishing the data-driven model: 1) the large capacity of HHTs leads to uncertainties in the control force; 2) in complex lines, there are unmodeled dynamics and time-varying characteristics in long time series that may significantly affect the train running trajectory. We thus develop a mechanism-driven and data-driven deep learning model for STPM. The mechanism model establishes the longitudinal dynamic relation and captures the force

Kexuan Xu and Qiang Liu are with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang, China (corresponding author: Q. Liu, e-mail: liuq@mail.neu.edu.cn).

characteristics. Data-driven model learns the estimation deviation caused by unmodeled dynamics and predicts the subsequent running state to compensate the mechanism model. Compared with other traditional mechanism-data hybrid modeling methods, this paper considers the effect of the disturbance dynamics of the operation environment and the complex physical characteristics in uncertainty force during the train running.

The remainder of this paper is organized as follows. Section II describes the longitudinal dynamics model of HHT. Section III analyzes the structure of the sliding window LSTM neural network and present the structure of the long and short-term memory mechanism hybrid model (LSTMMH) model. Section IV describes the experimental simulation. In the end, section V sets out the conclusions and briefly reviews the scope of future work.

II. Longitudinal Dynamic Model of HHTs

During the running of HHTs, due to the line gradient, curve, train body, etc., the train is subjected to various forces. Among them, the force perpendicular to the direction of train movement is called transverse force, and the one parallel to the direction of train movement is called longitudinal force. This section focuses on the kinetic analysis of the longitudinal forces and establishes the longitudinal kinetic model of the HHT in combination with the regulations on railway train traction calculation.

Refer to literature [17], mechanism model of HHT is established. The traction force, braking force, and resistance are considered. The corresponding differential equation is constructed according to the second law of Newtonian mechanics to reflect the force.

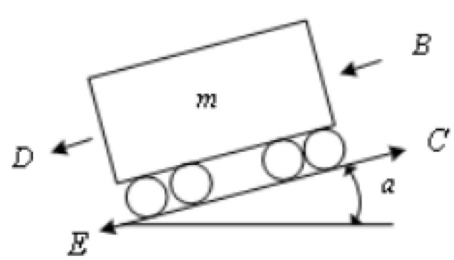

Figure 1. Force analysis of HHT.

The differential equation of the longitudinal dynamics of the HHT is as follows:

$$m\dot{v} = -B + C - D - E + f \tag{1}$$

where a denotes the slope angle of the line section. m is the train mass. $\dot{v}$ refers to the acceleration of the train. B refers to the resistance of the HHT. C is the traction force. D is the braking force. f represents the air resistance, wind resistance, adhesive force and unmodeled factors caused by the external environment on the train force during HHT running. B, C, and D can be calculated by the mechanism model to capture accurate force characteristics, while f needs to learn the environmental disturbance dynamics through the data-driven model.

III. Methodology

A. HHT mechanism-data hybrid model strategy

In order to derive the advantages of the mechanism-driven model containing explicit physical characteristics and the adaptability of the data-driven model, a LSTMMH modeling method is proposed. The LSTMMH model is shown in Figure 2. The train running data input is composed of a long time serie, and the length of the sliding window. The predicted values of the data-driven model are used as inputs to the mechanism model. The output of the mechanistic model and the predicted values compose the input of the LSTMMH model into the sliding window for the next moment.

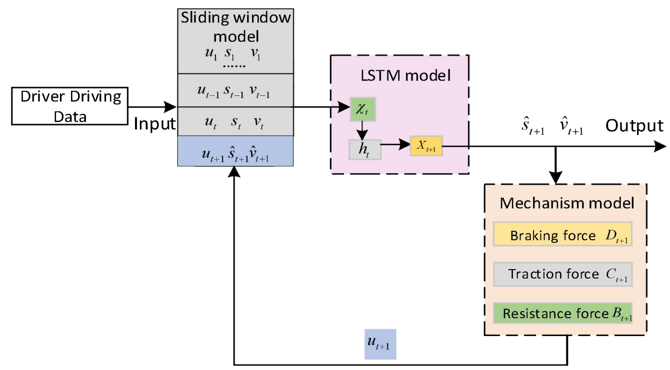

Figure 2. HHT mechanism-data hybrid model strategy.

