

# Convolutional neural networks for accurate measurement of train speed

Haitao Tian, Argyrios Zolotas  and Miguel Arana-Catania 

Proc IMechE Part F:  
J Rail and Rapid Transit  
2025, Vol. 239(10) 890–906  
© IMechE 2025



Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/09544097251367794  
[journals.sagepub.com/home/pif](https://journals.sagepub.com/home/pif)



## Abstract

In this study, we explore the use of Convolutional Neural Networks for improving train speed estimation accuracy, addressing the complex challenges of modern railway systems. We investigate three CNN architectures — single-branch 2D, single-branch 1D, and multiple-branch models — and compare them with the Adaptive Kalman Filter. We analyse their performance using simulated train operation datasets with and without Wheel Slide Protection activation. Our results reveal that CNN-based approaches, especially the multiple-branch model, demonstrate superior accuracy and robustness compared to traditional methods, particularly under challenging operational conditions. These findings highlight the potential of deep learning techniques to enhance railway safety and operational efficiency by more effectively capturing intricate patterns in complex transportation datasets.

## Keywords

Convolutional Neural Networks (CNNs), train speed prediction, signal data analysis, simulated datasets, machine learning in transportation, train control systems optimisation

Date received: 10 February 2025; accepted: 17 July 2025

## Introduction

Accurately estimating train speed is a critical aspect of modern railway systems, with far-reaching implications for safety, efficiency, and overall performance. Accurate speed data is essential for the Automatic Train Protection (ATP) system to calculate safe braking distances, enforce speed limits, prevent collisions, and activate emergency braking when necessary. Even minor errors in speed estimation could lead to catastrophic accidents, especially at high speeds.<sup>1</sup>

Beyond safety, accurate speed control plays a crucial role in energy efficiency, allowing trains to operate at optimal speeds and minimise unnecessary acceleration and braking.<sup>2,3</sup> This not only reduces energy consumption but also contributes to environmental sustainability in the transportation sector. Precise speed information is equally vital for maintaining schedules and ensuring punctuality, facilitating better coordination between trains, and optimising overall network efficiency.<sup>4</sup> Passenger comfort is another beneficiary of accurate speed control, as smooth acceleration and deceleration contribute to a pleasant riding experience.<sup>5</sup> Furthermore, infrastructure management is highly dependent on accurate speed data to plan maintenance activities and assess track conditions, as wear patterns are closely linked to train speeds.<sup>6</sup> Lastly, railways must adhere to strict speed limits and regulations, making accurate speed measurement necessary for demonstrating compliance and avoiding penalties.<sup>7</sup>

Despite its importance, train speed estimation faces numerous challenges that complicate the quest for accuracy.

The traditional method of measuring train speed through wheel revolutions is particularly problematic due to wheel wear and rolling contact fatigue.<sup>8,9</sup> As wheels decrease in diameter over time due to wear or reprofiling, speed calculations become increasingly inaccurate without frequent recalibration. This issue necessitates regular maintenance stops, introduces gradual drift in accuracy between calibrations, and can be affected by variations in wear between different wheels. Environmental factors pose another significant challenge, as train speed sensors are “susceptible to diverse operating conditions such as snow, rain, fog, tunnel, hilly region, slip, slide, etc.”<sup>10</sup> These conditions can interfere with various sensing technologies, impacting the reliability of speed measurements across different terrains and weather conditions.

The phenomena of wheel slip and slide during acceleration or braking, particularly in adverse weather, can lead to substantial errors in wheel-based speed measurements.<sup>11,12</sup> This challenge is compounded by the complexity of integrating data from multiple sensors, each with its own sampling rates and noise characteristics.<sup>13</sup> Real-time processing requirements add another layer of difficulty, as high-speed trains demand rapid, instantaneous speed estimation for safe

---

Faculty of Engineering and Applied Sciences, Cranfield University, Cranfield, UK

### Corresponding author:

Miguel Arana-Catania, Faculty of Engineering and Applied Sciences, Cranfield University, College Rd, Cranfield MK43 0AL, UK.  
Email: [humd0244@ox.ac.uk](mailto:humd0244@ox.ac.uk)

operation.<sup>14</sup> This necessitates efficient algorithms and processing capabilities to handle data from multiple sensors with minimal latency. Cost constraints also play a role, as highly accurate speed measurement systems may be prohibitively expensive for widespread deployment across entire rail networks. Moreover, any speed estimation system must be extraordinarily reliable and robust, capable of functioning consistently in the harsh railway environment, which includes vibrations, electromagnetic interference, and temperature variations.<sup>15</sup>

As railway systems continue to evolve, with trends towards higher speeds and increased automation, the importance of precise speed measurement will only grow. The integration of multiple sensor types, coupled with sophisticated algorithms and potentially machine learning (ML) approaches, promises to provide more robust and accurate speed estimates. Continued research and development in this area, leveraging advances in sensor technology, data processing, and artificial intelligence, will be crucial for ensuring the safety, efficiency, and performance of future rail transport systems. By addressing the challenges of accurate speed estimation, the railway industry can move towards more reliable, cost-effective, and technologically advanced solutions that support the ongoing evolution of train operations and safety systems.

A particularly relevant type of machine learning architecture is the Convolutional Neural Networks.<sup>16–19</sup> These are a class of neural networks particularly well-suited for processing data with spatial structures. In convolutional layers, neurons are connected only to a local region of the input data, known as the receptive field, which helps capture local features of the input. This local processing of the input significantly reduces the number of parameters in the network and lowers the model's complexity as compared with a traditional fully-connected neural network. Each convolutional kernel acts as a feature extractor, learning appropriate weights through training to detect specific features in the input. Multiple convolutional kernels can operate simultaneously, generating multiple feature maps, with each map capturing different features from the original input, which are then combined to produce the network output.

In this article we develop this latter proposal, applying machine learning approaches to train speed estimation. Currently, accurately predicting train speed predominantly relies on traditional sensor fusion methods, and the application of machine learning in this area remains relatively limited. This work develops a robust and adaptable methodology designed to enhance the accuracy and reliability of speed measurements under various operational conditions, including training data from challenging environments with low adhesion. Another key contribution is the exploration and rigorous evaluation of different Convolutional Neural Network (CNN) architectures, leading to the identification of the optimal model for speed estimation. The comparative analysis between CNN-based models and the Adaptive Kalman Filter (AKF), highlights the potential for machine learning to outperform conventional approaches. Ultimately, this research provides insights that could significantly improve train safety, efficiency, and operational performance across diverse railway environments.

To summarise, the contributions of this paper to the estimation of train speeds include:

- Developing a robust machine learning algorithm for accurate speed measurement under various conditions.
- Exploring and evaluating multiple CNN architectures.
- Demonstrating machine learning's potential to outperform traditional approaches like the Adaptive Kalman Filter.

### Related work

Multi-sensor fusion refers to the integration of data from multiple sensors to enhance the accuracy, reliability, and robustness of the system's performance. This approach is particularly beneficial in applications where single sensors may be susceptible to noise, errors, or limitations in measurement range. In the context of speed measurement, multi-sensor fusion leverages data from various sources, such as GPS, wheel speed sensors, and inertial measurement units (IMUs), to provide a more accurate and comprehensive estimation of the vehicle's speed. By combining these different data streams, the system can compensate for the weaknesses of individual sensors, such as the susceptibility of GPS to signal loss in urban canyons or the wheel speed sensor's sensitivity to skidding.<sup>20,21</sup> Techniques like Kalman filtering and particle filtering are commonly employed to fuse these sensor data effectively, producing a unified output that is more reliable than any single sensor's data. This methodology enhances the precision of speed estimation, making it invaluable for applications in autonomous driving, traffic management, and safety-critical systems where accurate speed information is essential.

*Sensor fusion for speed estimation.* Studies on speed estimation and vehicle localisation present a variety of methods and technologies, each with its unique contributions and drawbacks. A common thread across these studies is the utilisation of Kalman Filter techniques to enhance the accuracy of speed and position measurements, although the specific implementations and sensor integrations differ.<sup>22–24</sup>

Deng Z et al.<sup>25</sup> introduced the Covariance Intersection algorithm to integrate data from multiple sensors, including GPS, Inertial Navigation Systems (INS), and speed sensors. The AKF was used to adjust the measurement noise covariance matrix, thereby mitigating the effects of uncertain environmental factors. In the event of a GPS outage, the train's speed and position are estimated using the Velocity Sensor/INS scheme. Similarly, Ernest P et al.<sup>26</sup> tackled the railway vehicle localisation problem by fusing data from odometers and accelerometers using a Kalman Filter. However, the long-term accuracy of his method was heavily dependent on the precision of the odometer and correct detection of wheel slip, illustrating a common challenge in relying on wheel-based sensors.

In contrast, Tanelli M et al.<sup>27</sup> presented a state estimator algorithm that solely used data from hall-effect wheel encoders and a single-axis accelerometer. This method, notable for its low computational cost, could be implemented in commercial vehicle control units and was extensively tested under various driving conditions. Mei TX and Li H<sup>28</sup>

adopted a different approach by using inertial sensors on the railway vehicle bogie frame to measure dynamic responses and extract wheelset motion features. This method avoided issues related to wheel slip or slide but showed significant errors at low speeds, indicating a trade-off between sensor type and measurement accuracy.

Hu D et al.<sup>29</sup> introduced an extended Kalman Filter algorithm that fused data from multiple sensors, including accelerometers, gyroscopes, and wheel speed sensors, to estimate vehicle speed and road adhesion characteristics. Their preliminary results were promising, though largely based on simulations, underscoring the need for extensive field testing. Chu L et al.<sup>30</sup> combined fuzzy logic with Kalman Filters to estimate vehicle longitudinal speed, achieving an error margin within 4%. This hybrid approach demonstrated the potential benefits of integrating multiple estimation techniques.

Malvezzi M et al.<sup>31</sup> used tachometers and inertial measurement units (IMU) to improve speed and distance estimation on the Italian railway network, particularly under critical adhesion conditions. Their Kalman Filter-based method showed significant improvements over the existing algorithm in the Italian network, though it struggled with slow initial acceleration. This study emphasised the importance of algorithmic refinement to handle varying operational conditions.

Kim K et al.<sup>32</sup> proposed a two-stage federated Kalman Filter for high-speed train navigation systems, incorporating multiple sensors and an adaptive information-sharing algorithm to mitigate severe performance degradation from sensor errors. This approach highlighted the benefits of federated filters in managing sensor discrepancies and enhancing system robustness. Finally, Muniandi G and Deenadayalan E<sup>10</sup> introduced a probabilistic weighted fusion algorithm based on a nonlinear longitudinal train dynamic model. This method combined state estimates from distributed and sensor-specific extended Kalman Filters, using measurements from a diverse set of sensors to improve distance and speed estimations.

