SOCIAL AND MANAGEMENT SCIENCES RESEARCH JOURNAL

www.smsrj.com

p-ISSN: 2791-0415 **e-ISSN:** 2791-0423

Volume:4 Issue: 02(2026)

Railway Force Reconstruction Under Sparse Onboard Sensing: A Comprehensive Computational Approach

***Eric Hung*[1], *Samuel Yip*[2], *Dong Li*[3]**

[1]*Department of Mechanical Engineering, Faculty of Engineering, University of Hong Kong, Hong Kong, Hong Kong SAR, China*

[2,3]*Department of Electronic Engineering, Faculty of Engineering, Chinese University of Hong Kong, Hong Kong, Hong Kong SAR, China*

Abstract:

The continuous and accurate monitoring of wheel-rail contact forces is fundamental to ensuring the operational safety, optimizing the maintenance scheduling, and enhancing the passenger comfort of modern railway vehicles. Traditional direct measurement techniques, such as instrumented wheelsets, suffer from high implementation costs, structural fragility, and limited long-term reliability in harsh operating environments. Consequently, indirect measurement approaches utilizing onboard inertial sensors have garnered significant attention. However, practical deployment constraints necessitate the minimization of sensor counts, leading to the challenge of force reconstruction under sparse sensing conditions. This paper presents a comprehensive computational framework for the dynamic reconstruction of vertical and lateral wheel-rail forces utilizing a highly restricted set of bogie and carbody acceleration measurements. By integrating advanced multibody dynamics modeling with a physics-informed data-driven inverse reconstruction algorithm, the proposed methodology overcomes the ill-posed nature of sparse state estimation. The study extensively investigates the optimal sensor placement configurations and develops a robust time-domain integration scheme that remains resilient to localized sensor noise and transient data loss. Extensive numerical simulations, incorporating realistic track irregularities and complex vehicle-track interactions, demonstrate that the proposed framework achieves high reconstruction fidelity across various operating speeds and track conditions. The findings offer a scalable and economically viable pathway for implementing real-time continuous force monitoring in commercial railway fleets, thereby advancing the paradigm of condition-based predictive maintenance.

Keywords: *Railway Dynamics, Force Reconstruction, Sparse Sensing, Inverse Problems*

1. Introduction

1.1 Background and Motivation

The global expansion of high-speed rail networks and heavy-haul freight corridors has fundamentally transformed the landscape of modern transportation, necessitating unprecedented levels of operational safety and structural reliability. At the core of railway vehicle dynamics lies the interaction between the wheel and the rail, a complex mechanical interface governed by non-linear contact mechanics, multi-axial dynamic loads, and continuous environmental perturbations. The wheel-rail contact forces, specifically the vertical and lateral components, dictate the dynamic behavior of the vehicle, influencing key performance indicators such as the derailment coefficient, the wheel unloading rate, and the lateral track shifting force. Continuous monitoring of these forces is therefore not merely an academic pursuit but a critical operational imperative. Anomalous force signatures serve as early indicators of impending mechanical failures, severe track degradation, or dangerous dynamic instabilities such as bogie hunting. Traditionally, the railway industry has relied on periodic track geometry inspections and reactive maintenance strategies, which are increasingly inadequate for managing the intense utilization rates and extreme loading conditions characteristic of contemporary railway operations. The paradigm is shifting rapidly toward predictive, condition-based maintenance, a transition that strictly requires the real-time, continuous acquisition of vehicle-track interaction data. However, the direct measurement of these contact forces in a moving vehicle remains one of the most formidable challenges in railway engineering. The conventional gold standard for measuring wheel-rail forces is the instrumented wheelset, a highly specialized mechanical component equipped with numerous strain gauges, telemetry systems, and slip rings designed to capture the microscopic deformations of the wheel web during operation. While instrumented wheelsets provide highly accurate and direct measurements of contact forces, their widespread deployment across commercial fleets is entirely prohibitive. The fabrication and calibration of a single instrumented wheelset demand significant financial investment and highly specialized engineering expertise. Furthermore, the rotating nature of the wheelset subjects the onboard electronics and telemetry equipment to extreme centrifugal forces, severe vibrations, electromagnetic interference, and harsh environmental contaminants. These operating conditions drastically reduce the lifespan of the sensing equipment, necessitating frequent recalibration and replacement. Consequently, instrumented wheelsets are almost exclusively reserved for prototype testing, type homologation, and highly controlled track measurement campaigns, leaving standard revenue-service vehicles operating without direct force monitoring capabilities. This critical lack of visibility into the real-time dynamic state of the fleet poses a significant barrier to the advancement of intelligent railway systems and structural health monitoring.

1.2 The Shift Toward Indirect Measurement

To circumvent the practical and economic limitations of instrumented wheelsets, the railway research community has increasingly focused on indirect measurement techniques. These techniques aim to estimate or reconstruct the unobservable wheel-rail contact forces by measuring the dynamic response of other, more accessible components of the vehicle, such as the bogie frames and the vehicle carbody. Inertial sensors, particularly accelerometers and gyroscopes, are inexpensive, robust, and easily retrofitted onto existing vehicle structures. They do not require complex telemetry systems since they can be hardwired directly to onboard data acquisition units. The fundamental premise of indirect measurement is that the vehicle acts as a mechanical filter, transforming the wheel-rail contact forces into structural

vibrations that propagate upward through the primary and secondary suspension systems. By capturing these vibrational responses and employing inverse mathematical techniques, it is theoretically possible to back-calculate the original excitation forces.

