On robust detection of hunting on railway vehicles via multiple vibration sensors

K. Kritikakos, S.D. Fassois, J.S. Sakellariou
University of Patras, Department of Mechanical Engineering & Aeronautics,
Stochastic Mechanical Systems & Automation (SMSA) Laboratory,
26504 Patras, Greece
e-mail: **fassois@upatras.gr**

Abstract

The possibility of attaining enhanced rail hunting detection performance, primarily in terms of early detection and robustness to suspension faults or worn track conditions, via multi-sensor (vector) methods is considered. Towards this end two multi-sensor extensions of recently introduced single-sensor methods are postulated based upon Recursive Vector AutoRegressive (RVAR) modeling: A Vector Degree-of-Stochasticity (V-DS) based method and a Vector Damping Ratio (V-DR) counterpart. Their performance is systematically assessed via Monte Carlo simulations using a SIMPACK-based vehicle model and considering the True Positive Rate (TPR), the False Positive Rate (FPR), and the Detection Delay Time (DDT) with respect to the conventional hunting onset. The methods are shown to achieve impressive performance, with the V-DS reaching the ideal 100% TPR @ 0% FPR and sample mean DDT of $-1.1\,s$, and the V-DR method following closely. This performance drastically surpasses that of their single-sensor counterparts and a state-of-the-art method.

1 Introduction

Railway vehicle hunting (lateral instability) is a coupled lateral and yaw periodic limit cycle motion which is physically limited by the wheel flanges [1]. It occurs at a traveling speed v_h, referred to the hunting speed, which depends on the particular track conditions, but cannot be lower than a certain threshold, referred to as the critical speed v_{cr} [2]. The on-board early detection of hunting on railway vehicles is important for reasons associated with safety, ride comfort, and rolling stock life cycle.

For these reasons a number of vibration-based on-board hunting detection methods have been developed over the past several years [3–11]. They may be distinguished into 'standards-based' or 'dynamics-based'. 'Standards-based' methods follow available standards (see Table 1) and detect hunting as the bogie or carbody lateral vibration exceeds a certain limit under specific terms. They are simple, yet not necessarily best in terms of achievable performance, such as correct hunting detection, elimination of false alarms, and early detection. Furthermore, their performance may be further aggravated under certain conditions, such as worn tracks or suspension faults, suggesting insufficient robustness [12].

'Dynamics-based' methods attempt to exploit changes in the dynamics introduced by hunting via a proper dynamic quantity. As hunting is a limit cycle oscillation, the underlying dynamics is characterized by 'reduced damping' [6, 7], and the resulting vibration signal by 'structural simplicity' (for instance being close to a sinusoid) [8–11]. This condition also gives rise to 'high correlation/similarity' between two vibration signals [8], or may be viewed as a 'reduced entropy/stochasticity' state [9–11].

'Dynamics-based' methods may be further classified as single-sensor or multi-sensor based according to whether one or more sensors (typically accelerometers) are employed. A particular single-sensor-based method employs an entropy based dynamic quantity, with the hunting condition being associated with 'reduced entropy' in the bogie vibration signal [10]. Another is based upon Lyapunov exponent concepts, with the onset of hunting being associated with a Lyapunov exponent close to zero which signifies limit cycle oscillation and thus also 'reduced stochasticity' in the bogie or the axle-box vibration signal [11].

Table 1: Hunting detection standards [3–5].

Standard	Hunting Quantity	Sensor Location	Frequency Range (Hz)	Limit (m/s^2)
UIC-515	Peak Value	Bogie	[4–8]	$8^{\diamond}$
49 CFR 213	Root Mean Square (RMS)*	Bogie	[1–10]	3.92
AAR	Standard Deviation (SD)$^{\triangledown}$	Car-body	[1–15]	1.28

$^{\diamond}$The limit shall be exceeded for at least 6 consecutive times.

*The RMS is calculated over a 2 s period.

$^{\triangledown}$The SD is calculated over a 610 m traveled distance.

Multi-sensor-based methods employ two or more sensors in an effort to improve detection performance via additional information. One such method employs the spectral density information of the axle-box lateral and longitudinal vibration signals, with the hunting condition being associated with a spectral peak increase at the dominant frequency which signifies 'reduced damping' [6]. A second method employs an energy-type dynamic quantity and three sensors (two on the bogie and one on the axle-box), with hunting being associated with energy increase around the hunting frequency, thus also signifying 'reduced damping' [7]. Another method is based upon the maximum cross-correlation coefficient between bogie and car-body vibration signals at lags up to 1 s, with the onset of hunting being associated with 'high correlation/similarity' between the signals obviously due to the limit cycle oscillation [8]. A fourth method employs an entropy based dynamic quantity and three sensors (two on the bogie and one on the axle-box), with the hunting condition being associated with 'reduced entropy' [9]. Some multi-sensor-based methods [7, 9] operate within a Support Vector Machine (SVM) classification context, attempting to classify a selected dynamic quantity into various states such as normal (no hunting), small amplitude hunting, and (fully developed) hunting, and also employ fusion techniques (for instance based on the Dempster-Shafer theory [9]) for integrating the information from multiple sensors. It is mentioned that small-amplitude hunting [7, 9, 10] refers to the hunting condition in which the bogie lateral vibration amplitude surpasses the limit of $2\,m/s^2$ but remains lower than that of $8\,m/s^2$ required for full hunting according to the standards [3].

Despite the progress attained so far, the aforementioned methods are still characterized by certain performance limitations. These are related to their non-parametric nature, but also to the fact that estimation is based on signal portions within rectangular sliding windows; a strategy involving a compromise between adaptability (for which a short window is needed) and attainable accuracy (for which a long window is needed). Moreover, certain methods involve the subjective selection of a number of critical operational parameters via simulated or physical experiments [6, 7, 9, 11], which may also have its own difficulties [6]. Another issue is the computational complexity, which may, in certain cases, be high for online implementation [7, 11].

In addition, although of obvious importance, hunting detection robustness to suspension faults or worn track conditions has received very limited attention. The problem arising due to such conditions is that they may lead to hunting false alarms (in case of increased lateral vibration) or to missed hunting detection (in case of decreased lateral vibration). Also there is a lack of systematic and reliable assessments based on sufficiently high numbers of Test Cases within a Monte Carlo framework and the use of proper assessment metrics such as the True Positive Rate (TPR), that is the correct hunting detection rate, the False Positive Rate (FPR), that is the false hunting detection rate, or, else, the classification Accuracy (ACC), that is the correct decision rate, and the Detection Delay Time (DDT) associated with the hunting onset. Only the ACC metric is reported in a few recent studies, and specifically ACCs of 79.8% [9] and 97.2% [7] are reported for single-sensor-based methods and 94.6% [9], 98.7% [6] and 100% [7] for multi-sensor counterparts.

In an effort to overcome the aforementioned limitations, two novel single-sensor 'dynamics-based' methods have been postulated by the present authors [12, 13]. They are parametric, founded upon Recursive AutoRegressive (RAR) modeling of the scalar lateral car-body random vibration signal, with the first one being based on a Degree-of-Stochasticity measure (the S-DS method, with the prefix 'S' indicating that it is single-sensor-based) and the second on the lowest Damping Ratio associated with one of the hunting frequencies (the S-DR method). Their performance has been systematically assessed via Monte Carlo simulations and

Table 2: Main characteristics of the postulated and compared methods.

