Real-time train motion parameter estimation using an Unscented Kalman Filter

Cunillera, Alex; Bešinović, Nikola; van Oort, Niels; Goverde, Rob M.P.

DOI
10.1016/j.trc.2022.103794

Publication date
2022

Document Version
Final published version

Published in
Transportation Research Part C: Emerging Technologies

Transportation Research Part C 143 (2022) 103794

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Real-time train motion parameter estimation using an Unscented Kalman Filter

Alex Cunillera *, Nikola Bešinović, Niels van Oort, Rob M.P. Goverde

Department of Transport and Planning, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands

ARTICLE INFO

Keywords:
Railways
Train motion model calibration
Parameter estimation
Unscented Kalman Filter

ABSTRACT

Train movement dynamics are usually modelled by means of Newton's second law. The resulting dynamic equation can be very precise if the parameters that it depends on are determined accurately. However, these parameters may vary in time and show wide variations, making the calibration task nontrivial and jeopardizing the performance of a broad variety of applications in the railway industry: from timetable planning and railway traffic simulation to Driver Advisory Systems and Automatic Train Operation. In this article, the online train motion model calibration problem is addressed with a special focus on energy-efficient on-board applications. To this end, location and speed measurements are assumed to be available for a train running under normal operation conditions. A well-known real-time parameter estimation algorithm, the Unscented Kalman Filter, is combined with a driving regime calculator and a post-processing module in order to obtain bounds and statistics of parameters such as the maximum applied tractive effort and power, the applied brake rates, the cruise speed and the length of the final coasting and braking. The proposed framework is tested in a case study with real data from trains operating on the Eindhoven-'s-Hertogenbosch corridor in the Netherlands. Results obtained show that UKF is able to track the speed and location measurements and to estimate the parameters that model the running resistance in the dynamic equation. The proposed driving regime and the post-processing modules can determine the current regime accurately and give a deeper insight into the variations of the driving style, respectively.

1. Introduction

Railway digitization is boosting the sensorization of trains, tracks and signalling systems. Thousands of variables are continuously monitored, from passing times at a certain signal to the temperature in each car. Therefore, the amount of information that is measured, generated and stored is increasing dramatically and the possibility of exploiting the produced data opens new horizons for calibrating and improving current mathematical and expert rule-based methods and for devising new ones. As a consequence, data-driven algorithms are thus becoming increasingly popular in the railway science and industry (Ghofrani et al., 2018). Recent surveys of data-driven and big data analytics applications in railway systems can be found in De Martinis and Corman (2018) and Ghofrani et al. (2018). However, the combination of data-driven techniques and classical approaches show promising results in boosting the performance and accuracy of current methods.

On-board applications like Driver Advisory Systems (DAS) and Automatic Train Operation (ATO) are among the solutions that may benefit most from data-driven and hybrid approaches by obtaining more accuracy and even by allowing the implementation

* Corresponding author.
E-mail addresses: a.cunilleraperez@tudelft.nl (A. Cunillera), n.besinovic@tudelft.nl (N. Bešinović), n.vanoort@tudelft.nl (N. van Oort), r.m.p.goverde@tudelft.nl (R.M.P. Goverde).

https://doi.org/10.1016/j.trc.2022.103794
Received 9 December 2021; Received in revised form 19 May 2022; Accepted 5 July 2022
Available online 25 July 2022

of real-time and custom-train based solutions. DAS and ATO generally use a mathematical model based on Newton's second law in order to generate the reference speed profile to be followed, which includes the speed that the train should have at every point of the track along its movement. However, this train motion model has several input parameters that have to be carefully calibrated in order to reproduce train dynamics accurately.

Real-time train motion model calibration allows to obtain a more faithful representation of the current train characteristics for on-board applications. Nevertheless, the calibration is usually performed offline by means of the engine characteristics and the train running resistance parameters provided by rolling stock manufacturers or by analyzing historical data registered during tests and standard operation. However, railway operation is fundamentally stochastic, which lowers the effectiveness of the mentioned offline calibration approaches with respect to real-time calibration. This is due to the fact that train dynamics and its model parameters show spatial and time-dependent variations due to factors like the engine and train wear and faults, the adhesion condition, the weather and the driving style. Therefore, individual train tailored real-time calibration becomes a key element for accounting for the current operation conditions and the parameter variations that may jeopardize the accuracy of applications like DAS and ATO. For instance, an ATO that is not well calibrated could lead to altered punctuality rates, to trains stopping far from the allocated locations in the platform, to overspeeding with possible brake intervention from the automatic train protection system, reducing trust in the system by the driver, and to a decrease of the passenger comfort. Furthermore, such a real-time calibration must be performed on-board, which exerts extra pressure on the available on-board computational resources.

In this article, a real-time parameter estimation algorithm is used for determining the train motion model parameters in real time, aiming to develop a framework that can calibrate the input parameters of real-time trajectory optimization algorithms that are embedded in energy-efficient on-board applications like DAS or ATO. This algorithm, the Unscented Kalman Filter (UKF), is a well-known mathematical parameter estimator for nonlinear systems. It uses on-board location and speed measurements of a train running under normal, unperturbed operation conditions, as input data to estimate the train running resistance parameters. This algorithm also requires the applied tractive and brake efforts as input, which are calculated from the last estimations of the running resistance parameters and the gradient profile, so that the efforts match the train acceleration. Moreover, the algorithm is complemented by two modules. The first one is a driving regime recognition module that identifies the current driving regime, that is, whether the train is applying traction to accelerate, cruising, coasting or braking. The second one is a post-processing module that determines parameter bounds like the maximum tractive effort and power and driving style features like the applied brake rates, the cruise speed and the switching points between driving strategies. The mentioned train characteristics constitute most of the input parameters of on-board energy-efficient applications, where they are used to compute reference eco-driving trajectories and the associated driving advice or control policy. Consequently, the proposed real-time calibration framework may boost the performance of such energy-efficient applications by providing accurate estimates of their input parameters based on the current state of the train and the operation conditions. Furthermore, such accurate estimates may facilitate the implementation of such applications. In particular, the train running resistance parameters estimated by the UKF are directly related to the train energy consumption, since they model the main losses of kinetic energy of a train. Therefore, the proposed framework can be used to determine the energy consumption more precisely, which, combined with energy-efficient DAS or ATO, may lead to improved and more efficient railway operations.

The UKF is a filter that can be used for state and parameter estimation of nonlinear systems devised by Julier and Uhlmann (1997). In this technique, the covariance is propagated by means of a deterministic sampling procedure called Unscented Transformation. The state distribution is approximated by a Gaussian random variable, although it is represented by several carefully chosen sample points representing the mean and covariance. In each iteration, these points are propagated through the nonlinear system in order to calculate a new mean and covariance estimate. In this article, a UKF is used for estimating the train running resistance parameters, while the input data used in this filter are the train location and speed measurements. UKF has also been used in other transport modes, like vehicles (Antonov et al., 2011; Liu et al., 2021), aircraft (Khadilkar and Balakrishnan, 2011) and ships (Peng et al., 2019).

In railway systems, several approaches have already been considered for online train motion model parameter estimation, such as multi-innovation theory (Liu et al., 2018), multi-start gradient-based search (De Martinis and Corman, 2019), expectation maximization based on sliding windows (Jin et al., 2018), UKF (Howlett et al., 2004) and Gaussian sum theory combined with Extended Kalman Filters (Jin et al., 2017). However, most of them focus on determining a limited number of model parameters, such as the running resistance parameters (Liu et al., 2018; De Martinis and Corman, 2019; Howlett et al., 2004; Jin et al., 2017) and the adhesion coefficient (Jin et al., 2018). In contrast, the proposed framework, besides determining the three running resistance parameters and calculating the applied tractive and brake effort in real time, also aims to provide estimations of the maximum tractive power applied and to model driving style features like the applied brake rates and the switching points between driving regimes. These parameters are usually used as input of train trajectory optimization algorithms. Therefore, the proposed framework can improve the accuracy of such algorithms and the applications that usually embed them, like DAS and ATO. Furthermore, some of the mentioned techniques are applied for a specific driving regime, like coasting (Liu et al., 2018; Howlett et al., 2004) and braking (Jin et al., 2018), while only De Martinis and Corman (2019), Jin et al. (2017) and the proposed framework are operative in any driving regime. Moreover, UKF is also faster than expectation maximization (De Martinis and Corman, 2019), since that technique takes more than 6 s per iteration, while UKF can process a whole trajectory in less than a second. Last, when compared to EKF (Jin et al., 2017), UKF propagates the statistics of random variables nonlinearly, while EKF linearizes the process function that describes the evolution of such random variables and thus, in some cases UKF can be more accurate than EKF.

