Transportation Research Part C 19 (2011) 1243–1262

Contents lists available at ScienceDirect

Transportation Research Part C

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

On the assessment of vehicle trajectory data accuracy and application to the Next Generation SIMulation (NGSIM) program data

Vincenzo Punzo [a,*], Maria Teresa Borzacchiello [b,1], Biagio Ciuffo [b,2]

[a] Dipartimento di Ingegneria dei Trasporti, Università degli Studi di Napoli Federico II, Via Claudio, 21, 80125 Napoli, Italy
[b] Institute for the Environment and Sustainability, European Commission – Joint Research Centre, Via E. Fermi, 2749, 21027 Ispra (VA), Italy

ARTICLE INFO

Article history:
Received 29 July 2009
Received in revised form 23 December 2010
Accepted 23 December 2010

Keywords:
Vehicle trajectory data
Data accuracy
Analysis of error
Delta method
Trajectory consistency
Spectral analysis
Traffic flow theory
NGSIM

ABSTRACT

Trajectories drawn in a common reference system by all the vehicles on a road are the ultimate empirical data to investigate traffic dynamics. The vast amount of such data made freely available by the Next Generation SIMulation (NGSIM) program is therefore opening up new horizons in studying traffic flow theory. Yet the quality of trajectory data and its impact on the reliability of related studies was a vastly underestimated problem in the traffic literature even before the availability of NGSIM data. The absence of established methods to assess data accuracy and even of a common understanding of the problem makes it hard to speak of reproducibility of experiments and objective comparison of results, in particular in a research field where the complexity of human behaviour is an intrinsic challenge to the scientific method. Therefore this paper intends to design quantitative methods to inspect trajectory data. To this aim first the structure of the error on point measurements and its propagation on the space travelled are investigated. Analytical evidence of the bias propagated in the vehicle trajectory functions and a related consistency requirement are given. Literature on estimation/filtering techniques is then reviewed in light of this requirement and a number of error statistics suitable to inspect trajectory data are proposed. The designed methodology, involving jerk analysis, consistency analysis and spectral analysis, is then applied to the complete set of NGSIM databases.

1. Introduction

Trajectories drawn in a common reference system by all the vehicles on a road are the ultimate empirical data to investigate traffic dynamics. With such knowledge we can derive whatever information is sought about the physics of traffic phenomena on a road, including dynamics like car-following, lane-changing or gap-acceptance. This also explains the great interest that such data and related measurement techniques have attracted in the field of traffic flow theory (see e.g. Wu et al., 2003; Toledo et al., 2009).

Obtainable measurements of the trajectory of a vehicle consist of discrete observations of its position, equally spaced in time, that is the series of vehicle coordinates in the two- or three-dimensional space (coordinates refer to a specific point of the vehicle, e.g. the front bumper centre).

Remote-sensing and object tracking from video or photo cameras has been the main technique applied so far to gather such precious data. Cameras have been attached to aerial platforms as in Kometani and Sasaki (1959), Treiterer and Myers

* Corresponding author. Tel.: +39 081 7683948; fax: +39 081 7683946.
E-mail addresses: vinpunzo@unina.it (V. Punzo), maria-teresa.borzacchiello@jrc.ec.europa.eu (M.T. Borzacchiello), biagio.ciuffo@jrc.ec.europa.eu (B. Ciuffo).
[1] Tel.: +39 0332 785506; fax: +39 0332 786325.
[2] Tel.: +39 0332 786782; fax: +39 0332 785236.

0968-090X/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trc.2010.12.007

(1974), Xing (1995), Hoogendoorn et al. (2004), or mounted at fixed and elevated locations, as in Ozaki (1993) and Coifman et al. (1998). The latter layout, in particular, allows vehicle trajectories to be collected only along a limited stretch of road. Generally, traffic flow is not disturbed nor influenced at all by the measurement operations following this approach.

Coordinates of a vehicle moving on a road can also be obtained by applying Global Positioning System (GPS) technology. Given the high expected accuracy in detecting positions (up to millimetres in the differential kinematic mode), the very precise timing of the system and the high sampling frequency of the measures, GPS technology is currently expected to provide the most accurate trajectory measurements. However, only traffic dynamics within a limited number of equipped vehicles can be observed, and not their interactions with the surrounding traffic – see e.g. the car-following experiments on vehicle platoons carried out on test track by Gurusinghe et al. (2002), and on real roads by Punzo and Simonelli (2005) – and only experimental, rather than empirical data, can be gathered, i.e. the driving behaviour of participants in the experiment is likely to be somewhat influenced by the experiment itself.

Other technologies allow detailed information to be gathered on vehicle interactions, even though they do not provide the trajectories of individual vehicles in a common reference space and generally apply to a single pair of vehicles. The very first experiments of such a data collection scheme were carried out using wire-linked vehicle pairs on test tracks and real roads in the pioneering works on car-following by the General Motors research team (see Chandler et al., 1958; Herman and Potts, 1959) where the spacing and relative speed between two vehicles and the follower's speed and acceleration were measured.

More recently, similar experiments were carried out by means of up-to-date technologies. For example, by equipping a vehicle with an optical speedometer and microwave radar providing relative spacing and speed, Wu et al. (2003), monitored the distance keeping behaviour between the test vehicle, driven by a test driver, and a preceding vehicle unaware of the experiment for long stretches of a route, in variable environments and traffic conditions.

All these experiments led to considerable improvements in the knowledge of traffic phenomena. Yet the limited amount of available data, their very different accuracy, their dishomogeneity as regards traffic conditions, road layouts and the extension of the observed time–space domain, as well as the very limited sharing of such data within the community of researchers, have not allowed a common understanding on many aspects of the traffic flow theory to be established. For instance, model calibrations and comparisons against different sets of measurements often led to contrasting results in car-following models, as pointed out in Brackstone and McDonald (1999).

Therefore, in 2002, the US Department of Transportation – Federal Highway Administration, started the Next Generation SIMulation (NGSIM) program, aimed at the theoretical development of new behavioural models and at collecting detailed trajectory datasets to estimate the model parameters and validate them (US Department of Transportation – FHWA, 2008a). Major innovations are the "open source" approach in developing the new behavioural models and free internet access to the vast amount of data gathered within the program. Data consist of vehicle trajectories, wide-area detector data and supporting data for researching driver behaviour. Vehicle trajectory data, in particular, are collected following the first approach mentioned: several synchronized video cameras, mounted on top of high buildings adjacent to the roadway, record vehicles passing through the study area (see e.g. the study area of the I-80 survey in Fig. 1). Post-processing of images finally gives vehicle positions on the road section every tenth of a second. In Table 1 the length of the road sections covered and the number of cameras per NGSIM survey site is reported.[3]

The free availability to the research community of detailed NGSIM data is opening up new horizons in the study of traffic flow theory and in its applications. Many studies have already benefited from using the trajectory data to perform experimental analyses and support theoretical findings by means of calibration and validation against such data. However, of more than 30 studies which used NGSIM data between 2007 and 2008 for different purposes, only four studies raised the issue of the quality of trajectory data (Duret et al., 2008; Hamdar and Mahmassani, 2008; Herrera and Bayen, 2008; Thiemann et al., 2008) and only two applied filtering methods in order to discard measurement errors (Hamdar and Mahmassani, 2008; Thiemann et al., 2008).

Yet depending on the aim of the study, the potential impact of measurement errors on the results can be significant. For example, Ossen and Hoogendhoorn (2008) verified the effects on car-following model calibration of adding, to vehicle displacements, errors from different distributions accounting for various combinations of random and systematic components. They basically pointed out that measurement errors yield a considerable bias in the calibration results.

In general, the dynamics under study could be partially hidden by the effect of measurement errors, and comparison of results of studies making use of different sets of data, with unknown accuracy, can be really hazardous (as well as uncertain their reliability).