While most studies leveraged the Kalman Filter or its variants to enhance speed estimation accuracy, their approaches varied significantly in sensor integration and application contexts. The common challenge of sensor dependency and error management underscores the need for continued innovation and hybrid methodologies to achieve robust and reliable speed measurements across diverse railway and vehicular systems.

**Machine learning for speed estimation.** Recent advancements in speed estimation methods have seen a shift towards leveraging machine learning algorithms alongside conventional sensor fusion techniques. ML approaches, such as regression models, decision trees, random forests, neural networks, and deep learning models, are gaining popularity due to their ability to predict complex nonlinear relationships from extensive datasets without requiring explicit modelling of object dynamics.<sup>33–35</sup> Unlike traditional sensor fusion methods such as the Kalman Filter which rely on predefined models and assumptions about object behaviour, ML methods offer a more flexible framework. They can integrate data from various sensors installed on

moving objects, making them suitable for environments with nonlinear and non-Gaussian noise characteristics.<sup>36,37</sup>

Supervised machine learning, a subset of ML, dominates speed estimation tasks where known speed serves as labelled data. Models are trained to map sensor inputs (e.g., accelerometer readings, positional data) to speed labels, enabling accurate predictions in diverse scenarios. This approach contrasts with unsupervised machine learning, which discovers patterns without labelled data, and reinforcement learning, which optimises behaviour through interaction with the environment. Shafique MA and Hato E<sup>38</sup> explored SVM, AdaBoost, decision trees, and random forests to classify transportation modes, such as walking, bicycle, train or car, using accelerometer data, emphasising the impact of preprocessing and training data generation on prediction accuracy.

In another study, Wang KL and Xu J<sup>39</sup> proposed a deep CNN model for speed regression based on accelerometer data from wearable devices, achieving significant error reduction under different conditions. This study demonstrated the feasibility of deploying deep learning models on resource-constrained devices, reducing reliance on cloud-based processing. Moreover, advancements in neural network architectures have been pivotal. Hannink J et al.<sup>40</sup> employed deep CNNs to predict stride length in elderly patients, showcasing the efficacy of deep learning in extracting meaningful features from inertial sensor data. Similarly, Karlsson R and Hendeby G<sup>41</sup> introduced a CNN model using accelerometer and gyroscope data, outperforming traditional frequency analysis methods in automotive speed estimation tasks. These studies underline the superior performance of deep learning models in handling rapid speed changes and varying data rates, crucial for real-time applications.

Several works<sup>42,43</sup> further refined CNN architectures for speed detection during walking and running, leveraging dual-signal-channel models combining accelerometer and gyroscope inputs. Their approach surpassed previous benchmarks, illustrating that integrating sensor inputs at different model layers enhances accuracy and reduces model size. This contrasts with earlier works like Wang's, which solely relied on accelerometer data, thereby highlighting the evolution towards more sophisticated sensor fusion techniques within ML frameworks.

While traditional sensor fusion methods remain foundational in speed estimation, the integration of ML techniques has revolutionised the field by offering robustness to complex environments and superior predictive capabilities. ML's adaptability to diverse sensor inputs and its ability to learn from large datasets without explicit modelling of dynamics are key advantages. However, challenges such as parameter tuning and computational resources remain, underscoring the ongoing need for optimising ML models for real-world deployment.

## Dataset preparation

Obtaining comprehensive datasets of train speeds presents significant challenges due to the commercially sensitive nature of such information. To develop an accurate algorithm for train status estimation, it is crucial to acquire not

only precise train and wheel speeds but also GPS data and other sensor signals.

The datasets utilised in this article are extracted from the report “Dependable Speed Measurement for Improved Low Adhesion Braking”<sup>44</sup> using image-to-data conversion software. The report implements 19 simulations, of which 17 are suitable for this study, as they have signals from both the wheel and the GPS. Specifically, 13 simulations are conducted without Wheel Slide Protection (WSP) operation, while 4 include the WSP operation. Two simulations are excluded due to the absence of GPS speed data.

Figure 1(a) illustrates an example of the test figures presented in the report. The comparisons for each test are made between train speed, provided by the ground sensing radar; GPS speed, obtained from the INS input via GNSS antennas; wheel speed, inputted into the INS; and estimated speed, outputted from the INS. In this test, no odometer error is introduced and no WSP operation is simulated, resulting in the INS’s estimated speed being quite accurate, with only minor errors during the acceleration phase and at the highest speed. Thus, the prediction task is relatively straightforward in this scenario.

Another example, shown in Figure 1(b), demonstrates that during the acceleration phase, the estimated speed aligns with the train’s true speed. During the coasting phase, the estimated speed is slightly higher than the true train speed, with the GPS speed being approximately 1 m/s higher. At the beginning of the braking phase (190 s), a 20% error is introduced to the odometer input of the INS (red line). At this point, the estimated speed matches the train’s true speed for 20 seconds before it starts to drift towards the odometer speed. This example indicates that the prediction task can potentially fail if the odometer reading contains significant error or noise.

Figure 1(c) presents a test example where the WSP operation is simulated. During the braking phase, the train wheels exhibit behaviour similar to the ABS system in cars, continuously braking and releasing, causing significant oscillations in the wheel speed readings. In such situations, the train driver cannot accurately determine the train’s operational state by reading the wheel speed, which can be very unsafe under certain circumstances. The figure also shows that the INS’s speed estimation, which integrates wheel speed and GPS signals, displays some deviation. Accurately estimating train speed under these conditions becomes challenging, which is the primary issue addressed in this paper.

Four data channels are captured from the figures: time, wheel speed, GPS speed, and train speed, with the latter serving as the ground truth. Data from two figures are selected as test data - one without the WSP operation and one with the WSP operation. The remaining data are allocated for training and validation purposes. In total, the dataset comprises 5205 time stamps for training and validation, and 947 for testing.

The algorithm’s input is designed to incorporate not only the current speed information but also a specified number of historical speed data points. Assuming that the sampling frequency of the speed data is 1, then the input at a certain time step  $t$  goes from  $S_{t-n+1}$  to  $S_t$ , where  $S$  includes time, wheel speed, and GPS speed and  $n$  is the history length used

for each input. This approach enables the algorithm to learn from the evolving train status over time. Consequently, the size of the dataset may vary depending on the length of historical data considered. For instance, when using 10 as the history length, the training and validation set contains 5055 data points. This number decreases to 4905 when considering 20 as the history length, and further reduces to 4755 for 30 history length. Despite these variations, the overall data sizes remain relatively consistent.

To standardise the data, wheel speeds and GPS speeds are normalised by dividing them by 31.3877 m/s, which represents the highest speed value in the datasets. In real-world scenarios or with different datasets, other thresholds can be used.

The primary objective of this research is to identify an optimal CNN architecture that yields excellent prediction accuracy. The resulting model can be easily retrained to adapt to other datasets, ensuring its applicability across various scenarios.

## Methodology

This section introduces the methodologies employed in this article for speed estimation. The first part examines the AKF, a well-established technique in conventional speed estimation. This method leverages recursive estimation and dynamic modelling to provide accurate train speed predictions, particularly in scenarios characterised by varying noise levels and changing dynamics. The second part explores the machine learning approach, specifically CNN. As a subset of deep learning, CNN offers a powerful alternative by learning complex patterns from large datasets and adapting to diverse sensor inputs.

### Conventional methods for speed estimation

The Kalman Filter is a recursive estimation method, meaning that it can compute the current state estimate using only the estimate from the previous time step and the current observation.<sup>22</sup> Therefore, it does not need to retain a history of observations or estimates. Unlike most filters, the Kalman Filter is purely a time-domain filter. It does not require operating in the frequency domain followed by a conversion to the time domain, as is necessary for frequency-domain filters like low-pass filters.

The Kalman Filter model assumes that the true state at time  $k$  evolves from the state at time  $k - 1$  according to the following equation:

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k \quad (1)$$

where  $\mathbf{F}_k$  is the state transition model applied to the previous state  $\mathbf{x}_{k-1}$ .  $\mathbf{B}_k$  is the control-input model applied to the control vector  $\mathbf{u}_k$ .  $\mathbf{w}_k$  is the process noise, which is assumed to follow a multivariate normal distribution with zero mean and covariance matrix  $\mathbf{Q}_k$ , denoted as  $\mathbf{w}_k \sim N(0, \mathbf{Q}_k)$ .

At time  $k$ , a measurement  $\mathbf{z}_k$  of the true state  $\mathbf{x}_k$  satisfies the following equation:

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \quad (2)$$

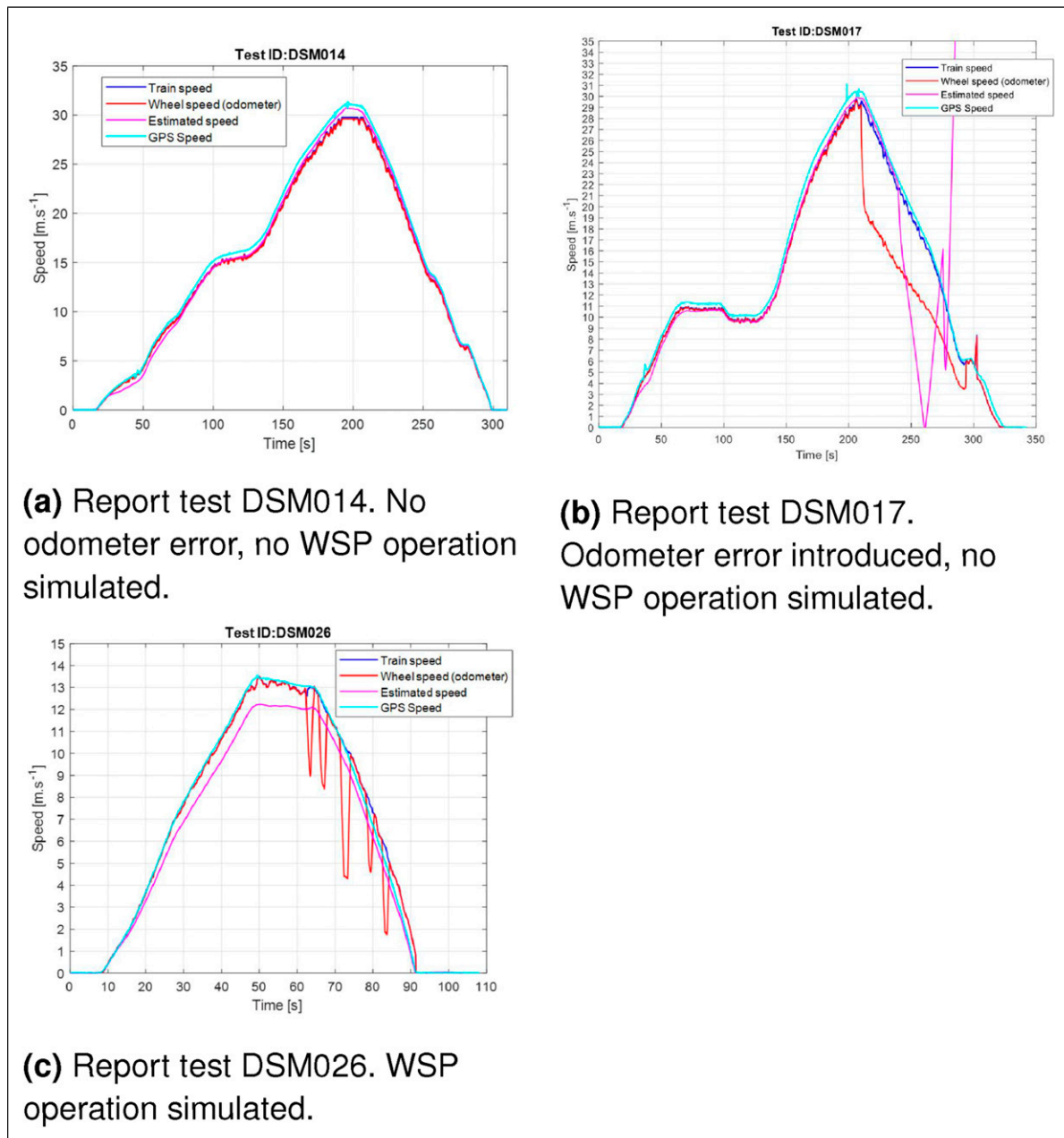


Figure 1. Dependable speed measurement for improved low adhesion braking report figures.

