ARCHIVES OF TRANSPORT *ISSN (print):* *0866-9546*
Volume 66, Issue 2, 2023 ***e-ISSN (online):*** *2300-8830*
DOI: *10.5604/01.3001.0016.3132*

RESEARCH ON SPEED CONTROL OF HIGH-SPEED TRAINS BASED ON HYBRID MODELING

Tao HOU[1], Li TANG[2], Hongxia NIU[3], Tingyang ZHAO[4]

[1, 2, 3, 4] *School of Automation and Electrical Engineering, Lanzhou Jiaotong University, Lanzhou, China*

Abstract:

With the continuous improvement of train speed, the automatic driving of trains instead of driver driving has become the development direction of rail transit in order to realize traffic automation. The application of single modeling methods for speed control in the automatic operation of high-speed trains lacks exploration of the combination of train operation data information and physical model, resulting in low system modeling accuracy, which impacts the effectiveness of speed control and the operation of high-speed trains. To further increase the dynamic modeling accuracy of high-speed train operation and the high-speed train's speed control effect, a high-speed train speed control method based on hybrid modeling of mechanism and data drive is put forward. Firstly, a model of the high-speed train's mechanism was created by analyzing the train's dynamics. Secondly, the improved kernel-principal component regression algorithm was used to create a data-driven model using the actual operation data of the CRH3 (China Railway High-speed 3) high-speed train from Huashan North Railway Station to Xi'an North Railway Station of "Zhengxi High-speed Railway," completing the mechanism model compensation and the error correction of the speed of the actual operation process of the high-speed train, and realizing the hybrid modeling of mechanism and data-driven. Finally, the prediction Fuzzy PID control algorithm was developed based on the natural line and train characteristics to complete the train speed control simulation under the hybrid model and the mechanism model, respectively. In addition, analysis and comparison analysis were conducted. The results indicate that, compared to the high-speed train speed control based on the mechanism model, the high-speed train speed control based on hybrid modeling is more accurate, with an average speed control error reduced by 69.42%. This can effectively reduce the speed control error, improve the speed control effect and operation efficiency, and demonstrate the efficacy of the hybrid modeling and algorithm. The research results can provide a new ideal of multi-model fusion modeling for the dynamic modeling of high-speed train operation, further improve control objectives such as safety, comfort, and efficiency of high-speed train operation, and provide a reference for automatic driving and intelligent driving of high-speed trains.

Keywords: *high-speed train, hybrid modeling, speed control, error compensation*

To cite this article:

Hou, T., Tang, L., Niu, H. X., Zhao, T. Y., (2023). Research on speed control of high-speed trains based on hybrid modeling. Archives of Transport, 66(2), 77-87. DOI: https://doi.org/10.5604/01.3001.0016.3132

Contact:
1) ht_houtao@163.com [https://orcid.org/0000-0002-7511-3013]; 2) tlhhxx_123@163.com [https://orcid.org/0000-0003-4029-9054]– corresponding author; 3) 369113896@qq.com [https://orcid.org/0000-0002-7732-5183]; 4) 1803600499@qq.com [https://orcid.org/0000-0002-2410-8732]

1. Introduction

As train speeds increase, the safety of trains operation becomes increasingly important. Hence, when constructing intelligent trains, stricter safety standards are demanded (Karolak, 2021). The research of high-speed train operation curve control and tracking method is highly relevant to attaining safe, efficient, and comfortable train operations. The current focus of study is on the modeling and control of the train speed tracking process, with the control premise being the development of an effective model description of the train operation process (Tan et al,. 2022). Due to the complexity of the high-speed train operating process, train operation control is affected by a number of variables, including vehicle performance, line conditions, environmental changes, and external interference. Creating an accurate mathematical model to explain the railway operation process is frequently challenging. This theoretical study issue has attracted a growing number of experts from around the world.

Related studies have shown that the model description methods for the high-speed train operation process can be generally characterized as mechanism modeling and data-driven modeling. There are two types of mechanism modeling: single-mass point modeling and multi-mass point modeling. Lian et al. (2020) constructed the state-space equations of high-speed trains based on the single mass point model to describe the dynamics of high-speed trains during operation, however, it differed significantly from the actual train operation. Zhang et al. (2022) established the train dynamics equation based on the running resistance generated during actual train operation; but the force analysis was insufficient and the interaction force between carriages was ignored. Multi-mass point modeling corresponds more closely to the actual train operation procedure. Hou et al. (2019) viewed the train as a "mass point chain" comprised of multiple mass points and applied fuzzy adaptive PID control to control the train's speed. Jia et al. (2020) created a multi-mass point model and employed nonlinear predictive control to follow the rolling stock's running curve. Mo et al. (2021) utilized a unit-shift multi-mass point model, which enabled the online control operation of trains and simplified the multi-mass point model calculation. The models shown above are derived from the empirical formula and have a distinct physical meaning. Yet, the formula parameters in the modeling and solving process are typically derived by experience, and some assumptions are made. It is difficult to accurately characterize the dynamic aspects of the train operation process using mechanism modeling. Further, the researchers conducted extensive research utilizing a data-driven model to characterize the train operation process. Using traction system sensor data, Jiang et al. (2020) suggested an ideal data-driven fault detection and diagnosis (FDD) method for solving the fault diagnostic problem of dynamic traction systems. Fu et al. (2022) did a comprehensive data mining and analysis of the motor system's data features. A combination method of improved principal component analysis and deep learning is proposed to handle the problem of identifying tiny faults in traction systems, hence improving train operating safety. Wang et al. (2022) changed the dynamic model of high-speed trains into a partial format dynamic linearized data model and suggested a model-free adaptive fault-tolerant control (PFDL-MFAFTC) algorithm to enhance the high-speed trains' adaptive fault-tolerant control capacity. Data-driven modeling solves the problems of mechanism modeling to a considerable extent, however the model has no physical significance, low interpretability and generalization ability, and stringent data source requirements.