Method	Principle	Sensor(s)
Vector **D**egree-of-**S**tochasticity (V-DS)	'reduced stochasticity'	car-body lateral & bogie yaw
Vector **D**amping **R**atio (V-DR)	'reduced damping'	
Scalar **D**egree-of-**S**tochasticity (S-DS) [12, 13]	'reduced stochasticity'	car-body lateral
Scalar **D**amping **R**atio (S-DR) [12]	'reduced damping'	
Vector **E**ntropy - **S**upport **V**ector **M**achine (V-E-SVM) [9]	'reduced entropy'	car-body lateral & bogie yaw

it has been shown that superior damage detection, reaching 100% TPR @ 0% FPR and low DDT (sample mean $\pm$ std of $2.6\,\mathrm{s} \pm 2.0\,\mathrm{s}$), is achievable. The methods have been also shown to achieve good robustness to suspension faults and worn track conditions.

Motivated by this work, the present study aims at examining whether or not further performance improvements, mainly in terms of early detection of hunting and robustness, may be attained by multi-sensor (vector) versions of the methods. Such versions, designated as Vector Degree-of-Stochasticity (V-DS) based and Vector Damping Ratio (V-DR), are presently postulated using non-stationary Recursive Vector AutoRegressive (RVAR) parametric modeling of the car-body lateral and bogie yaw acceleration signals; this selection being motivated by the nature of hunting as a coupled lateral and yaw motion of the vehicle.

The performance of the postulated methods is systematically and thoroughly assessed via extensive Monte Carlo simulations under four Scenarios (Scenarios A, B, C, D) using a SIMPACK-based [14] passenger vehicle model. Their performance is compared with that of their single-sensor (scalar) counterparts and a multi-sensor state-of-the-art method based on an Entropy-type quantity and Support Vector Machine (SVM) classification [9], which is presently referred to as the V-E-SVM-based method and operates on the exact same signals. The main characteristics of the postulated methods, along with those used in the comparisons, are summarized in Table 2.

The main questions posed in the study are:

(i) Are the postulated multi-sensor (vector) V-DS and V-DR parametric methods capable of surpassing their single-sensor (scalar) S-DS and S-DR counterparts in terms of Detection Delay Time (DDT) and robustness to suspension faults and worn track conditions?

(ii) How do the two postulated methods compare to each other in terms of achievable performance?

(iii) How does the performance of the postulated methods compare to that of a multi-sensor state-of-the-art Entropy based method using the exact same vibration signals?

(iv) Are the postulated methods suitable for real time implementation?

The rest of the study is organized as follows: The precise problem statement and the multi-sensor hunting detection methods are outlined in section 2. A brief overview of the SIMPACK-based vehicle model and the Monte Carlo simulations are provided in section 3. A comparative performance assessment of the methods via Monte Carlo simulations is provided in section 4, and concluding remarks are summarized in section 5.

2 Problem statement and the multi-sensor-based hunting detection methods

The hunting early detection problem may be posed as follows: At each time instant t, *given* the past and present vector (s-variate, that is of dimensionality s) random vibration signal samples[1] $\boldsymbol{y}[k]$ ($k = 1, 2, \ldots, t$;

[1]Column-vectors/matrices are designated via bold-face lower/upper case symbols.

with t and k designating normalized, by the sampling period, discrete time; $\dim \boldsymbol{y} = s$), detect the onset of hunting (*Operational Phase*). The method's operational parameters are selected off-line in an initial *Tuning Phase* using a set of q vector random vibration signals $\boldsymbol{y}_j[k]$ ($k = 1, 2, \ldots, N$; $j = 1, 2, \ldots, q$; with N designating the signal length in samples), with a number of them including the hunting onset.

2.1 The multi-sensor-based hunting detection methods

As already mentioned, two multi-sensor (vector) 'dynamics-based' hunting detection methods are postulated: A Vector Degree-of-Stochasticity (V-DS) based and a Vector Damping Ratio (V-DR) based. They are both founded upon parametric stochastic adaptive Recursive Vector AutoRegressive (RVAR) modeling of the car-body lateral and bogie yaw acceleration signals, as it is capable of instantaneously and effectively tracking the random and non-stationary signal characteristics, with non-stationarity due to the varying traveling speed and the potential hunting onset. An s-variate RVAR(n) model is of the form [15, pp. 366-367]:

$$\boldsymbol{y}[k] + \sum_{i=1}^{n} \boldsymbol{A}_i[k] \cdot \boldsymbol{y}[k-i] = \boldsymbol{e}[k], \qquad E(\boldsymbol{e}[k] \cdot \boldsymbol{e}^T[k]) = \boldsymbol{\Sigma}_e[k] \qquad (k = 1, 2, \ldots) \tag{1}$$

with ^{T} designating transposition, $\boldsymbol{y}[k] = [y_1[k] \ldots y_s[k]]^T_{s\times 1}$ the vector random vibration signal (presently $s = 2$; $y_1[k]$ the car-body lateral and $y_2[k]$ the bogie yaw acceleration), and $\boldsymbol{e}[k] = [e_1[k] \ldots e_s[k]]^T_{s\times 1}$ the innovations vector (one-step-ahead prediction error) signal that should be zero-mean serially uncorrelated (white) with time-dependent covariance matrix $\boldsymbol{\Sigma}_e[k]$. $\boldsymbol{A}_i[k]$ ($s \times s$) are the time-dependent AutoRegressive (AR) matrices, with n representing the AR order, and $E(\cdot)$ statistical expectation. At each time instant k, RVAR model estimation is accomplished via the Recursive Least Squares (RLS) algorithm [15, pp. 366-367].

2.2 A Vector Degree-of-Stochasticity (V-DS) based method

The Vector Degree-of-Stochasticity (V-DS) based method is founded upon the RVAR model's instantaneous determinant of the innovations (residuals) covariance matrix (det$\boldsymbol{\Sigma}_e$) over the instantaneous determinant of the vibration signal covariance matrix (det$\boldsymbol{\Sigma}_y$). It essentially constitutes a scalar measure of the degree of stochasticity of the vector signal, with 0 corresponding to determinism and 1 to a pure randomness (lack of serial correlation).

Within the *Operational Phase* of the method and at each time instant t, the innovations covariance matrix ($\boldsymbol{\Sigma}_e[t]$) is estimated via a rectangular sliding window acting on the M most recent innovation signal samples, that is[2]:

$$\widehat{\boldsymbol{\Sigma}}_e[t] = \frac{1}{M} \sum_{k=t-M+1}^{t} \widehat{\boldsymbol{e}}[k, k-1] \cdot \widehat{\boldsymbol{e}}^T[k, k-1] \tag{2}$$

with $\widehat{\boldsymbol{e}}[k, k-1]$ designating the vector innovations (residual) signal at time k recursively estimated using the RVAR model parameters at time $k-1$.

The covariance matrix of the measured vector vibration signal ($\boldsymbol{\Sigma}_y[t]$) is estimated as:

$$\widehat{\boldsymbol{\Sigma}}_y[t] = E\{\boldsymbol{y}[k] \cdot \boldsymbol{y}^T[k]\} = \sum_{j=0}^{\infty} \boldsymbol{G}_j[t] \cdot \widehat{\boldsymbol{\Sigma}}_e[t] \cdot \boldsymbol{G}_j^T[t] \tag{3}$$

with $\boldsymbol{G}_j[t]$ ($s \times s$) indicating the j-th matrix of the RVAR model impulse response function ($\boldsymbol{G}_0[t] := \boldsymbol{I}_s$) [16, pp. 40-41] is used.