where  $\mathbf{H}_k$  is the observation model that maps the true state space into the observed space, and  $\mathbf{v}_k$  is the observation noise, which is assumed to have zero mean and covariance matrix  $\mathbf{R}_k$ , denoted as  $\mathbf{v}_k \sim N(0, \mathbf{R}_k)$ .

The initial state and the noise terms at each time step  $\{\mathbf{x}_0, \mathbf{w}_1, \dots, \mathbf{w}_k, \mathbf{v}_1, \dots, \mathbf{v}_k\}$  are all assumed to be mutually independent.

Being a recursive estimator, the estimated current state is computed based only on the estimated state from the previous time step and the current measurement.

Two variables represent the state of the filter:  $\hat{\mathbf{x}}_{k|k}$  (the a posteriori state estimate at time k given observations up to and including time k) and  $\mathbf{P}_{k|k}$  (the a posteriori estimate covariance matrix, which measures the estimated accuracy of the state estimate).

The Adaptive Kalman Filter is an extension of the standard Kalman Filter designed to handle situations where the system model or noise characteristics are uncertain or

time-varying. The AKF dynamically adjusts filter parameters based on observed data, allowing it to adapt to changing conditions.

The key difference in the AKF lies in how it estimates and updates the process noise covariance  $\mathbf{Q}_k$  and measurement noise covariance  $\mathbf{R}_k$ . Two common approaches are covariance matching and maximum likelihood estimation. In the covariance matching approach, the innovation sequence is defined as:

$$\mathbf{v}_k = \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} \quad (3)$$

where  $\hat{\mathbf{x}}_{k|k-1}$  is the a priori state estimate. The theoretical covariance of the innovation is:

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k \quad (4)$$

where  $\mathbf{P}_{k|k-1}$  is the a priori estimate covariance

The actual covariance of the innovation is estimated using a moving window of size  $N$ :

$$\widehat{\mathbf{C}}_{v,k} = \frac{1}{N} \sum_{i=k-N+1}^k \mathbf{v}_i \mathbf{v}_i^T \quad (5)$$

The noise covariances are then adapted:

$$\mathbf{R}_k = \widehat{\mathbf{C}}_{v,k} - \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T \quad (6)$$

$$\mathbf{Q}_k = \mathbf{K}_k \widehat{\mathbf{C}}_{v,k} \mathbf{K}_k^T \quad (7)$$

where  $\mathbf{K}_k$  is the Kalman gain.

In the maximum likelihood approach, the following log-likelihood function is maximised:

$$L = -\frac{1}{2} \sum_{k=1}^N (\ln |\mathbf{S}_k| + \mathbf{v}_k^T \mathbf{S}_k^{-1} \mathbf{v}_k) \quad (8)$$

The resulting estimates of  $\mathbf{Q}_k$  and  $\mathbf{R}_k$  are then used in the standard Kalman Filter equations:

Prediction step:

$$\widehat{\mathbf{x}}_{k|k-1} = \mathbf{F}_k \widehat{\mathbf{x}}_{k-1|k-1} + \mathbf{B}_k \mathbf{u}_k \quad (9)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_k \mathbf{P}_{k-1|k-1} \mathbf{F}_k^T + \mathbf{Q}_k \quad (10)$$

Update step:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (11)$$

$$\widehat{\mathbf{x}}_{k|k} = \widehat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}_k \widehat{\mathbf{x}}_{k|k-1}) \quad (12)$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} \quad (13)$$

The AKF iteratively updates these equations along with the noise covariance estimation, enabling it to adapt to changing system dynamics and noise characteristics. This approach allows the filter to maintain optimal performance even when faced with uncertainties or variations in the system model or noise properties.

### CNN-based method for speed estimation

Convolutional Neural Networks<sup>16–19</sup> are a class of neural networks characterized by the inclusion of convolutional operations.<sup>45</sup> Originally developed for image recognition, CNNs also show strong performance in time-series regression tasks by capturing local temporal patterns across multiple input channels.

In this study, CNNs are used to estimate train speed from multi-channel time-series windows. Given an input window  $\mathbf{x} \in \mathbb{R}^{T \times C}$ , where  $T$  is the number of time steps and  $C$  is the number of input channels (e.g., GPS speed, wheel speed), the CNN aims to learn a nonlinear mapping:

$$v = f_\theta(\mathbf{x}) \quad (14)$$

where  $f_\theta$  represents the CNN model with trainable parameters  $\theta$ , and  $v$  is the estimated train speed.

The CNN architecture typically consists of stacked convolutional layers followed by nonlinear activation

functions, pooling layers for dimensionality reduction, and fully connected layers for regression. This design allows the model to automatically extract and integrate spatial-temporal features of the input data. The core of the convolutional layer is a small matrix called a convolution kernel or filter. It slides over the input matrix and performs a simple weighted summation operation at each position to extract local features from the input data and produce the feature map. Assuming the input of the neural network is a 2D matrix  $I$  and the convolution kernel is  $K$ , then the value at position  $(i, j)$  in the output feature map after the convolution operation is:

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(i + m, j + n) \cdot K(m, n) \quad (15)$$

The activation function plays a key role in transforming linear inputs into nonlinear outputs, thereby enhancing the network's expressive power. Without activation functions, a multi-layer neural network would degrade into a simple linear model, incapable of capturing the complex nonlinear relationships within the data. This work employs the Rectified Linear Unit (ReLU) as the activation function, which is a piecewise linear function mathematically expressed as

$$\text{ReLU}(x) = \max(0, x). \quad (16)$$

The neural network is trained by minimising a mean squared error loss  $L$ :

$$L = \frac{1}{N} \sum_{i=1}^N (y_i - \widehat{y}_i)^2 \quad (17)$$

where  $N$  is the number of samples in the mini-batch,  $y_i$  is the ground truth, and  $\widehat{y}_i$  is the inferred output. The minimisation process is mini-batch stochastic gradient descent.

To improve generalisation and avoid overfitting, dropout layers are used during training. Hyperparameters such as the number of convolutional blocks, kernel size, learning rate, and dropout rate are optimised using the Optuna<sup>1</sup> library, which applies Bayesian optimisation to efficiently explore the hyperparameter space. It employs the Tree-structured Parzen Estimator, a Bayesian optimisation-based algorithm, as its default hyperparameter tuning approach. Within this framework, kernel density estimation is utilised to model the probability density functions separately for the elite set, which comprises the better-performing hyperparameter configurations, and the non-elite set, consisting of the poorer-performing ones. This methodology allows Optuna to intelligently guide the search process towards regions of the hyperparameter space that are more likely to yield improved performance.

Three CNN architectures inspired by the works of Karlsson,<sup>41</sup> Wang,<sup>39</sup> and Seethi<sup>42</sup> are evaluated and adapted to our dataset, aiming to find the design that achieves the highest accuracy in speed estimation. Their specific details, together with the training details are presented in the following subsections.

**Single-branch 2D model.** The Single-branch 2D model proposed in<sup>41</sup> consists of three blocks, with each block containing a convolutional layer followed by a ReLU activation function and batch normalisation, succeeded by a 20% dropout layer, and culminating in a fully connected layer and a regression layer. The input to the network is  $N \times M$ , where  $M$  is the number of sensor signals used and  $N$  is the number of history time steps as explained in the Section titled Dataset preparation.

A larger input shape allows for the inclusion of more historical information, while a smaller input shape can track rapid changes and reduce computational complexity. The input shapes examined for this architecture are  $10 \times 3$ ,  $20 \times 3$ ,  $30 \times 3$ , and  $40 \times 3$ , each incorporating three signals: time, wheel speed, and GPS speed.

Another critical parameter in the design of CNNs is the number of layers or blocks. A model with too few blocks may underfit the data, failing to capture the underlying features effectively. Conversely, a model with too many layers may overfit the training data, learning noise and specific details that do not generalise well to new data. Determining the optimal number of blocks involves balancing model complexity, performance, and computational efficiency. The original architecture employs three blocks, in which the batch normalisation layers are put after the activation functions ReLU. Studies indicate that “adding the Batch Normalisation (BN) transform immediately before the nonlinearity” is “likely to produce activations with a stable distribution”.<sup>46</sup> In this work, the BN layer is placed before the activation function ReLU, deviating from the original architecture. A range of 1 to 20 blocks is provided to Optuna to identify the optimal configuration.

To simplify the optimisation search problem, the number of filters and the kernel size in each convolutional layer are kept consistent across all blocks. The range of filters evaluated spans from 8 to 64. Additionally, three kernel sizes are tested: (3,2), (5,2), and (7,2).

Three additional hyperparameters are also evaluated: the dropout rate, ranging from 0.0 to 0.5; the learning rate, spanning from  $10^{-5}$  to  $10^{-2}$  on a logarithmic scale; and the batch size, with values of 8, 16, 32, and 64. These hyperparameters are considered to assess their impact on the prediction accuracy.

**Single-branch 1D model.** The second architecture considered in this work is one adapted by<sup>39</sup> from a renowned deep CNN model originally developed for handwriting classification.<sup>47</sup> The 2-dimensional spatial convolution is replaced with 1-dimensional convolution to facilitate pattern recognition on time-series acceleration data, and the design parameters, including speed sample sizes, convolutional kernel sizes, and fully connected layer dimensions, are minimised to allow for a more compact hardware implementation.