In conclusion, the majority of the present dynamic models constructed around the train operating process have constraints such as low model accuracy, complex modeling, and sluggish solution speed, and it is imperative to search out new modeling techniques to solve the problem. In several domains, multi-model fusion modeling has become prevalent in response to the research trend of high modeling accuracy (Zhang et al,. 2021; Anifowose et al,. 2017; Kim et al,. 2021). The hybrid modeling method proposed in this paper is based on the fundamental design concept of the multi-model fusion modeling method, combining train dynamics theory with data driven by machine learning, utilizing the complementary advantages between single models to establish a fusion model, taking into account the physical significance of the mechanism model and data-driven data information, and then improving the modeling accuracy of high-speed trains. In addition, PID control is a frequently used classical control method, however it is better suited for linear systems (Zhang et al,. 2018). In this paper, fuzzy control and

PID control are coupled to reduce speed control error via PID parameter correction. Together with the benefits of predictive control for complicated control processes, the predictive fuzzy PID controller is intended to increase the accuracy and comfort of train speed control.

Lastly, utilizing the actual route data and train parameters, the efficacy of the proposed hybrid modeling method and compensation algorithm is validated by comparing the accuracy of train operation curve tracking under the mechanism model and the hybrid model under the condition that the predictive Fuzzy PID control algorithm is implemented.

2. Modeling of high-speed train dynamics mechanism

The EMUs of CRH3 are chosen as the object of study. The train's fixed formation and running characteristics are investigated. Using Newton's law, the running process of a high-speed train is analyzed dynamically, and the dynamic equation of the train under the resultant force is derived.

2.1. Traction/Braking characteristics

During the operation of a high-speed train, traction or braking force is delivered to maintain or change the train's operation. The application of traction or braking force varies depending on operational conditions (traction, braking, inertia, and cruising.) The formulae for traction and braking (Hou, 2015) for the CRH3 are Equation 1 and Equation 2, respectively.

$$F = \begin{cases} 300 - 0.2857v & 0 \leqslant v \leqslant 119\ km/h \\ \dfrac{8800 \times 3.6}{v} & 119 < v \leqslant 300\ km/h \end{cases} \tag{1}$$

$$B = \begin{cases} 59.8v & v \leqslant 5\ km/h \\ 300 - 0.2851v & 5 < v \leqslant 106.7\ km/h \\ \dfrac{8800 \times 3.6}{v} & 106.7 < v \leqslant 300\ km/h \end{cases} \tag{2}$$

2.2. Basic resistance

The train is affected by the elemental resistance in any operating condition. The empirical formula (Hou, 2015) for the primary resistance of a domestic CRH3 high-speed train rolling stock is shown in Equation 3.

$$w_0 = 39.65 + 33.4289v + 0.6048v^2 \tag{3}$$

Where, w_0 is the basic train resistance, v is the train operating speed.

2.3. Additional resistance

Additional resistance is an extra part of resistance when the train runs on memorable lines (such as curves, ramps, etc.), expressed as the sum of additional curve resistance w_r, ramp additional resistance w_b, and additional tunnel resistance w_c.

By analyzing the resistance during train operation and referring to Hou's research (2015), the train resistance of high-speed trains can be shown in Equation 4.

$$w = w_0 + w_r + w_b + w_c \tag{4}$$

2.4. Kinetic equations

The differential equation as in Equation 5 is used here to describe the dynamics of the high-speed train.

$$\begin{cases} \dfrac{ds}{dt} = v \\ m\dfrac{dv}{dt} = u(\eta, v) - w(s, v) \end{cases} \tag{5}$$

Where, η indicates the operating conditions. When $\eta = 1$, it indicates that the train is in traction condition. When $\eta = 0$, it indicates that the train is in idling condition. When $\eta = -1$, it indicates that the train is in braking condition. Considering that the cruising operation state is also composed of the aforementioned three working conditions, the running working condition is set here to contain three kinds (Ding, 2021), $\eta \in \{1,0,-1\}$. The control volume u is a function of the running working condition and running speed, the train resistance w is a function of the train position s and running speed, m is the train mass, and t is the train running time.

3. Mechanism and data-driven hybrid modeling

Due to the fact that the parameters of the equations in the mechanism modeling and solution process are often derived empirically and certain assumptions are made, there are discrepancies between the mechanism model and the real process of train operation. In light of the aforementioned issues, the mechanism model is utilized as the main body, and the data-driven model based on the real-time moving window

local weighted kernel principal component regression (RTMW-LWKPCR) technique is applied to the mechanism model as an error compensator. The residual error between the mechanism model result and the measured value is considered the target data (Liu et al., 2021), and the mechanism model is compensated and corrected by the output error compensation value to increase the overall model's accuracy. Figure 1 depicts the schematic of mechanism and data-driven hybrid modeling (Li et al., 2021).