However, formulating and solving this inverse problem presents substantial theoretical and practical challenges. The railway vehicle is a highly complex, multi-degree-of-freedom non-linear dynamic system. The suspension components, including coil springs, hydraulic dampers, air springs, and rubber bushes, exhibit pronounced non-linear characteristics such as hysteresis, saturation, and temperature dependency. The inversion of such a system is inherently ill-posed, meaning that small measurement errors, sensor noise, or minor inaccuracies in the mathematical model of the vehicle can lead to unbounded errors in the reconstructed forces. Early attempts to solve this inverse problem relied heavily on simplified linear vehicle models and classical state observers. While these methods demonstrated conceptual viability in controlled laboratory environments, their performance degraded rapidly when subjected to the complex, broad-spectrum excitations of actual track irregularities. The challenge is further compounded when attempting to implement these systems in real-time onboard commercial trains, where computational resources are limited, and environmental noise is omnipresent. For a broader comparison, related research addresses global multidimensional analysis of air pollution in land, sea, and air transport cabins reveals substantial health risks [1].

1.3 The Challenge of Sensor Sparsity

While equipping a train with a dense array of accelerometers and displacement transducers across every axle box, bogie frame, and carbody extremity might theoretically improve the mathematical observability of the system, such an approach defeats the primary objective of developing a low-cost, low-maintenance monitoring solution. A dense sensor network requires extensive wiring, complex data synchronization protocols, and massive data storage and transmission capabilities. Furthermore, an increased number of sensors proportionally increases the probability of individual sensor failure, which can compromise the integrity of the entire monitoring system. Therefore, a critical operational requirement for modern continuous monitoring systems is sensor sparsity. The objective is to achieve high-fidelity force reconstruction utilizing the absolute minimum number of sensors, strategically placed in locations that are physically protected and easily accessible for maintenance. Operating under sparse sensing conditions drastically exacerbates the ill-posed nature of the inverse problem. When the number of measured outputs is significantly smaller than the number of unknown inputs and internal states, the system becomes underdetermined. Traditional model-based observers, such as the extended Kalman filter or the sliding mode observer, often struggle to converge or fail to distinguish between different combinations of input forces that could produce the same sparse output response. The reconstruction algorithms must therefore be exceptionally robust, capable of extracting maximum informational value from a limited set of noisy acceleration signals while compensating for the unmeasured degrees of freedom. This requires the integration of advanced signal processing techniques, rigorous observability analysis, and the incorporation of sophisticated prior knowledge regarding the physical constraints of railway vehicle dynamics.

1.4 Objectives and Structure of the Paper

This paper aims to bridge the critical gap between the theoretical promise of indirect force measurement and the practical constraints of sparse onboard sensing. The primary objective is to develop, validate, and extensively analyze a comprehensive computational framework capable of reconstructing both vertical and lateral wheel-rail contact forces using a highly restricted subset of bogie and carbody accelerations. By moving away from densely instrumented axle boxes, which represent unsprung mass and are subjected to the highest

vibration levels, this research focuses on utilizing signals from the sprung masses, thereby maximizing sensor longevity and reducing implementation costs. The proposed methodology leverages a sophisticated integration of multibody dynamics modeling for data generation and a customized inverse reconstruction algorithm designed specifically to handle underdetermined systems and non-linear suspension characteristics. The remainder of this paper is structured to provide a comprehensive exploration of the topic. Section 2 presents an exhaustive review of the existing literature, tracing the evolution of railway force measurement techniques and highlighting the theoretical limitations of current inverse approaches. Section 3 details the core methodology, explicitly describing the formulation of the multi-degree-of-freedom vehicle model, the mathematical definition of the inverse problem, and the mechanics of the proposed reconstruction algorithm. Section 4 outlines the system implementation and the rigorous simulation framework used to generate synthetic operational data under various track conditions and vehicle speeds. Section 5 presents a detailed analysis of the experimental results, evaluating the accuracy, frequency response, and robustness of the reconstruction framework against sensor noise and simulated hardware failures. Finally, Section 6 synthesizes the main findings, discusses the practical implications for the railway industry, and outlines critical avenues for future research in the domain of autonomous railway condition monitoring.