Therefore the first necessary step to tackle such a problem is to define a methodology to inspect and quantify trajectory data accuracy. This is also needed to be able to develop more suitable trajectory estimation techniques. With this aim, in Section 2, the problem of estimating trajectory from discrete observations is analysed. First the structure of the error on point measurements and its propagation on the space travelled are investigated. In particular, under the specific assumption of independent errors on point measurements, evidence is given that the error propagated into the space travelled turns out to be a non-negative non-decreasing function of the number of observations. A consistency criterion for trajectories of following vehicles is therefore given. Literature on estimation/filtering techniques is thus reviewed in light of this consistency

[3] Note that the oldest surveys of the NGSIM program, namely "prototype" and "sample" datasets, are excluded from Table 1 as well as from the following analyses due to their lower accuracy.

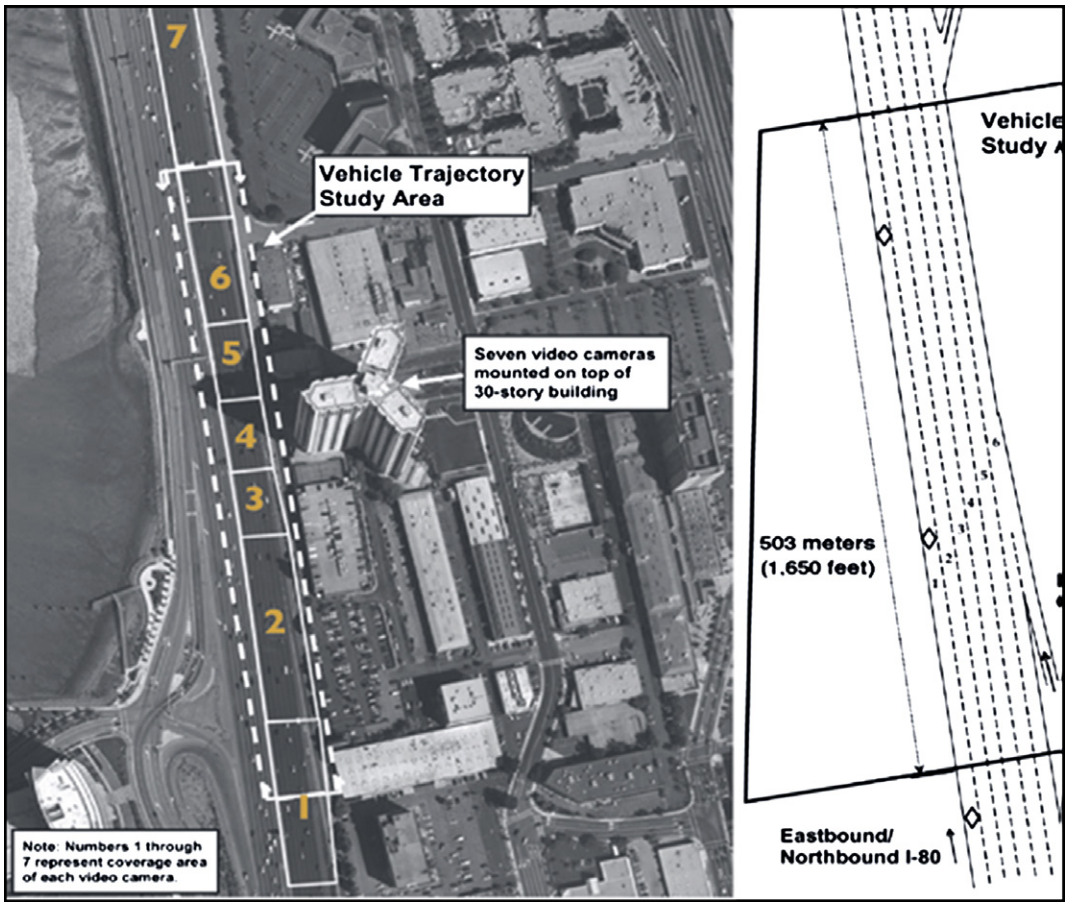

Fig. 1. Aerial photograph and schematic drawing of the I-80 study area. (*Source*: US Department of Transportation – FHWA, 2008b.)

Table 1
Main features of the NGSIM program collection sites.

Dataset name	I-80	US101	Lankershim	Peachtree
Road type	Freeway	Freeway	Arterial	Arterial
Site length (ft)	1650	2100	1600	2100
Site length (m)	503	640	488	640
Number of intersections			4*	5**
Number of cameras	7	8	5	8

* Signalized intersections.
** Un-signalized intersections.

requirement. In Section 3 a general methodology to inspect and quantify accuracy of trajectory data is then presented. The methodology drawn, involving analysis of jerks, consistency analysis and spectral analysis, is applied to the complete set of NGSIM databases in Section 4. The application of the methodology allowed us also to identify the main sources of data inaccuracy, illustrated in Section 4.6 along with some examples. Conclusions are then drawn and research prospects outlined.

2. The problem of estimating trajectory from discrete observations

Estimating vehicle trajectories from raw data, available through video cameras as well as GPS or radar devices, is far from simple. Data generally consist of a series of discrete observations equally spaced in time. Each observation comprises vehicle coordinates in two- or three-dimensional space (coordinates refer to a specific point of vehicles, e.g. the front bumper centre). Estimation aims to reconstruct, from sequential points, the "trajectory" of the observed vehicle, $\hat{s}_n(t)$, i.e. the function that gives the space travelled along its path by vehicle n until instant t (see also the definition for a one-dimensional

guideway by Daganzo, 1997) . From the trajectory function every characteristic of the vehicle motion can be derived, e.g. speed and acceleration functions. When the trajectories of vehicles in a given time–space domain are projected in a common reference system, also characteristics of traffic can be derived like e.g. time or space gaps, lane-changing durations, time to collisions, etc.

When gathering vehicle positions, whatever the technique used, the observed points happen to be dispersed in the neighbourhoods of the actual points and the path drawn on the observed positions inevitably fluctuates around the true path followed by the vehicle (see Fig. 2a for an example). Such measuring errors, when considering two consecutive vehicle positions, propagate into the value of the space travelled between these two points, that is to say into the estimated vehicle trajectory function, $\hat{s}_n(t)$:

(a) First, we may observe that the spaces travelled by a vehicle in two consecutive measurement intervals can differ greatly, up to unphysical values of this difference. This effect becomes clearly visible as a noise in $\hat{s}_n(t)$ that is even greater in the speed and acceleration functions since the differentiation process magnifies the noise whatever the numerical scheme adopted. This is the random component of the error in the space travelled.
(b) Secondly, the estimated cumulative space travelled by vehicle n until instant t, i.e. $\hat{s}_n(t)$, is generally greater than the true one. The evidence of a positive bias of $\hat{s}_n(t)$, i.e. the existence of a systematic component of the error as well, is given in Section 2.1, under the strict assumption of independent point measurement errors. Under such assumption the bias in the estimation of $\hat{s}_n(t)$ arises even assuming unbiased (i.e. zero mean) errors in the coordinates, being the effect of the error propagation. We will state the conjecture that such bias holds also for the error structures commonly underlying this type of measurements.

It is worth noting that, unlike the noise, the bias of $\hat{s}_n(t)$ is not self-evident if looking at the trajectory of a single vehicle, but it may become visible if one concurrently deals with the trajectories of a pair of following vehicles. Therefore in Section 2.2 it is shown how the use of the inter-vehicle spacing allows us to quantify this bias and define a criterion for verifying whether the estimated trajectories are biased. This criterion is referred to as *platoon consistency* and will also be the basis to calculate the bias of the trajectories in the NGSIM datasets (see Section 4). In Section 2.3 a brief review of the estimation/filtering methodologies currently applied to raw observations is presented. It is pointed out that most of these techniques, including the one employed within the NGSIM program, do not address the systematic component of the errors in the spaces travelled.

2.1. Bias in the space travelled

In this section, it is shown how errors in the measurement of coordinates are propagated into the space travelled and how this affects the trajectory estimation problem.