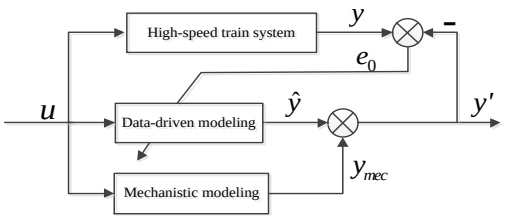

Fig. 1. Schematic diagram of mechanism and data-driven hybrid modeling

From the diagram, y is the actual outputs speed of the high-speed train; y' is the hybrid model output speed; y_{mec} is the mechanism model output speed; and $\hat{y}$ is the data-driven model compensation output speed. The relationship is as shown in Equation 6. In the data-driven model training process, the model output is made to approximately satisfy $y' = y_{mec} + \hat{y} = y$. When the error e_0 approaches zero, the training is completed, the error compensation model is obtained (Garca-matos et al., 2013), and the output speed of the high-speed train hybrid model is obtained.

$$y' = y_{mec} + \hat{y} \tag{6}$$

The input parameters of the high-speed train hybrid model are the speed of the train at the current moment, and the applied control force (tractive force/braking force) and train resistance, the control volume$u = \{v_1, F, B, w\}$, and the output parameters of the model are the speed of the high-speed train at the next moment, $y' = \{v_2\}$. The output of the data drive is the deviation of the actual high-speed train system from the speed of the mechanism model.

3.1. Local kernel principal component regression

Locally weighted kernel principal component regression (LWKPCR) is a feature learning of KPCR with the addition of correlation weights of input and output variables. The higher the correlation, the greater the impact on the predictive results (Ferrer et al., 2019; Liang, 2021).

(1) Calculate the local weighting strategy for each variable

Assume that the input space matrix is $\boldsymbol{X}_{M\times N} = [x_1, x_2, \cdots, x_n]$. Each sample point has M column vectors, which is expressed as $x_i = [x_{i1}, x_{i2}, \cdots, x_{i_m}]^T$, $x_i \in R^M$, $i = 1,2, \cdots N$. The output sample space can be represented as $\boldsymbol{Y}_{n\times 1} = [y_1, y_2, \cdots, y_n]$. There exists a mapping function $\phi : x \to \phi(x_i)$, $\phi(x_i) \in R^H$, $R^M \to R^H$, $H \gg M$. Then the covariance matrix $\boldsymbol{C}$ in the principal component analysis satisfies Equation 7.

$$\boldsymbol{C} = \frac{1}{N}\sum_{i=1}^{N}\left(\phi(x_i) - \frac{1}{N}\sum_{j=1}^{N}\phi\left(x_j\right)\right)\cdot \left(\phi(x_i) - \frac{1}{N}\sum_{j=1}^{N}\phi\left(x_j\right)\right)^T \tag{7}$$

Let $\varphi(x_i) = \phi(x_i) - \frac{1}{N}\sum_{j=1}^{N}\phi\left(x_j\right)$, so:

$$\boldsymbol{C} = \frac{1}{N}\sum_{i=1}^{N}\varphi\left(x_i\right)\left(\varphi\left(x_j\right)\right)^T \tag{8}$$

The correlation weights of the input and output variables are calculated using the Person coefficients. The equations are as in Equation 9 and Equation 10.

$$\rho_{x,y} = \frac{\sum(X-\bar{X})\,(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2\sum(Y-\bar{Y})^2}} \tag{9}$$

$$\rho_k = \frac{\left|\rho_{xy}^k\right|}{\sum_{k=1}^{m}\left|\rho_{xy}^k\right|}, k = 1,2,\cdots,m \tag{10}$$

Where, $\bar{X}$ is the mean value of input variables, $\bar{Y}$ is the mean value of output variables, $\rho_{x,y}$ is the Pearson correlation coefficient of input variables and output variables, ρ_k represents the correlation weights.

Then the local weighting strategy of each variable is Equation 11.

$$\begin{aligned} x_i^{\rho} &= x_i \cdot diag(\rho_1, \rho_2, \ldots, \rho_m) \\ &= [\rho_1 x_{i1}, \rho_2 x_{i2}, \ldots, \rho_m x_{im}]^T \end{aligned} \tag{11}$$

Where x_i is the i^{th} sample data, x_{im} is the m^{th} variable data of the i^{th} sample data.
Thus, the covariance matrix can be finally expressed as Equation 12.

$$\boldsymbol{C} = \frac{1}{N}\sum_{i=1}^{N} \varphi\left(x_i^{\rho}\right)\left(\varphi\left(x_i^{\rho}\right)\right)^{T} \tag{12}$$

(2) Calculate the data weighted projection
Let $\boldsymbol{\lambda}$ be the eigenvalue matrix of $\boldsymbol{C}$ and $\boldsymbol{A}$ be the corresponding eigenvector matrix of $\boldsymbol{C}$. That is:

$$\boldsymbol{CA} = \boldsymbol{\lambda A} \tag{13}$$

Since the mapping function $\phi(x)$ is implicit, there is no explicit form of expression, and it is impossible to find $\boldsymbol{A}$ and $\boldsymbol{\lambda}$ directly. Therefore, the Gaussian kernel function is chosen, and the kernel technique is introduced.

$$\boldsymbol{K}^{\rho}(i,j) = \left(\varphi\left(x_i^{\rho}\right)\right)^{T} \varphi\left(x_j^{\rho}\right) = e^{\frac{-\left\|x_i^{\rho}-x_j^{\rho}\right\|^2}{2\sigma^2}} \tag{14}$$

The feature vector a can be expressed by a linear combination of $\varphi\left(x_i^{\rho}\right)$ as Equation 15.