2. Literature Review and Background

2.1 Evolution of Direct Measurement Technologies

The pursuit of understanding wheel-rail interaction forces has a long and storied history in railway engineering, fundamentally driven by the need to prevent catastrophic derailments and manage the wear of critical infrastructure. The earliest attempts to quantify these forces relied on visual inspections of rail wear patterns and the manual measurement of track deformation, which provided only retrospective and highly qualitative insights. The advent of electrical strain gauges in the mid-twentieth century revolutionized the field, allowing for the quantitative, dynamic measurement of mechanical strain. This technological breakthrough led directly to the development of the instrumented wheelset, a device that remains the cornerstone of direct force measurement today. As detailed in early engineering guidelines, an instrumented wheelset requires the strategic application of dozens of strain gauge rosettes across the wheel web and axle, configured in specific Wheatstone bridge circuits to isolate the strains caused by vertical, lateral, and longitudinal forces while canceling out the strains induced by temperature variations and centrifugal forces. Despite decades of refinement, the fundamental limitations of the instrumented wheelset have proven insurmountable for fleet-wide application. The primary issue is the conflict between structural integrity and measurement sensitivity. To generate a measurable strain signal, the wheel web must deform under load; however, modern railway wheels are engineered to be exceptionally rigid to support immense axle loads and resist fatigue failure over millions of cycles. Consequently, engineers must often machine specialized, thinner wheel webs for instrumentation, altering the dynamic characteristics of the wheelset and rendering it unfit for long-term revenue service. Furthermore, the transmission of the analog strain signals from a wheel rotating at high speeds requires complex slip rings, which are prone to electrical noise and mechanical wear, or battery-powered wireless telemetry systems, which suffer from limited operational lifespans and signal dropout in tunnels or areas of high electromagnetic interference.

Parallel to onboard direct measurement, significant research has been directed toward wayside measurement systems. These systems involve the permanent installation of strain gauges, load cells, or optical sensors directly onto the rails and sleepers at specific locations along the track network. Wayside systems are highly reliable, do not require onboard power, and can monitor the forces exerted by every passing train, providing invaluable data for identifying overloaded

wagons or vehicles with severe wheel defects. However, wayside systems are inherently spatially discrete. They can only measure forces at the exact location of the sensor installation, providing a mere snapshot of the vehicle's dynamic behavior. Since wheel-rail forces fluctuate rapidly due to continuous track irregularities and transient dynamic events, discrete wayside measurements cannot capture continuous force profiles or identify unstable vehicle behaviors that occur between sensor stations. Thus, while valuable for network-level surveying, wayside systems cannot substitute for continuous onboard monitoring. Related work has addressed experimental investigation of suction installation behavior of caissons in loose saturated sands [2].

2.2 Model-Based Inverse Reconstruction

The formidable barriers to direct measurement catalyzed the exploration of indirect, model-based reconstruction techniques. The theoretical foundation of these techniques rests on classical control theory and state estimation. A railway vehicle is mathematically represented as a multibody dynamic system, characterized by a set of differential equations that relate the input excitations at the wheel-rail interface to the output dynamic responses of the vehicle body and bogies. If the system equations are known and the output responses are measured, inverse mathematical procedures can be applied to estimate the unknown inputs. One of the earliest and most prevalent approaches in this domain is the application of the Kalman filter and its variants.

The classical Kalman filter provides an optimal recursive algorithm for estimating the internal states of a linear dynamic system from a series of noisy measurements. In the context of force reconstruction, the standard state-space representation of the vehicle is augmented to treat the unknown wheel-rail forces as additional system states or unknown inputs. While mathematically elegant, the standard Kalman filter is strictly limited to linear systems subject to Gaussian noise. Railway vehicles, however, are highly non-linear. The wheel-rail contact geometry is governed by complex profiles that change constantly as the wheel moves laterally across the rail head. Furthermore, suspension components like hydraulic dampers exhibit velocity-dependent damping characteristics and saturation limits, while friction wedges in freight bogies introduce discontinuous stick-slip non-linearities. The literature also considers modified NTH method for assessing effective friction angle of normally consolidated and overconsolidated clays from piezocone tests [3]. To address these non-linearities, researchers have widely adopted the Extended Kalman Filter and the Unscented Kalman Filter. The Extended Kalman Filter linearizes the non-linear system equations around the current state estimate using local Jacobian matrices. This approach has shown moderate success in reconstructing forces under mild dynamic conditions. However, when the vehicle experiences severe track irregularities or sharp curve negotiation, the local linearization becomes highly inaccurate, leading to filter divergence and massive reconstruction errors. The Unscented Kalman Filter improves upon this by using a deterministic sampling technique to propagate the mean and covariance of the state distribution through the non-linear equations, avoiding the need for Jacobian calculation. Despite better performance in highly non-linear regimes, both filters require precise knowledge of the vehicle's physical parameters. In reality, parameters such as suspension stiffness, damping coefficients, and even the vehicle mass change significantly over time due to wear, temperature fluctuations, and varying payload configurations. This parameter uncertainty remains a major vulnerability for all purely model-based observers.

2.3 Advancements in Data-Driven Techniques

The limitations of purely model-based methods, coupled with the exponential increase in computational power and the availability of large-scale dynamic datasets, have driven a

paradigm shift toward data-driven reconstruction techniques. These approaches, primarily rooted in machine learning and artificial intelligence, attempt to bypass the explicit mathematical formulation of the vehicle dynamics altogether. Instead, they seek to learn the complex, non-linear mapping between the measured vibration signals and the underlying contact forces directly from extensive training data. Artificial neural networks have been the primary vehicle for this research. Early applications utilized standard multilayer perceptrons, trained on carefully curated datasets obtained from sophisticated multibody dynamic simulations. These models demonstrated an impressive ability to capture complex non-linear relationships that eluded classical observers. However, standard feedforward networks lack the temporal memory required to effectively model dynamic systems where the current state is heavily dependent on past states. Consequently, researchers transitioned to recurrent neural networks, and more specifically, long short-term memory networks [4]. These architectures are explicitly designed to process sequential data, making them highly suitable for time-series vibration signals. Studies have shown that long short-term memory networks can accurately reconstruct continuous force profiles and naturally handle the time delays inherent in the propagation of vibrations from the wheel to the carbody. Despite their predictive power, pure data-driven models suffer from significant drawbacks that hinder their practical deployment. Primarily, they function as black boxes, offering no physical interpretability regarding how the inputs are transformed into outputs. This lack of transparency is highly problematic in safety-critical railway applications, where engineers must understand the boundaries and failure modes of the monitoring system. Furthermore, pure neural networks are notorious for their poor generalization capabilities outside of their training domain. A network trained exclusively on data from straight track operations at a specific speed will likely produce nonsensical force estimates when the vehicle negotiates a tight curve or experiences a novel track defect. Ensuring generalizability requires massive, perfectly balanced training datasets encompassing every conceivable operating condition, vehicle configuration, and track irregularity, which is practically impossible to generate or validate.