Let $P_i = (x_i, y_i)$ be the true position of a vehicle at instant i, and $\delta x_i \sim G(\mu_{x_i}, \sigma^2_{x_i})$ and $\delta y_i \sim G(\mu_{y_i}, \sigma^2_{y_i})$ the components of the measurement error along x and y, where G stands for a generic distribution (we refer to the bi-dimensional case only for the sake of simplicity). The observed position will therefore be $P_i^{obs} = (x_i + \delta x_i, y_i + \delta y_i)$. Let us assume that the space travelled by the vehicle between instants i and $i+1$ is equal to the Euclidean distance of points P_i and P_{i+1}. The error resulting from such an approximation can be generally neglected at a measurement sampling frequency of 10 Hz (which is that adopted in the NGSIM program, for example) as, usually, the higher the bend curvature and hence the difference between the arc and the chord subtended the lower is vehicle speed, that is the arc length travelled in 0.1 s. Adopting this approximation, for the sake of simplicity, the error on the space travelled by the vehicle between instants i and $i+1$ may be evaluated as:

$$\begin{aligned}\Delta_{i,i+1}(\delta x_i, \delta x_{i+1}, \delta y_i \delta y_{i+1}) &= \overline{P_i^{obs} P_{i+1}^{obs}} - \overline{P_i P_{i+1}} \\ &= \sqrt{[(x_i + \delta x_i) - (x_{i+1} + \delta x_{i+1})]^2 + [(y_i + \delta y_i) - (y_{i+1} + \delta y_{i+1})]^2} \\ &\quad - \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}\end{aligned} \tag{1}$$

The mean of the random variable $\Delta_{i,i+1}$, that is the sought bias in the space travelled, can be estimated using the Delta method (Oehlert, 1992). According to this method, if $X_1, \ldots, X_n$ are n continuous random variables with known means $m_1, \ldots, m_n$ and variances $s_1^2, \ldots, s_n^2$ and $\varphi(\cdot)$ is a real, continuous and differentiable function, the mean of the random variable $Y = \varphi(X_1, \ldots, X_n)$, can be evaluated as follows:

$$E[Y] \approx \varphi(m_1, \ldots, m_n) + \frac{1}{2}\sum_{r=1}^{n}\left[\left(\frac{\partial^2 \varphi}{\partial X_r^2}\right)_{m_1,\ldots,m_n} \cdot s_r^2\right] + \sum_{r=1}^{n-1}\sum_{j=r+1}^{n}\left[\left(\frac{\partial^2 \varphi}{\partial X_r \partial X_j}\right)_{m_1,\ldots,m_n} \cdot E[(X_r - m_r)(X_j - m_j)]\right] \tag{2}$$

which, in the hypothesis that $X_1, \ldots, X_n$ are independent, becomes:

$$E[Y] \approx \varphi(m_1, \ldots, m_n) + \frac{1}{2}\sum_{r=1}^{n}\left[\left(\frac{\partial^2 \varphi}{\partial X_r^2}\right)_{m_1,\ldots,m_n} \cdot s_r^2\right] \tag{3}$$

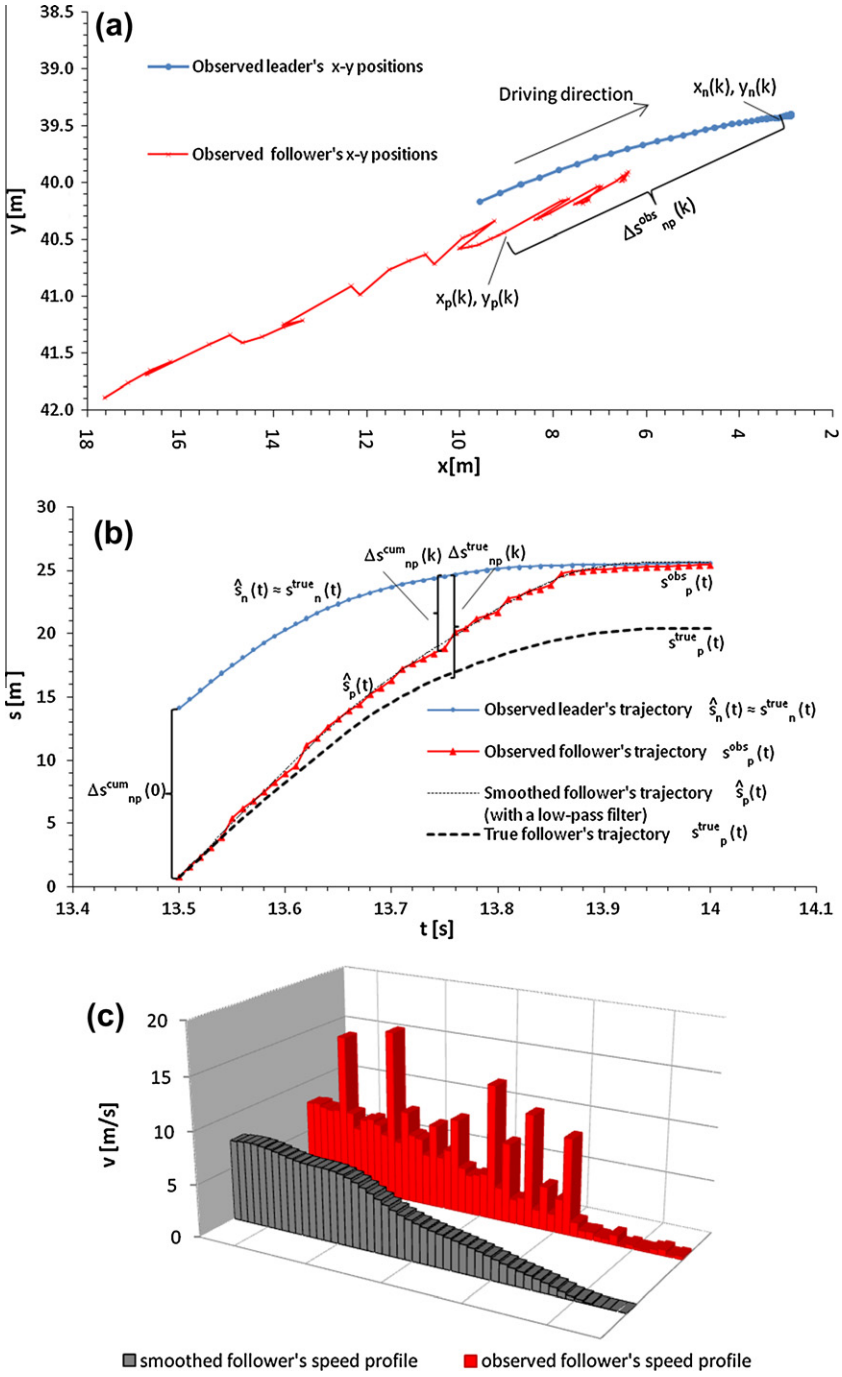

Fig. 2. (a) Absolute positions in the (x–y) space of two following vehicles; (b) observed, smoothed and true trajectories of two following vehicles; (c) observed and smoothed speed profile of the follower. Data source Punzo et al. (2005).

Thus, for $Y = \varphi(X_1, \ldots, X_n) = \Delta_{i,i+1}(\delta x_i, \delta x_{i+1}, \delta y_i, \delta y_{i+1})$ given by (1), let us assume that $\delta x_i, \delta x_{i+1}, \delta y_i, \delta y_{i+1}$ are independent random variables, i.e. that Eq. (3) holds. In addition, let us assume that $\delta x_i, \delta x_{i+1}, \delta y_i, \delta y_{i+1}$ are also identically distributed with mean m and variance s^2 : $\delta x_i, \delta x_{i+1}, \delta y_i, \delta y_{i+1} \sim G(m, s^2)$. Eq. (3) therefore yields the following compact expression:

$$E[\Delta_{i,i+1}] \approx \varphi(m, \ldots, m) + \frac{1}{2}s^2\left(\sum_{j=0}^{1}\frac{\partial^2\Delta_{i,i+1}}{\partial^2\delta x_{i+j}} + \sum_{j=0}^{1}\frac{\partial^2\Delta_{i,i+1}}{\partial^2\delta y_{i+j}}\right)_{m,\ldots,m} = \frac{s^2}{\sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}}$$

which represents the sought systematic component of the error made in the evaluation of the space travelled between instants i and $i+1$.