$$a = \sum_{i=1}^{N} \alpha_i^{\rho}\, \varphi\left(x_i^{\rho}\right) \tag{15}$$

Where, α_i^{ρ} is the linear combination coefficient of $\varphi\left(x_i^{\rho}\right)$, denoted as $\alpha^{\alpha} = \left[\alpha_1^{\rho}, \alpha_2^{\rho}, \ldots, \alpha_N^{\rho}\right]^{T}$. Multiplying $\varphi\left(x_i^{\rho}\right)^{T}$ left to the left and right sides of Equation 13, it can be obtained as follows:

$$\boldsymbol{\lambda}\varphi\left(x_i^{\rho}\right)^{T}\boldsymbol{A} = \varphi\left(x_i^{\rho}\right)^{T} CA \tag{16}$$

Then substituting Equation 15 into Equation 16:

$$\boldsymbol{\lambda}\left(\varphi\left(x_i^{\rho}\right)^{T}\sum_{i=1}^{N}\alpha_i^{\rho}\,\varphi\left(x_i^{\rho}\right)\right) = \varphi\left(x_i^{\rho}\right)^{T}\cdot \left(\frac{1}{N}\sum_{i=1}^{N}\varphi\left(x_i^{\rho}\right)\left(\varphi\left(x_i^{\rho}\right)\right)^{T}\sum_{i=1}^{N}\alpha_i^{\rho}\,\varphi\left(x_i^{\rho}\right)\right) \tag{17}$$

Then substituting Equation 14 for the kernel function into Equation 17, the formula was written as:

$$N\boldsymbol{\lambda}\alpha^{\rho} = \boldsymbol{K}\alpha^{\rho} \tag{18}$$

The above equation shows that α^{ρ} is the eigenvector of the kernel matrix $\boldsymbol{K}$. The kernel matrix $\boldsymbol{K}$ and the eigenvectors of the kernel matrix α^{ρ} can be calculated by the kernel function given earlier. normalizing a to be $a^{T}a = 1$, the formula can be written as:

$$\begin{aligned} a^{T}a &= \left(\alpha^{\rho}\varphi(x^{\rho})\right)^{T}\alpha^{\rho}\varphi(x^{\rho}) \\ &= (\alpha^{\rho})^{T}\boldsymbol{K}\alpha^{\rho} = N\boldsymbol{\lambda}(\alpha^{\rho})^{T}\alpha^{\rho} \end{aligned} \tag{19}$$

The weighted projection of $\varphi\left(x_i^{\rho}\right)^{T}$ is calculated as follows:

$$\begin{aligned} e_i &= \varphi\left(x_i^{\rho}\right)^{T} a = \varphi\left(x_i^{\rho}\right)^{T}\sum_{j=1}^{N}\alpha_j^{\rho}\,\varphi\left(x_j^{\rho}\right) \\ &= \sum_{j=1}^{N}\alpha_j^{\rho}\,\varphi\left(x_j^{\rho}\right)^{T}\varphi\left(x_j^{\rho}\right) \\ &= \sum_{j=1}^{N}\alpha_j^{\rho}\,\boldsymbol{K}\left(x_i, x_j\right) = \boldsymbol{K}(x_i,:)\alpha^{\rho} \end{aligned} \tag{20}$$

(3) Calculating the kernel principal component matrix
Firstly, the eigenvalue matrix λ is arranged in descending order, and the first τ principal components are used to describe the original data sample. Secondly, the least squares regression is performed on the descending data to predict the velocity, and the eigenvalues are the number of principal components. The corresponding data are the first τ columns of λ, denoted as $Q_{\tau} = [q_1, q_2, \cdots, q_{\tau}]$. The eigenvectors corresponding to the first τ eigenvalues Q_{τ} are denoted as $\alpha_{\tau}^{p} = \left[\alpha_1^{p}, \alpha_2^{p}, \cdots, \alpha_{\tau}^{p}\right]$. The equation for projecting the data set into a high-dimensional space is Equation 21.

$$\boldsymbol{z} = \boldsymbol{K}\alpha_m^{\rho} \tag{21}$$

Where **z** denotes the kernel principal component matrix.
(4) Calculating the regression coefficient matrix
Multiple linear regression is performed on the projected data, and the regression coefficient matrix is calculated as Equation 22.

$$\theta = \left(\boldsymbol{z}^{T}\boldsymbol{z}\right)^{-1}\boldsymbol{z}^{T}\boldsymbol{Y} \tag{22}$$

The output sample space $\boldsymbol{Y}$ will need to be normalized. Thus, the output of the data sample, which is the value of the velocity error compensation, is Equation 23.

$$\hat{y} = \boldsymbol{z}\theta + \bar{y} \tag{23}$$

Where $\bar{y}$ is the mean value of the output variable.

3.2. Immediate moving window

Due to the complicated and ever-changing external environment of the actual train operating process, ongoing usage of the historical model would inevitably lower the accuracy of the model's calculation outputs as a result of the constant introduction of new working conditions. To address this issue, the real-time moving window technique was proposed to respond to actual working conditions by continuously updating the error compensation model to ensure real-time model updating.

Assuming that the data samples in the i^{th} moment window are$\{X_{wi}, Y_{wi}\}$. Firstly, the model is trained on the sample data in the current window, and the mean μ and standard deviation$\sigma$are calculated. Secondly, the model is updated or not according to the changes in the mean and standard deviation of the following sample data compared with the samples in the current window. When the mean and standard deviation of the following sample data do not change much, the previous training model is extended for predictive, and the model is not updated. When the change is significant, the window is shifted, and the locally weighted kernel principal component regression model is updated (Liang, 2021). The algorithm flow and working process are shown in Figure 2 and Figure 3, respectively.