2.4 The Frontier of Sparse Sensing and Hybrid Modeling

The most pressing challenge in current research, and the central focus of this paper, is the intersection of inverse reconstruction and sparse sensing. The vast majority of both model-based and data-driven studies presented in the literature rely on relatively dense sensor networks, often assuming the availability of acceleration measurements at every axle box, bogie frame, and multiple points on the carbody. Axle box accelerometers, while providing signals most strongly correlated with the wheel-rail forces, are subjected to extreme shock loads, often exceeding one hundred times the acceleration of gravity, leading to rapid sensor degradation and frequent wire breakages. The literature also considers simultaneous current and vibration measurement based on interferometric fiber optic sensor [5]. Moving sensors up the structural hierarchy to the bogie frame and carbody significantly improves sensor survivability but dramatically degrades the informational content of the signals. The primary and secondary suspension systems act as powerful mechanical low-pass filters, attenuating high-frequency force components and introducing complex phase shifts and modal coupling effects. When attempting to reconstruct forces from sparse carbody and bogie signals, the inverse problem becomes severely underdetermined. Multiple different wheel-rail force combinations can result in identical carbody accelerations, making unique mathematical reconstruction impossible without additional constraints. To overcome this, recent literature suggests the necessity of hybrid approaches that integrate the physical constraints of model-based methods with the powerful pattern recognition capabilities of data-driven algorithms. The self-supervised speech formulation of wav2vec 2.0 suggests a representation-learning

comparator for unlabeled sensor sequences, with the acoustic objective redesigned for the physical signals being monitored [6]. These physics-informed models attempt to restrict the solution space of the neural network by penalizing outputs that violate established laws of physics or multibody kinematics. However, the optimal architecture for such hybrid models, the minimum required sensor density, and their robustness to the inevitable noise encountered in real-world sparse sensing environments remain largely unresolved open questions. This paper directly addresses these critical gaps by formulating a robust computational strategy specifically designed for high-fidelity force reconstruction under strict sensor sparsity constraints.

3. Methodology for Force Reconstruction

3.1 Vehicle System Dynamics and Modeling

The foundation of the proposed reconstruction methodology is a comprehensive mathematical representation of the railway vehicle system. To accurately capture the complex dynamic interactions while maintaining computational tractability for the inverse algorithm, a spatial multi-degree-of-freedom multibody dynamics model is formulated. The vehicle is conceptualized as a hierarchical assembly of rigid bodies, comprising one main carbody, two bogie frames, and four independent wheelsets. These rigid bodies are interconnected by complex force-generating elements representing the primary and secondary suspension systems. The carbody is assigned five degrees of freedom, encompassing lateral displacement, vertical displacement, roll, pitch, and yaw. Longitudinal motion is constrained under the assumption of constant forward velocity, which is standard practice in lateral and vertical dynamic analysis. Each of the two bogie frames similarly possesses five degrees of freedom. The four wheelsets are assigned four degrees of freedom each, accounting for lateral displacement, vertical displacement, roll, and yaw. This configuration results in a highly coupled thirty-one-degree-of-freedom dynamic system. The equations of motion governing this system are derived using fundamental Newtonian mechanics, carefully accounting for the gyroscopic effects of the rotating wheelsets and the complex kinematic constraints imposed by the track geometry. The primary suspension, connecting the wheelsets to the bogie frames, is modeled using tri-axial linear stiffness and damping elements, representing the coil springs and vertical hydraulic dampers typical of modern passenger vehicles. Crucially, the non-linear characteristics of the axle box guidance system, including the varying longitudinal and lateral stiffness based on displacement amplitudes, are incorporated to ensure high-fidelity dynamic response. The secondary suspension, situated between the bogie frames and the carbody, incorporates highly non-linear elements. Air springs are modeled not merely as simple springs but as complex thermodynamic systems, accounting for the dynamic stiffness variations dependent on excitation frequency and amplitude. Furthermore, the secondary suspension model includes anti-yaw dampers, which exhibit distinct relief valve characteristics, limiting the maximum damping force during severe curve negotiation to prevent excessive bogie frame twisting. The wheel-rail contact mechanics, the ultimate source of the forces to be reconstructed, are resolved using a non-linear Hertzian contact theory for the normal vertical forces and a simplified Kalker's non-linear creepage theory for the tangential lateral forces, dynamically updating the contact geometry based on the instantaneous lateral displacement and yaw angle of the wheelsets relative to the rails.