Reflecting on the existing online train motion model parameter estimation literature, UKF was first used estimating two of the three running resistance parameters in Howlett et al. (2004). The algorithm was verified from simulated speed and location data

of a train coasting on a level and straight track. Nevertheless, in this article we extend the UKF algorithm by estimating all the running resistance parameters and we validate the results obtained using real location and speed measurements. Furthermore, the addition of a driving regime recognizer and a module to obtain parameter bounds and statistics gives this framework an added value with respect to the existing online parameter estimation algorithms, since the proposed framework constitutes the online parameter estimation algorithm with the broadest and most complete scope in the current scientific literature.

The contributions of this article are:

- Proposing an accurate real-time train motion model parameter estimation framework.
- Estimating in real time the input parameters of on-board energy-efficient railway applications like DAS and ATO from train location and speed measurements.
- Showing in a case study with real data that the proposed framework can be used effectively for model calibration in railways.
- Calculating bounds and obtaining statistics of driving style parameters and applied maximum tractive power and brake rates.

The remainder of the article is organized as follows. In Section 2 the train motion model is introduced and some of the main sources of variability and uncertainty of the model input parameters are described. The proposed online parameter estimation algorithm is presented in Section 3. Its performance is assessed in a case study, and the results obtained are discussed in Section 4. Last, final remarks are outlined in Section 5.

2. Train motion model

The train motion model that aims to reproduce the dynamics of a train can be described by means of Newton's second law,

$$\frac{dv}{dt} = f_t(v) - f_b(v) - r(v) - g(s), \tag{1}$$

where v stands for the train speed (in m/s), s for its location (in m) and t for the time (in s), which in this case is the independent variable. Moreover, f_t is the mass-specific applied tractive effort (in N/kg), $f_t = F_t/(\gamma m)$, where F_t is the tractive effort (in N), γ is the rotating mass factor (dimensionless) and m the train mass (in kg), f_b is the mass-specific applied brake effort (in N/kg), $f_b = F_b/(\gamma m)$, where F_b is the brake effort (in N), and $g(s)$ is the mass-specific term (in N/kg) that accounts for the effects of the track geometry on the train dynamics. $r(v)$ is the mass-specific running resistance, which can be described by means of Davis equation:

$$r(v) = r_0 + r_1 v + r_2 v^2, \tag{2}$$

where r_0, r_1 and r_2 are nonnegative parameters (in N/kg, N/(kgm/s) and N/(kg(m/s)^2), respectively).

Most of the mentioned parameters show time and spatial variations and they are influenced by wear and external factors. The applied tractive effort f_t is often limited by two variables that also show variations with respect to the manufacturer specifications: the maximum tractive effort and power. The maximum tractive effort that a train engine can apply is limited by factors like adhesion, which is the friction between the drive wheels and the rail, and engine wear. In turn, adhesion depends on factors like wheel and rail wear and weather. Yet a train cannot apply this maximum effort at any speed. At higher speeds, the tractive effort is limited by the electric power available and the maximum power of the engine, which in turn may show variations due to the engine fault record and wear.

In manually-driven trains, the driving style also produces variations in the running speed, the acceleration and deceleration rates and the coasting and braking points, which are the locations where the train neither applies traction nor brake and where it starts braking, respectively. Moreover, under normal operation conditions, the brake effort f_b is limited by weather conditions and wear, by Automatic Train Protection systems, by comfort constraints and by the use of predetermined brake curves and rates. The train driver may also apply a brake rate lower than the maximum one that is limited by the train characteristics, and some variations might be observed when comparing the brake rate applied in different runs. Therefore, these driving-induced variations constitute an extra source of parameter uncertainty that differs from the physical sources of variation mentioned earlier. Driving-induced variations may also be the target of online parameter estimation frameworks, as they are of special relevance for on-board energy-efficient applications like DAS. The observed driving-induced variations could be introduced as input to a DAS in order to produce a driving advice adapted to the particularities and preferences of the current driver, maximizing the probability of compliance in following the driving advice, guaranteeing the efficacy of the energy-efficient algorithm and increasing the overall satisfaction and confidence of train drivers with respect to the DAS.

Running resistance parameters r_0, r_1 and r_2 are usually computed by manufacturers by means of phenomenological equations, computational fluid dynamics (Rochard and Schmid, 2000) or by performing run-down tests (Lukaszewicz, 2007). However, their values are significantly influenced by external factors like wind (Trivella et al., 2020), the presence of tunnels, and the train and rails wear. In particular, r_2 has a strong dependence on the train head geometry. Since energy consumption is strongly related to these parameters, most efforts on train motion model calibration have been devoted to monitoring them.

Even though train mass is not made explicit in Eq. (1), since the parameters in the mentioned equation are considered mass-specific, it is also a source of parameter uncertainty in the train motion model. A train tare is usually well-determined, while train load varies at stations due to passengers boarding and disembarking and freight loading and unloading, although it is easier to calculate the load mass of freight trains in comparison to passenger trains. Moreover, in the case of diesel-powered locomotives the fuel mass decreases during the journey. However, using mass-specific parameters allows to reduce in one the number of parameters to be estimated and to keep the linearity of the model with respect to the parameters, while considering the effects of the mass variability in the dynamics.

Infrastructure managers publish periodically a simplified description of the geometry of tracks and platforms, consisting of the gradients, curves and cant. Depending on the level of detail of the track geometry description, it usually shows some uncertainties that may affect the performance of the train motion model. For instance, the track is usually divided in intervals that contain several different gradients. Nevertheless, the track geometry description only includes the most representative or restrictive gradients in each track interval, as this description usually focuses on safety and operation targets. Furthermore, the gradients are depicted as piecewise constant, and they are sometimes made continuous by means of parabolic links. Curves are described in a similar way, in terms of the radius of curvature. Overall, this constitutes a simplified description of the existing track geometry. Therefore, the difference between the real track geometry and the description provided by infrastructure managers constitutes a source of error that may lead to inaccuracies in the term $g(s)$ of the train motion model and the reproduced train dynamics. However, in the proposed monitoring algorithm we assume that the track geometry description used is perfectly accurate, as this helps to keep the model simple and the effect of any mismatch between the real track geometry and its description might be up to some extent mitigated by overfitting the running resistance parameters.

3. Methodology

A parameter estimation algorithm has been developed to monitor the variability of the train motion model parameters in real time. The proposed algorithm is formed by three modules. First, a UKF module estimates the running resistance parameters. On-board location and speed measurements are used as input for this module along with the calculated tractive or brake effort needed to match the observed train acceleration. Eq. (1) is used to calculate the mentioned effort using the last running resistance parameters produced by the UKF. Next, the train acceleration is evaluated from the filtered speed and it is utilized along the calculated tractive and brake effort in a driving regime detection module for determining whether the train is applying traction to accelerate, holding a constant speed, coasting or braking. Last, the results obtained are post-processed online in a feature extraction module in order to determine parameters like the maximum tractive effort and power, brake rates, switching points between driving strategies and cruise speeds. Fig. 1 shows a flowchart of the proposed algorithm. Although the proposed methodology is validated using real data from a main line train in Section 4, it could be adapted to any other railway system, provided that the internal parameters of the three modules are fine-tuned to adapt the framework to the particularities of each railway system. The remainder of this section is organized as follows. The UKF module is described in Section 3.1, the driving regime recognition module is outlined in Section 3.2, and Section 3.3 describes the post-processing module.

3.1. Unscented Kalman Filter (UKF)

The UKF is a state observer that can perform parameter estimation of nonlinear systems by including the mentioned parameters in the state vector. This technique considers the system dynamics and the measurement procedure to be stochastic and models the corresponding noises as Gaussian white noise. The UKF evaluates the nonlinear evolution of the state statistics by means of the Unscented Transformation (UT), which is a method for approximating the statistics of a random variable that is transformed nonlinearly. In UT, the state statistics are thus expressed by means of a special set of points, the sigma points, that are chosen so that their mean and covariance are equal to the state respective values. Then, to calculate the evolution of the state statistics these sigma points are propagated nonlinearly through the dynamical system and the measurement function that maps the state vector into the measured variables. After that, the mean and covariance of the state at the next time step are calculated as weighted sums of the statistics of the propagated sigma points. The main advantages of UT is that it offers a computationally efficient way of calculating the state statistics of nonlinear systems, with an accuracy up to the third moment when the state statistics is symmetric. Furthermore, the sampling is deterministic, so it offers more consistent results than Monte Carlo sampling methods.

The performance of the UKF can be described as follows. First, at every iteration, sigma points of the previous state are calculated and propagated through the system dynamics model, and the mean and covariance of the state at the current time are computed by means of a weighted sum of the propagated sigma points. Second, the propagated sigma points are mapped through the measurement function into the sigma points in the measurement space, which are used to obtain the expected value of the measured variables and the associated covariance. Last, the cross-covariance of the state and the measurements is calculated to obtain the Kalman gain, which is used to perform the estimation and to compute its covariance. This section aims to give an insight into the UKF.