We use the LSTMMH to predict the train running speed at the next moment. The mechanism model is used to analyze the train forces in detail and reflect the real-time control force performance. We consider a time step t, the three-dimensional matrix u,s,v of actual driver running data is used as the input of the LSTM model. v is the running speed of HHT, s is the running mileage, and u is the control force. The LSTM model is to obtain the predicted values of the speed and mileage at the next moment. The predicted value u_{t+1} is obtained at the next moment through the mechanism model, and then the inputs of the LSTM model are formed with $\hat{s}_{t+!}, \hat{v}_{t+1}$. Also the outputs of the model are $\hat{s}_{t+!}$ and $\hat{v}_{t+1}$ at the next moment.

B. Sliding window LSTM-based trajectory prediction

Recurrent neural networks (RNNs) are used for time serie prediction. The connections between nodes form a directed graph along a time series. This makes it a time-dynamic behavior. However, it is only suitable to short time serie, and can not process the effective information of long time serie, resulting in a decrease in accuracy.

As a variant of RNN, LSTM can overcome the problem that RNN cannot handle long distance dependencies [18]. The key to LSTM is the memory unit, also known as the cell state. Cell states can transmit information throughout the sequence and decide when to discard and update information. Through the gating mechanism, LSTM can control the flow of information. The main doors of the LSTM include the memory gate, the input gate, and the output gate.

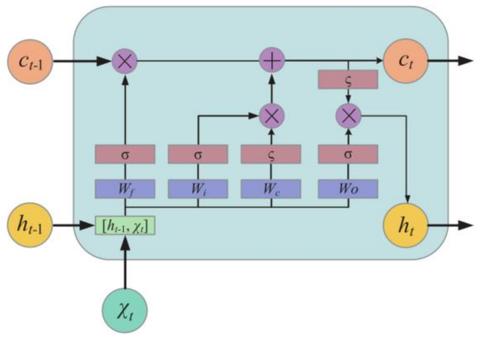

Figure 3. LSTM neural network structure.

In this paper, the LSTM neural network is chosen as the data-driven model whose basic structure is shown in Figure 3. For the convenience of description, we consider one time step t. Let $\chi_t = [v_{1:t}, s_{1:t}, u_{1:t}]^{\mathrm{T}}$ be the current state vector. Let $h_t = [\hat{v}_t, \hat{s}_t]$ be the output of the next moment. We adopt the LSTM structure, which is composed of three layers. The first is used to receive the running speed, mileage and control force as input characteristics χ_t. The second is the recurrent hidden layer composed of memory blocks involving memory cells and adaptive gating units, and the last layer is the output layer to obtain v_t, s_t.

Single-step forecasting inputs all actual values and predicts only one subsequent value. Multi-step prediction predicts multiple subsequent values, but the prediction error of multi-step prediction accumulates with the number of steps increases. In this paper, multi-step-ahead prediction (MSP) is conducted to obtain a long-term prediction sequence. Multi-step-ahead prediction is realized by performing the one-step-ahead prediction iteratively. One-step-ahead prediction means that the LSTMMH model only predicts one set of data at a time, and combines it with the output of the mechanism model to form the input data of the LSTMMH model at the next time. The sliding window length is set to L. The prediction is iteratively achieved to obtain long-term time serie. We assume the current period is i predicted from $i+1$.