The bias in the estimation of the space travelled by vehicle n until instant k, $\epsilon_n(k)$, that is the mean of the difference between the estimated and true trajectory functions is therefore:

$$\varepsilon_n(k) = E[\hat{s}_n(k) - s_n(k)] = E\left[\sum_{i=1}^{k-1} \Delta_{i,i+1}\right] = s^2 \sum_{i=1}^{k-1} \frac{1}{\sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2}} \tag{4}$$

Eq. (4) shows that, in case of IID measurement errors on vehicle points' coordinates, the space travelled by the vehicle and estimated from these points is biased. Such bias is a non-negative, non-decreasing function of the number of observations.

This result holds also for errors that are only independent, i.e. not identically distributed, merely yielding a less compact formula than Eq. (4).

To quantify this bias and evaluate the Delta method approximation, Eqs. (1) and (4) have been applied to synthetic trajectories and compared. In fact, the (x, y) positions of a vehicle along a straight road have been generated with a frequency of 10 Hz in four scenarios differing in the vehicle speed profile (the number of points generated in each scenario varying according to such speed profiles). In particular, the speed profiles in the first three scenarios were constant, while the fourth was the actual speed profile of the first vehicle gathered in the field experiment referred to as 25/02B in Punzo et al. (2005) (also applied to construct Fig. 2). Finally, 100,000 different noisy paths were generated per each scenario by adding circular Gaussian errors, $\delta_x = \delta_y \sim N(0, 100\ \text{cm}^2)$, to the original coordinates. The bias in the space travelled given by (4) was therefore compared to the average of those calculated through (1) over the 100,000 noisy paths.

Table 2 shows the results. It can be noted that the bias given by (4) is very close to that numerically calculated through (1) (note that Eq. (4) cannot be applied in the last scenario, as the vehicle speed, and thus the distance between successive observations gets null values in some points, making the error calculated by Eq. (4) raising to infinite values). Furthermore, the bias is sensibly affected by the vehicle speed as, given the same distance travelled, both the number of terms in the summation, i.e. the number of measurement points, and the values of such terms increase with the lowering of speed.

Results of Table 2 show that, in the hypothesis of IID errors and in case of heavily congested conditions, the bias in the estimation of travelled space is all but negligible.

It is worth noting that the gathering techniques and the algorithms usually applied generally introduce autocorrelation in the measurement errors. Therefore, our conjecture is that a positive bias in the travelled space is likely to occur even with the error structures usually encountered in the practice that, generally, are mean reverting. In this sense results of Table 2 have to be intended only as indicative and are mainly useful to understand which order of magnitude this bias may reach and how this magnitude is influenced by the vehicle speed.

Finally, it is worth noting that, if the (x, y) coordinates of a vehicle are projected on the lane alignment, the resulting longitudinal component of the total space travelled (i.e. the projected trajectory) is not affected by the bias just described. This suggests that in all the cases where it is acceptable to confuse the actual vehicle path with its projection on the lane alignment, as is the case of "pure" car-following studies, the projection of the coordinates on the lane alignment is the basic way to eliminate the bias in the space travelled. Alternatively, to eliminate the bias, it is possible to project the noisy points along an estimate of the actual vehicle path (obtained, for instance, by filtering the unphysical curvatures given by the noisy points by means of a low-pass filter).

2.2. Consistency of vehicle trajectories

When focusing on two following vehicles, the errors arising in the estimation of both the trajectory functions can alter their inter-vehicle spacing, even up to unphysical values. This is what happens in the example of Fig. 2 where, due to the noisy observed positions of the following vehicle (see Fig. 2a), at a stop, its estimated trajectory function overlaps the leader's (see Fig. 2b) i.e. the vehicles' front bumpers overlap while, in reality, the follower is still safely behind the leader. Such a situation would make unfeasible the use of data for car-following model calibration, for example.

The previous considerations suggest that the spacing between a vehicle pair can be used to quantify the error in the estimated trajectories. In fact, independently of the trajectory estimation, the inter-vehicle spacing in one instant can be measured directly from the positions of the vehicles in that instant, for example assuming that it is the length of the road section separating the vehicles. Under such assumption it can be formally defined as:

Table 2
Bias in the space travelled along paths 1809 m long.

Scenario	Average speed (m/s)	SD of speed (m/s)	Bias given by (1)* (m)	Bias given by (4) (m)
1	27.78	0	2.35	2.35
2	13.9	0	9.40	9.40
3	5.7	0	56.63	55.70
4	5.7	4.5	184.8	–

* Values obtained applying (1) to all the points of each path and averaging over the 100,000 paths.

$$\Delta s_{np}^{obs}(k) = c(\overline{P_{n,k}^{obs}}) - c(\overline{P_{p,k}^{obs}}) \tag{5}$$

where $\Delta s_{np}^{obs}(k)$ stands for the *observed* inter-vehicle spacing between vehicles n and p at time k, and $c(\overline{P_{j,k}^{obs}})$, with $j = n, p$, represents the station of vehicle j at time k, that is the length of the lane alignment from the road origin to the point $\overline{P_{j,k}^{obs}} = (\overline{x_{j,k}^{obs}}, \overline{y_{j,k}^{obs}})$, projection on the lane alignment of the actual vehicle position $P_{j,k}^{obs} = (x_{j,k}^{obs}, y_{j,k}^{obs})$.

When vehicles travel along a straight line the computation $\Delta s_{np}^{obs}(k)$ becomes trivial, as it can be straightforwardly obtained from the coordinates:

$$\Delta s_{np}^{obs}(k) = \sqrt{\left(x_{n,k}^{obs} - x_{p,k}^{obs}\right)^2 + \left(y_{n,k}^{obs} - y_{p,k}^{obs}\right)^2}$$

Anyway, it is worth noting that the calculation of $\Delta s_{np}^{obs}(k)$ needs only the knowledge of the positions of the vehicles pair in the instant k and therefore this measure is affected only by measurement errors made in that single instant.

On the other hand, inter-vehicle spacing can also be derived from the trajectory functions of a vehicle pair. This calculation may be referred to as *cumulative* inter-vehicle spacing, $\Delta s_{np}^{cum}(k)$, where the adjective *cumulative* recalls that it is calculated as the difference of the cumulative space travelled by the vehicles (in Fig. 2b it is the vertical distance between the trajectory functions of the two following vehicles):

$$\Delta s_{np}^{cum}(k) = \Delta s_{np}^{cum}(0) + \hat{s}_n(k) - \hat{s}_p(k) \tag{6}$$

where $\hat{s}(\cdot)$ is the vehicle trajectory function estimated from discrete observations[4] and $\hat{s}_j(0) = 0$, with $j = n, p$.

It is worth noting that, for two vehicles travelling along the lane alignment, the two previous different measures of the inter-vehicle spacing must be equal. However, in practice, the calculations, given by Eqs. (5) and (6), always assume different values, as vehicles do not strictly travel along the lane alignment and, above all, measurement errors come into play differently in (5) and (6). $\Delta s_{np}^{obs}(k)$ is calculated by means of the observed positions in the sole instant k (see Fig. 2a), while $\Delta s_{np}^{cum}(k)$ entails the estimation of trajectories $\hat{s}(\cdot)$ along the entire interval $[0, k]$ (see Fig. 2b) and hence is affected by a number of measurement errors.

Thus calculations given by (5) are definitely more accurate than those given by (6) and may be used as independent reference measurements in the trajectory estimation process. A criterion to verify if trajectories are consistent is thus straightforward:

$$\Delta s_{np}^{obs}(k) \approx \Delta s_{np}^{cum}(k) \quad \forall k \tag{7}$$

which is what we refer to as *platoon consistency*. It is worth noting that criterion (7) is valid whatever the measurement error structure is.

Eq. (7) can be adopted to constrain the estimation of the trajectories of a pair of vehicles to respect the inter-vehicle spacing measured through (5) that is the concept underlying the estimation method proposed in Punzo et al. (2005). Obviously, when trajectory data are used for studying traffic characteristics and not the motion of individual vehicles, respect of this requirement is absolutely necessary.