Let $x^\top = (s, v, r_0, r_1, r_2)^\top$ be the train state vector of size L, including the running resistance parameters to be estimated and Δt the sampling rate of the location and speed measurements. Since measuring is a time-discrete process, the system evolution can also be considered discrete and evaluated in time steps equal to the measurement sampling rate Δt, so that $x[k] \equiv x(k\Delta t)$, for $k = 0, 1, 2....$ Considering that the evolution of the location and speed is determined by Eq. (1) and neglecting the fact that the running resistance parameters can evolve due to nonmodelled dynamics, the discretization of the evolution of x can be described as follows:

$$x[k] = f(x[k-1], u[k-1]) + v[k-1] \tag{3}$$

where $f(x[k-1], u[k-1])$ is the process function, $u[k] = f_t[k] - f_b[k]$ is the control variable, which corresponds in this case to the tractive and brake efforts and $v[k-1]$ accounts for the noise in the process at the $(k-1)$-th time step, modelled as a Gaussian random variable of zero mean and covariance matrix Q.

$$f(x[k], u[k]) = x[k] + (\Delta t v[k] + 0.5(\Delta t)^2 a[k], \Delta t a[k], 0, 0, 0)^\top \tag{4}$$

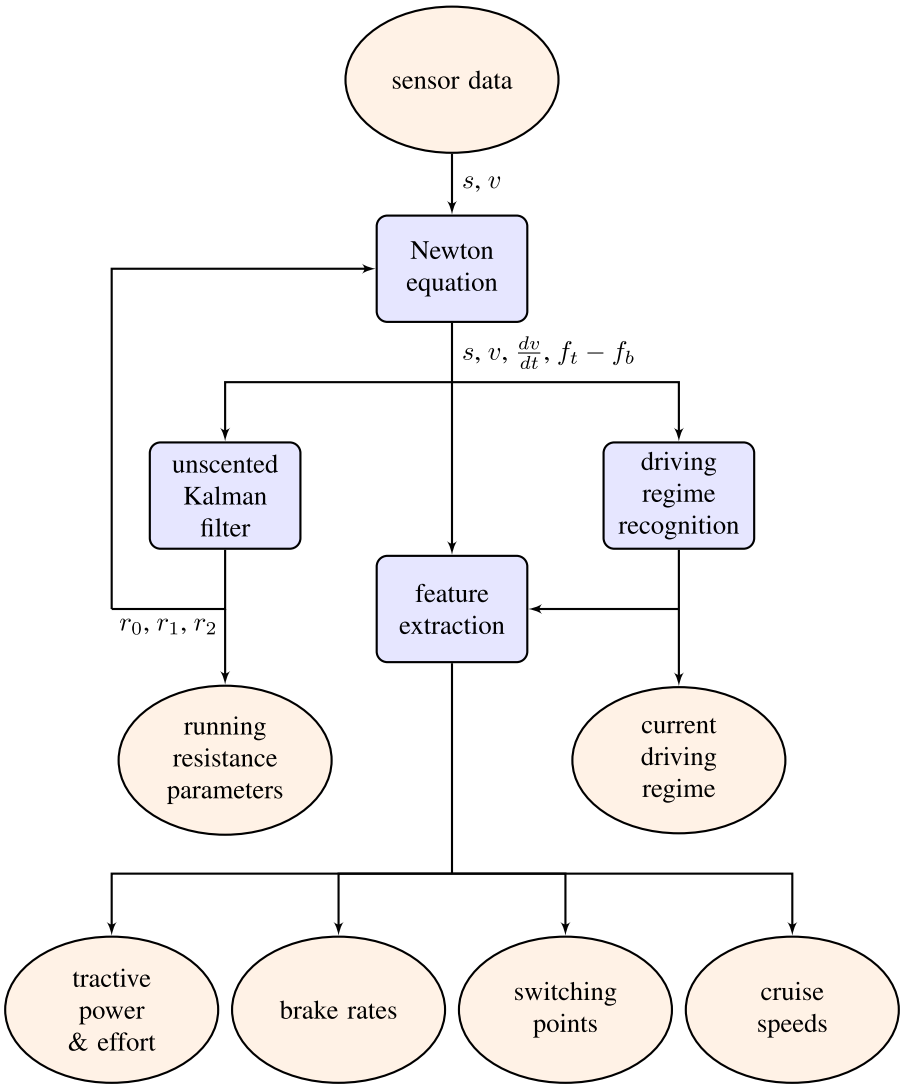

Fig. 1. Flowchart of the proposed algorithm.

where $a[k]$ is the acceleration term

$$a[k] = u[k] - r_0[k] - r_1[k]v[k] - r_2[k]v[k]^2 - g(s[k]). \tag{5}$$

Let $z[k]$ be the measurements at the time step k. The measurement function h is in this case linear, since measurements for the train location and speed are available in this case,

$$z[k] = (s[k], v[k])^\top = h(x[k]) + \omega[k], \tag{6}$$

where $\omega[k]$ is the measurement noise, which is again considered to be Gaussian with zero mean and covariance matrix R.

As mentioned earlier, in the UKF the state distribution is represented by a vector x consisting of L Gaussian random variables. At time step $k-1$ it has mean $\bar{x}[k-1]$ and covariance $P[k-1]$. In order to describe its statistics, UT requires to calculate a minimum of $2L+1$ sigma points $X_i[k-1]$, $i = 0, \dots, 2L$ to be able to capture the state statistics, with L the number of variables of the state. In this article, van der Merwe's scaled sigma point allocation is used (Van der Merwe, 2004), which is one of the existing ways of calculating sigma points (Menegaz et al., 2015),

$$\begin{aligned} X_0[k-1] &= \bar{x}[k-1], \\ X_i[k-1] &= \bar{x}[k-1] + (\sqrt{(L+\lambda)P[k-1]})_i, \quad i = 1, \dots, L, \\ X_i[k-1] &= \bar{x}[k-1] - (\sqrt{(L+\lambda)P[k-1]})_{i-L}, \; i = L+1, \dots, 2L, \end{aligned} \tag{7}$$

where $\lambda = \alpha^2(L+\kappa) - L$ is a scaling parameter, α and κ are parameters that are used to regulate the spread of the sigma points around the mean and the weights of each sigma point when calculating the statistics of the observed variables. The square root of a matrix is in this case performed as a Cholesky factorization and $(\sqrt{(L+\lambda)P[k-1]})_i$ corresponds to the ith column of matrix $\sqrt{(L+\lambda)P[k-1]}$. In order to obtain the prior distribution of the state at time step k, the corresponding sigma points $Y_i[k]$ are calculated by propagating the sigma points $X_i[k-1]$ through the system dynamics function in Eq. (8),

$$Y_i[k] = f(X_i[k-1], u[k-1]),\ i = 0, \dots, 2L. \tag{8}$$

Then, the mean and covariance of the prior distribution of the state x at time step k can be calculated by means of the UT as a weighted sum of the propagated sigma points, being w_i^m and w_i^c be the sigma weights associated to each sigma point for the calculation of the mean and covariance, respectively,

$$\begin{aligned} w_0^m &= \tfrac{\lambda}{L+\lambda}, \\ w_0^c &= \tfrac{\lambda}{L+\lambda} + 1 - \alpha^2 + \beta, \\ w_i^m &= w_i^c = \tfrac{1}{2(L+\lambda)}, \qquad i = 1, \dots, 2L, \end{aligned} \tag{9}$$

where β is a parameter usually used to incorporate prior knowledge of the state distribution. Therefore,

$$\bar{x}[k] \approx \sum_{i=0}^{2L} w_i^m Y_i[k], \tag{10}$$

$$P_x[k] \approx \sum_{i=0}^{2L} w_i^c (Y_i[k] - \bar{x}[k])(Y_i[k] - \bar{x}[k])^\top + Q. \tag{11}$$

However, the UKF corrects the prior statistics of the state by means of the information obtained from the measured variables. To this end, the propagated sigma points $Y_i[k]$ are mapped into the measurement space through the measurement function h.

$$Z_i[k] = h(Y_i[k]), \tag{12}$$

Analogously, the expected value of the measured variables and the associated covariance can be calculated through the UT,

$$\bar{z}[k] \approx \sum_{i=0}^{2L} w_i^m Z_i[k], \tag{13}$$

$$P_z[k] \approx \sum_{i=0}^{2L} w_i^c (Z_i[k] - \bar{z}[k])(Z_i[k] - \bar{z}[k])^\top + R. \tag{14}$$

Last, the Kalman gain $K[k]$ is calculated from the cross-covariance function $P_{xz}[k]$,

$$P_{xz}[k] \approx \sum_{i=0}^{2L} w_i^c (Y_i[k] - \bar{x}[k])(Z_i[k] - \bar{z}[k])^\top, \tag{15}$$

$$K[k] = P_{xz}[k] P_z[k]^{-1}. \tag{16}$$

Therefore, the new state estimate $\tilde{x}[k]$ and its covariance $P[k]$ can be calculated from the prior covariances, the Kalman gain and the expected value of the measured variables,

$$\tilde{x}[k] = \bar{x}[k] + K[k](z[k] - \bar{z}[k]), \tag{17}$$

$$P[k] = P_x[k] - K[k] P_z[k] K[k]^\top. \tag{18}$$

Even though the accuracy of the estimations does not depend significantly on the initial state of the system, it is specially sensitive to its initial covariance (Wan and Van Der Merwe, 2000). Moreover, parameters α, β and κ have to be carefully chosen in order to maximize the accuracy of the UT. The methodology used for selecting these parameters will be covered in Section 4.