Hunting detection is based on the following decision making rule:

$$\begin{aligned} D[t] := \max_{k\in[1,\,t)} (\det\widehat{\boldsymbol{\Sigma}}_e[k]/\det\widehat{\boldsymbol{\Sigma}}_y[k]) - (\det\widehat{\boldsymbol{\Sigma}}_e[t]/\det\widehat{\boldsymbol{\Sigma}}_y[t]) \geq l_{lim} &\Longrightarrow \text{Hunting} \\ \text{otherwise} \qquad &\Longrightarrow \text{No Hunting} \end{aligned} \tag{4}$$

[2]The hat designates estimators/estimates.

which detects the hunting onset via a considerable reduction in the ratio det$\mathbf{\Sigma}_e[t]$/det$\mathbf{\Sigma}_y[t]$ from its (historical) maximum value under the normal (no hunting) condition. l_{lim} designates a user-selected detection threshold (limit).

Tuning of the method's operational parameters n (AR order), λ (forgetting factor), M (sliding window length), and l_{lim} (detection threshold) is based on Bayesian Optimization [17] in the initial, off-line, *Tuning Phase*, with the objective being the minimization of the distance of the resulting ROC curve from the upper left (0,1) point; in case of ideal detection performance (zero distance), the objective is the minimization of the mean Detection Delay Time (DDT).

2.3 A Vector Damping Ratio (V-DR) based method

The method is founded upon the fact that the vehicle vibration under hunting condition is heavily dominated by the periodic (limit cycle) oscillation frequency f_h and its first two harmonics $3f_h$ and $5f_h$ [18]. f_h is an inherent feature of the vehicle dynamics and its estimation is performed via signals under hunting during the method's initial *Tuning Phase*.

At each time instant t, within the method's *Operational Phase*, hunting detection is based on the lowest damping ratio ($\zeta_{min}[t]$) associated with a mode with natural frequency within the Δ frequency range:

$$\Delta := (f_h \pm \delta) \cup (3f_h \pm \delta) \cup (5f_h \pm \delta) \tag{5}$$

with δ designating a range parameter, while the natural frequencies and damping ratios are obtained from the estimated RVAR model [19].

Hunting is then detected by the following decision making rule:

$$\begin{array}{ll} D[t] = \widehat{\zeta}_{min}[t] < l_{lim} & \Longrightarrow \text{Hunting} \\ \text{otherwise} & \Longrightarrow \text{No Hunting} \end{array} \tag{6}$$

with l_{lim} designating a user-selected detection threshold (limit).

Like with the previous method, tuning of the operational parameters n (AR order), λ (forgetting factor), δ (range parameter), and l_{lim} (detection threshold) is based on the Bayesian Optimization in the (off-line) *Tuning Phase*.

3 The simulation model and the Monte Carlo simulations

3.1 The simulation model

The simulations are based on a developed 42 degree-of-freedom high-fidelity Simpack-based [20] vehicle model. This encompasses 15 rigid bodies, including a car-body, two bogie frames, four wheel-sets, and eight axle-boxes. The bogie frames are connected with the wheel-sets via rubber springs, while the car-body is connected with the bogie frames via airsprings and lateral dampers as shown in Figure 1. The modeling of the rubber springs and airsprings is based on linear spring and damping elements in the three coordinate directions (x, y, z), while that of lateral dampers is based on nonlinear damping elements in the lateral (y) direction [21]. The wheel/rail profiles follow the S1002/UIC60 standard, while the track random irregularity profiles (in the vertical, lateral, and cross-level directions) follow the ERRI B176 standard [22]. Further information on the vehicle model is available in [12].

3.2 The Scenarios, the Monte Carlo simulations and the vibration signals

Hunting detection performance assessment is based on four Scenarios (Scenarios A, B, C, and D; Table 3), with indicative bogie lateral vibration acceleration signals for each one shown in Figure 2.

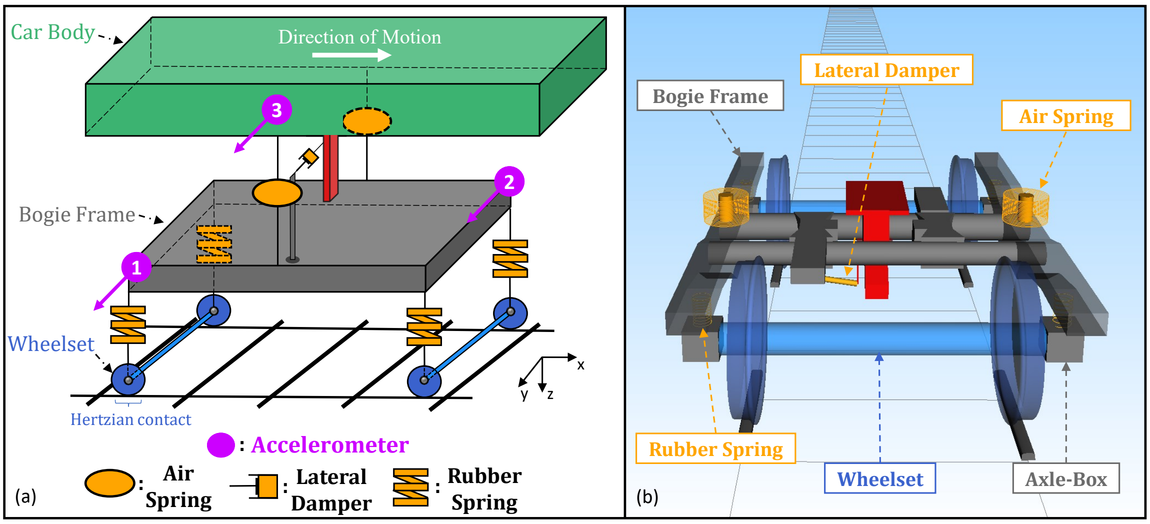

Figure 1: (a) Simplified schematic representation of the half-vehicle model with the accelerometer sensors measuring in the lateral (y) direction, and (b) a Simpack-based representation of the bogie frame.

Table 3: The Scenarios and signal details.

Scenario	Condition	Track/Vehicle Health State	Description	Hunting Speed (v_h; km/h)	Critical Speed (v_{cr}; km/h)
A	Hunting	Nominal	Nominal track/vehicle	210	110
B	Hunting	Faulty	Faulty suspension$^{\triangledown}$	200	115
C	Hunting	Faulty	'Slightly' worn track*	180	110
D	No Hunting	Faulty	'Heavily' worn track*	–	110
Signal details (all Scenarios): Length 2 700 samples (135 s); sampling frequency $f_s = 20\,Hz$.					

All Scenarios refer to motion on a tangent track.
*Track irregularity standard deviation (SD) increased by 40% ('slightly' worn track) and 145% ('heavily' worn track) in the vertical (V), lateral (L) and cross-level (CL) directions (Nominal track SD = 1.5 mm (V); 1 mm (L); 1.4 mm (CL)).
$^{\triangledown}$Stiffness and damping coefficient reduction by 20% in the primary suspension of the leading bogie (all four corners).

Scenarios A, B, and C are employed in order to investigate the methods' hunting detection performance under realistic vehicle/track health states (details in Table 3). For this to be achieved the vehicle speed is linearly increased till that causing hunting initiation (the hunting speed v_h). This speed, which is higher than the critical speed v_{cr}, depends on the track irregularity [2] and is calculated as in [23]. The critical speed v_{cr} is, for each health state of the vehicle (nominal/faulty suspension), computed via the Stichel's numerical method [24]. The hunting onset time is conventionally defined as the time instant of the first contact between the wheel flange and the rail gauge [12]. As is evident from Figure 2(a-c), Scenarios A, B, and C correspond to the so-called small-amplitude hunting [7, 9, 10] as the bogie lateral vibration amplitude surpasses the $2\,m/s^2$ limit but remains significantly lower than the $8\,m/s^2$ threshold required by the standards for full hunting [3] (also see Table 1).