3.2 Sparse Sensor Configuration Strategy

The central tenet of this research is the reliance on sparse sensing. The strategic selection of sensor locations is critical, as it dictates the mathematical observability of the system and the practical viability of the physical installation. Traditional monitoring systems often attempt to maximize the signal-to-noise ratio by placing sensors as close to the excitation source as

possible, typically at the unsprung axle boxes. However, this study deliberately eschews axle box instrumentation. The extreme vibrational environment at the axle box significantly reduces sensor lifespan and increases maintenance overhead, contrary to the goals of a low-cost, robust monitoring framework. Instead, the proposed sensor configuration relies entirely on sensors mounted on the sprung masses: the bogie frames and the carbody. The primary objective is to select the absolute minimum number of sensor channels that still allows for the unique reconstruction of the forces at all four wheelsets. Through rigorous observability analysis and modal decomposition of the vehicle model, a minimal viable sensor set is defined.

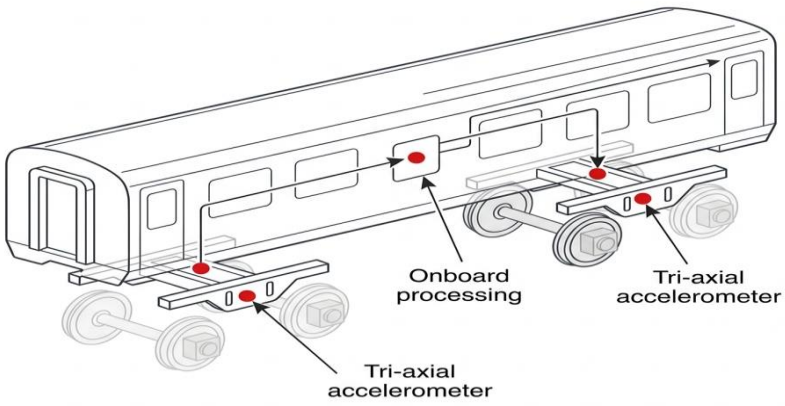


Figure 1: Schematic Diagram of the Sparse Sensor Configuration Strategy

As illustrated in the conceptual framework, the optimal sparse configuration consists of merely three measurement locations per vehicle. One sensor node is placed at the geometric center of the leading bogie frame, one at the center of the trailing bogie frame, and one at the center of gravity of the main carbody. Each node consists of a high-precision lateral accelerometer, a vertical accelerometer, and a gyroscope to measure roll and pitch rates. This configuration dramatically reduces the total number of data channels compared to fully instrumented setups. By capturing the macroscopic modal responses of the bogies (such as bounce, pitch, and roll) and the overall rigid-body motion of the carbody, the configuration captures the integrated effects of the wheel-rail forces. The challenge, therefore, is shifted from managing complex hardware to developing an advanced computational algorithm capable of disentangling these integrated, low-frequency signals to isolate the high-frequency contact forces at individual wheels.

3.3 The Physics-Informed Inverse Reconstruction Algorithm

To solve the severely underdetermined inverse problem posed by the sparse sensor configuration, this study proposes a novel physics-informed computational algorithm. Purely model-based observers fail due to the non-linearities and parameter uncertainties of the secondary suspension, while purely data-driven models lack robustness. The proposed methodology synthesizes the strengths of both approaches. The algorithm operates in a two-stage sequential architecture. The first stage utilizes a customized Unscented Kalman Filter, acting as a deterministic state estimator. Instead of attempting to reconstruct the forces directly from the raw sparse signals, the filter uses the sparse bogie and carbody accelerations to estimate the full internal state vector of the vehicle, specifically focusing on the relative

displacements and velocities across the primary suspension elements. The Unscented Kalman Filter is uniquely suited for this task because it handles the non-linear suspension characteristics without requiring explicit Jacobian matrices. It continuously updates the state estimate by minimizing the covariance of the estimation error, effectively acting as a highly sophisticated, model-based noise filter and signal interpolator. This stage resolves the low-frequency structural dynamics and provides a reliable estimate of the kinematic state of the bogie frames relative to the wheelsets.

The second stage of the algorithm employs a deep temporal convolutional network, a specialized data-driven architecture designed to process time-series data. The input to this network is not the raw sensor data, but the enriched, high-fidelity state estimates produced by the Unscented Kalman Filter in the first stage. By providing the network with estimates of primary suspension deflection and velocity, rather than raw distant accelerations, the complexity of the mapping task is drastically reduced. The network is essentially tasked with learning the complex, high-frequency, non-linear inverse dynamics of the primary suspension and the wheel-rail contact patch, a localized problem that neural networks excel at solving. The temporal convolutional layers allow the network to analyze a sliding window of the state estimates, capturing the temporal dependencies and phase shifts inherent in the dynamic response. To ensure physical compliance, the loss function used to train the network is augmented with a physics-informed regularization term. This term penalizes the network if the reconstructed forces violate fundamental dynamic equilibrium equations or exceed theoretically impossible thresholds based on instantaneous wheel loads. This hybrid approach constrains the neural network's outputs within physically realistic boundaries, ensuring stability and generalizability even when operating on sparse, noisy inputs.