3.2. Driving regime recognition module

In this article, we determine the driving regime in real time according to two criteria: the calculated applied tractive effort and brake, taking into account the current location and speed, and the train acceleration (De Martinis and Corman, 2019). First, we calculate the applied effort from the speed measurements. To cope with the noise present in those measurements, we define a tolerance of ± 0.025 N/kg in the applied effort in order to determine the coasting regime. Then, if the applied effort is higher than the upper bound, the train is applying traction, and if it is lower than the lower bound, it is braking. This bound constitutes approximately 5% of the maximum applicable tractive effort, according to the engine characteristics provided by the manufacturer. The lower bound for zero traction has been determined using the same value for simplicity, and both bounds have proven to determine correctly the driving regime on a predominantly flat track, as will be shown in the next section. Furthermore, the proposed framework should be able to distinguish whether a train is applying traction, brake or coasting even in steep uphill or downhill track section by means of this procedure, provided that the track description is sufficiently accurate. Moreover, if a train is exerting tractive

or brake effort, the observed acceleration is considered: if it is between ± 0.06 m/s^2, the train is cruising, otherwise it is applying traction to accelerate or braking, respectively. This bound neglects small speed variations up to 2 km/h between consecutive speed measurements, mitigating thus the impact on the cruising regime recognition of measurement errors, small driving variations and inaccuracies in the track description.

3.3. Feature extraction module

The results obtained from the UKF are post-processed in order to determine parameters like the maximum tractive effort and power that the engine can apply, the applied brake rates and information regarding the different driving regimes. To this end, the driving regime recognition module takes an essential role in extracting some data from the speed profiles. The extracted features are specially relevant for calibrating algorithms like train trajectory optimizers and for assessing the variations in manual driving styles.

Maximum tractive power and effort can mainly be observed when applying traction to accelerate. At low speeds the maximum tractive effort is assumed to decrease linearly with the speed or to be approximately constant. Here, it is calculated by fitting a linear model to the filtered tractive effort time series when the train is continuously accelerating between 10 km/h and 30 km/h. The maximum tractive power is computed by taking the mean of the filtered tractive effort divided by the train speed when the train is continuously accelerating between 50 km/h and 130 km/h. However, in the case of diesel-powered locomotives, the maximum tractive power is not unique, since each engine notch can apply a different maximum tractive power. This leads to jumps in the hyperbolic part of the tractive effort curves. This could be addressed, for instance, by monitoring the notch changes. Nevertheless, this multiple notch approach will not be considered in this article since here only data from electric-powered trains is used. The maximum tractive power and effort calculated by means of this procedure might not correspond to the real maximum capability of the engine. A lower adhesion condition may lead to wheel slip, lowering the effective applicable tractive effort below the engine capability. The overhead power may also influence the applied maximum tractive power, for example, when several trains depart from a station at the same time, the available catenary power is lower, reducing the acceleration rate. Besides physical sources of variation, driving-induced maximum tractive power and effort can also be observed, for instance, when a train driver accelerates gently to avoid disturbing the passengers' comfort or to compromise the integrity of the composition of freight trains. Therefore, although the mentioned observations may not represent the real traction capability of the train, they can still be considered as input of energy-efficient on-board applications like DAS to reduce their impact on the expected arrival time.

To calculate the brake rates, the fact that a train might not decelerate with a constant brake rate is taken into consideration. *Automatische TreinBeïnvloeding Eerste Generatie* (ATB-EG, in English, Automatic Train Protection - First Generation) is the most-widely installed automatic train protection system in the Dutch railway network. Besides high speed lines, the maximum speed limit in the Dutch railway network is 140 km/h. If there is no hazard ahead and the permanent speed limit of the track allows so, ATB-EG may show a Green signal, allowing the train to run with a maximum speed of 140 km/h. In the event of a yellow signal, ATB-EG may show 4 different signals, Yellow13, Yellow8, Yellow6 and Yellow, requiring the train to decelerate until reaching a speed below 130 km/h, 80 km/h, 60 km/h and 40 km/h, respectively. Moreover, the train driver is required to decelerate with a minimum deceleration rate until reaching the target speed, although the brake rate might be different for each type of yellow signal. If the driver fails to do so, ATB-EG intervenes and brakes the train. Therefore, in this article five speed steps are defined to build a deceleration model: over 130 km/h, 130 km/h–80 km/h, 80 km/h–60 km/h, 60 km/h–40 km/h and 40 km/h–0 km/h, and the mean observed deceleration rate in each speed step is calculated. Therefore, the proposed brake rate model may also be suitable for calculating mean deceleration rates in each ATB-EG speed step in the event of a yellow signal.

The values of the observation window of the maximum tractive effort and power and the speed steps related to ATB-EG are selected to match the particularities of the Dutch railway network and the rolling stock considered in the case study presented in the next section. Therefore, these values should be updated in any application in order to adjust the Feature extraction module to the expected braking behaviour and the tractive performance of the considered train.

Last, this module also calculates statistics of the cruise speed and the length of the final brake and of the previous coasting phase. To this end, the information obtained in the driving recognition module is utilized in the feature extraction module to calculate the final coasting and braking times. Moreover, the cruise speed during a cruising regime is easily determined by calculating the median of the filtered speeds along the mentioned regime.

4. Results and discussion

4.1. Case study

In this article, 67 runs of the same rolling stock unit in the Dutch railway network, departing from Eindhoven Centraal to 's-Hertogenbosch between March 2020 and December 2020, are considered. The mentioned speed profiles have been filtered, discarding the perturbed ones, that is, eliminating those that show an unexpected deceleration or stop due to a yellow or red signal. The considered unit is a manually-driven double-deck Intercity train with six carriages (VIRM-VI) operated by NS, the main railway operator in the Netherlands.

The available data for each run consists of GPS location and speed measurements, both of them sampled every 10 s. The GPS measurements are mapped into the track in order to calculate the mean gradient affecting the train at that time, taking into account its length. The track is 32103 m long and is predominantly flat, with small, but nonsteep uphill and downhill gradients. The speed is

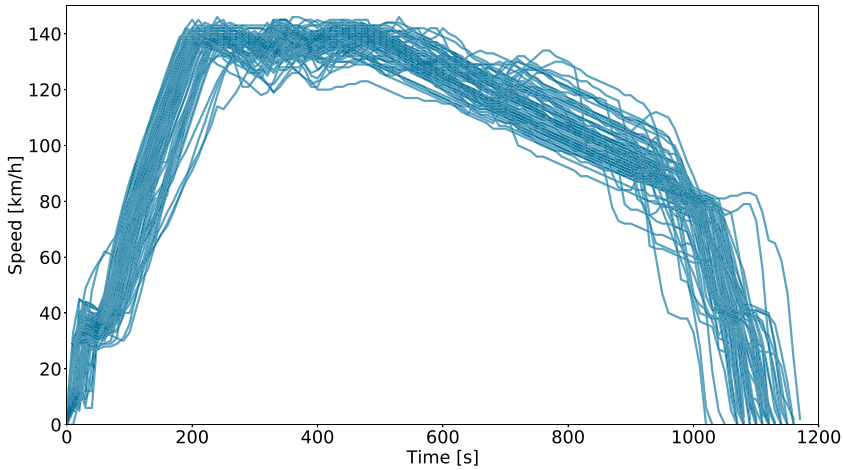

Fig. 2. The 67 considered speed profiles in the case study from Eindhoven Centraal to 's Hertogenbosch.

measured in km/h, with an accuracy of 1 km/h. Fig. 2 shows the speed profile of all the considered trajectories. Overall, a pattern can be distinguished in most of them consisting of an initial acceleration phase, cruising at a certain speed and coasting before braking. The overall maximum speed is 140 km/h save near the departure and arrival locations, where the speed limit is equal to 40 km/h. Moreover, variability is observed in the speed profiles due to the manual driving, not only in the running time between the two stations, but also in the cruising speed, the coasting and braking points and in the acceleration and deceleration rates when braking. However, the deceleration rates when coasting are significantly similar for all the speed profiles. Even though the maximum speed of the track is 140 km/h, in some cases the train runs with higher speed. The variability of the speed profiles and the driving style will be studied later in this section.