The extracted HHT running data are divided into two parts: the measurement sequence $HI_{1:i} = [h_{i+1}, h_{i+2}, \cdots, h_{i+n}]$, and the prediction sequence $\hat{H}_{1:i} = (\hat{h}_{i+1}, \hat{h}_{i+2}, \cdots, \hat{h}_{i+n})$.

$$\hat{h}_{i+1} = g(\hat{h}_{i-L+1}, \hat{h}_{i-L+2}, \cdots, \hat{h}_i) \quad (2)$$

$g(\cdot)$ is a nonlinear function. The $\hat{h}_{i+2}$ is predicted using the same method, but using the output of Eq. (2) as the input:

$$\hat{h}_{i+2} = g(\hat{h}_{i-L+2}, \hat{h}_{i-L+3}, \cdots, \hat{h}_{i+1})^T \quad (3)$$

where $\hat{h}_{i+1}$ is the predicted value of measurement sequence at cycle $i+1$. It is worth noting that in the sliding window input module, to capture the dynamic characteristics of the time series, the length of the sliding window needs to be continuously filled to the maximum length L. The processing of performing pre-padding operation on the input sequence is shown in Figure 4.

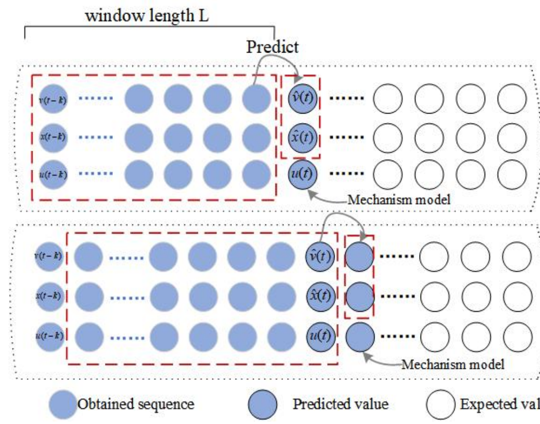

Figure 4. Slide-window pre-fill operation of input sequences.

C. LSTM mechanism hybrid module

Based on the standard LSTM, our LSTMMH considers the unique nature of the STPM and makes a number of improvements. Specifically, the LSTMMH model includes a mechanism module, a data-driven module, a sliding window input module, and a LSTM mechanism hybrid module. The LSTM mechanism hybrid module captures the input characteristics of the control commands through the collected line operation data. The three-dimensional matrix u, s, v extracted from the running state is used as the input to the LSTM model processed through hidden layers of forgetting gates, input gates, and output gates to obtain $\hat{s}, \hat{v}$.

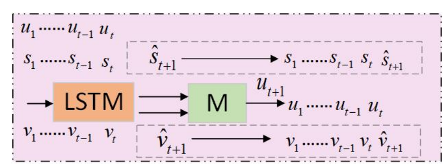

Figure 5. LSTMMH neural input and output structure.

The calculation of the locomotive traction is as follows:

$$C_T = \begin{cases} 760 & v \in [0,5] \\ 760\text{-(v-5)} \times 228/65 & v \in (5,65] \\ 9600 \times 3.6/v & v \in (65,100] \end{cases} \quad (5)$$

The calculation of the locomotive braking force is as follows:

$$D_T = \begin{cases} 461 \times v/5 & v \in [0,3] \\ 461 & v \in (3,75] \\ 9600 \times 3.6/v & v \in (75,100] \end{cases} \quad (6)$$

where C_T is the locomotive traction, kN. D_T is the locomotive braking force, kN. v refers to locomotive running speed, km/h.

The calculation of locomotive operation resistance is:

$$w_1 = 1.2 + 0.0065v + 0.000279v^2 \quad (7)$$

$$w_{11} = 0.92 + 0.0048v + 0.000125v^2 \quad (8)$$

where w_1 is the basic resistance per unit of the locomotive, N/kN. w_{11} is the basic resistance per unit of carriages, N/kN.

The calculation of basic train running resistance is:

$$W_1 = 0.001 \times (w_1 + w_{11})mg \tag{9}$$

where W_1 is the basic resistance of train running, kN. 0. 001 is the unit conversion coefficient.