Scenario D (Figure 2(d)) is employed for investigating the methods' ability to avoid false alarms despite the increased lateral vibrations arising from the motion on a 'heavily' worn track under normal (no hunting) condition. For this to be achieved the vehicle speed is increased to a selected speed below the *critical* (v_{cr}). Notice that the small-amplitude hunting limit is surpassed in this case too, but, unlike the previous Scenarios, not by the required frequency needed for the declaration of small-amplitude hunting.

Evidently, the above Scenarios are particularly challenging for hunting detection as they: (a) Are characterized by the harder-to-detect small-amplitude (instead of full) hunting, and (b) include suspension faults or slightly/heavily worn track which may lead to false alarms or to missed hunting detection.

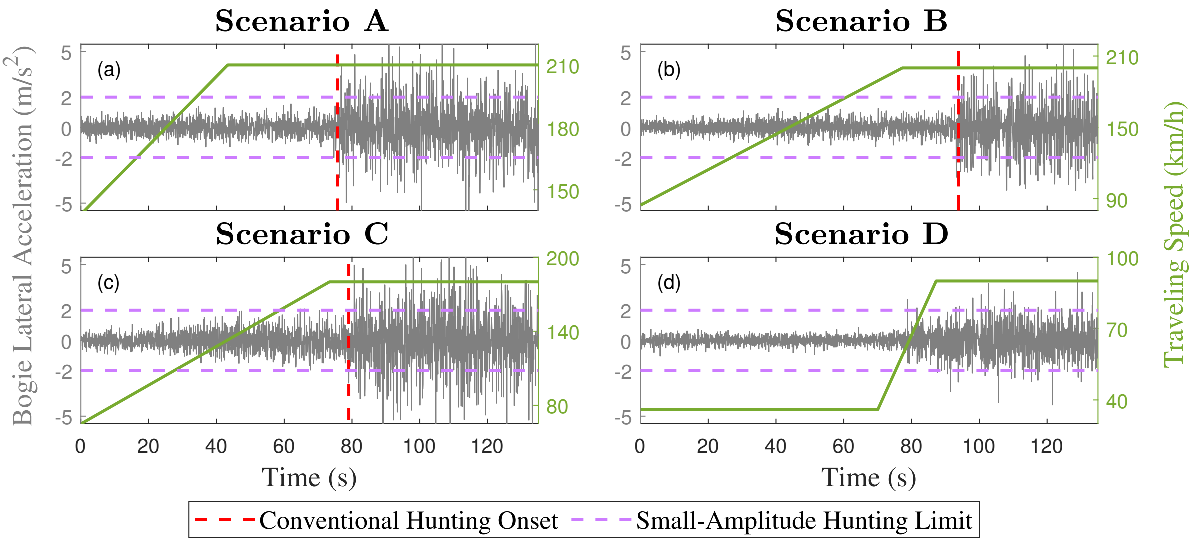

Figure 2: Indicative Test Cases from each Scenario: Bogie lateral acceleration from sensor 1 and vehicle speed profile for (a) Scenario A, (b) Scenario B, (c) Scenario C, (d) Scenario D. (The dashed vertical line designates the conventional hunting onset, while the dashed horizontal line the small-amplitude hunting limit.)

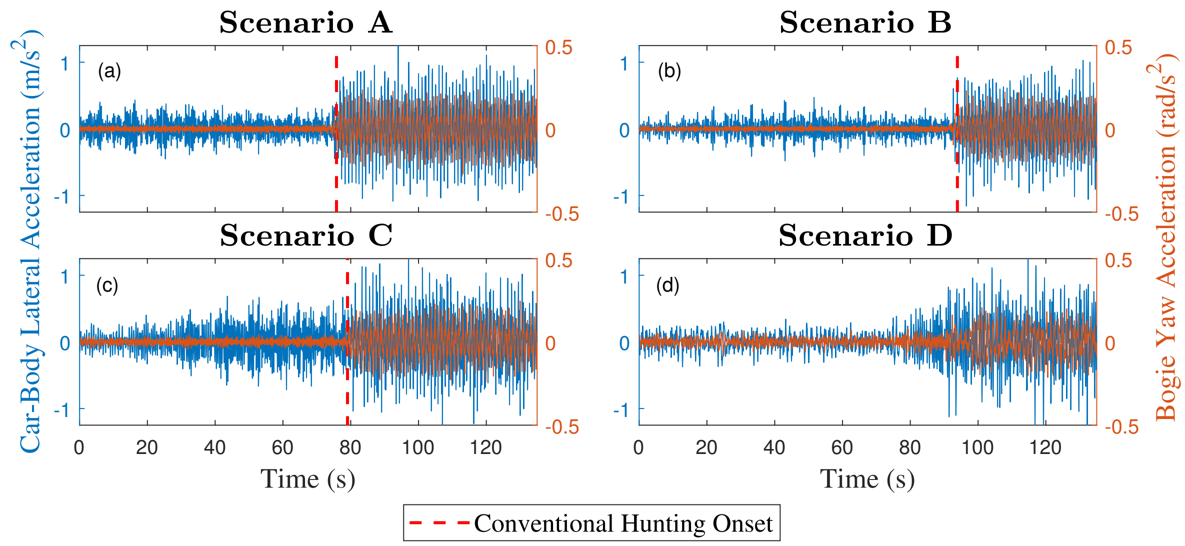

Figure 3: Car-body lateral and bogie yaw acceleration signals for indicative (a) Scenario A, (b) Scenario B, (c) Scenario C, (d) Scenario D Test Cases. (The dashed vertical line designates the conventional hunting onset.)

Monte Carlo simulations (Test Cases) are relying on distinct random realizations of track irregularities, thus leading to different track irregularity realizations in each Test Case. All methods are based on the measurement of three lateral acceleration signals: two on the leading bogie frame and one at the car-body (Figure 1(a)). Based on the two bogie frame vibration signals, the bogie yaw (angular) acceleration is deduced [8] and is used for hunting detection along with the car-body lateral acceleration signal. Indicative signals are, for each Scenario, presented in Figure 3. Evidently, all vibration signals are non-stationary due to the varying traveling speed and the (potential) hunting onset. Further information on the signals is provided in Table 3.

4 Performance assessment via Monte Carlo simulations

The performance assessment of the methods is based on Monte Carlo simulations. In addition, two types of comparisons are presented:

(i) Comparisons with the single-sensor (scalar) versions of the postulated methods (referred to as S-DS-based and S-DR-based) [12, 13] operating on the car-body lateral vibration signal, and,

(ii) a state-of-the-art multi-sensor (vector) method employing an Entropy based quantity and Support Vector Machine classification [9] (presently referred to as V-E-SVM-based) operating on the exact same vibration signals with the postulated methods.

4.1 The assessment procedure and metrics

Hunting detection performance is assessed via Receiver Operating Characteristic (ROC) curves [25] and Detection Delay Time (DDT) histograms [12]. ROC curves depict the True Positive Rate (TPR) or correct hunting detection rate versus the False Positive Rate (FPR) or false alarms rate. The DDT histograms provide the time of hunting detection with respect to the conventional hunting onset.

For Scenarios A, B, and C Test Cases, in which hunting occurs, the focus is on its proper and prompt detection. The first false alarm, or lack thereof, is recorded up to $5\ s$ prior to the conventional hunting onset (as hunting actually initiates prior to its conventional onset), and only the first correct detection, or lack thereof, is recorded thereafter. For Scenario D Test Cases, in which no hunting occurs, false alarms (only the first, or lack thereof, in each Test Case) are recorded. The reader is referred to Table 4 for additional details.