Once consistent estimations of trajectories in the sense given by (7) have been obtained, speeds and accelerations have to be calculated respecting the basic equations of motion as well[5]:

$$\hat{s}(k) = \hat{s}(0) + \int_0^k \hat{v}(t)dt \tag{8a}$$

$$\hat{v}(k) = \hat{v}(0) + \int_0^k \hat{a}(t)dt \tag{8b}$$

where $\hat{v}(t)$ and $\hat{a}(t)$ are the estimates of the speed and acceleration functions. Eqs. (8) are referred to as *internal consistency* by Toledo et al. (2007), who claim that it may be more useful to evaluate internal consistency than platoon consistency. Actually, the respect of Eqs. (8) ensures that the differentiation of $\hat{s}(t)$ does not introduce additional errors, but it does not tell us if the estimated $\hat{s}(t)$ functions are unbiased, that is, if the actual platoon dynamics are preserved.

By contrast, platoon consistency can easily embrace internal consistency by including Eqs. (8) in (6) and then in (7). For example, considering the first of Eqs. (8) yields:

$$\Delta s_{np}^{obs}(k) \approx \Delta s_{np}^{cum}(0) + \int_0^k \hat{v}_n(t)dt - \int_0^k \hat{v}_p(t)dt \tag{9}$$

which is the original form in which the problem of trajectories' consistency was stated in Punzo et al. (2005).[6]

As mentioned in Section 2.1, if the point coordinates are projected on the lane alignment, the resulting space travelled is unbiased (being affected only by the measurement errors in the first and last positions of the vehicle). Eq. (6) becomes the

[4] From a practical point of view, $\Delta s_{np}^{cum}(0)$ in Eq. (6) can be calculated through Eq. (5).

[5] The estimation process, of course, does not necessarily have to follow this order.

[6] In June 2007, the TRB delivered a corrected version of the paper, due to an editing error in the original paper in the comment to Eq. (1) – just about trajectories' consistency. Please contact the corresponding author to get a copy of the amended paper.

Table 3
Estimation techniques applied in the literature.

	Averaging	Smoothing	Kalman filtering	Kalman smoothing
Coordinates (x, y)	Hamdar and Mahmassani (2008)		Ervin et al. (1991)	
Trajectory, $s(\tau)$	Ossen and Hoogendhoorn (2008), Thiemann et al. (2008)	Punzo et al. (2005), Toledo et al. (2007)		Ma and Andreasson (2006)
Inter-vehicle spacing, $\Delta s(\tau)$			Punzo et al. (2005)	Ma and Andreasson (2006)
Speed, $v(\tau)$	Hamdar and Mahmassani (2008), Thiemann et al. (2008)	Brockfeld et al. (2004), Punzo et al. (2005), Lu and Skabardonis (2007), Xin et al., 2008	Punzo et al. (2005)	Ma and Andreasson (2006)
Acceleration, $a(\tau)$	Thiemann et al. (2008)	Xin et al. (2008)		

same of (5) and platoon consistency by (7) is strictly observed. In spite of this, Eq. (9) still proves to be useful for inspecting the accuracy of trajectories as it allows us to ascertain the presence of errors in the estimated speeds and to quantify them by looking at their effect on platoon dynamics. Section 4.3 and the examples in Section 4.6 clarify this point.

2.3. Background about estimation/filtering techniques

If not properly accounted for, the inconsistencies in the estimated vehicle speeds and accelerations are not generally negligible in practical applications. This is shown in Section 4 as regards the NGSIM data. Therefore to gain deeper insight into the problem of estimating vehicle trajectory data with particular concern for the issue of estimating consistency, a brief review of the filtering techniques applied so far is presented below.

According to Rayner (1971), *filtering* is just a general name given to the process of systematically modifying data which are arranged in a sequence or array. *Averaging* is one form of *smoothing*, which in turn is one form of filtering. Following this very basic consideration a number of techniques so far applied to the estimation/filtering problem of vehicle positions is reported in Table 3. Columns are arranged in increasing order of complexity from Averaging to Kalman Smoothing while rows point at the variable to which each technique was applied.

In general, it can be said that smoothing removes the high-frequency and, in part, the medium-frequency disturbances from the data. Essentially, it is a low-pass filter, since it allows the low frequencies through. In other words, when applied to position data it can remove the random component of the error, but not the bias, unless specifically conceived for. In the example of Fig. 2, a low-pass filter namely a Butterworth filter is applied to smooth the trajectory function $\hat{s}_n(t)$ of the follower. The filter effectiveness in removing the noise of the $\hat{s}_n(t)$ function (see Fig. 2b) and also of the $\hat{v}_n(t)$ function (see Fig. 2c) is manifest. However, such a filter preserves the integral function of $\hat{s}_n(t)$, that is the cumulative space travelled and hence the bias. This can be verified comparing the "smoothed follower's trajectory" with the "true follower's trajectory" in Fig. 2b.

In general, all the techniques in Table 3 catalogued as "Averaging" and "Smoothing" respond to this behaviour. Attention must thus be paid since they may remove some detailed information while leaving part of the noise. The most basic technique among those referred is the simple moving average. In Ossen and Hoogendhoorn (2008) it is applied to synthetic noisy data, while Hamdar and Mahmassani (2008) apply a mono-dimensional Gaussian Kernel moving average to eliminate discontinuities in lateral x-coordinate and speeds from NGSIM data. They report minor discontinuities after smoothing. Thiemann et al. (2008) use a symmetric exponential moving average filter to smooth NGSIM data and propose to address the problem of noise magnification after differentiation, first by differentiating positions to speeds and accelerations and then by smoothing all three variables. It is worth noting that the moving average generally needs a sensitivity analysis to find windows and weights that provide the best trade-off between the damping of the noise and the losing of high-frequency parts of original data.

Among those who apply smoothing techniques, instead, Brockfeld et al. (2004) calculate the accelerations from speeds from Differential GPS data by means of a Savitzky–Golay smoothing filter, with a second-order polynomial over one-second windows. Instead, Toledo et al. (2007) apply a locally weighted regression technique to derive instantaneous speeds and accelerations in the central point of the moving window. They report good results of the method in terms of internal consistency as given by Eqs. (8).

Subsequently, Xin et al. (2008) employ a bi-level optimization structure that attempts to minimize measurement errors and preserve internal consistency in estimating speeds and accelerations from position data (from video recordings). They attain better results, in terms of internal consistency, compared to locally weighted regression.

As already stated, none of the techniques mentioned so far address the problem of the platoon consistency. To tackle this problem, Punzo et al. (2005) apply a non-stationary Kalman filter to DGPS data, introducing inter-vehicle spacing in model and measurement equations. The results are compared to those from the applications of a Butterworth filter to $\hat{v}_n(t)$ and of a

locally weighted regression technique to $\hat{s}_n(t)$. Comparisons point out the significant inconsistency of the trajectories filtered by the two last methods against those by Kalman filtering.

Earlier, Kalman filtering had been applied in a different form, to position data, by Ervin et al. (1991), who designed a stationary filter to process (x, y) coordinates and yaw. They report results of application on noisy synthetic data and no mention to the consistency of the estimated trajectory is given. Subsequently, Ma and Andreasson (2006) apply a Kalman smoother first to displacements and speeds of an instrumented vehicle and then to inter-vehicle spacing from an observed vehicle (measured by means of a laser beam scan). However, due to the particular experimental setting no information about platoon consistency of estimates could be derived.

Finally, in the NGSIM project, from the information reported and the analysis of the available dataset, it is possible to argue what follows. From the image processing system implemented, the position of each single vehicle is directly reported with a frequency of 10 Hz in a global geographical reference system. Such coordinates are then projected in a local coordinates system which follows the road monitored. In particular, one of the axes of this local system overlaps the road alignment (named "LocalY" as will be shown in the remainder), whereas the other is perpendicular. The speed provided is likely to be then calculated from the "LocalY" using a locally weighted regression[7] (that, as already mentioned, does not address the platoon consistency).