The CNN architecture begins with an input tensor of shape (1,100,6), where 100 is the sequence length, and 6 is the number of input signals. The first convolutional layer processes this input and produces 32 feature maps. This is followed by a MaxPooling operation with a pool size of  $1 \times 5$ , which reduces the sequence length to 20 while retaining the same quantity of feature maps, resulting in a tensor of

shape (1,20,32). A second convolutional layer is applied, maintaining the tensor shape at (1,20,32). Another MaxPooling operation with a pool size of  $1 \times 5$  further reduces the sequence length to 4, yielding a tensor of shape (1,4,32). This output is then flattened and passed through a fully connected layer with 128 units, followed by a dropout layer. The subsequent fully connected layer with 32 units produces the final output tensor, which is then transformed into a single output speed value.

In the hyperparameter study for this architecture, four distinct input shapes are evaluated: (1,10,3), (1,20,3), (1,30,3), and (1,40,3). Each input shape is assessed to determine its impact on the performance. The number of blocks, where a block is defined as a convolutional layer followed by a MaxPooling layer, varies from 1 to 20. The kernel sizes considered range from 2 to 10. Other hyperparameters, including the number of filters, dropout rate, learning rate, and batch size, are maintained within the ranges previously established during the optimisation process for the previous model.

**Multiple-branch model.** The final architecture evaluated and optimised is the Multiple-branch model.<sup>42</sup> A notable distinction of this structure from the previous two is that each sensor signal is processed independently through separate convolutional branches before the features are concatenated. This approach offers the advantage of mitigating noise introduced by signal fusion, thereby simplifying the feature extraction process. The concatenation operation integrates features passed by the global max-pooling layers from multiple branches, and outputs to two dense layers followed by a dropout layer respectively. Finally, a single neuron output layer presents the estimated speed.

Optuna searches the optimal input shape of each branch within four options: (10,1), (20,1), (30,1) and (40,1). A block in this architecture is defined as two consecutive convolutional layers, with the second layer having twice the number of filters as the first. This design ensures that each block's structure and the number of convolutional kernels remain similar to the original architecture. The number of blocks being searched ranges from 1 to 3, which is narrower than the range considered in the previous two architectures due to the presence of multiple channels.

The structure following the convolutional layers remains consistent with the original architecture. The two dense layers contain 64 and 32 neurons, respectively, considering that the sample size in the reference is 156, whereas in this work, it is only 40 at most, which is significantly smaller. The ranges of number of filters, dropout rate and learning rate remain the same with the Single-branch 2D model and the Single-branch 1D model.

## Results

### Adaptive Kalman Filter

The AKF introduced in the section titled Conventional methods for speed estimation is applied to predict train speeds on test DSM014 (without WSP operation), and test DSM029 (with WSP operation). The results in [Figure 2](#)

provide a baseline for comparison with the subsequent results obtained from the CNN models.

The AKF algorithm demonstrates a good estimation of train speeds when the speed varies smoothly, particularly at lower speeds, as exemplified in the first 150 s of Figure 2(a) and the first 55 s of Figure 2(b). However, Figure 2(c) reveals that the error fluctuations become more pronounced when the train speed exceeds 20 m/s, occasionally exceeding 1 m/s. The AKF algorithm encounters challenges when dealing with high train speeds.

After 55 s in Figure 2(b), when the WSP operation is simulated, the AKF's prediction accuracy rapidly deteriorates, reaching a peak of 2.25 m/s. The presence of WSP operation is associated with a higher Root Mean Square Error (RMSE) of 0.5274 compared to 0.4832 in its absence. The wheel slip scenario introduces nonlinear dynamics and uncertainties, which the AKF may not fully account for in its estimation process, even if the scene is in an overall low-speed state.

### CNN optimal hyperparameter search

In this section are presented the results for the hyperparameter optimisation for the three CNN architectures studied. These results are required to define the final CNN architectures, which will lead to the machine learning results presented in the next section.

The hyperparameter search spaces are summarised in Table 1.

For each architecture, a comprehensive study is conducted wherein a model is trained for 60 epochs in each trial, employing an 80/20 split ratio for the training and validation data. Throughout the training process, the performance of the model is monitored, and the best validation loss achieved during each trial is recorded.

The search history of optimal hyperparameters and their importance for the three architectures are plotted in Figure 3. Pruned trials are excluded from the plot, and the y-axis is set to a logarithmic scale. In all three cases, we can observe a general downward trend in the objective value as

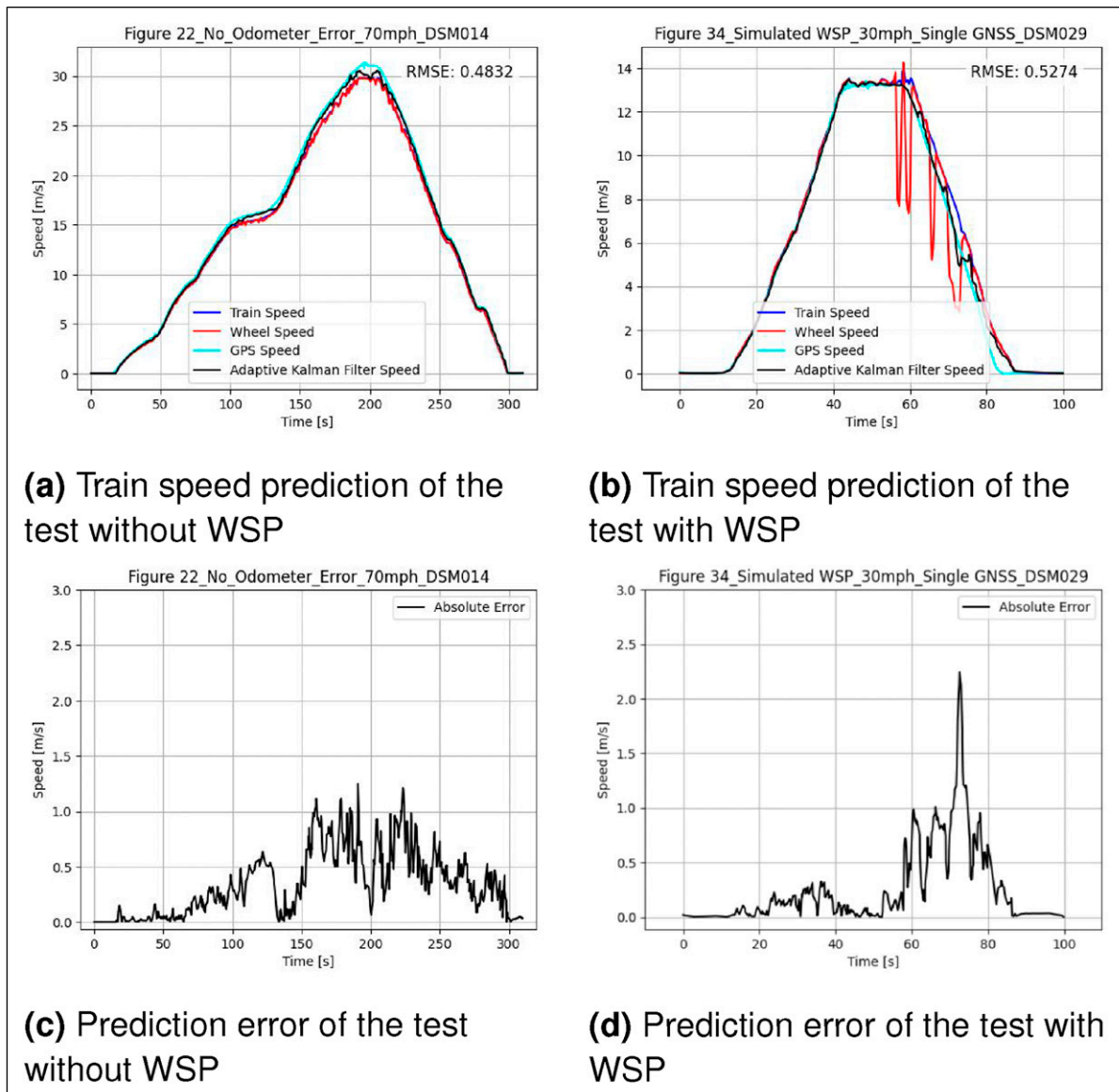
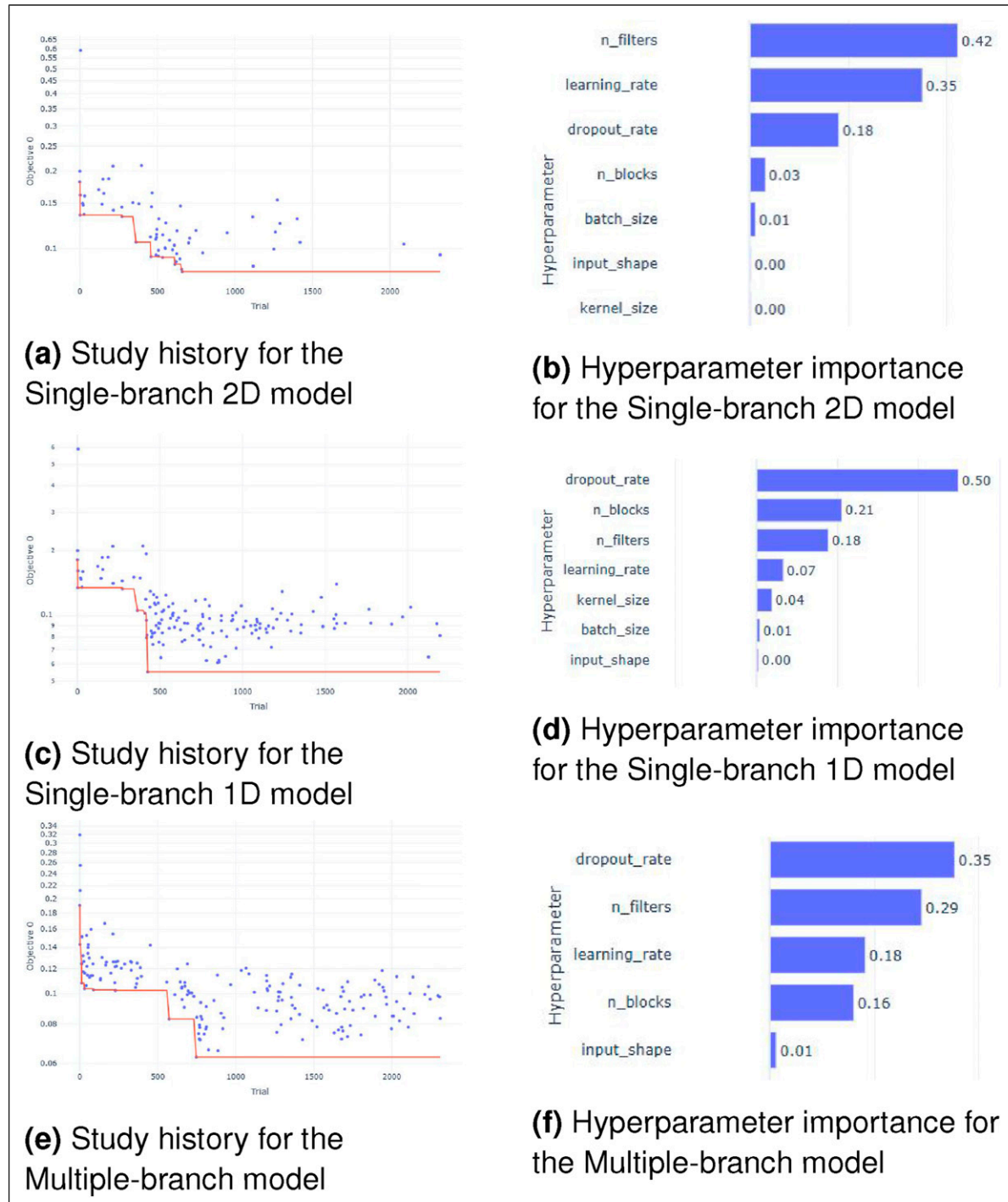


Figure 2. Train speed predictions and errors by the AKF.