4.2 The Tuning Phase

20 Test Cases per Scenario (80 in total; see Table 4) are employed for determining the methods' operational parameters during the *Tuning Phase* via a Bayesian Optimization procedure (details in subsection 2.2). The optimized operational parameters and estimation details are, for each method, provided in Table 5.

4.3 The Operational Phase and comparative performance assessment

Performance assessment is based on 75 Test Cases per Scenario (300 in total; see Table 4).

Performance for an indicative Scenario A Test Case. Detection results by all methods for an indicative Scenario A (hunting) Test Case are provided in Figure 4, where both the hunting conventional onset and the detection time (by each method) are designated via a vertical dashed and a solid line, respectively. Evidently, all methods achieve correct hunting detection. The postulated V-DS-based method achieves very early (prior to the conventional onset) detection of hunting which also is quite earlier than achieved by its V-DR-based

Table 4: Numbers of Test Cases in the Tuning/Operational Phases and performance assessment rules.

Scenario	Number of Test Cases		Performance Assessment Rules
	Tuning Phase*	Operational Phase*	
A	20	75	Only the first false alarm, or lack thereof, is recorded up to $5\ s$ prior to the hunting onset. Only the first correct detection, or lack thereof, is recorded thereafter.
B	20	75	
C	20	75	
D	20	75	Only the first false alarm, or lack thereof, is recorded for the complete time duration.

*The Test Cases in the Tuning Phase are always different from those in the Operational Phase.

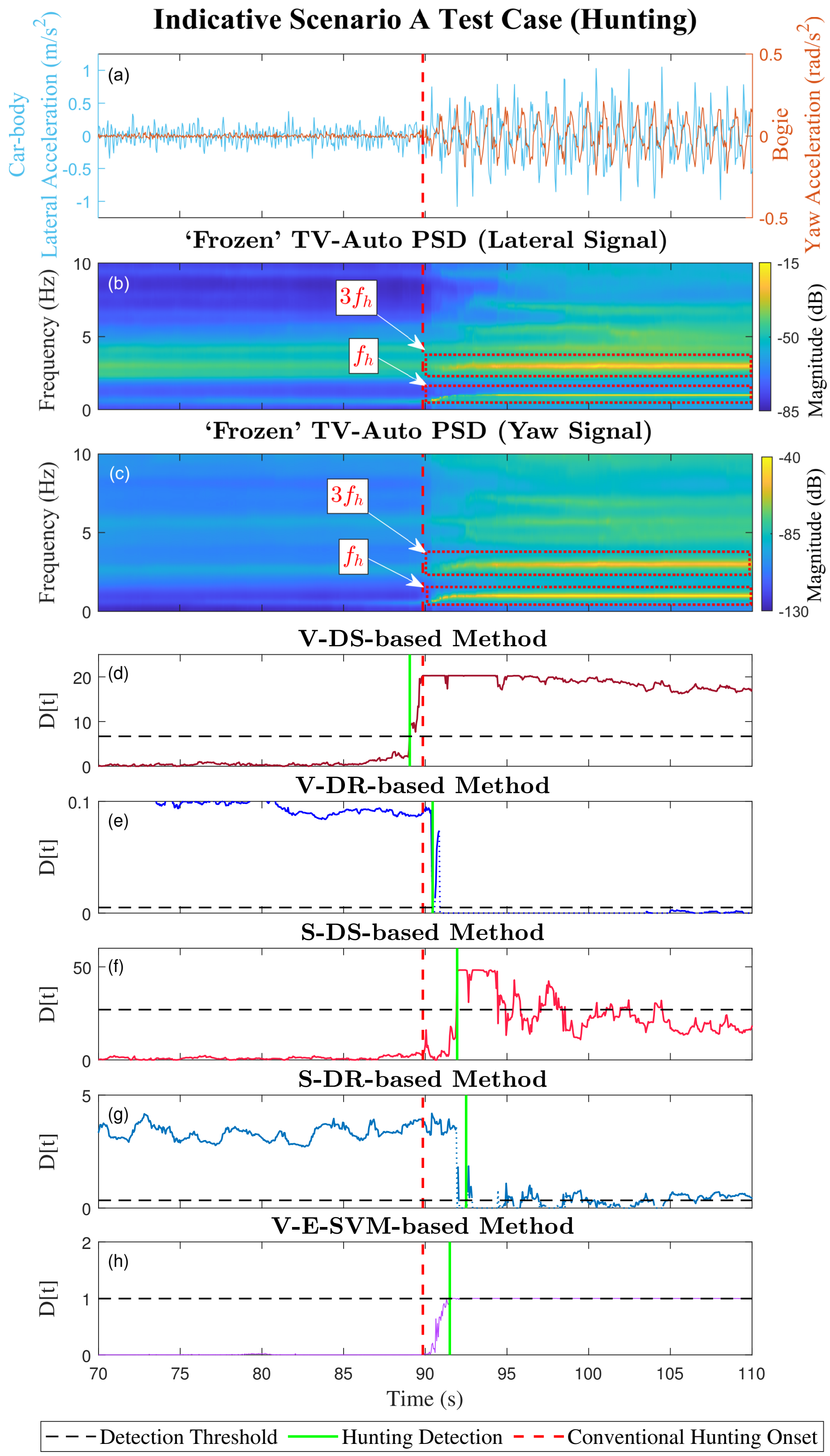

Figure 4: Hunting detection results for an indicative Scenario A Test Case by each method. (a) The car-body lateral and bogie yaw acceleration signals; (b, c) RVAR(16) model based 'frozen' TV-PSDs (f_h is the main hunting frequency and $3f_h$ its first harmonic); (d) V-DS-based detection; (e) V-DR-based detection; (f) S-DS-based detection; (g) S-DR-based detection; (h) V-E-SVM-based detection. (The dotted portions of certain curves designate omitted signal samples due to RVAR model instability.)

Table 5: The selected operational parameters and estimation details.

Method	Operational Parameter	Search Area	Selected Value
V-DS-based$^\diamond$	AR order (n)	[1, 30]	27
	Forgetting factor (λ)	[0.980, 0.999]	0.9975
	Window length (M)	[50, 1 000]	994
V-DR-based*	AR order (n)	[1, 30]	16
	Forgetting factor (λ)	[0.980, 0.999]	0.9989
	Range parameter (δ)	[0.001, 0.8]	0.3550
S-DS-based$^\diamond$	AR order (n)	[1, 50]	38
	Forgetting factor (λ)	[0.980, 0.999]	0.9967
S-DR-based*	AR order (n)	[1, 50]	38
	Forgetting factor (λ)	[0.980, 0.999]	0.9981
	Range parameter (δ)	[0.001, 0.8]	0.2358
V-E-SVM-based	Embedding dimension (m)	[1, 6]	4
	Noise tolerance (r)	[0.01, 0.8]	0.3682
	Window length (M)	[50, 800]	386
	Kernel type	{Linear, Polynomial, Gaussian}	Linear
	Kernel scale (γ)	$[10^{-5}, 10^5]$	1 770
	Box constraint (C)	$[10^{-5}, 10^5]$	35 569

Bayesian optimization details

Objective function: Distance of the ROC-curve from the (0,1) point (in case that this distance is zero, the objective function is the mean DDT); Acquisition function: expected-improvement-per-second-plus; Objective function evaluations: 200; Number of initial evaluation points: 4; MATLAB function: bayesopt.m

RVAR model estimation details

Estimation method: Recursive Least Squares (RLS); Initial parameter vector $\widehat{\boldsymbol{\theta}}[0] = \mathbf{0}$, and covariance matrix $\boldsymbol{P}[0] = 10^4\boldsymbol{I}$.