Due to the long formation of HHTs, the length of the train needs to be considered. The calculation formula of the additional resistance per unit running is:

$$w_2 = \frac{1}{L_1}\left[\sum(i_i l_{ii}) + 600\sum\frac{l_{ri}}{R_i}\right] \tag{10}$$

where w_2 is the additional resistance per unit running, N/kN. 600 is the resistance coefficient of the calculated curve. L_1 is the length of HHT, m. i_i is the slope of the first ramp covered by the train (0.1%). l_{ii} is the length of the first ramp covered by the train, m. R_i is the radius of the first curve covered by the train, m. l_{ri} is the length of the i th curve covered by the train, m.

The additional resistance during train running is：

$$W_2 = 0.001 \times w_2 mg \tag{11}$$

$$B_T = W_1 + W_2 \tag{12}$$

where B_T is the resistance force, kN. W_2 is the additional resistance, kN. w_2 is the additional resistance per unit running of the train, N/kN. g is the gravitational acceleration, 9.8 N/kg.

$u_{t+1}, \hat{s}_{t+1}, \hat{v}_{t+1}$ can similarly form the three-dimensional matrix and as the input of LSTM through the sliding window.

IV. Experimental Study

A. Comparison of prediction results

To verify the effectiveness of the LSTMMH-based speed trajectory prediction method, we uses actual data to simulate. In this paper, the sliding window length of the train dataset is 90 and the test dataset is 10. The test running dataset of HHT speed, mileage and control force on locomotive are collected per second to form a input matrix of 3390 × 3. The experimental HHT departed from kilometer mark 107km, ran 57km, and ended the running experiment at kilometer mark 164km.

Figures 6-8 show the speed trajectory predictions of the mechanistic model in reference [19], LSTM in reference [20], and LSTMMH, respectively. The black line is the actual running trajectory, and the red one is the model prediction trajectory. Figures 9-11 show the speed prediction errors of these models. As seen from these figures, the prediction performance of the mechanism model is the worst. The prediction is significantly degraded. The prediction error of the LSTM model ranges between [-2, 2], while the error of the LSTMMH model ranges between [-1.5,1].

In addition, the trajectory predicted by the LSTM has a small amplitude of oscillation. In particular, it is more likely to lead to a larger prediction error when the train changes at the operation condition switch point (for example, when the train changes from the traction condition to the coasting condition or from the traction condition to the braking condition). From a detailed analysis, the reason is that there is a time delay in air brake wave transmission in the train braking process, while the the ordinary LSTM model cannot learn these physical characteristics. Therefore, our LSTMMH model can overcome the above shortcomings by calculating the control force to decide the operation condition.

Moreover, to demonstrate the effectiveness of the proposed LSTMMH, four common metrics including root mean squared error (RMSE), mean absolute error (MAE), mean squared error (MSE), and R-square (R^2), are adopted to evaluate the performance. The results are as shown in TABLE I. The four columns on the left are four metrics of the predicted mileage s , and the right ones are the four metrics predicted speed v . The following conclusions can be drawn. The metrics of the mechanistic model still perform the worst compared with the LSTM and LSTMMH models. Compared with LSTM and LSTMMH, there is a large gap between RMSE, and MAE, MSE and R^2 also have deviations. It can be seen that the prediction of LSTM is easily affected by outliers, and the prediction performance will be affected due to the lack of physical characteristics during the operation condition transformation. In summary, the LSTMMH significantly outperforms the other models in these four performance metrics.