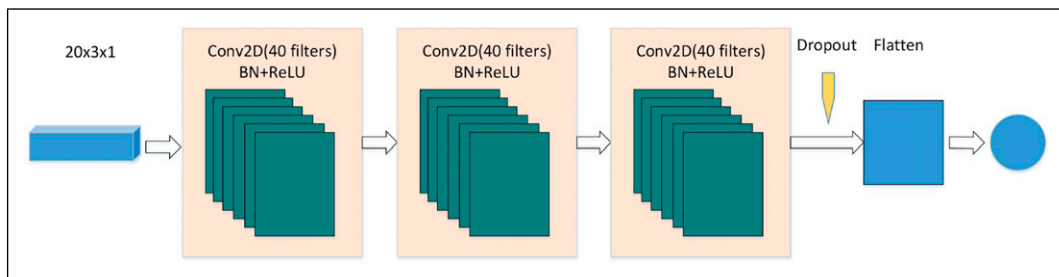
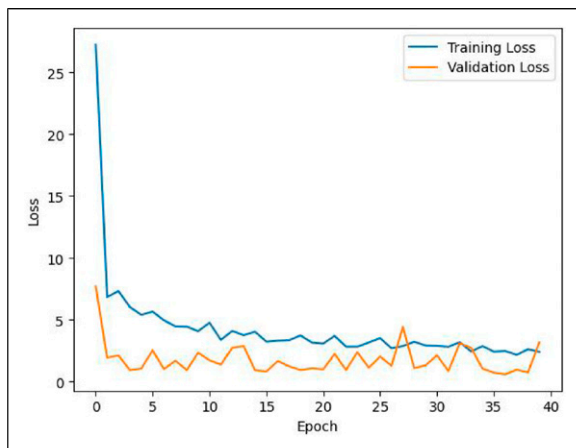
**Table 1.** Hyperparameter ranges of CNN architectures.

	Single-branch 2D	Single-branch 1D	Multiple-branch
Input shape	(10,3,1), (20,3,1) (30,3,1), (40,3,1)	(10,3), (20,3) (30,3),(40,3)	(10,1), (20,1) (30,1), (40,1)
Number of blocks	1–20	1–20	1–3
Number of filters	8–64	8–64	8–64
Kernel size	(3,2), (5,2), (7,2)	2–10	2–10
Dropout rate	0–0.5	0–0.5	0–0.5
Learning rate	$10^{-5}$ – $10^{-2}$	$10^{-5}$ – $10^{-2}$	$10^{-5}$ – $10^{-2}$
Batch size	8, 16, 32, 64	8, 16, 32, 64	8, 16, 32, 64

**Figure 3.** Optimal hyperparameter search. In the left figures, the blue dots are completed trials and the red line represents the best validation loss at each point of the search. The right figures show the relative importance of each hyperparameter.

**Table 2.** Optimal hyperparameters of three architectures searched.

	Single-branch 2D	Single-branch 1D	Multiple-branch
Total trials	2429	2196	2314
Pruned trials	2363	1994	2137
Completed trials	66	202	177
Input shape	(20,3,1)	(10,3)	(30,1)
Number of blocks	3	4	2
Number of filters	40	53	46
Kernel size	(7,2)	2	2
Dropout rate	4.9e-5	8.8e-3	1.9e-4
Learning rate	1.7e-4	2.0e-3	1.8e-3
Batch size	8	8	32

**Figure 4.** Structure of the Single-Branch 2D model.**Figure 5.** Training and validation loss of the Single-branch 2D model with optimal hyperparameters.

the number of trials increases, indicating that the hyperparameter optimisation process is successfully finding better configurations over time. The red line in each graph represents the minimum validation loss value achieved so far, showing a step-wise improvement as better solutions are found.

The search histories of the three architectures exhibit similar patterns, characterised by significant drops in the objective value within the first 500 trials. The best values in all three cases stabilise after approximately 750 trials, with no further improvements observed thereafter.

In [Figure 3\(a\)](#), the number of filters ( $n_{\text{filters}}$ ) is the most important hyperparameter for the Single-branch 2D model, with an importance score of 0.42. This is closely followed by the learning rate at 0.35. The dropout rate also plays a significant role with a score of 0.18. The number of blocks

( $n_{\text{blocks}}$ ) and batch size have much lower importance scores, while input shape and kernel size appear to have negligible impact on the model's performance for this architecture. [Figure 3\(c\)](#) presents a different distribution of importance. The dropout rate takes the lead as the most crucial hyperparameter with a score of 0.50. The number of blocks and the number of filters follow with moderate importance at 0.21 and 0.18 respectively. Other hyperparameters show a very small impact, suggesting minimal to no importance for the Single-branch 1D model. In [Figure 3\(e\)](#), it can be seen that for the Multiple-branch model, the dropout rate and the number of filters are key factors, with the learning rate and the number of blocks contributing moderately.

[Table 2](#) summarises the optimal hyperparameters identified through the search within the defined spaces. In each case, over 2000 trials were conducted, with the majority of these trials being automatically pruned by the algorithm due to their large validation losses.

### Model training and results

In this section, the previous optimal hyperparameters are employed to train the three architectures. Subsequently, the models are applied to the test data to evaluate the predictive performance of each architecture.

**Single-branch 2D model.** The Single-branch 2D model is trained with the hyperparameters listed in [Table 2](#), and the structure plotted in [Figure 4](#). The model is trained for 40 epochs. The training process, as depicted in [Figure 5](#), shows a downward trend as the model learns. The training loss curve maintains a relatively steady decline, which

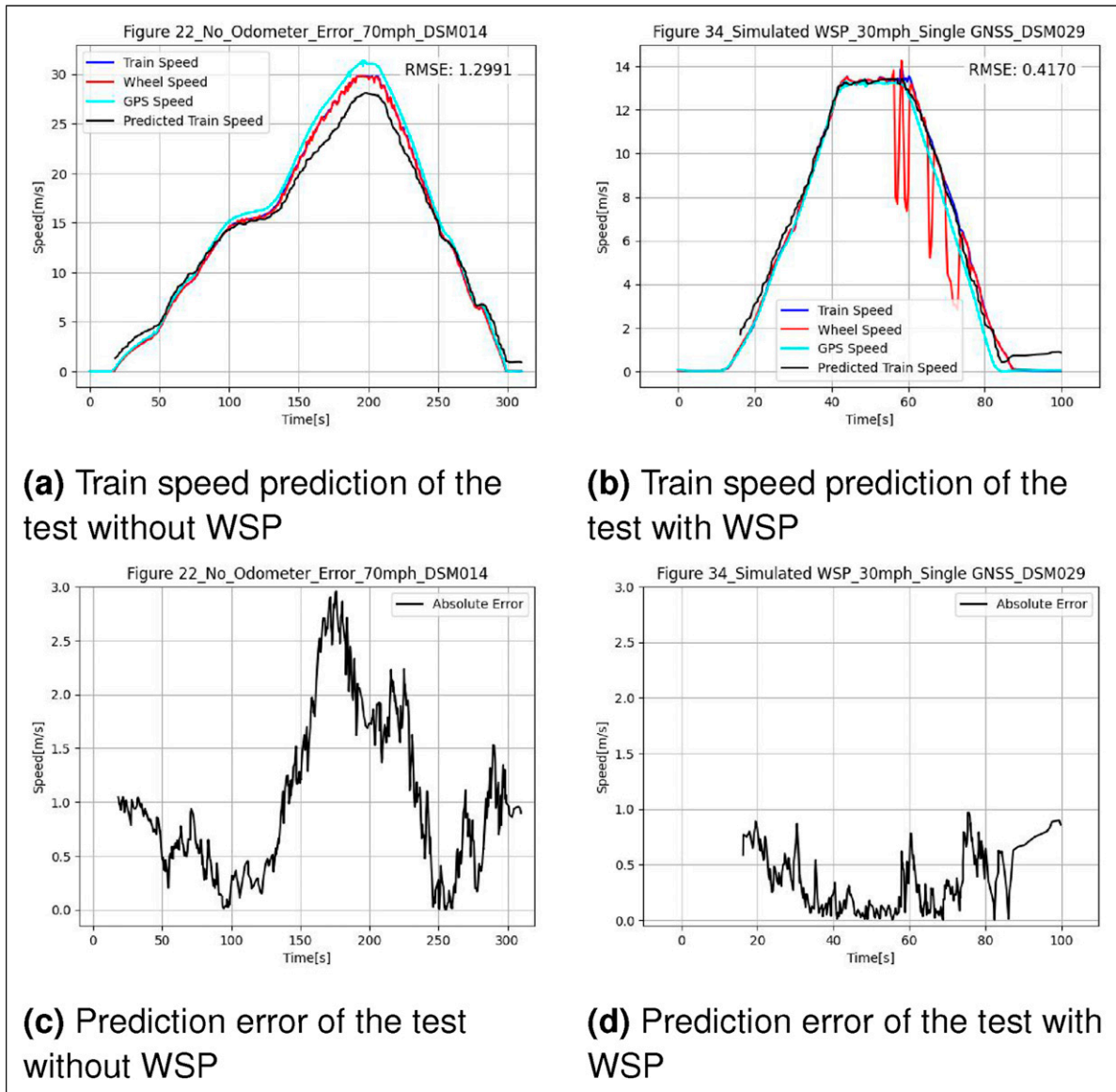


Figure 6. Train speed predictions and errors by the Single-branch 2D model.