$^\diamond$Number of impulse response function samples (L): 500

*Dominant hunting frequency (f_h): 1 Hz.

counterpart. Yet, both achieve earlier detection than their single-sensor (S-DS-based and S-DR-based) counterparts, but also than the state-of-the-art V-E-SVM-based method.

Cumulative (all Test Cases) performance for each Scenario and robustness to suspension faults / worn tracks. Cumulative (all Test Cases) detection results, separately for each Scenario, are provided in terms of True Positive Rate (TPR), False Positive Rate (FPR), and Detection Delay Time (DDT), in Figure 5. It is to be recalled that Scenario A refers to nominal vehicle and track, while Scenario B to faulty vehicle suspension, Scenario C to 'slightly' worn track, and Scenario D to 'heavily' worn track (this one with no hunting; see Table 3). Evidently, the two postulated (V-DS-based and V-DR-based) methods achieve ideal TPRs/FPRs for all four Scenarios and the best overall DDTs; these are only somewhat affected by the 'slightly' worn track condition (Scenario C) suggesting their robustness to the various conditions (Figure 5(a-d)). On the other hand, their scalar counterparts exhibit generally higher DDTs, which are also increased by the 'slightly' worn track condition (Scenario C), while the S-DS-based method also yields 16% FPR for Scenario D (Figure 5(e-h)). Finally, the V-E-SVM-based method also exhibits 4% FPR for Scenario D and generally higher DDTs which are also increased under the 'slightly' worn track condition (Scenario C; see Figure 5(i-j)).

Cumulative performance for all Scenarios and Test Cases. The cumulative over all Scenarios and Test Cases performance of each method is presented via ROC curves in Figure 6. Evidently, the postulated V-DS-based and V-DR-based methods achieve perfect performance (100% TPR @ 0% FPR), with the V-DS-based method achieving significant performance improvement over its single-sensor (S-DS-based) counterpart. The state-of-the-art V-E-SVM-based method achieves somewhat inferior (100% TPR @ 1% FPR) performance than that of the postulated methods.

The Detection Delay Times (DDTs) for each method are provided in Figure 7 in terms of histograms for Scenarios A, B and C using each method's optimal ROC-based threshold (see Figure 6). Evidently, thep-

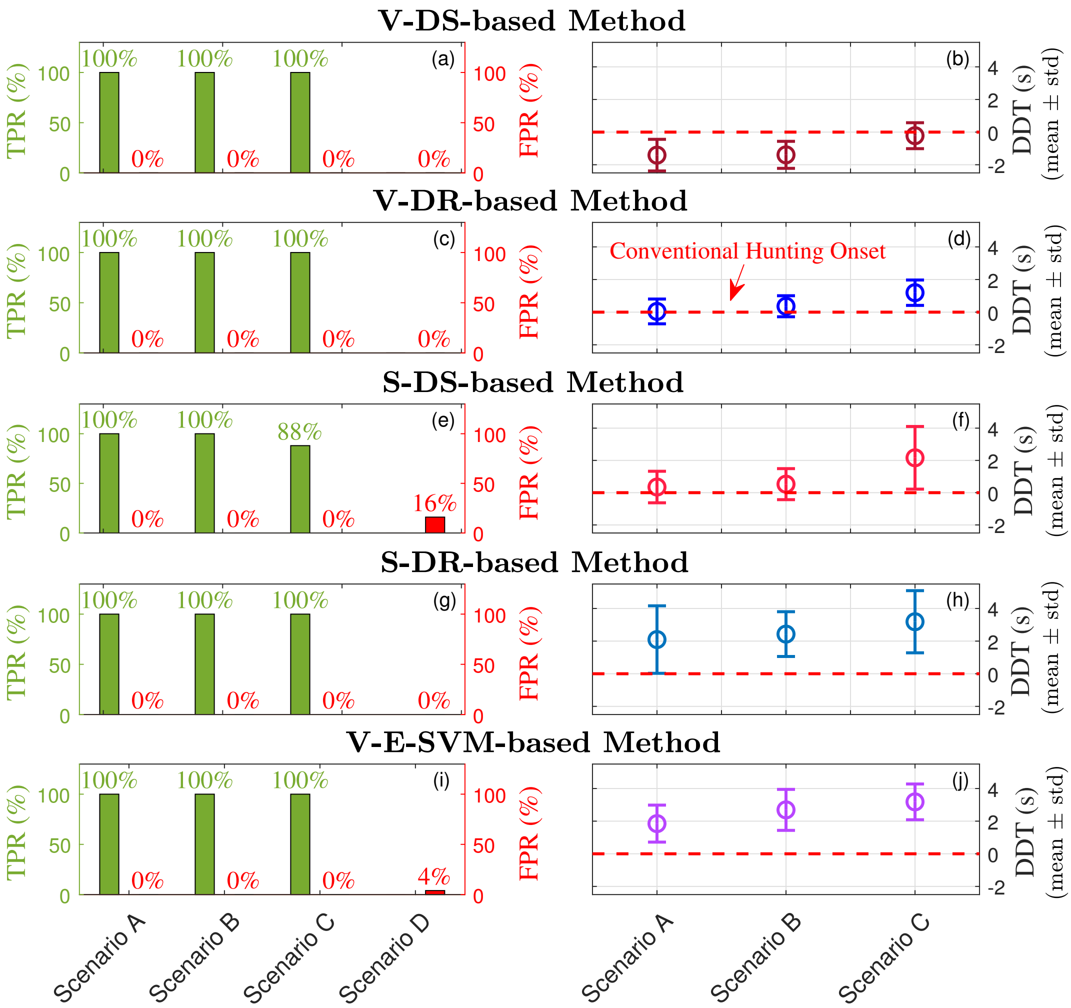

Figure 5: Performance and robustness assessment for all methods under each Scenario: True Positive Rates (TPRs), False Positive Rates (FPRs), and Detection Delay Times (DDTs) for the (a,b) V-DS-based, (c,d) V-DR-based, (e,f) S-DS-based, (j,h) S-DR-based, and (i,j) V-E-SVM-based method. (75 Test Cases per Scenario.)

ostulated V-DS-based method achieves the best DDTs, with the V-DR-based method following with performance similar to that of the single-sensor version of the former (S-DS-based). Finally the S-DR-based and V-E-SVM-based methods lag further behind.

Performance and CPU times - overview. A bird's eye view of each method's performance (all Test Cases in all Scenarios) is presented in Figure 8 in terms of True Positive Rate (TPR) and Detection Delay Times (DDTs; sample mean $\pm$ std) for 0% False Positive Rate (FPR). Evidently, the two postulated (V-DS-based and V-DR-based) methods, as well as the single-sensor version of the latter (S-DR-based), achieve ideal (100%) TPRs. Yet, the V-DS-based also is a clear leader in terms of achieving the lowest DDT and consistently detecting hunting prior to its conventional onset.

Mean CPU times for all methods are presented in Figure 9 (2 500 executions) using each method's selected operational parameters (see Table 5). The experiments are conducted on a desktop computer with Matlab© R2020A (programming language), AMD Ryzen™ 5 2600X @ 3.6 GHz (CPU), 32 GB DDR4 @ 3 GHz (RAM). The mean CPU time of all methods are satisfactory in the sense that they are below the sampling period T_s ($= 50$ ms), implying their suitability for real-time operation. Yet it is noticed that the price paid by the V-DS-based method for achieving the highest overall performance is the highest mean CPU time.