4.2. Unscented Kalman Filter

The tractive and brake effort needed to match the observed acceleration or deceleration for each location and speed measurement is calculated by means of Eq. (1), thus generating an effort time series for each speed profile. It is used along the location and speed measurements and the track height profile as input for the UKF algorithm. The train is assumed to be loaded at 35% of its seat capacity and the rotatory mass coefficient is considered constant and equal to 6% of the empty mass of the train, $\gamma = 1.06$, which corresponds to the value provided by the manufacturer. In order to estimate the running resistance parameters, they are considered as state variables along with the train location and speed, as described in Section 3.1. The evolution of the state variables at each time step is therefore described by Eqs. (3) and (4), where $\Delta t = 10$s, matching the measurement sampling rate. Taking into account the short computation time of the algorithm, the considered UKF implementation could be able to cope with smaller sampling rates, which might lead to more accurate estimates. However, subsampling the available data led to less accurate results, particularly in the Driving regime recognition and the Feature extraction modules. Therefore, considering the data available, we used the smallest sampling rate in this case study. However, no process noise is considered due to the difficulty of quantifying it without using ground data of the running resistance parameters and the train location and speed, which is unavailable in this case study, so the term $\nu[k-1]$ is equal to zero at every time step. However, we have noticed that small additive Gaussian white noise in the acceleration may lead to slightly richer dynamics of the estimates. We have also observed that either stronger acceleration noise or small noise in the running resistance parameters' equations lead to wrong estimates, namely negative estimates of the resistance parameters and unrealistic values of the applied effort. The UKF calculates the prior distribution of the state variables at the next time step, which is updated by means of the information obtained from the location and speed measurements at that time step. The standard deviation of the measurement noise of the location is considered to be 4 m and 1 km/h in the case of the speed.

Nevertheless, the UKF parameters have to be carefully selected in order to produce accurate estimations. These parameters are van der Merwe's scaled sigma point allocation parameters α, β and κ, and the initial covariance matrix of the state vector. This parameter selection has been performed according to two criteria: First, by maximizing the total sum of the logarithms of the conditional likelihood of the measurements with respect to the estimated location and speed and their respective estimated variances, assuming that the distribution of these variables is Gaussian (Scardua and Da Cruz, 2017). Second, by fine-tuning the obtained parameters, considering that the running resistance parameters are nonnegative. UKF was found to be very sensitive to the initialization values of the covariances of the measurements and the running resistance parameters. For instance, an increase in two orders of magnitude in one of the initial covariances of any running resistance parameter led to wrong running resistance parameter estimates that violated their nonnegativity and to calculating the applied effort calculation erroneously. Table 1 shows the parameters used in the UKF.

Fig. 3 shows the results obtained by the UKF for a single speed profile. The upper plot shows the measured speed profile as a blue dotted line and the estimated speed is shown as a red line. It can be observed that the speed profile is accurately reproduced by the

Table 1
UKF initialization parameters. α, β and κ are van der Merwe's scaled sigma point allocation parameters and $\sigma_{s,\text{ini}}$, $\sigma_{v,\text{ini}}$, $\sigma_{r_0,\text{ini}}$, $\sigma_{r_1,\text{ini}}$ and $\sigma_{r_2,\text{ini}}$ are the initial standard deviation of the location, speed, r_0, r_1 and r_2, respectively.

Parameter	Value	Parameter	Value
α	0.001	β	2
κ	−2	$\sigma_{s,\text{ini}}$	1 m
$\sigma_{v,\text{ini}}$	0.0018 km/h	$\sigma_{r_0,\text{ini}}$	10^{-6} N/kg
$\sigma_{r_1,\text{ini}}$	10^{-8} N/(kgm/s)	$\sigma_{r_2,\text{ini}}$	10^{-9} N/(kg(m/s)^2)

UKF. Moreover, the four mentioned driving regimes, namely acceleration, cruising, coasting and braking, can be distinguished and they take place in this order, being coasting the longest phase. The second and third plots show in red the calculated acceleration and deceleration rates for every measurement and the calculated tractive and brake effort, respectively. The dashed lines represent the maximum deceleration rate, -0.66m/s^2, and the speed-dependent maximum applicable tractive effort. This boundary has been calculated by means of the manufacturer values of the maximum applicable tractive effort and power, and assuming that for low speeds the maximum tractive effort is constant. It is observed that the applied tractive effort exceeds the mentioned theoretical boundary at the end of the acceleration regime. This will be analyzed later when discussing the results obtained from the feature extraction module. Next, along the constant speed regime a first interval where the acceleration is zero is observed, followed by a second interval in which the train reaccelerates briefly to gain speed and coasts when the train speed is higher than the target cruise speed. This exemplifies two of the most frequently used cruising techniques in manual driving: holding a constant speed and oscillating around a target speed by applying traction and coasting cycles. Then, the acceleration is slightly negative along the coasting regime, where the applied effort is approximately zero, save for small deviations due to noise in the measurements. Therefore, a small tolerance around zero traction and brake is required in the driving recognition module to determine the coasting regime. Last, during the final brake, it can be observed that in this case the train deceleration is lower than the 59% of the maximum deceleration rate, so the driver decelerates the train progressively. The running resistance parameters are shown in the last three plots. The UKF estimations are depicted in red, while the theoretical values are represented by a horizontal blue dashed line. It can be observed that the three parameters show similar dynamics. Moreover, the theoretical values are used as the initial state in the UKF, the running resistance parameters start decreasing at the beginning of the trajectory, which indicates that the theoretical values overestimate the real time-variant parameter values. For lower speeds, the parameters increase again, which might hint an inverse proportional relation between the resistance parameters and the speed. Although there is a tunnel at the entrance of an intermediate station, Best, we did not observe any influence on the estimated running resistance, probably due to the large cross-section of the tunnel. However, further studies would be required to assess the performance of the proposed framework on detecting the extra aerodynamic drag when running through tunnels and its impact on the running resistance.

Fig. 4 shows a histogram of the estimated values of r_0, r_1 and r_2 from all the speed profiles. The manufacturer value is represented by a vertical dashed line. The negative skewness of the distributions shows that the statistics here are biased due to the usage of the manufacturer value as the initial condition of the UKF algorithm. However, besides the outliers, parameter r_0 is bounded between 0.0072 N/Kg and 0.00738 N/Kg. The manufacturer value is near the mentioned upper bound and the distribution is biased due to its usage as the initial value of the UKF. Therefore, the manufacturer value overestimates the value of r_0. The other two running resistance parameters, r_1 and r_2, show similar statistics.

Fig. 5 shows a stacked area plot of each of the terms of the running resistance, where the parameters r_0, r_1 and r_2 are those obtained by the UKF for the speed profile shown in Fig. 3. In turn, the resistance due to the track geometry is depicted in magenta, in the same scale as the running resistance, to allow for comparison. The term r_0 is represented in dark blue, the term that depends linearly on the speed is represented in lilac and the quadratic term, in light blue. It may be observed that the variations of the constant term do not produce significant variations in the total running resistance, so this term may be consider constant in each interstation. The linear term accounts for up to the 11% of the total running resistance, only 2% below the constant contribution. However, Howlett et al. (2004) claim that this term contribution is small compared to the other two, and neglect this parameter and overestimate parameters r_0 and r_2 in order to overfit the total running resistance. Moreover, Bešinović et al. (2013) do not consider the linear term since it produces a negligible variation in the running times (Lukaszewicz, 2001). It may be concluded then, that even though the linear term of the running resistance can be sometimes neglected, this cannot be done systematically for any type of rolling stock, at least when energy consumption and accuracy are the main focus of the parameter estimation. At higher speeds, the quadratic term accounts for up to the 76% of the total running resistance, being therefore the most relevant term for energy consumption. As a consequence, most efforts have to be dedicated to estimating r_2.

4.3. Driving recognition module

The driving regime recognition algorithm described in Section 3.2 is used to assess the variability of the speed profiles with respect to the driving style. To this end, the cruising speed, and the coasting and braking times are calculated for each speed profile. The driving recognition module estimates the current driving regime for each measurement sample by means of the ruleset defined in Section 3.2. Fig. 6 shows the segmentation of the speed profile represented in Fig. 3 in terms of the different driving regimes. It may be observed that the reaccelerations part of the cruise regime is not fully considered as cruising, but short traction, coasting and brake phases are also determined along the mentioned cycles.