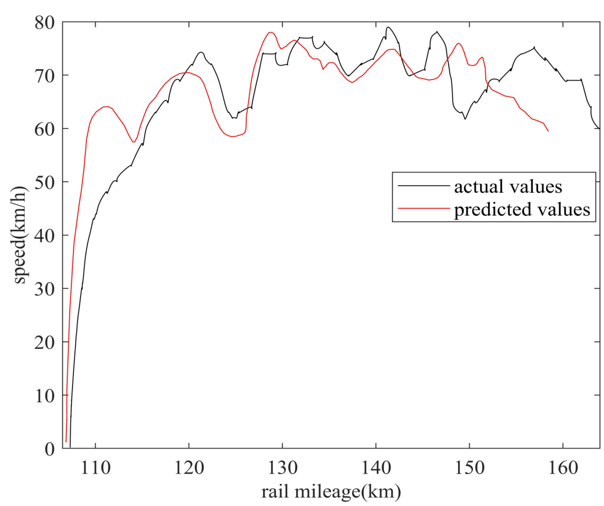

Figure 6. Speed trajectory prediction of the mechanistic model.

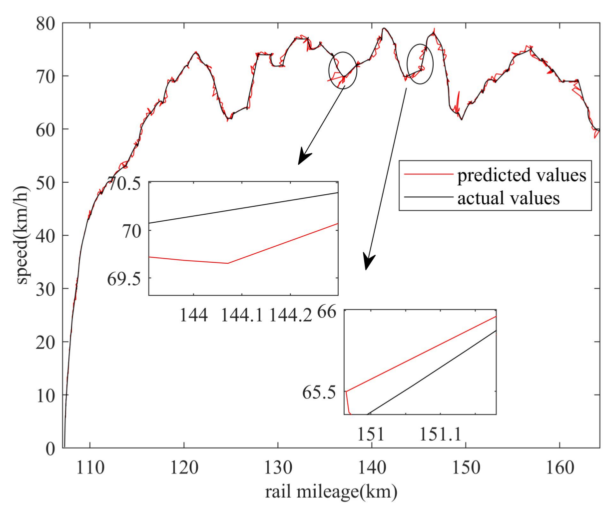

Figure 7. Speed trajectory prediction of the LSTM model.

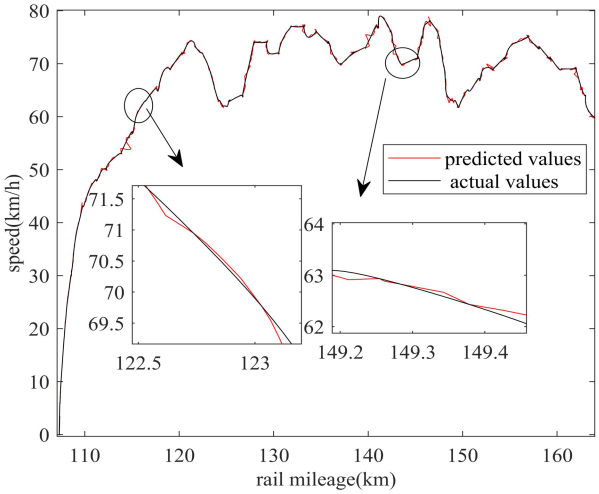

Figure 8. Speed trajectory prediction of the LSTMMH model.

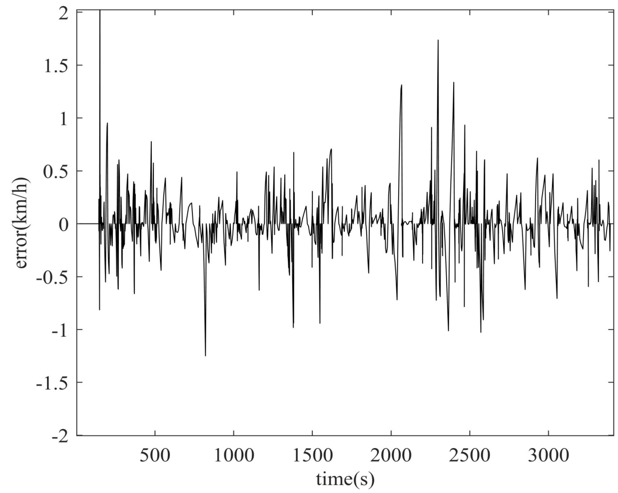

Figure 10. LSTM-based model prediction error.