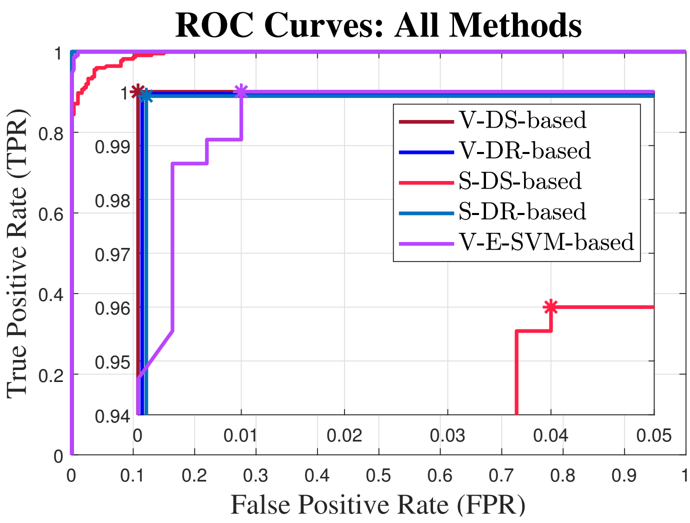

Figure 6: Aggregate (all Scenarios included) hunting detection performance results for the V-DS-based, V-DR-based, S-DS-based, S-DR-based, and V-E-SVM-based methods via ROC curves. The asterisk (∗) designates the optimal operating point on each ROC curve (75 Test Cases per Scenario).

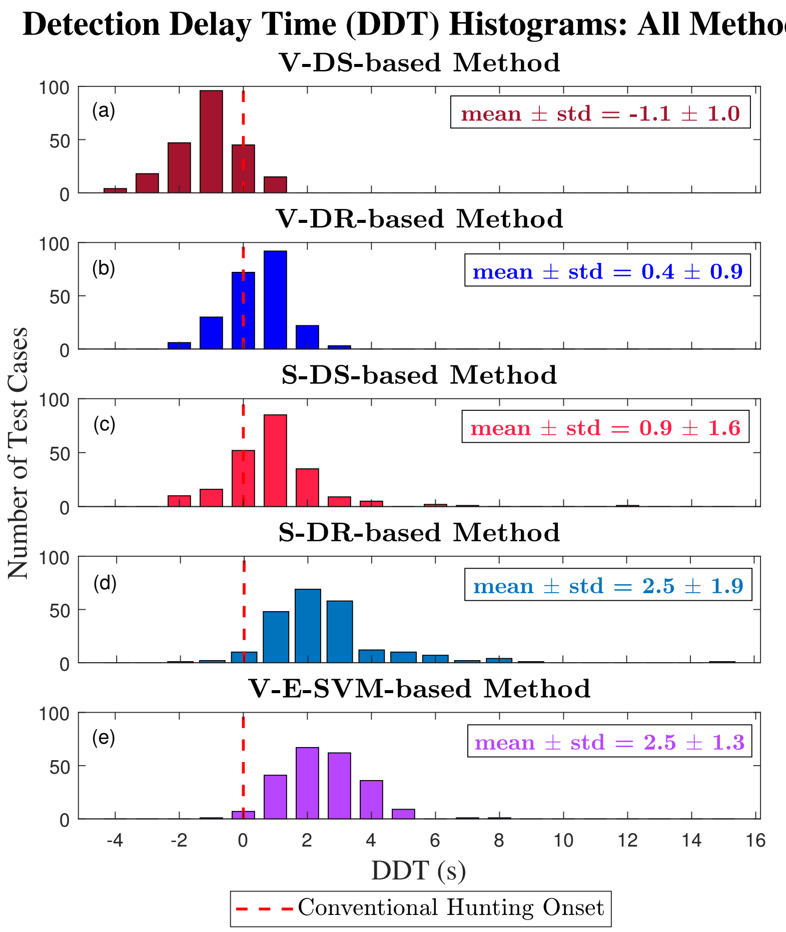

Figure 7: Detection Delay Time (DDT) histograms for the (a) V-DS-based, (b) V-DR-based, (c) S-DS-based, (d) S-DR-based, and (e) V-E-SVM-based method under Scenarios A, B, and C (75 Test Cases per Scenario; detection threshold based on the optimal operating point on each ROC curve).

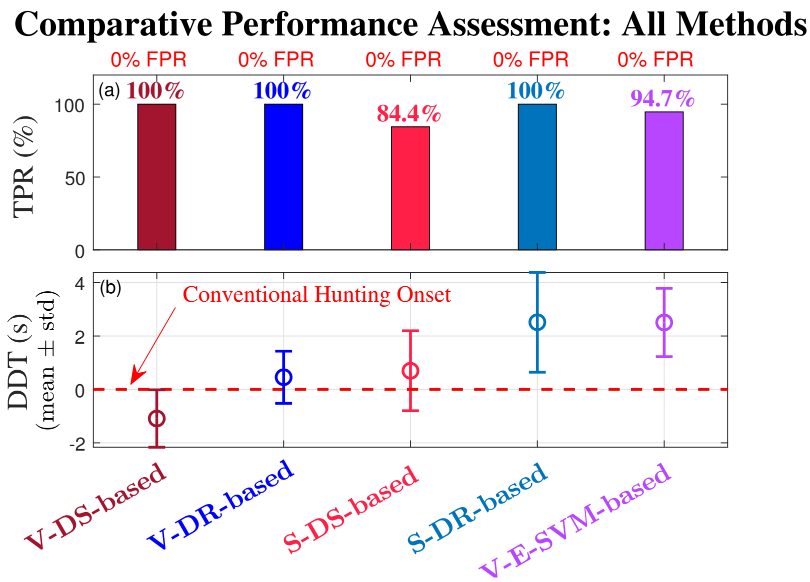

Figure 8: Overview of each method's performance: (a) True Positive Rate (TPR) and (b) Detection Delay Time (DDT) for 0% False Positive Rate (FPR) (all Scenarios, 75 Test Cases per Scenario; ROC-based thresholds are used for each method).

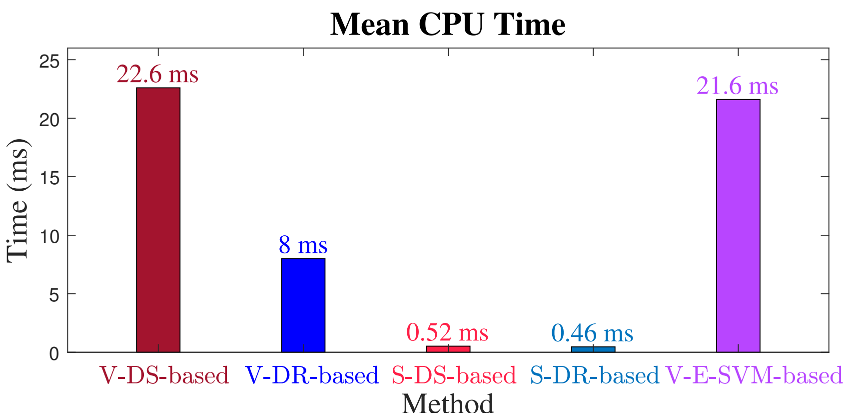

Figure 9: Mean CPU time for the V-DS-based, V-DR-based, S-DS-based, S-DR-based, and V-E-SVM-based method (2 500 executions).