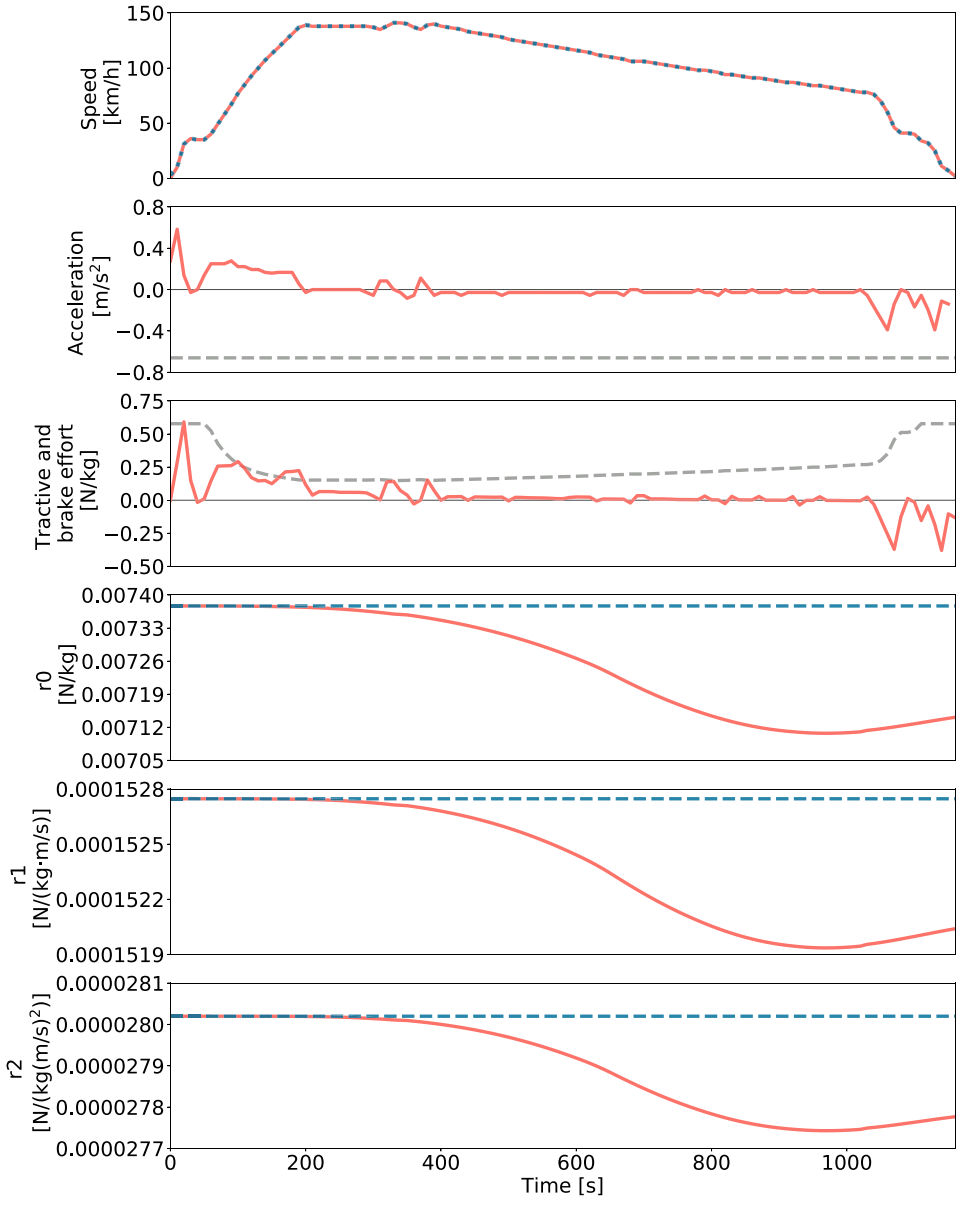

Fig. 3. Example of the performance of the UKF on a single speed profile.

For this trajectory, the train applies traction to accelerate in 16.4% of the total running time and it cruises at a constant speed during 17% of the time. In this interstation, the train coasts during 57% of the trajectory, and it brakes during the remaining 9.3%.

This module allows for obtaining statistics on driving style features like the cruising speed and the length of the final coast and brake regimes. Fig. 7 (left) shows a histogram of the observed cruise speeds in the 67 considered train runs. The cruise speed distribution is clearly skewed towards 140 km/h, which corresponds to the maximum speed of the track. This speed is exceeded in 12% of the speed profiles. Fig. 7 (center) shows a histogram of the lengths of the coast regime before the final brake. The final coast regime length can vary significantly, from 2.5 to 10 min. Statistics on the final brake are shown in Fig. 7 (right). Even though the length of the final brake regime is usually considerably shorter than the coast phase, they show comparable variations, since the coefficient of variation of the brake times is 20%, while the coefficient of variation of the coast times is 30%.

4.4. Feature extraction

The feature extraction module is used to obtain statistics of the brake deceleration rates for each of the ATB-EG speed steps. Fig. 8 shows the observed brake rates from all the considered speed profiles in blue, while the mean brake rates in each speed step

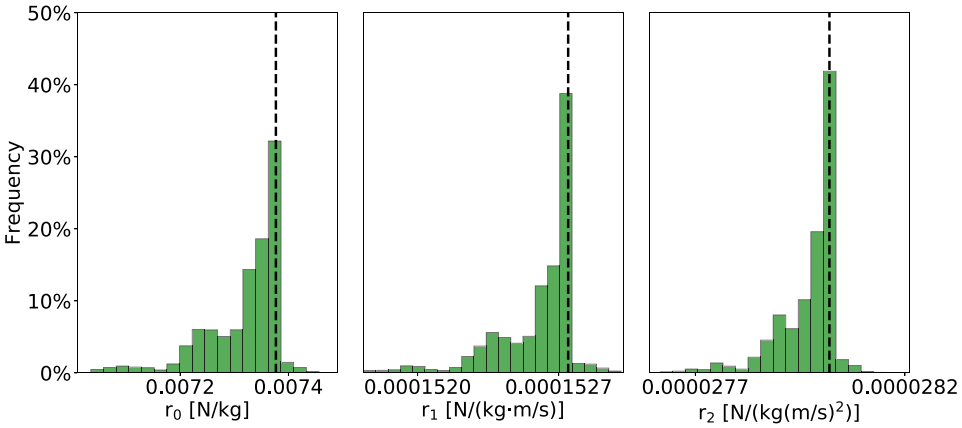

Fig. 4. Histogram of the occurrence of parameters r_0, r_1 and r_2 in the 67 speed profiles processed by the UKF. Their manufacturer values are represented by vertical black dashed lines.

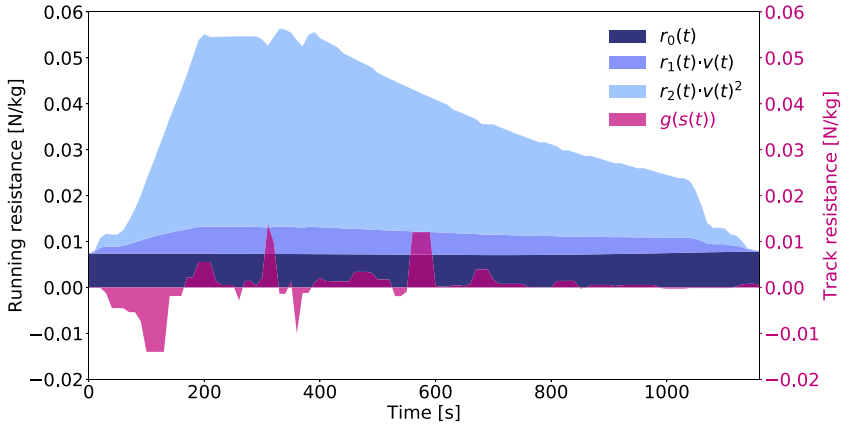

Fig. 5. Stacked area plot of the contribution to the running resistance $r(v)$ of each of the terms that constitute it, namely the constant term r_0, the linear term, $r_1 v$ and the quadratic term $r_2 v^2$. The resistance due to the track geometry is represented in the right axis.