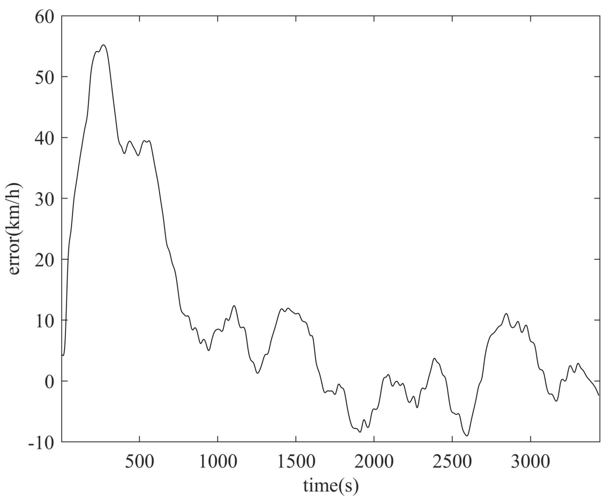

Figure 9. Mechanism model prediction error.

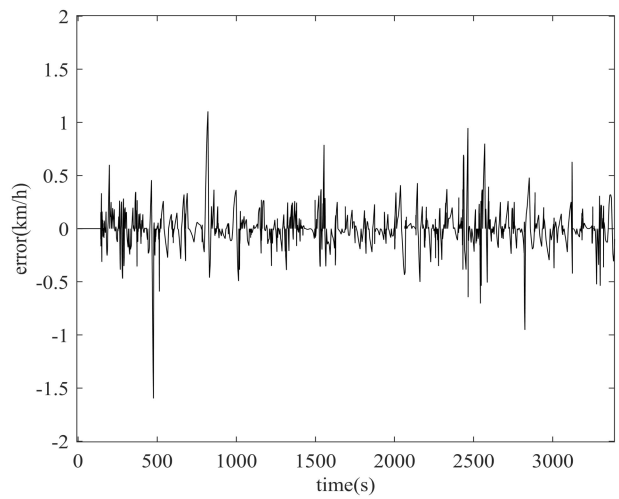

Figure 11. LSTMMH-based model prediction error.

TABLE I. COMPARISON OF PREDICTION PERFORMANCE RESULTS OF MECHANISTIC, LSTM, AND LSTMMH MODELS.

Method	**$RMSE_S$**	**MAE_S**	**MSE_S**	**R^2_S**	**$RMSE_V$**	**MAE_V**	**MSE_V**	**R^2_V**
Mechanism model [19]	30.0853	5.7772	46.6385	0.8361	60.3461	8.1485	56.0676	0.5495
LSTM [20]	6.9293	0.0912	0.0254	0.9996	16.9974	0.2413	0.1259	0.9997
LSTMMH	**3.2024**	**0.0342**	**0.0054**	**0.9998**	**7.3390**	**0.0939**	**0.0277**	**0.9999**

V. CONCLUSIONS

In order to establish an accurate STPM for HHTs to predict the speed trajectory over a long running time serie, this study proposes a new LSTMMH model. We combine a mechanism-driven model and a data-driven deep learning model (LSTM) into a hybrid model. While the mechanism-driven model can predict the force of HHTs, the data-driven model can obtain the unmodeled dynamics during running. The actual data verify that the LSTMMH has higher accuracy in trajectory prediction and robustness compared to traditional LSTM and mechanism-driven model.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (61991401, 62161160338).

REFERENCES

[1] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182-197, 2002.

[2] S. Li, L. Yang, and Z. Gao, "Distributed optimal control for multiple high-speed train movement: an alternating direction method of multipliers," *Automatica*, vol. 112, no. 5, pp. 108646, 2020.

[3] K. Fadhloun, H. Rakha, A. Loulizi, and A. Abdelkefi, "Vehicle dynamics model for estimating typical vehicle accelerations," *Transportation Research Record*, vol. 2491, no. 1, pp. 61–71, 2015.