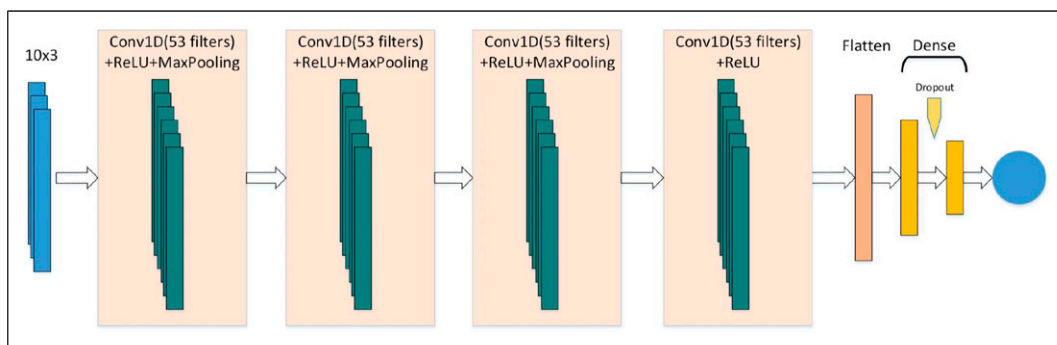


Figure 7. Structure of the Single-branch 1D model.

means the model is continuously improving its predictions on the training data. The validation loss curve, while also showing an overall decrease, experiences spikes at certain points because of a non-smooth loss landscape.

Figure 6 illustrates the true train speed, GPS speed, wheel speed, and the estimated train speed by the Single-branch 2D

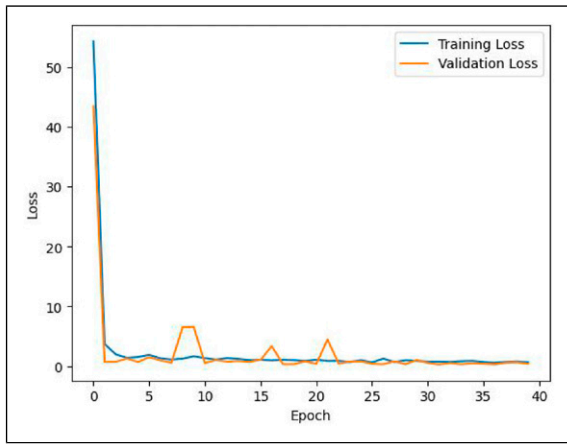
model for test cases DSM014 and DSM029, with a braking curve from 31 m/s and 13.5 m/s respectively. The model does not provide speed estimates in the first few seconds as it requires 20 consecutive time steps of speed data to generate predictions. In Figure 6(a) and (c) for the test DSM014, where the wheel speed aligns closely with the train speed, the

predicted speed is slightly higher than the ground truth until the train accelerates to 15 m/s, with an error not exceeding 1 m/s. However, when the train speed ranges from 15 m/s to 31 m/s, the prediction error reaches up to 3 m/s during both the

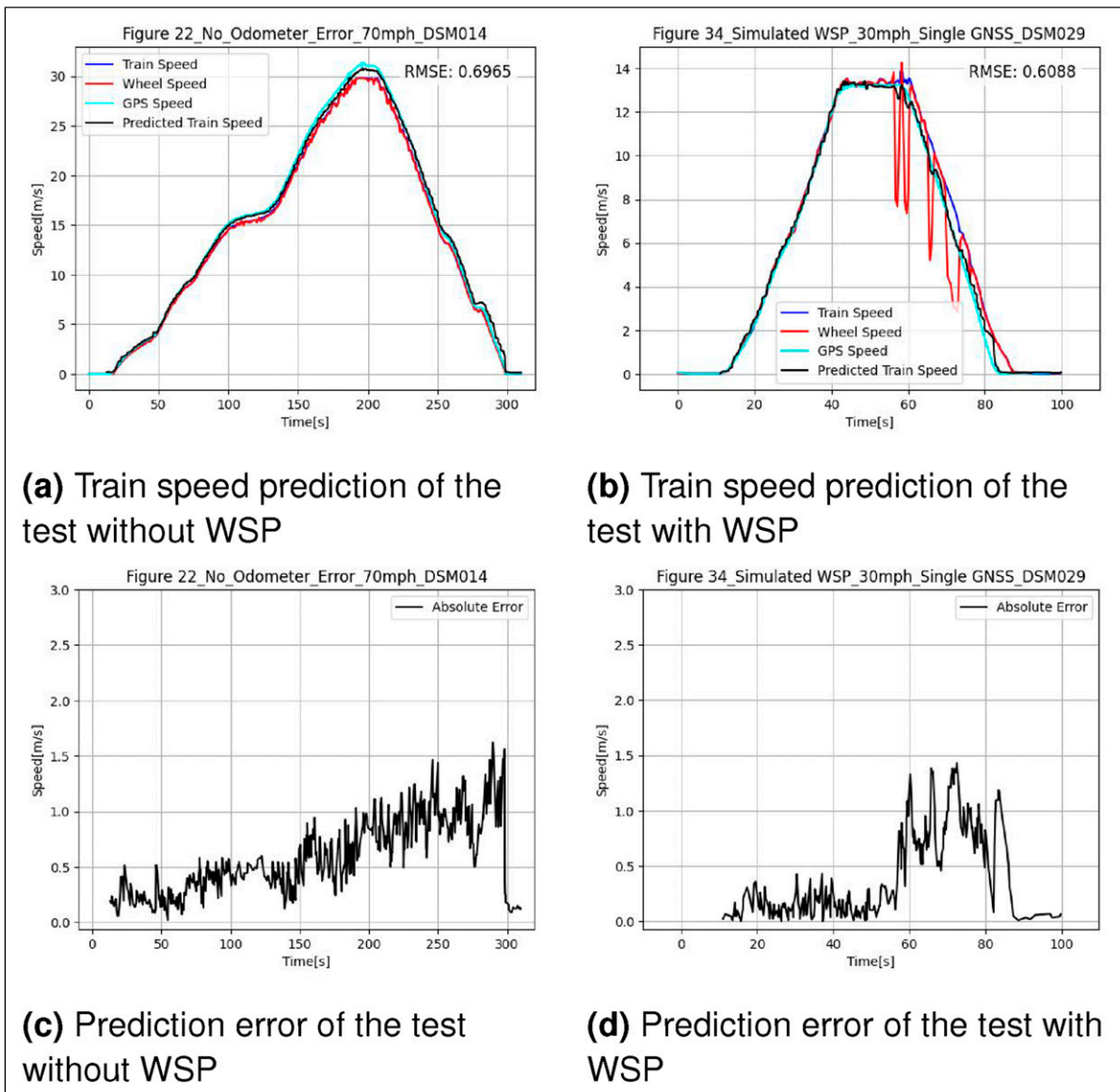
acceleration and braking phases. The error reduces to less than 1.5 m/s in the final stage, but a residual error remains even when the train is at a complete stop. The prediction has a root mean square error of around 1.3 across the entire profile.

The speed profile of the test DSM029 in Figure 6(b), which includes the simulated WSP operation, is supposed to be more challenging to predict due to the significant discrepancies introduced by the wheel speed signal. Surprisingly, the network produces a better estimation with an error of no more than 1 m/s and a RMSE of 0.4170, even during the braking phase when the wheel speed deviates substantially from the ground truth. Similar to the test DSM014, the network ultimately predicts a non-zero speed when the train comes to a stop.

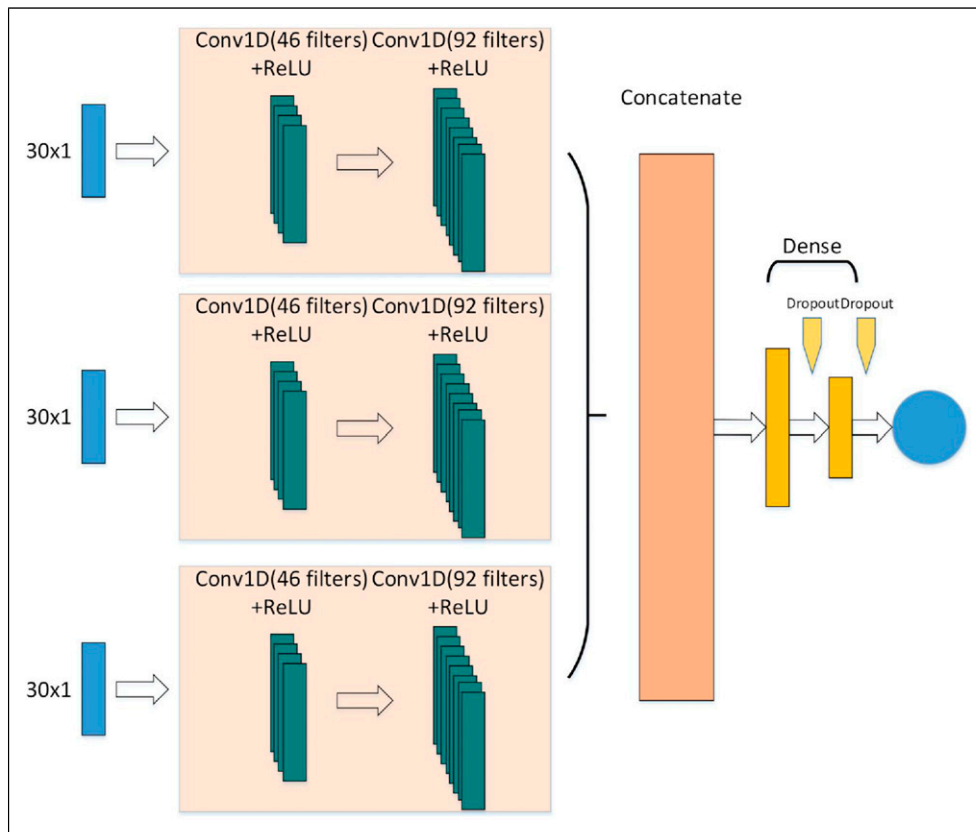
*Single-branch 1D model.* By training the optimal hyperparameters and structure of the Single-branch 1D model plotted in Figure 7, the training and validation losses illustrated in Figure 8 are obtained. It is clear that this model converges rapidly and exhibits significantly greater stability compared to the Single-branch 2D model. The loss curves also suggest that 40 epochs are sufficient, as additional training epochs do not further reduce the validation loss.



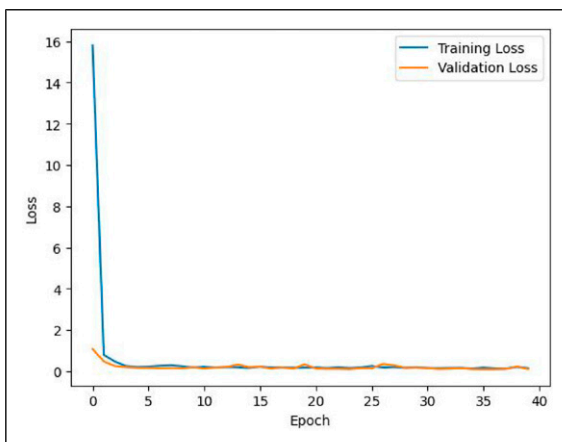
**Figure 8.** Training and validation loss of the Single-branch 1D model with optimal hyperparameters.