3. Methodology

In this section the methodology for analysing trajectory data, later applied to NGSIM data, is introduced. In particular, three kinds of analysis are presented: the first examines the jerk values from trajectory data in order to check acceleration feasibility; the second aims at verifying both the platoon and internal consistency of data; while the third involves the examination of spectral frequencies on speed, acceleration and jerking.

3.1. Jerk analysis

Vehicle acceleration plays an important role in studying dynamics of autonomous vehicles as well as of traffic streams. The most obvious way of verifying acceleration data is to inspect its distribution over the entire dataset. Unfeasible extreme values or an anomalous shape of the distribution are clear indications of problems in the data collection and estimation process.

Apart from the distribution of the acceleration values, inspection of their temporal dynamics gives an important indication of data quality. Such analysis is done here by using the jerking factor, j [m/s^3], namely the derivative of the acceleration, which physically represents the variations of acceleration in time. Jerking values in reality attain 10 m/s^3 (32.80 ft/s^3) as an absolute value (Martinez and Canudas-de-Wit, 2007), even though comfortable values do not exceed 2 m/s^3 (Schultz and Rilett, 2005, considered values around 3 m/s^3 (9.84 ft/s^3) as acceptable for applications of traffic flow micro-simulation).

Both an analysis of extreme jerk values and of its sign variation is herein proposed. In particular, values exceeding the threshold of 15 m/s^3 (49.20 ft/s^3), which is appreciably higher than the comfort threshold, are considered not mechanically feasible. Furthermore, the number of times that jerking reverses its sign in a one-second time window is analysed.

It is claimed that more than one sign inversion in a one-second time window is not physically consistent with human and mechanical observed responses (Jagacinski and Flach, 2003), or common sense. Indeed, similar behaviour would involve sudden inversion in the acceleration slope beyond the physical possibilities of human/vehicle systems.

The indicators suggested are therefore:

- the percentage of observations over the entire dataset with jerk values higher than the threshold of ± 15 m/s^3;
- maximum and minimum jerk values in the whole dataset;
- the percentage of one-second windows (i.e. 10 observations) over the entire dataset containing more than one sign inversion of the jerk.

3.2. Consistency analysis

Platoon and internal consistency analyses are largely based on the discussion reported in Section 2.2. The former aims to verify whether the inter-vehicle spacing drawn by the trajectories estimated for a pair of vehicles is consistent with the actual one, while the latter whether or not the differentiation in the vehicle trajectory function yields consistent speeds and accelerations.

A first measure of platoon inconsistency is:

$min_k(\Delta s_{np}^{cum}(k))$: the minimum cumulative inter-vehicle spacing that vehicle pairs attain during car-following.

This yields an obvious indirect measurement of the bias in trajectories. Indeed, when its value, at least for one instant, decreases below a threshold of 5 m (16.4 ft) (assumed here as the average vehicle length), this means that, according to estimated trajectories, a collision occurred between the vehicles, thus of course not verifying platoon consistency.

[7] Such information is not available in any public document but derives from a personal communication with a team member of Cambridge Systematics, Inc.

The previous measure can be used for calculating the number and the percentage – over the whole dataset – of vehicle pairs with unphysical inter-vehicle spacing. Similar measures are the number and percentage of vehicle pairs with negative inter-vehicle spacing that are even more clear indicators of bias.

A direct measure of the bias in the estimated trajectories of the vehicle pair (n, p) at instant k can be immediately derived by criterion (7):

$$\epsilon_{np}(k) = \Delta s_{np}^{cum}(k) - \Delta s_{np}^{obs}(k) \tag{10}$$

Some of the statistics that may be obtained from $\epsilon_{np}(k)$, applied as well below, are:

- the minimum bias for the vehicle pair (n, p):
  $min_k(\epsilon_{np}(k))$;
- the *maximum bias* for the vehicle pair (n, p):
  $max_k(\epsilon_{np}(k))$;
- the *mean bias* for the entire dataset:
  $$\mu(\epsilon) = \frac{1}{MN_{np}} \sum_{np=1}^{M} \sum_{k=1}^{N_{np}} (\epsilon_{np}(k))$$

where N_{np} stands for the number of observations available for the (n, p) pair and M for the total number of pairs in the dataset. In congested traffic conditions, this value is expected to be equal to zero as the bias of the single vehicle trajectory enters the calculation twice and with opposite signs, once when the vehicle is a leader and the other when it is a follower;

- the average *Root Mean Square Error* (RMSE) among all the vehicle pairs in the dataset, which gives an estimation of the error variance, amplifying larger errors with respect to smaller ones:
  $$\mathrm{RMSE} = \frac{1}{M} \sum_{np=1}^{M} \sqrt{\frac{1}{N_{np}} \sum_{k=1}^{N_{np}} (\epsilon_{np}(k))^2}$$
- the average *Root Mean Square Percentage Error* (RMSPE) among all the vehicle pairs in the dataset, a non-dimensional index yielding the normalised bias with respect to the values of the variables concerned:
  $$\mathrm{RMSPE} = \frac{1}{M} \sum_{np=1}^{M} \sqrt{\frac{1}{N_{np}} \sum_{k=1}^{N_{np}} \left(\frac{\epsilon_{np}(k)}{\Delta s_{np}^{obs}(k)} \right)^2}$$

As instead regards internal consistency, in the sense given by Eqs. (8), it has to be evaluated on individual vehicles and not on vehicle pairs like platoon consistency. However, the measures may be the same as platoon consistency.

3.3. Spectral analysis

Spectral analysis allows in-depth understanding of data and the system that produced them (see e.g. Rayner, 1971). Where the inputs are discrete finite observations sampled by a continuous function, like vehicle positions from its space trajectory, Discrete Fourier Transform (DFT) gives a continuum density spectrum. This spectrum provides information on the portion of the signal lying within each given frequency band (indeed the calculation of the spectrum involves the fitting by least squares of sinusoidal curves of different frequencies to the set of data). Therefore, analysis of the spectrum can tell whether there are separable generating processes, as it accounts for different frequency components. For example, this is the case of a noise with frequency components higher than those of the original signal.

A major advantage of spectral analysis is that differentiation in the time domain transforms to multiplication in the frequency domain:

$$F[v(f)] = 2\pi f \cdot F[s(f)]$$

where $F[v(f)]$ and $F[s(f)]$ stand for the Fourier transforms respectively of speed and trajectory functions, and f is the frequency. Hence the higher the frequency, the more amplified are the corresponding signal components, while lower frequency components are damped down.

Since human/vehicle responses are unlikely to have a frequency exceeding 1 Hz, whilst the noise has general higher frequency components, noise is expected to be amplified in the differentiation process. Therefore from noisy position data, when not adequately filtered, unphysical (and inconsistent) speed and acceleration functions may well be obtained. For this reason, spectral analysis is fundamental in the analysis of vehicle trajectory data accuracy as well as in the choice and design of appropriate filters e.g. to design the most appropriate cut-off frequency and the frequency response of the filter.

4. Analysis of NGSIM trajectory data

In this section the analyses carried out on the NGSIM datasets are described. After a general description of the datasets under study and the data used, details of the application of the methodology of analysis from Section 3 and the correspond-

ing results are presented. It is worth noting that the analysis focuses only on the NGSIM numerical databases, that is on the accuracy of the trajectories reported there. To verify whether all the vehicles in the traffic flow were actually detected and correctly included in the databases (e.g. not duplicated) was out of the scope of the present work, even though these occurrences might impact the quality of the trajectory data.

4.1. Data description

Ten datasets were chosen for the analyses among those available within NGSIM program: three 15-min datasets from the Interstate 80 in Emeryville, California, indicated in Tables 4–9 as I80-1, I80-2 and I80-3; three from the US101 freeway in Los Angeles, California, (US101-1, US101-2, US101-3); two from Lankershim Boulevard in Los Angeles, California (Lank-1, Lank-2) and two from Peachtree Street in Atlanta, Georgia, (Peach-1, Peach-2). In all the tables the survey date and road type are also specified. Average traffic conditions during the surveys were very congested for the two freeways, and fairly congested for the two arterial roads (detailed information can be found in the "data analysis reports" at the website of the US Department of Transportation – FHWA, 2008a).