3.3. RTMW-LWKPCR algorithm

The main steps of the RTMW-LWKPCR algorithm are shown in Table 1.

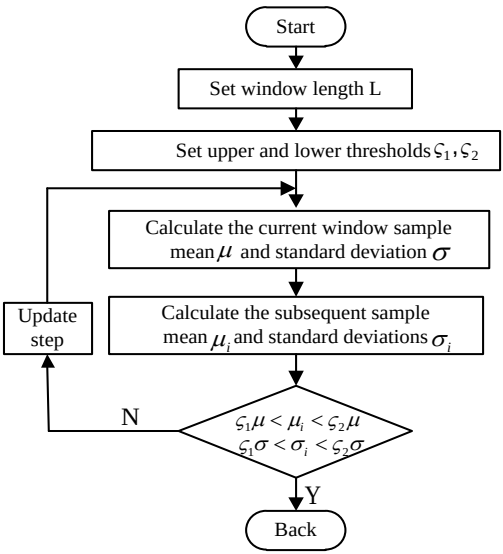

Fig. 2. Algorithm flow of real-time moving window

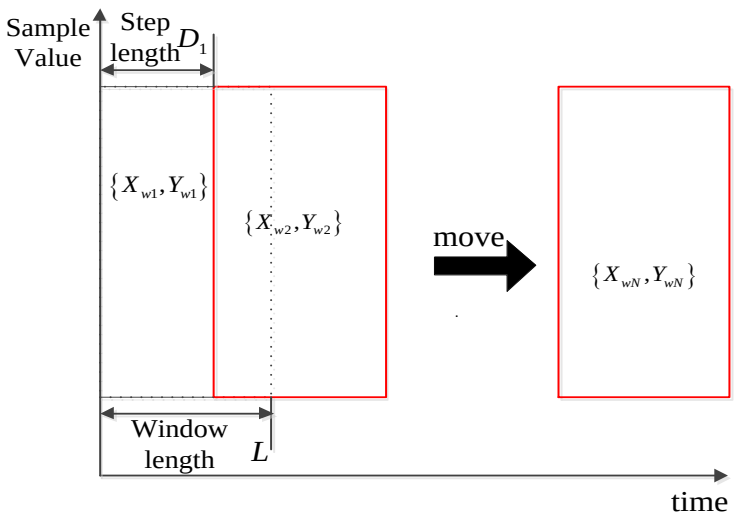

Fig. 3. Vehicle speed profile obtained in the RDE test

Table 1. RTMW-LWKPCR algorithm steps

Algorithm	**RTMW-LWKPCR**
Input	Sample matrix$X = [x_1, x_2, \cdots, x_n]^T$, window length$L$ (number of model samples), upper step sizeD_{max}, scale factor ς_1 and ς_2.
Output	Predict the speed error compensation value$\hat{y}$

Process	(1) Calculation weighting strategy: use Equation 9-11 to calculate the weighting of each variable;
	(2) Take out Lsamples, calculate their meanμ and standard deviationσ;
	(3) Computational data-weighted projection: use Equation 20 to project data into a high-dimensional space to make it a linear separable data;
	(4) Calculate the kernel principal component matrix: obtained using Equation 21;
	(5) Calculate the regression coefficient matrix: obtained using Equation 22;
	(6) Calculate the predicted value: use Equation 23 to calculate the predicted value;
	(7) Take out a new sample$i$
	While $L + i < D_{max}$
	If $\varsigma_1\mu < \mu_i < \varsigma_2\mu$ and $\varsigma_1\sigma < \sigma_i < \varsigma_2\sigma$
	Go back to step 6;
	Else
	Go back to step 2;
	End while

4. Predictive Fuzzy PID control algorithm

To verify the efficacy of the hybrid modeling and compensating algorithm, the Fuzzy PID algorithm is chosen to construct the controller. The deviation and deviation transformation rate of the optimal input value and the given input value obtained by the predictive controller are used as the input of the fuzzy PID controller, after which the three parameters of the PID controller are adjusted online in accordance with the fuzzy control principle, and the output is then applied to the controlled object in order to determine the input value of the predictive controller. Figure 4 demonstrates its structure: y_r is the target speed; y_q is the predicted speed; y_p is the feedback corrected speed; y is the actual outputs speed.

The method chosen for the simulations is the dynamic matrix control algorithm in this paper. After applying a control increment Δu(k) to the high-speed train at moment k, the P predicted speed values for the future moment (Hou et al., 2020) are given by:

$$y_{PQ}(k) = y_0(k) + \alpha\Delta u_Q(k) \tag{24}$$

Where, $y_{PQ}(k)$ is the predicted speed at P future moments, $y_0(k)$ is the initial predictive value of the system, α is the predictive model vector of the high-speed train for the step response, $\Delta u_Q(k)$ is Q continuous control increments.