**Figure 9.** Train speed predictions and errors by the Single-branch 1D model.



**Figure 10.** Structure of the Multiple-branch model.



**Figure 11.** Training and validation loss of the Multiple-branch model with optimal hyperparameters.

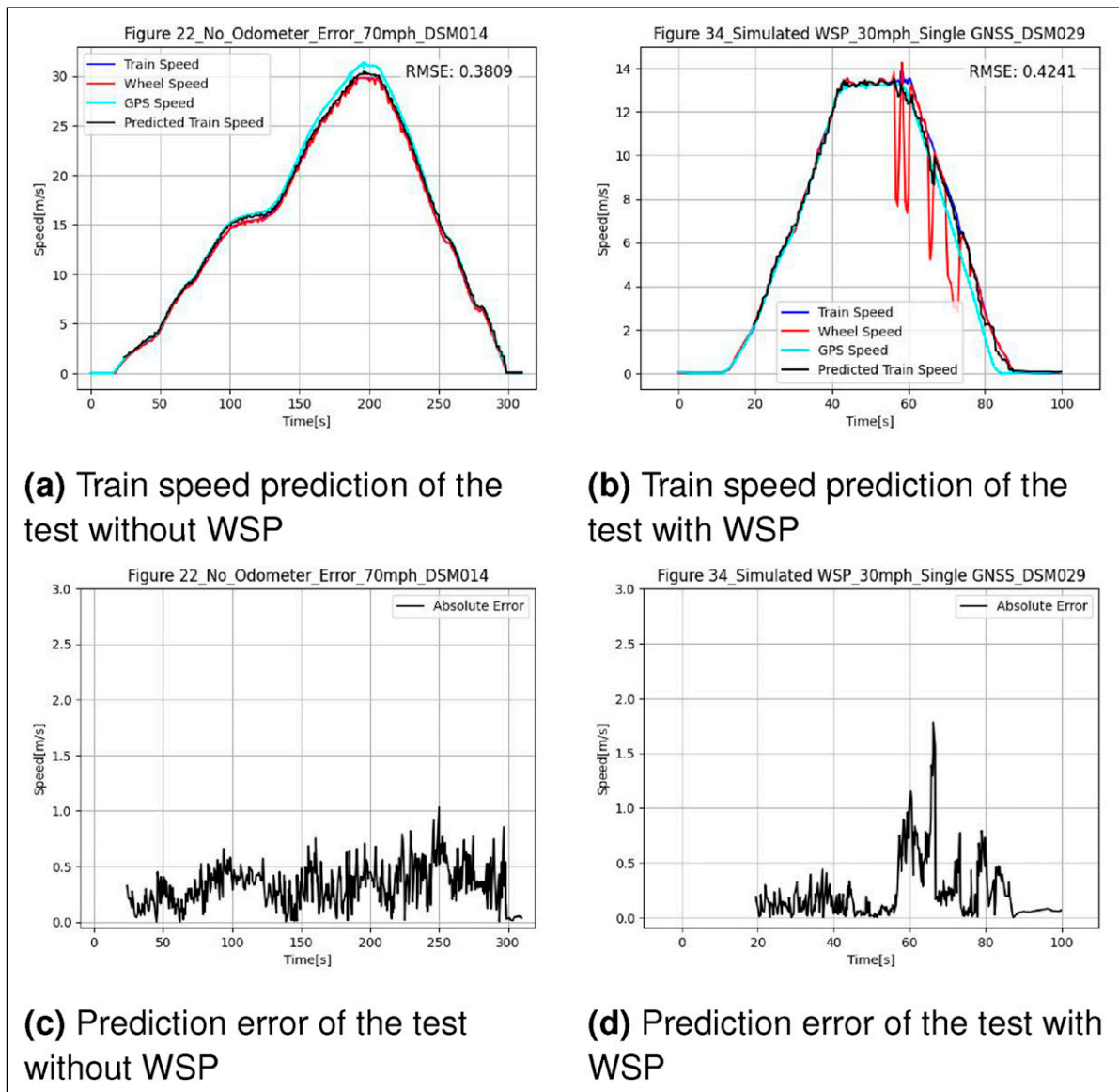
Figure 9 illustrates the model's predictions for two tests. Compared to the previous Single-branch 2D model, this model performs quite well on DSM014 without the WSP operation, closely matching the true train speed throughout the entire profile with a maximum error of around 1.5 m/s and errors less than 1 m/s during the acceleration phase. The RMSE is reduced from 1.2991 to 0.6965. For DSM029 with the WSP operation, the errors are all below 1.5 m/s, with errors during the acceleration and constant speed phases smaller than 0.5 m/s. Notably, the model accurately predicts the zero-speed state after the train stops, which is more aligned with real-world conditions. Although the RMSE of this test is higher than that of the Single-branch 2D model and the prediction errors increase

during the braking phase, the model demonstrates excellent performance during the acceleration and constant speed phases.

**Multiple-branch model.** The optimal structure of the Multiple-branch model is plotted in Figure 10.

Figure 11 presents the losses of the model during its training. The flat curves indicate that the model has achieved a good fit on the training data and that its performance on the validation set is stable. It significantly outperforms the Single-branch 2D model, and the validation loss exhibits fewer anomalies compared to the Single-branch 1D model. Similar to the previous model, 40 epochs of training are sufficient to obtain stable results.

Figure 12 illustrates this model's predictions on tests without and with the WSP operation. The RMSE for both tests is notably low. As shown in Figure 12(a) and (c), the estimated speed closely matches the train's true speed throughout the entire test, with errors not exceeding 1 m/s in both the acceleration and braking phases. The model's prediction accuracy does not significantly decrease at high speeds. In Figure 12(b), before the introduction of the WSP operation, the prediction error remains below 0.5 m/s. The rapid changes in wheel speed increase the error slightly, but it generally stays below 1 m/s, with only one peak exceeding 1.5 m/s. The model's predictions for the train speed in both tests have very low RMSE values, respectively 0.3809 and 0.4241. Overall, the model based on the multiple-branch structure demonstrates high prediction accuracy and stability. This can be attributed to its independent processing of signals from different channels, which allows for the aggregation of all signal features and mitigates the impact of a single channel's failure on the final prediction outcome.



**Figure 12.** Train speed predictions and errors by the Multiple-branch model.

## Discussion

The results presented in the Section titled Model training and results indicate that the Single-branch 2D model demonstrates a relatively bad prediction performance. Its predictions exhibit substantial errors and high variability, particularly under smooth, low-speed train conditions. Although this model shows a lower RMSE with the WSP operation simulated, the reliability and stability of this architecture require further enhancement. Processing multiple signals as an image does not seem to be an effective approach. In contrast, the Single-branch 1D model provides more stable predictions during smooth acceleration and deceleration phases. However, it exhibits higher errors during the WSP operation simulation. Compared to the previous two models, the Multiple-branch model, which trains three separate signals separately, achieves the lowest prediction error under the scenario without WSP. When WSP operation simulated, its prediction accuracy and stability are slightly inferior to the Single-branch 2D model, but still comparable. The Multiple-branch model

demonstrates the most robust performance among the three architectures.

The conventional speed estimation method - AKF - exhibits commendable prediction accuracy and stability, surpassing both the Single-branch 2D and the Single-branch 1D models. However, the CNN-based Multiple-branch model demonstrates greater potential. It achieves a smaller RMSE and lower maximum errors compared to the AKF. Further improvements in prediction accuracy and reductions in error peaks can be achieved by incorporating additional signals and expanding the dataset.

It is worth to note that several limitations that could impact the generalisability and robustness of the results. The work is based on simulated data. The dataset used is extracted from figures in an existing report, which introduces certain inaccuracies and limitations in the data. The size of the dataset is also constrained, which likely influenced the training efficacy and predictive performance of the CNN models. The current study focuses on only three channels of signal data. While these channels provide valuable information, the exclusion of other potentially

relevant channels, such as acceleration and braking status, may have limited the model's robustness and reliability. Incorporating additional channels can enrich the dataset, allowing the CNN models to capture a broader range of operational dynamics. In addition, exploring the application of the model in certain channel signal loss scenarios can also improve the robustness of the system. These factors underscore the need for more accurate and comprehensive datasets in order to optimise the models further and achieve better prediction accuracy.

## Conclusions

This study demonstrates the potential of Convolutional Neural Networks in the prediction of train speeds using signal data, offering a promising approach for advancing train control systems. The CNN models developed in this work were able to identify and learn patterns from the provided signal data, highlighting their applicability in the field of railway operations. Among the three CNN designs, the Multiple-branch model exhibits excellent performance in the scenario without WSP and maintains good predictive accuracy with WSP operation simulated. Overall, this model outperforms the other two CNN architectures and the Adaptive Kalman Filter in terms of robustness and accuracy across different train operating conditions.

The study's findings have important implications for railway safety and efficiency. By providing more accurate and robust speed estimates, especially in adverse conditions, these CNN-based models can enhance the performance of critical systems such as Automatic Train Protection and improve overall train control.

Future work can focus on validating these models with real-world operational data across a wider range of conditions and train types. Further explorations may also include the integration of these CNN models into existing train control systems and the investigation of their real-time performance and computational requirements.

## Acknowledgements

The work stems from an MSc Applied AI programme individual research project outcome supported by RSSB. The authors thank M. Carmichael for his invaluable guidance and support throughout the MSc project. The authors also thank the project reviewers for the useful discussions and comments during the work.

## Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

## ORCID iDs

Argyrios Zolotas  <https://orcid.org/0000-0002-2829-1298>

Miguel Arana-Catania  <https://orcid.org/0000-0003-4277-0292>

## Note

1. <https://optuna.org/>.