For each of the survey locations, microscopic information regarding the motion of vehicles on the road collected every 0.1 s is reported in "vehicle trajectory data" files. According to the "trajectory data dictionary", the data used for the purposes of analysis, among those available in the vehicle trajectory data files, were:

- "Vehicle ID: vehicle identification number (ascending by time of entry into section)";
- "FrameID: identification number of the 0.1-s temporal frame (ascending by start time)";
- "Total Frames: total number of 0.1-s temporal frames in which the vehicle appears in the data set";
- "Local X: lateral (X) coordinate of the front-centre of the vehicle – perpendicular to the median of the road concerned. Vehicles travelling on the east side of the median have positive Local X values, while those travelling on the west side of the median have negative Local X values";
- "Local Y: longitudinal (Y) coordinate of the front-centre of the vehicle along the median of the road concerned. The starting point is at the southern boundary of the study area";
- "Vehicle Velocity: instantaneous velocity of vehicle (feet/s)";
- "Vehicle Acceleration: instantaneous acceleration of vehicle (feet/s square)";
- "Preceding Vehicle: vehicle identification number of the lead vehicle in the same lane. A value of '0' represents no preceding vehicle";
- "Spacing: distance between the front-centre of a vehicle and the front-centre of the preceding vehicle".

4.2. Jerk analysis

As already pointed out by Thiemann et al. (2008), the distribution of accelerations surveyed in NGSIM datasets is not realistic (see Fig. 3).

Table 4
Error statistics of NGSIM trajectories – jerk analysis.

Date	April 2005			June 2005			June 2005		November 2006	
Dataset Road type	I80-1 Freeway	I80-2	I80-3	US101-1 Freeway	US101-2	US101-3	Lank-1 Arterial	Lank-2	Peach-1 Arterial	Peach-2
% Of jerk values > abs (15 (m/s^3))	16.9	11.2	10.6	10.9	7.7	6.5	17.4	14.6	7.3	10.8
Maximum jerk (m/s^3)	59	51	51	49	47	48	76	77	35	39
Minimum jerk (m/s^3)	−57	−49	−49	−40	−36	−37	-68	-69	-33	-38
% Of 1 s windows with more than 1 sign inversion of the jerk	90.5	85.1	86.3	89.5	87.4	84.1	94.3	93.3	74.6	84.8

Note: 15 m/s^3 = 49.20 ft/s^3; 50 m = 164 ft; 5 m = 16.4 ft; 1 m = 3.28 ft.

Table 5
Error statistics of NGSIM trajectories – platoon consistency (general indicators).

Date	April, 2005			June, 2005			June, 2005		November, 2006	
Dataset Road type	I80-1 Freeway	I80-2	I80-3	US101-1 Freeway	US101-2	US101-3	Lank-1 Arterial	Lank-2	Peach-1 Arterial	Peach-2
Total no. of vehicle pairs	4232	3597	3442	4662	3171	2794	3385	4012	4301	4023
No. vehicle pairs with spacing < 50 m	3615	3202	3222	3596	2725	2458	1835	2511	1815	1286
%. Vehicle pairs with spacing < 50 m	85.4	89.0	93.6	77.1	85.9	88.0	54.2	62.6	42.2	32.0
Mean observation period of vehicle pairs (s)	25	29	30	23	33	38	17	20	11	9
The longest observation period for a vehicle pair (s)	95	183	171	92	117	123	154	158	80	89

Note: 15 m/s^3 = 49.20 ft/s^3; 50 m = 164 ft; 5 m = 16.4 ft; 1 m = 3.28 ft.

Table 6
Error statistics of NGSIM trajectories – platoon consistency (spacing indicators).

Date	April, 2005			June, 2005			June, 2005		November, 2006	
Dataset Road type	I80-1 Freeway	I80-2	I80-3	US101-1 Freeway	US101-2	US101-3	Lank-1 Arterial	Lank-2	Peach-1 Arterial	Peach-2
No. of vehicle pairs with unphysical inter-vehicle spacing (crash) (<5 m)	206	283	398	143	158	140	103	235	78	61
% Vehicle pairs with unphysical inter-vehicle spacing (crash) (<5 m)	5.7	8.8	12.4	4.0	5.8	5.7	5.6	9.4	4.3	4.7
No. of vehicle pairs with negative inter-vehicle spacing (<0 m)	11	3	7	0	1	1	5	14	3	3
% Vehicle pairs with negative inter-vehicle spacing (<0 m)	0.3	0.1	0.2	0.0	0.0	0.0	0.3	0.6	0.2	0.2
Minimum bias (m)	−22.50	−5.75	−4.42	−1.34	−2.56	−0.95	−27.62	−25.78	−18.28	−19.45
Maximum bias (m)	22.48	3.54	3.39	3.43	2.54	3.77	49.38	59.26	18.78	11.02
Mean bias – με (m)	−0.04	−0.03	−0.01	−0.01	−0.01	−0.01	0.10	0.10	0.02	0.02
% Vehicle pairs j with $\mu\varepsilon_j$ > abs(1 (m))	7.9	1.1	0.3	0.1	0.0	0.0	23.5	20.2	3.5	4.6
RMSE	0.38	0.11	0.07	0.07	0.06	0.06	1.15	0.93	0.23	0.24
RMSPE (%)	2.8	1.0	1.2	0.5	0.7	0.4	8.2	9.2	2.6	36.6
% Vehicle pairs j with $RMSPbias_j$ > 10%	6.6	0.8	1.3	0.4	0.2	0.2	16.8	16.6	4.8	5.5

Note: 15 m/s3 = 49.20 ft/s3; 50 m = 164 ft; 5 m = 16.4 ft; 1 m = 3.28 ft.

Table 7
Error statistics of NGSIM trajectories – internal consistency (general indicators).

Date	April, 2005			June, 2005			June, 2005		November, 2006	
Dataset Road type	I80-1 Freeway	I80-2	I80-3	US101-1 Freeway	US101-2	US101-3	Lank-1 Arterial	Lank-2	Peach-1 Arterial	Peach-2
Total no. of vehicles	1724	1415	1241	1992	1532	1297	1210	1230	1213	1105
Mean observation period of vehicle (s)	61	74	84	53	68	81	58	73	40	35
The longest observation period for a vehicle (s)	117	194	180	101	124	133	189	218	238	242

Note: 15 m/s^3 = 49.20 ft/s^3; 50 m = 164 ft; 5 m = 16.4 ft; 1 m = 3.28 ft.

Table 8
Error statistics of NGSIM trajectories – internal consistency (travelled space/speed consistency).

DATE	April, 2005			June, 2005			June, 2005		November, 2006	
Dataset Road type	I80-1 Freeway	I80-2	I80-3	US101-1 Freeway	US101-2	US101-3	Lank-1 Arterial	Lank-2	Peach-1 Arterial	Peach-2
Minimum error (m)	−22.49	−8.34	−11.52	−8.60	−2.53	−4.07	−92.92	−88.65	−54.96	−47.45
Maximum error (m)	4.69	3.48	3.39	2.02	0.84	1.12	519.42	569.23	691.19	207.60
Mean error (m)	−0.51	−0.08	−0.05	−0.03	−0.02	−0.02	−1.65	0.36	3.22	3.35
% Vehicles j with $\mu\varepsilon_j$ > abs(1(m))	20.1	3.5	1.0	0.2	0.1	0.1	76.1	76.9	76.6	80.3
RMSE	0.74	0.25	0.17	0.11	0.09	0.09	7.79	5.83	4.22	4.38
RMSPE (%)	1.6	1.2	1.0	0.4	0.3	0.4	396.7	431.5	1302.6	506.4
% Vehicles j with $RMSPE_j$ > 10%	1.7	0.6	0.5	0.1	0.0	0.1	35.3	35.8	62.7	71.9

Note: 15 m/s^3 = 49.20 ft/s^3; 50 m = 164 ft; 5 m = 16.4 ft; 1 m = 3.28 ft.