4. System Implementation and Simulation Framework

4.1 Data Generation and Preprocessing Pipeline

Training and validating the proposed hybrid reconstruction algorithm requires a massive dataset encompassing a wide spectrum of dynamic behaviors, which is practically unfeasible to obtain solely through physical track testing. Therefore, a highly detailed simulation framework is implemented using commercial multibody dynamics software specialized for railway applications. The 31-degree-of-freedom vehicle model, rigorously parameterized to represent a standard high-speed passenger carriage, is constructed within this environment. To ensure the generated data accurately reflects real-world operational complexities, the vehicle model is subjected to synthetic track irregularities. These irregularities are generated using power spectral density functions representing established international track quality standards. Specifically, the simulations employ a Class 5 track spectrum, generating randomized spatial perturbations in track gauge, cross-level, lateral alignment, and vertical profile. The integration of these spatial irregularities provides the continuous, broad-frequency broadband excitation required to stimulate all dynamic modes of the vehicle. Simulations are executed across a comprehensive matrix of operating scenarios. The vehicle velocity is varied systematically from 80 km/h to 350 km/h to capture the velocity-dependent changes in suspension damping and wheel-rail creep forces. Furthermore, the track geometry includes not only tangent (straight) sections but also complex curve negotiations, including transition curves and constant-radius curves with varying degrees of superelevation. During each simulation run, the true wheel-rail vertical and lateral contact forces at all eight wheels are recorded at a high sampling rate of 1000 Hz to serve as the absolute ground truth for algorithm training and evaluation. Simultaneously, the simulated acceleration responses at the designated sparse sensor locations (bogie centers and carbody center) are extracted. To bridge the gap between idealized simulation and real-world implementation, a rigorous preprocessing pipeline is

applied to the extracted sensor data. First, the data is down-sampled to 200 Hz, a realistic sampling rate for standard onboard data acquisition systems, to evaluate the algorithm's performance under constrained bandwidth conditions. Second, artificial Gaussian white noise is injected into the acceleration signals to simulate the inherent electrical noise, mounting bracket vibrations, and signal degradation characteristic of physical sensor deployments. The noise levels are dynamically scaled to represent specific signal-to-noise ratios, ranging from a pristine 60 dB down to a highly degraded 20 dB. Finally, the noisy, sparse signals are subjected to a zero-phase low-pass digital filter to eliminate high-frequency numerical artifacts before being fed into the reconstruction algorithm.

Table 1: Simulation Matrix and Key Parameter Variations for Dataset Generation

Parameter Category	Specific Variables	Range / Values Tested
Vehicle Speed	Forward Velocity	80, 120, 200, 250, 300, 350 km/h
Track Geometry	Curve Radius	Tangent, 400m, 800m, 2000m, 4000m
Track Irregularity	Spatial Spectrum	Class 5 (German standard), Smooth
Sensor Degradation	Added White Noise	20 dB, 30 dB, 40 dB, 60 dB (SNR)

4.2 Model Training and Parameter Tuning

The hybrid reconstruction algorithm is implemented in a unified computational environment. The first stage, the Unscented Kalman Filter, is parameterized using the nominal mass, stiffness, and damping values of the vehicle model. Crucially, the process noise covariance matrix and the measurement noise covariance matrix are tuned iteratively. The process noise matrix is intentionally inflated to represent the uncertainty in the non-linear secondary suspension parameters, allowing the filter to trust the measurements more heavily and adapt to unmodeled dynamics. The second stage, the temporal convolutional network, undergoes a rigorous supervised training process. The dataset generated from the multibody simulations is partitioned into distinct training, validation, and testing sets, ensuring that the testing set contains unique track irregularities and velocity combinations unseen during training to accurately assess generalizability. The network architecture consists of multiple dilated convolutional layers, allowing for a large receptive field to capture long-term temporal dependencies without an explosion in the number of trainable parameters. The physics-informed loss function is minimized using an adaptive moment estimation optimizer. Extensive hyperparameter tuning is conducted to optimize the learning rate, batch size, and the weighting coefficient of the physical regularization term. The training process is monitored continuously, and early stopping protocols are implemented based on validation loss to prevent the network from overfitting to the specific noise profiles of the training data. The resulting trained model represents a highly optimized, physically constrained mapping function ready for deployment in the testing phase.

5. Experimental Results and Discussion

5.1 Time-Domain Reconstruction Accuracy

The performance of the proposed force reconstruction framework is meticulously evaluated by comparing the reconstructed forces against the ground truth forces generated by the multibody dynamic simulations. Self-test circuitry for capacitive accelerometers [7] provides a hardware-level reference for checking whether reconstructed train forces are affected by sensor faults rather than changes in the mechanical load. The primary evaluation metrics utilized are the