[4] H. Ji, Z. Hou, and R. Zhang, "Adaptive iterative learning control for high-speed trains with unknown speed delays and input saturations," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 1, pp. 260–273, 2016.
[5] X. Zhuan and X. Xia, "Optimal scheduling and control of heavy haul trains equipped with electronically controlled pneumatic braking systems," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 6, pp. 1159–1166, 2007.
[6] K. Gao, Z. Huang, J. Wang, J. Peng, and W. Liu, "Decentralized control of heavy-haul trains with input constraints and communication delays," *Control Engineering Practice*, vol. 21, no. 4, pp. 420–427, 2013.
[7] L. Zhang and X. Zhuan, "Development of an optimal operation approach in the MPC framework for heavy-haul trains," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 3, pp. 1391–1400, 2014.
[8] X. Liu, J. Xun, B. Ning, and C. Wang, "Braking process identification of high-speed trains for automatic train stop control," *ISA Transactions*, vol. 111, no. 4 , pp. 171–179, 2021.
[9] Q. Zhang, F. Chen, T. Zhang, and Z. Yuan, "Intelligent prediction and characteristic recognition for joint delay of high speed railway trains," *Acta Automatica Sinica*, vol. 45, no. 12, pp. 9-19, 2019.
[10] D. wang, W. Ma, L. Li, Y. Yang, and C. Xiang, "Estimating traction energy consumption of high-speed trains based on BP neural network," *Journal of Central South University (Science and Technology)*, vol. 66, no. 4, pp. 1104-1110, 2017.
[11] K. Min, D. Kim, J. Park, and K. Huh, "RNN-based path prediction of obstacle vehicles with deep ensemble," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10252–10256, 2019.
[12] Y. Meng, S. Yu, J. Zhang, J. Qin, Z. Dong, G. Lu, and H. Pang, "Hybrid modeling based on mechanistic and data-driven approaches for cane sugar crystallization," *Journal of Food Engineering* , vol. 257, no. 9, pp. 44-55, 2019.
[13] X. Hu, L. Xu, X. Lin, and M. Pecht, "Battery lifetime prognostics," *Joule*, vol. 4, no. 2, pp. 310-346, 2020.
[14] Y. Jie, T. Chai, C. Luo, and Y. Wen, "Intelligent demand forecasting of smelting process using data-driven and mechanism model," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9745-9755, 2019.
[15] Y. Xu, X. Wu, X. Zhong, D. Zhao, M. Sorrentino, J. Jiang, C. Jiang, X. Fu, and X. Li, "Mechanism model-based and data-driven approach for the diagnosis of solid oxide fuel cell stack leakage," *Applied Energy*, vol. 286, no. 31, pp. 116508, 2021.
[16] J. Dai, N. Chen, X. Yuan, W. Gui, and L. Luo, "Temperature prediction for roller kiln based on hybrid first-principle model and data-driven MW-DLWKPCR model," *ISA Transactions*, vol. 98, no. 19, pp. 403-417, 2020.
[17] H. Yang, H. Ma, Y. Fu, and Z. Li, "Brake strategy for heavy-haul train based on multi-particle model," *Journal of Beijing University Technology*, vol. 46, no. 8, pp. 169–179, 2020.
[18] Y. Zuo, Y. Wu, G. Min, C. Huang, and K. Pei, "An intelligent anomaly detection scheme for micro-services architectures with temporal and spatial data analysis," *IEEE Transactions on Cognitive Communications and Networking*, vol. 6, no. 2, pp. 548-561, 2020.
[19] H. Yu, Y. Huang, M. Wang, Z. He, X. Meng, and Y. Li, "Research on operating strategies of heavy haul train based on genetic algorithm," *Journal of the China Railway Society*, vol. 42, no. 7, pp. 110-116, 2020.
[20] L. Zhen, T. Tang, and C. Gao, "Long short-term memory neural network applied to train dynamic model and speed prediction," *Algorithms*, vol. 12, no. 8, pp. 173, 2019.