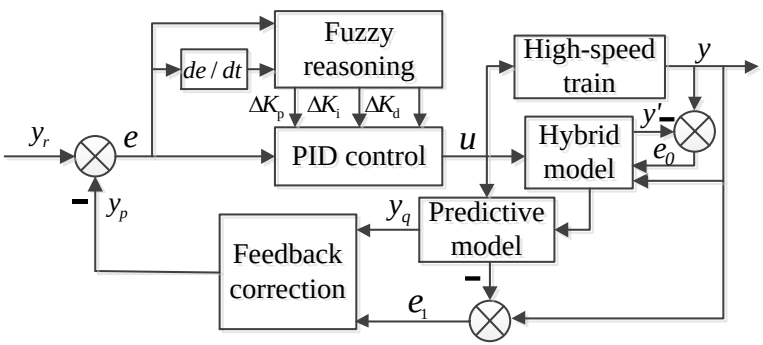

Fig.4. Diagram of predictive Fuzzy PID control system

The predictive error $e_1$is the difference between the predicted speed y_q and the train's operating speed $y(k+1)$ at moment k+1.

$$e_1(k+1) = y(k+1) - y_q(k+1,k) \tag{25}$$

Then the corrected predicted velocity is shown in Equation 26.

$$y_p(k+1) = y_q(k+1) + he_1(k+1) \tag{26}$$

Where, h is the correction vector.

In the Fuzzy PID control algorithm, the Fuzzy subset defines the train speed tracking error and error change rate as {NB(negative large), NM(negative medium), NS (negative small), ZE(zero), PS(positive small), PM (positive medium), PB(positive large)}. The Gaussian distribution type is chosen for the affiliation function of the input quantity, and the triangular type with higher sensitivity is used for the affiliation function of the output quantity. The fuzzy reasoning adopts Mamdani inference, and anti-fuzzification adopts the center of gravity method.

The specific Fuzzy rules are shown in Table 2. The parameters are corrected as shown: $K_P' = K_P + \Delta K_P$, $K_i' = K_i + \Delta K_i$, $K_d^{'} = K_d + \Delta K_d$.

5. System simulation and result analysis

The actual line data (Hou, 2015) from Huashan North Station to Xi'an North Station of Zhengxi High-speed Railway is utilized as the baseline data to validate the effectiveness of the hybrid modeling and algorithm suggested in this paper. The simulation tests are conducted on a Windows 10 system with MATLAB R2019b as the simulation environment. Certain train parameters and line data are displayed in Tables 3 and 4, respectively.

Figure 5 shows the target speed profile of the CRH3 train obtained from a one-way run on this simulated test line, with a one-way run time of 1680 s.

During the simulation, the parameters of the RTMW-LWKPCR algorithm for the hybrid model were set as follows: the window length was set to 200 s/step, the kernel function width parameter was 3, the scale factor parameters $\varsigma_1 = 0.8$, $\varsigma_2 = 1.2$, the maximum step size D_{max}s/step. The $\pm 5\%$ speed error of the actual line data is used as the data sample. The speed error compensation values of the data-driven model based on the normalized RTMW-LWKPCR algorithm are shown in Figure 6.

In the predictive Fuzzy PID (PF PID) control, the theoretical domains of the train speed tracking error e and the error rate e_c of change are [-1, 1], [-6, 6], respectively; the theoretical domains of the control quantities are [-3, 3]; the quantization factors$K_e = 3$, $K_{ec} = 0.5, K_u = 1$; the initial parameter$K_p = 1.2$, $K_i = 0.5$, $K_d = 1$; the correction factor $\Delta K_p = 2.5$, $\Delta K_i = 3$, $\Delta K_d = 3$. The step response in the DMC control basically reaches stability at about 30 s. The values of the parameters of the model are taken as the sampling period $T = 0.5$s, the modeling time domain $N = 50$, the control time domain $Q = 2$, and the optimization time domain $P = 16$, respectively (Yang et al., 2022).

The velocity profile tracking under the same predictive Fuzzy PID controller control is executed independently for the mechanism and hybrid models, and the simulation results are depicted in Figure 7. Local enlargement is depicted in Figures 8 and 9. As shown in Figure 7, the speed curve under the mixed model corresponds with the target velocity curve of the solid line, and the model and algorithm achieve the tracking control of the actual speed-time curve to the target speed-time curve with a satisfactory tracking effect. Figure 10 shows the comparison of speed control errors under the two models.

Table 2. The Fuzzy control rules of ΔK_P,ΔK_i,ΔK_d

e	e_c						
	NB	NM	NS	ZE	PS	PM	PB
NB	PS/PB/NB	NS/PB/NB	NM/PB/NM	NB/PM/NM	NM/PM/NS	PM/PS/ZE	PS/ZE/ZE
NM	PS/PB/NB	NS/PB/NM	NS/PB/NM	NM/PM/NS	NS/PM/NS	NM/ZE/ZE	ZE/ZE/ZE
NS	ZE/PB/NM	NS/PM/NM	ZE/PM/NS	NS/ZE/NS	NS/ZE/ZE	NS/ZE/PS	ZE/PS/PS
ZE	ZE/PM/NM	ZE/PS/NS	ZE/PS/PS	ZE/ZE/NS	ZE/ZE/ZE	ZE/NS/PS	ZE/NS/PS
PS	ZE/PS/NS	PS/ZE/ZE	PS /ZE/ZE	PS/NS/PS	ZE/NS/PS	PS/NM/PM	PM/NM/PM
PM	ZE/ZE/ZE	PS/ZE /ZE	PS/NS/PS	PM/NM/PM	PS/NM/PM	PS/NM/PM	PB/NB/PB
PB	ZE/ZE/ZE	PS /NS/ZE	PM/NM/PM	PB/NM/PM	PS/NB/PB	PM/NB/PB	PB/NB/PB

Table 3. Basic train parameters

Train parameters	Parameter value
Train length (m)	200.67
Crew mass (t)	536
Maximum traction power (kW)	8800
Maximum braking power (kW)	8000
Maximum operating speed (km/h)	350