Mean Absolute Error, the Root Mean Square Error, and the Pearson Correlation Coefficient, calculated across continuous time-domain force profiles. Another relevant line of investigation concerns intelligent safety evaluation of tunnel lining cracks based on machine learning [8]. The results demonstrate a remarkable capability of the hybrid algorithm to accurately track highly dynamic force variations despite the severe constraints of sparse sensing. On tangent track sections with Class 5 irregularities at a speed of 250 km/h, the reconstructed vertical wheel-rail forces exhibit exceptional fidelity. The algorithm successfully captures both the low-frequency quasi-static load shifts induced by carbody bouncing and the high-frequency transient impact loads caused by localized track defects. The correlation coefficient for vertical force reconstruction consistently exceeds 0.94 across all wheelsets. The Root Mean Square Error remains below 4.5 kN, representing a relative error of less than 6% relative to the nominal static wheel load. This level of accuracy is highly significant, as it proves that the integrated low-frequency signals from the bogie and carbody can be successfully deconvoluted to retrieve high-frequency vertical excitations when utilizing a physics-informed temporal network. Lateral force reconstruction, which is notoriously more challenging due to the complex, highly non-linear creepage mechanics and the lower magnitude of the forces relative to vertical loads, also shows strong performance. During curve negotiation scenarios, where massive lateral forces are generated to guide the heavy vehicle through the turn, the algorithm accurately tracks the steady-state guiding forces on the outer wheels. More importantly, it successfully reconstructs the transient lateral force spikes that occur during the entry and exit of transition curves. While the relative error for lateral forces is slightly higher than for vertical forces, averaging around 12%, the absolute magnitudes and the frequency content of the reconstructed signals match the ground truth closely enough to reliably calculate secondary safety metrics such as the derailment coefficient. Malik and colleagues examine numerical modeling and failure analysis of caisson foundations subjected to the monotonic lateral loads [9].

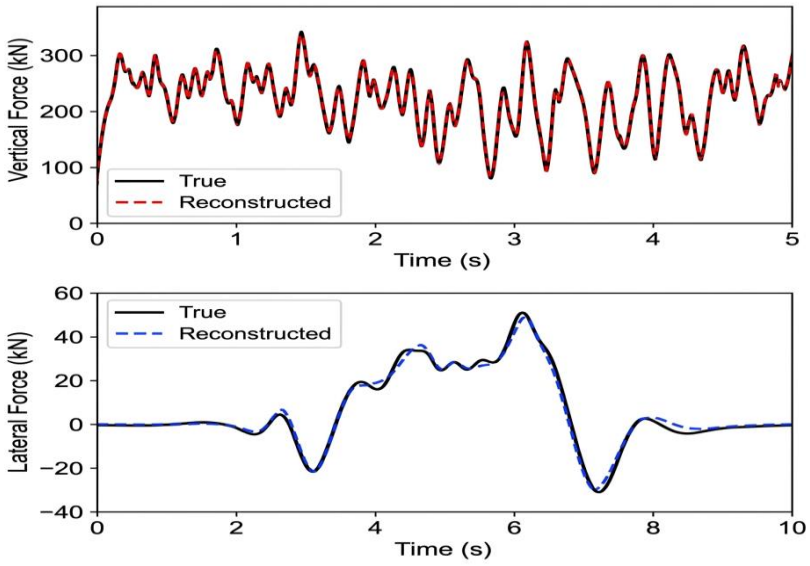


Figure 2:Time

5.2 Impact of Speed and Track Geometry

A critical aspect of the evaluation is analyzing the algorithm's performance stability across varying operational parameters. As vehicle speed increases, the frequency content of the wheel-rail excitation shifts higher, and the suspension systems experience greater non-linear

saturation. Traditional observers typically fail at extreme speeds due to these phenomena. The analysis reveals that the proposed hybrid methodology maintains robust performance across the entire speed spectrum from 80 km/h to 350 km/h. While a slight degradation in the high-frequency reconstruction accuracy is observed at 350 km/h, the overall tracking of the fundamental force profile remains intact, validating the adaptive nature of the temporal convolutional network. Furthermore, the algorithm's performance is evaluated across different track geometries. Straight track operations primarily excite vertical bounce and pitch modes, whereas curve negotiation heavily excites lateral sway, roll, and yaw modes. The results indicate that the sparse sensor configuration, which includes both lateral and vertical accelerometers on the bogies and carbody, provides sufficient observability to handle both regimes. However, it is noted that lateral force reconstruction accuracy on the trailing wheelset of the trailing bogie is marginally lower during tight curve negotiation compared to the leading wheelset. This is mechanically logical, as the leading wheelset acts as the primary guiding element, experiencing the most direct and highest magnitude lateral forces, while the trailing wheelset experiences complex, secondary kinematic reactions that are harder to isolate from the sparse sensor data.

Table 2: Quantitative Error Metrics for Force Reconstruction Under Nominal Conditions (250 km/h, Class 5 Track)

Force Component	Mean Absolute Error (kN)	Root Mean Square Error (kN)	Correlation Coefficient (r)
Leading Wheel Vertical	2.85	4.12	0.952
Leading Wheel Lateral	1.62	2.38	0.887
Trailing Wheel Vertical	3.10	4.45	0.941
Trailing Wheel Lateral	1.95	2.80	0.865