## References

1. Wang J. *Safety theory and control technology of high-speed train operation*. London: Academic Press, 2018.
2. Peng Y, Lu S, Chen F, et al. Energy-efficient train control incorporating inherent reduced-power and hybrid braking characteristics of railway vehicles. *Transport Res C Emerg Technol* 2024; 163: 104626. DOI: [10.1016/j.trc.2024.104626](https://doi.org/10.1016/j.trc.2024.104626).
3. Davoodi M, Jafari Kaleybar H, Brenna M, et al. Energy management systems for smart electric railway networks: a methodological review. *Sustainability* 2023; 15(16): 12204. DOI: [10.3390/su151612204](https://doi.org/10.3390/su151612204).
4. Wang X, Tang T, Su S, et al. An integrated energy-efficient train operation approach based on the space-time-speed network methodology. *Transport Res E Logist Transport Rev* 2021; 150: 102323. DOI: [10.1016/j.tre.2021.102323](https://doi.org/10.1016/j.tre.2021.102323).
5. Peng Y, Zhou J, Fan C, et al. A review of passenger ride comfort in railway: assessment and improvement method. *Transportation Safety and Environment* 2022; 4(2): tdac016. DOI: [10.1093/tse/tdac016](https://doi.org/10.1093/tse/tdac016).
6. Sedghi M, Kauppila O, Bergquist B, et al. A taxonomy of railway track maintenance planning and scheduling: a review and research trends. *Reliab Eng Syst Saf* 2021; 215: 107827. DOI: [10.1016/j.ress.2021.107827](https://doi.org/10.1016/j.ress.2021.107827).
7. Lu C. A discussion on technologies for improving the operational speed of high-speed railway networks. *Transportation Safety and Environment* 2019; 1(1): 22–36. DOI: [10.1093/tse/tdz003](https://doi.org/10.1093/tse/tdz003).
8. Ekberg A, Åkesson B and Kabo E. Wheel/rail rolling contact fatigue – probe, predict, prevent. *Wear* 2014; 314(1): 2–12. DOI: [10.1016/j.wear.2013.12.004](https://doi.org/10.1016/j.wear.2013.12.004).
9. Shi H, Wang J, Wu P, et al. Field measurements of the evolution of wheel wear and vehicle dynamics for high-speed trains. *Veh Syst Dyn* 2018; 56(8): 1187–1206. DOI: [10.1080/00423114.2017.1406963](https://doi.org/10.1080/00423114.2017.1406963).
10. Muniandi G and Deenadayalan E. Train distance and speed estimation using multi sensor data fusion. *IET Radar, Sonar Navig* 2019; 13(4): 664–671. DOI: [10.1049/iet-rsn.2018.5359](https://doi.org/10.1049/iet-rsn.2018.5359).
11. Moaveni B, Fathabadi FR and Molavi A. Supervisory predictive control for wheel slip prevention and tracking of desired speed profile in electric trains. *ISA Trans* 2020; 101: 102–115. DOI: [10.1016/j.isatra.2020.01.011](https://doi.org/10.1016/j.isatra.2020.01.011).
12. Namooano B, Ruiz-Carcel C, Emmanouilidis C, et al. Data-driven wheel slip diagnostics for improved railway operations. *IFAC-PapersOnLine* 2022; 55(19): 103–108. DOI: [10.1016/j.ifacol.2022.09.191](https://doi.org/10.1016/j.ifacol.2022.09.191).
13. Zhan X, Mu ZH, Kumar R, et al. Research on speed sensor fusion of urban rail transit train speed ranging based on deep learning. *Nonlinear Eng* 2021; 10(1): 363–373. DOI: [10.1515/nleng-2021-0028](https://doi.org/10.1515/nleng-2021-0028).
14. Zhou H, Wan Y, Ye H, et al. Real-time energy-efficient optimal control of high-speed electric train. *Control Eng Pract* 2021; 112: 104825. DOI: [10.1016/j.conengprac.2021.104825](https://doi.org/10.1016/j.conengprac.2021.104825).
15. Yaghoubi H. *High-speed rail*. Rijeka: IntechOpen, 2019. DOI: [10.5772/intechopen.76549](https://doi.org/10.5772/intechopen.76549).

16. Homma T, Atlas LE and Marks RJ. An artificial neural network for spatiotemporal: application to phoneme classification. In: Proceedings of the 1st International Conference on Neural Information Processing Systems. Cambridge, MA, USA: MIT Press. 31–40.
17. Denker J, Gardner W, Graf H, et al. Neural network recognizer for hand-written zip code digits. *Adv Neural Inf Process Syst* 1988; 1: 323–331.
18. LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1989; 1(4): 541–551.
19. LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. *Proc IEEE* 1998; 86(11): 2278–2324.
20. Liu Z, Liu J, Xu X, et al. DeepGPS: deep learning enhanced GPS positioning in urban canyons. *IEEE Trans Mobile Comput* 2024; 23(1): 376–392. DOI: [10.1109/TMC.2022.3208240](https://doi.org/10.1109/TMC.2022.3208240).
21. Zhao T, Tao H and Hongxia N. Research on detection and correction of high-speed train idling and skidding based on improved fuzzy neural network. *Automat Control Comput Sci* 2024; 58(3): 274–288. DOI: [10.3103/S0146411624700160](https://doi.org/10.3103/S0146411624700160).
22. Govaers F. *Introduction and implementations of the Kalman filter*. Norderstedt: BoD – Books on Demand, 2019.
23. Cunillera A, Bešinović N, van Oort N, et al. Real-time train motion parameter estimation using an unscented Kalman filter. *Transport Res C Emerg Technol* 2022; 143: 103794. DOI: [10.1016/j.trc.2022.103794](https://doi.org/10.1016/j.trc.2022.103794).
24. Pichlík P and Zděnek J. Train velocity estimation by extended Kalman filter. In: 2016 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Ploiesti, Romania, 2 July, 2016. DOI: [10.1109/ECAI.2016.7861182](https://doi.org/10.1109/ECAI.2016.7861182).
25. Deng Z, Song H, Huang H, et al. Multi-sensor based train localization and data fusion in autonomous train control system. In: 2020 Chinese Automation Congress (CAC), Shanghai, 6–8 November, 2020. DOI: [10.1109/CAC51589.2020.9327825](https://doi.org/10.1109/CAC51589.2020.9327825).
26. Ernest P, Mazl R and Pfeucil L. Train locator using inertial sensors and odometer. In: IEEE Intelligent Vehicles Symposium, 2004, Parma, 14–17 June, 2004. DOI: [10.1109/IVS.2004.1336497](https://doi.org/10.1109/IVS.2004.1336497).
27. Tanelli M, Savaresi SM and Cantoni C. Longitudinal vehicle speed estimation for traction and braking control systems. In: 2006 IEEE Conference on Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, Munich, 4–6 October, 2006. DOI: [10.1109/CACSD-CCA-ISIC.2006.4777080](https://doi.org/10.1109/CACSD-CCA-ISIC.2006.4777080).
28. Mei TX and Li H. Measurement of vehicle ground speed using bogie-based inertial sensors. *Proc Inst Mech Eng F J Rail Rapid Transit* 2008; 222(2): 107–116. DOI: [10.1243/09544097JRR154](https://doi.org/10.1243/09544097JRR154).
29. Hu D, Zong C and Na H. Research on information fusion algorithm for vehicle speed information and road adhesion property estimation. In: 2009 International Conference on Mechatronics and Automation, Changchun, 12 April, 2009. DOI: [10.1109/ICMA.2009.5246209](https://doi.org/10.1109/ICMA.2009.5246209).
30. Chu L, Chao L, Zhang Y, et al. Design of longitudinal vehicle velocity observer using fuzzy logic and Kalman filter. In: Proceedings of 2011 International Conference on Electronic & Mechanical Engineering and Information Technology, Harbin, Heilongjiang, 12–14 August, 2011. DOI: [10.1109/EMEIT.2011.6023772](https://doi.org/10.1109/EMEIT.2011.6023772).
31. Malvezzi M, Vettori G, Allotta B, et al. A localization algorithm for railway vehicles based on sensor fusion between tachometers and inertial measurement units. *Proc Inst Mech Eng F J Rail Rapid Transit* 2014; 228(4): 431–448. DOI: [10.1177/0954409713481769](https://doi.org/10.1177/0954409713481769).
32. Kim K, Kong SH and Jeon SY. Slip and slide detection and adaptive information sharing algorithms for high-speed train navigation systems. *IEEE Trans Intell Transport Syst* 2015; 16(6): 3193–3203. DOI: [10.1109/TITS.2015.2437899](https://doi.org/10.1109/TITS.2015.2437899).
33. Jordan MI and Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science* 2015; 349(6245): 255–260. DOI: [10.1126/science.aaa8415](https://doi.org/10.1126/science.aaa8415).
34. Choi RY, Coyner AS, Kalpathy-Cramer J, et al. Introduction to machine learning, neural networks, and deep learning. *Transl Vis Sci Technol* 2020; 9(2): 14. DOI: [10.1167/tvst.9.2.14](https://doi.org/10.1167/tvst.9.2.14).
35. Alpaydin E. *Machine learning. The MIT press essential knowledge series*. Cambridge, MA: The MIT Press, 2021.
36. Blasch E, Pham T, Chong CY, et al. Machine learning/artificial intelligence for sensor data fusion—opportunities and challenges. *IEEE Aero Electron Syst Mag* 2021; 36(7): 80–93. DOI: [10.1109/MAES.2020.3049030](https://doi.org/10.1109/MAES.2020.3049030).
37. Brena RF, Aguilera AA, Trejo LA, et al. Choosing the best sensor fusion method: a machine-learning approach. *Sensors* 2020; 20(8): 2350. DOI: [10.3390/s20082350](https://doi.org/10.3390/s20082350).
38. Shafique MA and Hato E. Use of acceleration data for transportation mode prediction. *Transportation* 2015; 42(1): 163–188. DOI: [10.1007/s11116-014-9541-6](https://doi.org/10.1007/s11116-014-9541-6).
39. Wang KL and Xu J. A speed regression using acceleration data in a deep convolutional neural network. *IEEE Access* 2019; 7: 9351–9356. DOI: [10.1109/ACCESS.2019.2890967](https://doi.org/10.1109/ACCESS.2019.2890967).
40. Hannink J, Kautz T, Pasluosta CF, et al. Mobile stride length estimation with deep convolutional neural networks. *IEEE J Biomed Health Inform* 2018; 22(2): 354–362. DOI: [10.1109/JBHI.2017.2679486](https://doi.org/10.1109/JBHI.2017.2679486).
41. Karlsson R and Hendeby G. Speed estimation from vibrations using a deep learning CNN approach. *IEEE Sens Lett* 2021; 5(3): 1–4. DOI: [10.1109/LESENS.2021.3055528](https://doi.org/10.1109/LESENS.2021.3055528).
42. Seethi VDR and Bharti P. CNN-Based speed detection algorithm for walking and running using wrist-worn wearable sensors. In: 2020 IEEE International Conference on Smart Computing, Virtual, 14–17 September, 2020. DOI: [10.1109/SMARTCOMP50058.2020.00064](https://doi.org/10.1109/SMARTCOMP50058.2020.00064).
43. Bharti P, De D, Chellappan S, et al. HuMAN: complex activity recognition with multi-modal multi-positional body sensing. *IEEE Trans Mobile Comput* 2019; 18(4): 857–870. DOI: [10.1109/TMC.2018.2841905](https://doi.org/10.1109/TMC.2018.2841905).
44. Brant J. *Dependable speed measurement for improved low adhesion braking*. Huddersfield: University of Huddersfield, 2021. Technical Report IRR Report 110/223.

45. Ketkar N and Moolayil J. Convolutional neural networks. In: Ketkar N and Moolayil J (eds). *Deep Learning with Python: Learn Best Practices of Deep Learning Models with PyTorch*. Berkeley, CA: Apress, 2021, pp. 197–242. DOI: [10.1007/978-1-4842-5364-9\\_6](https://doi.org/10.1007/978-1-4842-5364-9_6).
46. Ioffe S and Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France, 7–9 July, 2015.
47. Géron A. Hands-on machine learning with scikit-learn, Keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems. In: *Data science / machine learning*. 3 ed. Sebastopol, CA: O'Reilly, 2023.