Table 4. Line data

Line parameters	Parameter value
Total length of the line (km)	121.0612
Maximum speed limit (km/h)	310
Intermediate speed limit (km/h)	200
Running time (Second)	1680

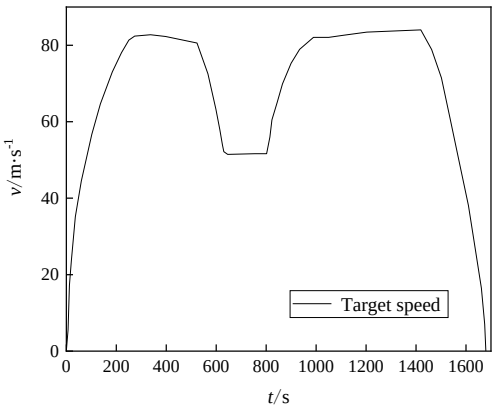

Fig. 5. Diagram of target speed curve

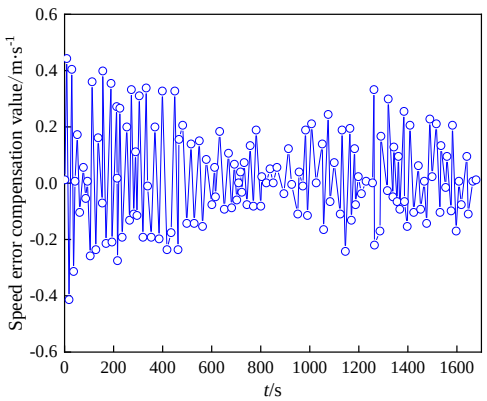

Fig. 6. Error compensation diagram of data-driven model

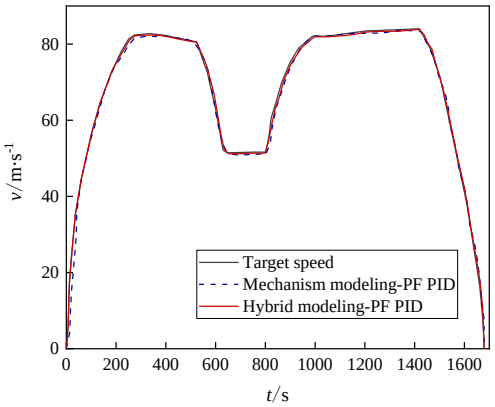

Fig. 7. Comparison of speed control curves under two models

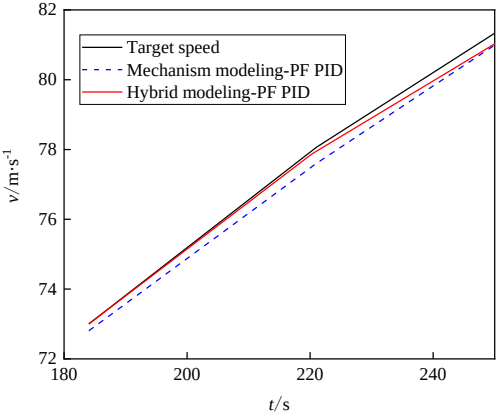

Fig. 8. Speed control curves of local amplification 1

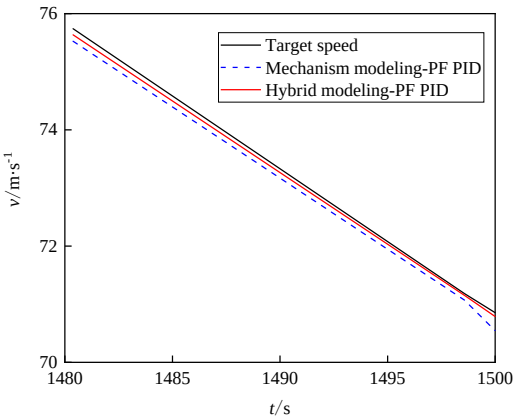

Fig. 9. Speed control curves of local amplification 2

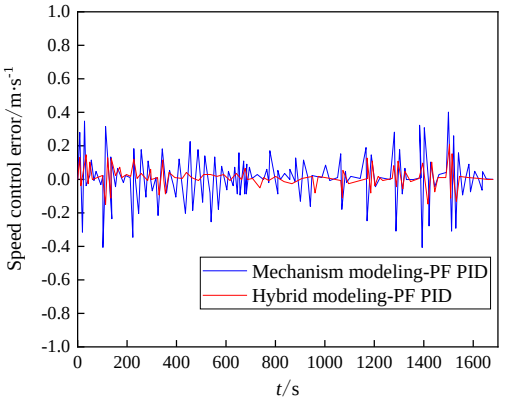

Fig. 10. Comparison of speed control error under two models

Analyzing the simulation results, the error range of high-speed train speed control based on the single mechanism model in Fig.10 is ±0.45m/s, and the maximum error of speed control is 0.4373 m/s; however, the error range of high-speed train speed control based on the hybrid model is ±0.15m/s, the maximum error of speed control is 0.2328 m/s.

6. Conclusion

On the basis of the fundamental concept of multi-model fusion, a hybrid model of high-speed train mechanism and data-drive is created. The simulation verification and result comparison demonstrate that, compared to the single mechanism model, the maximum control error of the speed control based on the hybrid model is reduced by 46.76 %, and the average speed control error is reduced by 69.42 %, which proves the hybrid model and compensation algorithm's effectiveness. In addition, the real-time window ensures that the model is updated in real-time, hence enhancing the model's precision. In conclusion, the suggested model and compensation algorithm significantly minimize speed control error and raise the speed control precision of high-speed trains by 13.75 %. The simulation results demonstrate that the hybrid model and compensation algorithm suggested in this paper are successful. The innovative points of this study are:

(1) A hybrid modeling method based on mechanism and data-driven is proposed, which increases modeling accuracy and speed control effect and overcomes the problem of single modeling's low accuracy.