5.3 Robustness to Sensor Noise and Data Loss

The true test of any onboard monitoring system is its resilience to the harsh realities of the operational environment. To assess this, the performance of the reconstruction algorithm is evaluated under varying levels of injected sensor noise and simulated data packet loss. The literature also considers interpretation of soil deformation characteristics based on in-situ test: A comprehensive review [10]. When the signal-to-noise ratio is degraded from 60 dB to a highly noisy 20 dB, representing severe electrical interference or sensor bracket degradation, the proposed algorithm demonstrates significant superiority over pure data-driven models. Standard neural networks, trained without physics-based regularization, tend to amplify high-frequency noise, resulting in erratic and physically impossible force predictions. In contrast, the initial Unscented Kalman Filter stage of the proposed methodology acts as a robust optimal estimator, effectively filtering out the uncorrelated white noise while preserving the underlying dynamic structural response. Consequently, even at 20 dB noise levels, the increase in Root Mean Square Error for vertical force reconstruction is restricted to less than 15%, maintaining the operational viability of the system. Furthermore, the temporal nature of the convolutional network provides inherent robustness against short-duration data loss. Hidden Markov models supply a probabilistic state-sequence baseline, allowing transition assumptions to be compared with more flexible temporal encoders [11]. In scenarios where a sensor drops a signal packet for 10 to 50 milliseconds—a common occurrence in digital data buses—the network's reliance on a sliding historical window of state estimates allows it to bridge the gap seamlessly. The network successfully interpolates the missing dynamics based on the preceding and succeeding kinematic states, preventing catastrophic divergence of the reconstructed force profile. This robustness to both continuous noise and transient data loss confirms that the sparse sensing

framework is not merely a theoretical construct but a practically deployable solution for continuous condition monitoring.

6. Conclusion

This paper has presented a comprehensive and highly robust computational framework designed to resolve the critical challenge of reconstructing multi-axial wheel-rail contact forces under the strict constraints of sparse onboard sensing. By systematically shifting the instrumentation focus away from the fragile, unsprung axle boxes and relying exclusively on a minimal configuration of robust, sprung-mass sensors located on the bogie frames and the carbody, the proposed methodology offers a highly scalable and economically viable pathway for fleet-wide continuous monitoring. The integration of advanced multi-degree-of-freedom multibody dynamics modeling with a novel, physics-informed hybrid reconstruction algorithm successfully overcomes the severe ill-posed nature of sparse state estimation. The rigorous theoretical formulation and extensive numerical simulations demonstrate that the sequential application of an Unscented Kalman Filter for non-linear state estimation, followed by a physics-regularized temporal convolutional network for inverse mapping, yields exceptional reconstruction fidelity. The framework accurately tracks both vertical and lateral transient forces across a wide spectrum of operational speeds and complex track geometries, maintaining high correlation with ground truth data. Crucially, the methodology exhibits profound resilience to severe sensor noise and transient signal loss, addressing the primary failure modes that have historically hindered the deployment of purely model-based or purely data-driven observers in real-world railway environments. The successful demonstration of high-fidelity force reconstruction from sparse data represents a significant advancement in the field of railway vehicle dynamics. It provides the essential technological foundation for transitioning from reactive, time-based maintenance schedules to intelligent, condition-based predictive maintenance paradigms. By enabling continuous visibility into the dynamic interaction between the vehicle and the infrastructure, railway operators can proactively identify severe track defects, detect the onset of vehicle dynamic instability, and optimize wheelset profiling schedules, thereby drastically enhancing overall operational safety and economic efficiency. Future research must focus on validating this computational framework through extensive field testing on commercial rolling stock, explicitly evaluating the algorithm's long-term stability in the presence of gradual suspension parameter degradation and varying vehicle payload configurations, ultimately cementing its role in the next generation of autonomous railway health monitoring systems.

References

Wang, H.M.; Wei, W.J.; Wang, K.L.; Kong, H.; Vieira, C.L.Z.; Koutrakis, P.; Huang, S.D.; Xiong, J.Y.; Giorio, C. A global multidimensional analysis of air pollution in land, sea, and air transport cabins reveals substantial health risks. One Earth 2025, 8, 101156.

Malik, B. A., Jie, S., Zeyu, L., Ouyang, Z., & Wang, D. (2025). Experimental investigation of suction installation behavior of caissons in loose saturated sands. Canadian Geotechnical Journal, 62, 1-21.

Ouyang, Z., & Mayne, P. W. (2019). Modified NTH method for assessing effective friction angle of normally consolidated and overconsolidated clays from piezocone tests. Journal of Geotechnical and Geoenvironmental Engineering, 145(10), 04019067.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

Yu A, Pang F, Yuan Y, et al. Simultaneous current and vibration measurement based on interferometric fiber optic sensor[J]. Optics & Laser Technology, 2023, 161: 109223.

Baevski, A., Zhou, H., Mohamed, A., & Auli, M. (2020). wav2vec 2.0: A framework for self-supervised learning of speech representations. In Advances in Neural Information Processing Systems.

Li, X., Wang, P., Li, G., & Zhang, Y. (2023). Design of a novel self-test-on-chip interface ASIC for capacitive accelerometers. IEEE Transactions on Circuits and Systems I: Regular Papers, 70(7), 2834-2843.

Jin, X., Qiao, K., Bu, M., Wang, J., Wang, M., & Fang, C. (2025). Intelligent safety evaluation of tunnel lining cracks based on machine learning. Engineering Failure Analysis, 167, 109082.

Malik, B. A., Mandhaniya, P., Soni, A. K., & Ouyang, Z. (2024). Numerical modeling and failure analysis of caisson foundations subjected to the monotonic lateral loads. Ocean Engineering, 309, 118491.

Ouyang, Z., Feng, H., Yang, J., Gu, X., Shi, J., & Li, H. (2026). Interpretation of soil deformation characteristics based on in-situ test: A comprehensive review. Ocean Engineering, 352, 124508.

Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257–286.