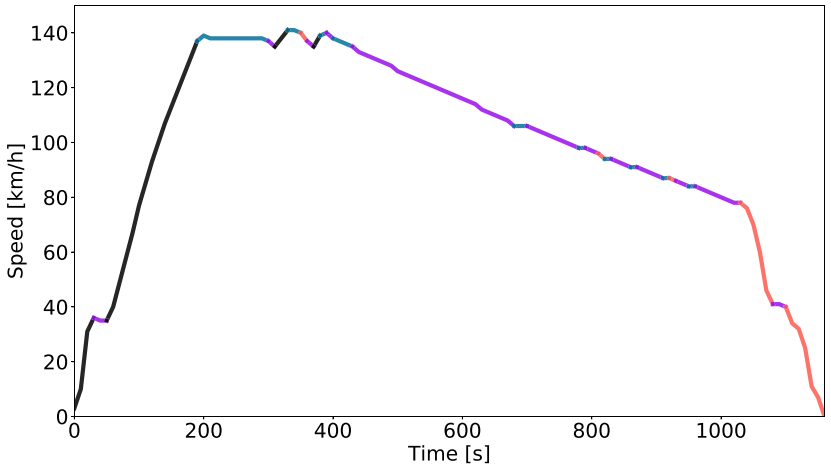

Fig. 6. Driving regimes of the speed profile shown in Fig. 3. The traction regime is depicted in black, cruising in blue, coasting in purple and braking in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

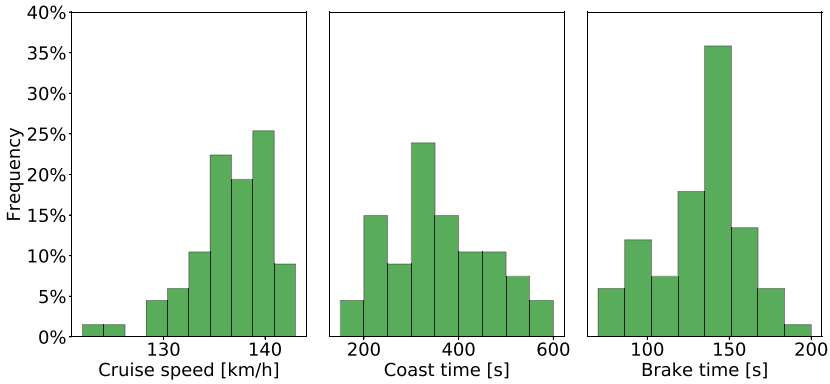

Fig. 7. Left: Histogram of the observed cruise speeds. Center: Histogram of the observed length of the coast regime before the final brake. Right: Histogram of the observed length of the final brake regime.

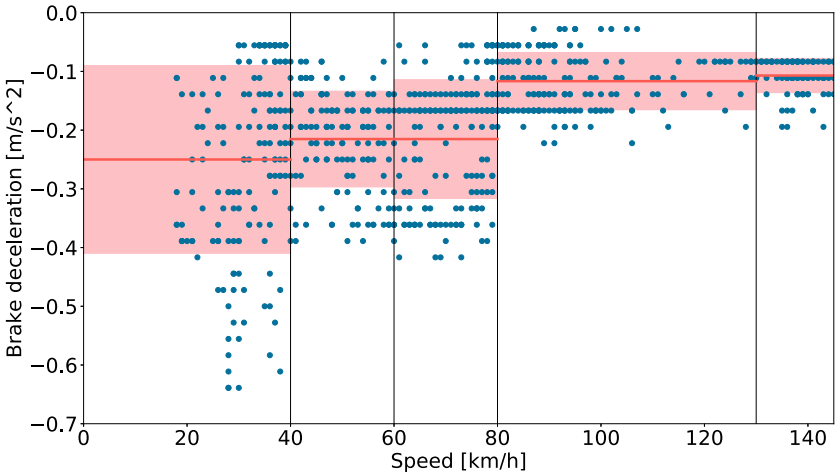

Fig. 8. Observed brake rates and mean brake rate in each ATB-EG speed step.

are represented in red, and the associated standard deviations are depicted as a shadow red area. Each speed step is separated by a vertical black line. A large variability is observed in the brake rates produced by the manual driving, however, the lower the train speed, the higher the mean applied brake rate. This is linked to the fact that brake systems usually provide a larger brake effort at lower speeds and that drivers are encouraged to decelerate gradually, since braking gently increases passengers comfort. Moreover, the variability is larger for lower speeds. The wider variance of the brake decelerations at speeds lower than 40 km/h might be due to the fact that ATB-EG does not associate a target deceleration rate for this speed step and that the train has to prepare to stop to a standstill at the stopping point of the track, which requires a certain level of accuracy. At larger speeds, drivers also tend to start braking gradually, therefore the initial deceleration rates are usually lower than the final ones. It is noteworthy that the ATB-EG speed steps between 40 km/h and 80 km/h and those speed steps higher than 80 km/h respectively share a similar braking behaviour, including the mean observed brake rate.

The maximum applicable tractive effort and power take a relevant role in railway applications like train trajectory optimization. An accurate estimation of these two quantities is essential for calculating accurately the running time and energy consumption of a train. However, estimating maximum bounds for the tractive effort and power from speed measurements can be complex. The information on whether the train driver is applying maximum traction or not is not available in this case study. Moreover, in the case of manual driving, the time in which the driver usually applies maximum traction is limited. When a train is at a standstill, train drivers tend to apply traction gently in order to guarantee the passengers comfort. Therefore, the first tractive effort measurements or calculations have to be discarded. Moreover, the maximum applicable tractive effort is usually modelled as constant for low speeds, and proportional to the inverse of the speed at higher speeds. The maximum applicable tractive effort in these two traction steps only coincides at a certain transition speed, which, according to the manufacturer values of the maximum tractive effort and power of the engine, in this case takes place around 38 km/h. However, the transition speed may vary depending on many factors such as the adhesion condition, the available electric power and the engine condition. Therefore, to guarantee that the tractive effort measurements and calculations used to calculate the mentioned upper bounds belong to the right tractive step, those tractive effort

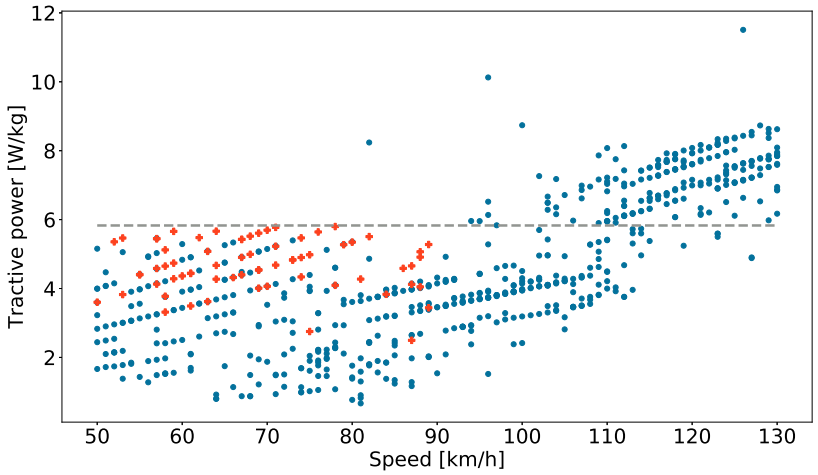

Fig. 9. Tractive power diagram of the initial acceleration phase of the 67 considered speed profiles.

calculations that take place around the transition speed have to be discarded. In addition, in this case study there is a 40 km/h speed limit that affects the train trajectory when applying traction to accelerate. Train drivers tend to reach such a speed limit by reducing the applied traction. Once the train is running past the mentioned speed limit, train drivers usually reaccelerate gently, that is to say, they do not apply maximum traction immediately. Therefore, in the maximum tractive effort and power calculations the calculated tractive effort for speeds below 10 km/h and between 30 km/h to 50 km/h are discarded.

In the case of the maximum tractive effort, due to the small speed window and the small amount of data available due to the low measurement rate, few data are available. Moreover, the low accuracy of the speed measurements, which is determined within an error of 1 km/h, introduces extra noise in the calculation of the tractive effort, producing a large dispersion in the set of calculated efforts at low speeds. Therefore, there is not enough data available in order to calculate relevant statistics of the maximum applied tractive effort. Nevertheless, in the case of the maximum applied tractive power, there is enough data available. Again, the data close to the maximum speed of the track, 140 km/h, is discarded. Fig. 9 shows the calculated tractive power for each speed measurement corresponding to the initial traction regime of each of the 67 considered train speed profiles. The calculated tractive power is represented by blue dots, and the value provided by the train manufacturer for the maximum tractive power is represented by a horizontal grey dashed line. Moreover, the observed maximum tractive power of each trajectory between 50 km/h and 90 km/h are represented by red crosses. Two clouds of points are observed. One in the lower left part of the figure, below the horizontal grey dashed line and for speeds below 115 km/h, and another in the upper right part of the figure, above the manufacturer value, for higher speeds. Furthermore, a line pattern in both clouds is observed due to the speed discretization with an accuracy of 1 km/h. However, all trajectories allocate tractive power points in both clouds. Therefore, a transition may take place in a certain point in the track. The difference in the calculated tractive power between both clouds is high enough so that it cannot be produced due to an inaccurate gradient profile of the track. Therefore, the observed transition may be due to the presence of several power stations, supplying energy with less power to the mentioned part of the track near 's Hertogenbosch station than to the next part of the track. Moreover, for the lower speeds, the train is departing from the station, where the available electric power can be lower than expected due to the fact that there are several trains using the available electric power. This explains the difference in tractive power and the fact that the transition does not take place at a specific speed, since, as can be observed in Fig. 2, there is a wide variability in the passing speeds in each point of the track. Moreover, there is a large variability in the observed maximum applied tractive power for lower speeds, although the manufacturer value constitutes in this case an accurate upper bound. The mentioned variability is probably due to the manual driving style, but nonhuman sources of maximum applied tractive power like variations in the available electric power at the station and train load may also be secondary sources of variability, which relevance could be assessed by analyzing data from ATO-driven trains.