5 Concluding remarks

The possibility of attaining enhanced hunting performance, primarily in terms of early detection and robustness to suspension faults and worn tracks, via multi-sensor (vector) parametric methods and the use of the car-body lateral and bogie yaw acceleration signals has been examined. Towards this end multi-sensor (vector) extensions, referred to as the Vector Degree-of-Stochasticity (V-DS) based and the Vector Damping Ratio (V-DR) based methods, of our recently introduced [12, 13] single-sensor schemes have been postulated based upon parametric, sequential in time, non-stationary Recursive Vector AutoRegressive (RVAR) modeling.

The performance of the postulated multi-sensor methods has been systematically assessed via extensive Monte Carlo simulations under four Scenarios using a SIMPACK-based passenger vehicle model and the use of the True Positive Rate (TPR), the False Positive Rate (FPR), and the Detection Delay Time (DDT) with respect to conventional hunting onset. Critical comparisons with their single-sensor (scalar) counterparts (the S-DS-based and S-DR-based methods) as well as a state-of-the-art V-E-SVM-based method [9] have been also performed.

The main conclusions of the study may be summarized as follows (compare with the corresponding questions posed in the Introduction):

1. The postulated multi-sensor (vector) parametric methods achieve impressive performance and drastically surpass their single-sensor (scalar) counterparts in terms of Detection Delay Time (DDT) and robustness to suspension faults and worn tracks. They reduce the sample mean DDT from 2.6 s (S-DR-based) to -1.1 s (V-DS-based) and exhibit improved robustness, primarily to worn tracks.

2. The best postulated (V-DS-based) method achieves 100% TPR @ 0% FPR and DDT of -1.1 s $\pm$ 1.0 s (sample mean $\pm$ std), with the V-DR-based being a close follower.

3. The state-of-the-art V-E-SVM-based method achieves clearly inferior performance (94.7% TPR @ 0% FPR and DDT with sample mean $\pm$ std of 3.9 s $\pm$ 2.6 s) than that of the postulated methods.

4. The real-time (on-board) implementation of the methods is facilitated by their relatively low computational complexity.

Acknowledgements

The study has been partly funded by the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH-CREATE-INNOVATION (The MAIANDROS project with code T1EDK-01440), and partly by the Hellenic Foundation for Research and Innovation (HFRI) under the 3rd Call for HFRI PhD Fellowships (Fellowship Number 6803).

References

[1] O. Polach and I. Kaiser, “Comparison of methods analyzing bifurcation and hunting of complex rail vehicle models,” *Journal of Computational and Nonlinear Dynamics*, vol. 7, no. 4, 2012, 041005.

[2] O. Polach, “Application of nonlinear stability analysis in railway vehicle industry,” in *Non-smooth Problems in Vehicle Systems Dynamics*, P. Grove Thomsen and H. True, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 15–27.

[3] International Union of Railways, *Passenger rolling stock–Trailer bogies–Running gear–General provisions applicable to the components of trailers bogies*, 2nd ed., 2003, standard UIC Code 515-1. [Online]. Available: https://uic.org

[4] Federal Railroad Administration, *Track Safety Standards*, 2011, standard 49 CFR 213. [Online]. Available: https://www.govinfo.gov/app/collection/cfr

[5] N. Wilson, R. Fries, M. Witte, A. Haigermoser, M. Wrang, J. Evans, and A. Orlova, “Assessment of safety against derailment using simulations and vehicle acceptance tests: a worldwide comparison of state-of-the-art assessment methods,” *Vehicle System Dynamics*, vol. 49, no. 7, pp. 1113–1157, 2011.

[6] R. Kulkarni, A. Qazizadeh, M. Berg, U. Carlsson, and S. Stichel, “Vehicle running instability detection algorithm (VRIDA): A signal based onboard diagnostic method for detecting hunting instability of rail vehicles,” *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 236, no. 3, pp. 262–274, 2022.

[7] J. Ning, M. Fang, W. Ran, C. Chen, and Y. Li, “Rapid multi-sensor feature fusion based on non-stationary kernel JADE for the small-amplitude hunting monitoring of high-speed trains,” *Sensors*, vol. 20, no. 12, 2020, 3457.

[8] J. Sun, E. Meli, W. Cai, H. Gao, M. Chi, A. Rindi, and S. Liang, "A signal analysis based hunting instability detection methodology for high-speed railway vehicles," *Vehicle System Dynamics*, vol. 59, no. 10, pp. 1461–1483, 2021.

[9] J. Ning, Q. Liu, H. Ouyang, C. Chen, and B. Zhang, "A multi-sensor fusion framework for detecting small amplitude hunting of high-speed trains," *Journal of Vibration and Control*, vol. 24, no. 17, pp. 3797–3808, 2018.

[10] J. Ning, W. Cui, C. Chong, H. Ouyang, C. Chen, and B. Zhang, "Feature recognition of small amplitude hunting signals based on the MPE–LTSA in high-speed trains," *Measurement*, vol. 131, pp. 452–460, 2019.

[11] Y. Zeng, W. Zhang, and D. Song, "A new strategy for hunting alarm and stability evaluation for railway vehicles based on nonlinear dynamics analysis," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, vol. 234, no. 1, pp. 54–64, 2020.

[12] K. Kritikakos, S. D. Fassois, and J. S. Sakellariou, "Robust early detection of hunting on railway vehicles: single-sensor based methods and critical assessment," 2022, under preparation for publication.

[13] K. Kritikakos, S. D. Fassois, J. S. Sakellariou, I. Chronopoulos, A. Deloukas, I. A. Iliopoulos, G. Leoutsakos, I. Tountas, and G. Vlachospyros, "On the problem of on–board early detection of hunting on rail vehicles: an exploratory study," *IFAC-PapersOnLine*, vol. 54, no. 20, pp. 191–197, 2021, Modeling, Estimation and Control Conference MECC 2021.

[14] C. Wiedemann, "State-of-the-art railway vehicle design with multi-body simulation," *Journal of Mechanical Systems for Transportation and Logistics*, vol. 3, no. 1, pp. 12–26, 2010.

[15] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Prentice Hall PTR, 1999.

[16] R. S. Tsay, *Multivariate time series analysis: with R and financial applications*, 1st ed. John Wiley & Sons, 2014.

[17] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'12. Red Hook, NY, USA: Curran Associates Inc., 2012, pp. 2951–2959.

[18] X. J. Gao, Y. H. Li, and Y. Yue, "The resultant bifurcation diagram method and its application to bifurcation behaviors of a symmetric railway bogie system," *Nonlinear Dynamics*, vol. 70, no. 1, pp. 363–380, 2012.

[19] S. D. Fassois, "MIMO LMS-ARMAX identification of vibrating structures—part I: the method," *Mechanical Systems and Signal Processing*, vol. 15, no. 4, pp. 723–735, 2001.

[20] Simpack, Dassault Systemes, 2019. [Online]. Available: https://www.3ds.com/

[21] S. Bruni, J. Vinolas, M. Berg, O. Polach, and S. Stichel, "Modelling of suspension components in a rail vehicle dynamics context," *Vehicle System Dynamics*, vol. 49, no. 7, pp. 1021–1072, 2011.

[22] A. Haigermoser, B. Luber, J. Rauh, and G. Gräfe, "Road and track irregularities: measurement, assessment and simulation," *Vehicle System Dynamics*, vol. 53, no. 7, pp. 878–957, 2015.

[23] W. Zhai and K. Wang, "Lateral hunting stability of railway vehicles running on elastic track structures," *Journal of Computational and Nonlinear Dynamics*, vol. 5, no. 4, pp. 1–9, 2010.

[24] H. True, "Multiple attractors and critical parameters and how to find them numerically: the right, the wrong and the gambling way," *Vehicle System Dynamics*, vol. 51, no. 3, pp. 443–459, 2013.

[25] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861 – 874, 2006.