(2) Due to the real-time nature of high-speed train operation, the real-time moving window is designed to continually track and update the model's state in real-time, thereby enhancing the model's accuracy.

In this paper, a portion of the modeling of high-speed trains is examined, and phased research results are presented. Yet, there are still issues that require additional research. In the next phase of the research, the alterations in internal forces and running lines (ramps, curves, etc.) in the workshop will be analyzed in order to improve the design of the model error compensation algorithm and optimization control method, thereby advancing the train speed control objective.

Acknowledgements

This work was supported by Natural Science Foundation of Gansu Province(Grant No. 21JR7RA321, Grant No. 22JR5RA358).

References

[1] Karolak, J. (2021). Interface and connection model in the railway traffic control system. *Archives of Transport*, 58(2), 137-147.

[2] Tan, C., Li, Y. Q. (2022). Adaptive braking control for high-speed trains with input time delays. *Journal of Railway Science and Engineering*, 19(04), 1071-1080.

[3] Lian, W. B., Liu, B. H., Li, W. W., et al. (2020). Automatic operation speed control of high-speed train based on ADRC. *Journal of the China Railway Society*, 42(01), 76-81.

[4] Zhang, W. J., Cao, B.W., Liu, Y. F., et al. (2022). Operation control method for medium speed maglev trains based on fractional order sliding mode adaptive neural network. *China Railway Science*, 43(2), 152-160.

[5] Hou, T, Guo, Y. Y., Niu, H. X. (2019). Research on speed control of high-speed train based on multi-point model. *Archives of Transport*, 50(02), 35-46.

[6] Jia, C., Xu, H. Z., Wang, L. S. (2020). Nonlinear model predictive control for automatic train operation based on multi-point model. *Journal of Jilin University(Engineering and Technology Edition)*, 50(05), 1913-1922.

[7] Mo, X. T., Wang, X. Q., Liang, X. R., et al. (2021). Speed tracking control for high-speed trains with Fuzzy RBF neural network. *Modern Computer*, 27(26), 1-7+14.

[8] Jiang, B., Chen, H. T., Yi, H., et al. (2020). Data-driven fault diagnosis for dynamic traction systems in high-speed trains. *Scientia Sinica*(*Informationis*), 50(04), 496-510.

[9] Fu, C. X., Zhao, T. (2022). Fault identification based on data-driven method for traction control systems in high-speedtrain. *Microcomputer Application*, 38 (10), 138-141.

[10] Wang, H., Liu, G. F., Hou, Z. S. (2022). Data-driven model-free adaptive fault tolerant control for high-speed trains. *Control and Decision*, 37(5), 1127-1136.

[11] Zhang, M. X., Liu, H. C., Wang, M., et al. (2021). Intelligence hybrid modeling method and applications in chemical process. *Chemical Industry and Engineering Progress*, 40(4), 1765-1776.

[12] Anifowose, F. A., Labadin, J., Abdulraheem, A. (2017). Hybrid intelligent systems in petroleum reservoir characterization and modeling: the journey so far and the challenges ahead. *Journal of Petroleum Exploration and Production Technology*, 7(1), 251-263.

[13] Kim, J., Chen, M., Han, J. J., et al. (2021). The development of leak detection model in subsea gas pipeline using machine learning. *Journal of Natural Gas Science and Engineering*, 94, 104134.

[14] Hou, T. (2015). Speed control study of multi-mode intelligent control based on multi-information fusion and filter on high-speed train. Lanzhou: Lanzhou Jiaotong University.

[15] Ding, P. (2021). A multiple point-mass podel based high-speed train adaptive speed tracking control scheme. Nanchang: East China Jiaotong University.

[16] Liu, G., Yuan, Z. Y., Chen, L., et al. (2021). Preliminary study on application of hybrid modeling method in oil and gas pipeline networks. *Oil & Gas Storage and Transportation*, 40(09), 980-990.

[17] Li, J. X., Zhou, D. J., Xiao, W., et al. (2019). Hybrid Modeling of Gas Turbine based on Neural Network. *Journal of Engineering for Thermal Energy and Power*, 34(12), 33-39.

[18] Garca-matos, J. A., Sanz-bobi, M. A., Sola, A. (2013). Hybrid Model-based Fault Detection and Diagnosis for the Axial Flow Compressor of a Combined-cycle Power Plant. *Journal of Engineering for Gas Turbines and Power*, 135(5), 054501.

[19] Ferrer, S., Mezquita, A., Aguilella, V. M., et al. (2019). Beyond the energy balance: Exergy analysis of an industrial roller kiln fring porcelain tiles. *Applied Thermal Engineering*, (150), 1002-1015.

[20] Liang, Y. Y. (2021). Research on predictive of firing zone temperature in roller kiln based on mechanism and data hybrid driven. Guangzhou: Guangdong University of Technology.

[21] Hou, T., Guo, Y.Y., Chen, Y., Yang, H.K. (2020). Study on speed control of high-speed train based on multi-point model. *Journal of Railway Science and Engineering*, 17(02), 314-325.

[22] Yang, H. K., Hou, T., Chen Y. (2022). Research on optimal control of high speed trains based on predictive Fuzzy PID control algorithm. *Railway Transport and Economy*, 44(8),130-136.