5. Conclusion

A train motion model based on Newton's second law is the core of railway applications like energy-efficient train trajectory optimization, DAS and ATO. This model can reproduce train dynamics accurately when its internal parameters are carefully tuned. However, this parameter estimation is usually difficult, since they may vary stochastically in time due to a broad diversity of causes, from the current weather and adhesion condition to the mechanical wear of the train and the driving style of the driver. Therefore, an accurate real-time train motion model parameter estimation is essential to guarantee the performance and energy-efficiency of on-board railway applications and operations.

A real-time train motion model parameter estimation framework consisting of 3 modules has been devised. First, an Unscented Kalman Filter (UKF) module uses location and speed measurements as input data along the calculated traction and brake efforts

needed to match the observed acceleration to estimate the train running resistance parameters. Second, the observed acceleration and the calculated tractive and brake efforts are used in a driving recognition module to distinguish the driving regime at each measurement time. To this end, four driving regimes are defined: applying traction to accelerate, cruising, coasting and braking. Last, a feature extraction module is utilized to obtain statistics of the applied brake rates and the maximum applied power of the engine.

The proposed framework has been tested in a case study using real data, showing that the UKF is able to track the location and speed measurements and to provide an estimation of the running resistance parameters. The driving regime recognition module determines the current driving regime successfully. Moreover, some statistics have been computed on the observed cruise speeds, and the coast and brake lengths. Furthermore, the impact of manual driving on the applied brake rates and the maximum applied tractive power has been studied by means of the feature extraction module. In particular, the observed brake rates have been modelled in terms of the ATB-EG speed steps, showing that drivers tend to brake gently. From a statistical point of view, in this case three speed-dependent braking behaviours are observed: one for speeds higher than 80 km/h, another for speeds between 40 km/h and 80 km/h, and the last one for speeds lower than 40 km/h.

The proposed framework has a limitation regarding the availability of traction and brake efforts measurements, which are required to obtain an accurate estimation of the running resistance parameters. Even though calculating the applied tractive and brake efforts based on the observed acceleration allows to obtain approximate values of the running resistance parameters, it would smooth part of the running resistance dynamics, compromising then the performance of the UKF. Moreover, the driving recognition module could be updated in order to enhance its accuracy, and the information obtained could be used in real time to improve the performance of the parameter estimation module. Furthermore, a higher measurement sampling rate would improve the accuracy of the proposed framework.

This research opens new possibilities in railway operation data analysis. The case study included in this article could be extended to demonstrate that the proposed framework is also able to calibrate train dynamics of several rolling stock units and to detect the existing parameter variability within the different units. This would allow to demonstrate the importance of individual train calibration in the railway industry and academia. The performance of the proposed framework could be studied also for different types of trains: from high speed trains to light rail and metro. To address this, the internal parameter configuration of the framework should be retuned based on the characteristics of those railway systems. In addition, the spatial variability of parameters could be assessed by comparing the results obtained from trains running on different lines, which may lead to determining the impact of different track maintenance conditions on the dynamics of a train. The proposed framework could also allow to study temporal parameter variability. For example, data spanning a whole year could be used to study seasonal parameter variations, like the ones due to different temperature, weather, adhesion conditions and train wear. Moreover, data could be classified in an hourly basis and the impact of a tight timetable and delays on manual driving style could be assessed. What is more, if the data used is linked to the train driver, individual driving styles statistics could be computed and the differences between them could be studied.

CRediT authorship contribution statement

Alex Cunillera: Conceptualization, Methodology, Software, Data curation, Visualization, Writing – original draft. **Nikola Bešinović:** Writing – review & editing, Supervision. **Niels van Oort:** Writing – review & editing, Supervision. **Rob M.P. Goverde:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the Dutch railway undertaking Nederlandse Spoorwegen (NS) for funding the research and the data provided.

References

Antonov, S., Fehn, A., Kugi, A., 2011. Unscented Kalman filter for vehicle state estimation. Veh. Syst. Dyn. 49 (9), 1497–1520.

Bešinović, N., Quaglietta, E., Goverde, R.M.P., 2013. A simulation-based optimization approach for the calibration of dynamic train speed profiles. J. Rail Transp. Plan. Manag. 3 (4), 126–136.

De Martinis, V., Corman, F., 2018. Data-driven perspectives for energy efficient operations in railway systems: Current practices and future opportunities. Transp. Res. C 95, 679–697.

De Martinis, V., Corman, F., 2019. Online microscopic calibration of train motion models: towards the era of customized control solutions. In: RailNorrköping 2019. 8th International Conference on Railway Operations Modelling and Analysis (ICROMA), Norrköping, Sweden, June 17th–20th, 2019. Vol. 69. Linköping University Electronic Press, pp. 917–932.

Ghofrani, F., He, Q., Goverde, R.M.P., Liu, X., 2018. Recent applications of big data analytics in railway transportation systems: A survey. Transp. Res. C 90, 226–246.

Howlett, P., Pudney, P., Vu, X., 2004. Estimating train parameters with an unscented Kalman filter. In: Proceedings of the Fifth Asia Pacific Industrial Engineering and Management Systems Conference. Vol. 34. p. 13.

Jin, Y., Xie, G., Qian, F., Zhang, C., 2017. Online parameters identification of high speed train based on Gaussian sum theory. In: 2017 12th IEEE Conference on Industrial Electronics and Applications. ICIEA, IEEE, pp. 1493–1498.

Jin, Y., Xie, G., Zang, Q., Fan, L., Wen, T., Zhu, L., 2018. Modeling of train braking based on environment and online identification of time varying parameters. In: 2018 International Conference on Control, Automation and Information Sciences. ICCAIS, IEEE, pp. 451–455.

Julier, S.J., Uhlmann, J.K., 1997. New extension of the Kalman filter to nonlinear systems. In: Signal Processing, Sensor Fusion, and Target Recognition VI. Vol. 3068. International Society for Optics and Photonics, pp. 182–193.

Khadilkar, H., Balakrishnan, H., 2011. A multi-modal unscented Kalman filter for inference of aircraft position and taxi mode from surface surveillance data. In: 11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, Including the AIAA Balloon Systems Conference and 19th AIAA Lighter-than. p. 7051.

Liu, Y.J., Dou, C.H., Shen, F., Sun, Q.Y., 2021. Vehicle state estimation based on unscented Kalman filtering and a genetic-particle swarm algorithm. J. Inst. Eng. (India) Ser. C 1–23.

Liu, X., Ning, B., Xun, J., Wang, C., Xiao, X., Liu, T., 2018. Parameter identification of train basic resistance using multi-innovation theory. IFAC-PapersOnLine 51 (18), 637–642.

Lukaszewicz, P., 2001. Energy Consumption and Running Time for Trains (Doctoral dissertation, Doctoral Thesis). Railway Technology, Department of Vehicle Engineering, Royal Institute of Technology, Stockholm.

Lukaszewicz, P., 2007. Running resistance-results and analysis of full-scale tests with passenger and freight trains in Sweden. Proc. Inst. Mech. Eng. F J. Rail Rapid Transit 221 (2), 183–193.

Menegaz, H.M.T., Ishihara, J.Y., Borges, G.A., Vargas, A.N., 2015. A systematization of the unscented Kalman filter theory. IEEE Trans. Automat. Control 60 (10), 2583–2598.

Peng, X., Zhang, B., Rong, L., 2019. A robust unscented Kalman filter and its application in estimating dynamic positioning ship motion states. J. Mar. Sci. Technol. 24 (4), 1265–1279.

Rochard, B.P., Schmid, F., 2000. A review of methods to measure and calculate train resistances. Proc. Inst. Mech. Eng. F J. Rail Rapid Transit 214 (4), 185–199.

Scardua, L.A., Da Cruz, J.J., 2017. Complete offline tuning of the unscented Kalman filter. Automatica 80, 54–61.

Trivella, A., Wang, P., Corman, F., 2020. The impact of wind on energy-efficient train control. EURO J. Transp. Logist. 100013.

Van der Merwe, R., 2004. Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models (Doctoral dissertation). OGI School of Science & Engineering at OHSU.

Wan, E.A., Van Der Merwe, R., 2000. The unscented Kalman filter for nonlinear estimation. In: Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium. Cat. No. 00EX373, IEEE, pp. 153–158.