As regards extreme acceleration values, these were seemingly cut-off in the phase of data post-processing at a threshold whose value depends on the dataset (3.42 m/s^2 (11.21 ft/s^2) for US101 and I-80; 4.83 m/s^2 (15.81 ft/s^2) for Lankershim, 3.74 m/s^2 (12.27 ft/s^2) for Peachtree). This drastic remedy is expected to be a major source of the important (internal) inconsistency between speed and acceleration functions observed in the NGSIM datasets (see Section 4.4).

Temporal acceleration dynamics was inspected by means of jerks, as explained in Section 3.1. Datasets were processed first to calculate the jerking values at each instant for all the vehicles. Secondly, for each vehicle the records were scanned to find the one-second windows with at least one sign inversion and the number of sign inversions of each one of these windows. The search was made avoiding the overlapping of time windows.

Results of the analysis are reported in Table 4, in terms of the indicators introduced in Section 3.1. In particular, it is noteworthy that the percentage of jerk values higher than 15 m/s^3 (49.20 ft/s^3) ranges from 6.5% (US101-3) to 17.4% (Lank-1) and that every vehicle has at least one observation with jerk values over the threshold, thus confirming the guess on the accelerations. Maximum and minimum jerk values reach very unreliable values in all the databases, even though values from Lankershim datasets outlie the others (with a minimum of −69 m/s^3 (226 ft/s^3) and a maximum of 77 m/s^3 (253 ft/s^3)).

Table 9
Error statistics of NGSIM trajectories – internal consistency (speed/acceleration consistency).

Date	April, 2005			June, 2005			June, 2005		November, 2006	
Dataset	I80-1	I80-2	I80-3	US101-1	US101-2	US101-3	Lank-1	Lank-2	Peach-1	Peach-2
Road type	Freeway			Freeway			Arterial		Arterial	
Minimum error (m/s)	−28.72	−23.96	−23.99	−19.43	−17.93	−17.58	−13.89	−16.51	−18.04	−16.37
Maximum error (m/s)	25.18	17.72	13.11	17.24	14.25	12.34	19.74	18.72	11.70	14.52
Mean error (m/s)	0.44	0.23	−0.06	−0.09	−0.09	0.02	0.23	0.36	0.17	−0.15
% Vehicles j with $\mu\varepsilon_j$ > abs(1(m/s))	50.9	40.6	37.1	24.6	14.6	19.4	62.9	60.7	24.9	28.1
RMSE	1.94	1.42	1.29	0.95	0.74	0.90	2.56	2.50	1.21	1.28
RMSPE (%)	313	103	107	24	43	52	1318	1783	1538	981
% Vehicles j with $RMSPE_j$ > 10%	76.2	79.6	79.5	36.1	46.1	66.1	94.0	96.0	47.7	53.7

Note: 15 m/s^3 = 49.20 ft/s^3; 50 m = 164 ft; 5 m = 16.4 ft; 1 m = 3.28 ft.

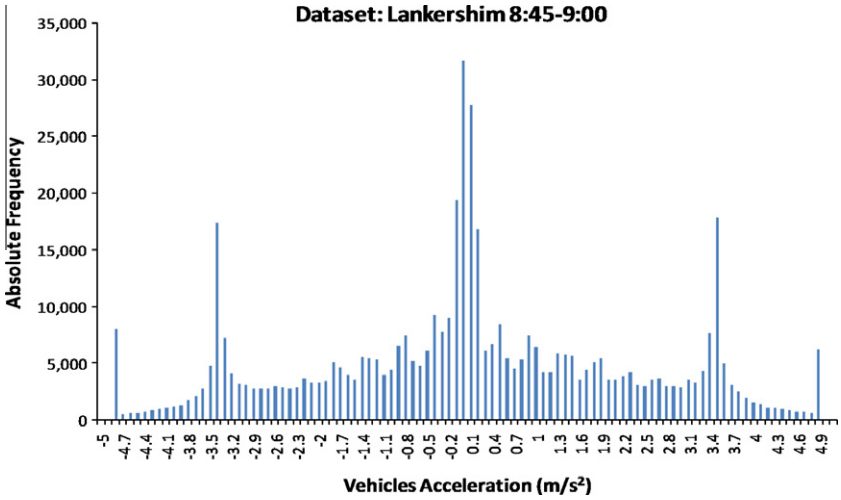

Fig. 3. Acceleration frequency of one of the Lankershim datasets.

As regards the number of sign inversions, impressively, each dataset has more than 74% of the one-second windows with more than one inversion in the jerk sign. It seems that all the indicators are correlated as the Lankershim datasets present the uppermost values of all the measures.

4.3. Platoon consistency analysis

It is worth noting first that, in the current analysis of NGSIM data, platoon consistency was evaluated by means of Eq. (9). This was a safe assumption so as not to present extremely scant results. If platoon consistency had been verified according to criterion (7) using the very noisy coordinate values (i.e. calculating travelled spaces directly from the point coordinates by means of a piecewise linear approximation of the vehicle path), this would have entailed very poor platoon consistency results. Moreover as users prevalently apply "Local Y" and "Vehicle Velocity" data, results on point coordinates would have been less interesting.

Indeed the terms of Eq. (9) were determined as follows: $\Delta s_{np}^{obs}(k)$ was calculated as the difference between the "Local Y" values of the preceding and following vehicles, which is how it is defined in Section 2.2 (see Eq. (5)). Instead, the integrals on the right hand side of Eq. (9) were calculated following a trapezoidal rule using values from the "Vehicle Velocity" field. As in NGSIM datasets the "Vehicle Velocity" is calculated applying a local regression procedure to the "Local Y" and, as the space travelled given by the "Local Y" is unbiased (according to the discussion at the end of Section 2.1), the just described application of Eq. (9) actually gives the error in the inter-vehicle spacing due to the "internal" inconsistency in the speeds of two vehicles.

Thus in order to evaluate such platoon consistency, for each vehicle we determined the time interval during which the corresponding preceding vehicle does not change: in other words, vehicle pairs were extracted from NGSIM datasets. In so doing, two kinds of problems were recognised. The first is related to records lacking information about "Spacing", that were discarded for safety reasons albeit storing information on corresponding leader and follower vehicles (in this case the records considered for the vehicle pair concerned were the first portion before the interruption). The second occurs when

for some preceding vehicles in the dataset there is no record corresponding to its "Vehicle ID" number, thus not allowing use of the speed concerned. For these records the vehicle pair was discarded.

From an operational point of view, only vehicle pairs with a maximum spacing below 50 m (164 ft) were selected. Indeed, these pairs are more likely to be in car-following situations and, above all, the errors in spacing are likely to be higher percentagewise for these pairs, hence producing the greatest impact in applications like car-following model development and calibration. In Table 5, some descriptive information about the databases, before and after vehicle pair selection, is reported. In particular:

- *The total number of vehicle pairs*. This is the total number of vehicle pairs in each dataset before pair selection, approximately ranging from 2800 to 4700 vehicles.
- *The number of vehicle pairs with maximum spacing below 50 m* and the corresponding percentage over the total number of pairs, which is obviously correlated with the traffic conditions, being higher for the most congested datasets (94% for I80-3 whilst only 32% for Peach-2).
- *The mean observation period of vehicle pairs* that gives an indication of the duration of trajectories in the datasets. This is rather small, as it ranges from 9 to 38 s.
- *The longest observation period for a vehicle pair*. This measure gives an idea of the longest car-following situation one can expect in a dataset. The highest durations are of course in the congested I-80-2/3 datasets and are around 3 min.

As regards platoon consistency, results of the analysis are summarized in Table 6. In particular, the following statistics are meaningful:

- *Number (and percentage) of vehicle pairs with unphysical inter-vehicle spacing*. This is one of the most impressive results. The mean value among all the datasets is almost 7%. For I-80-3 this value exceeds 12%, probably due to the preponderance of congested traffic conditions. Moreover, there are a few vehicle pairs with even *negative inter-vehicle spacing*.
- *Minimum and maximum bias.* These measures are found to reach as high as 59 m (196.80 ft) in absolute value, and are higher for arterial streets than for freeways, with the only inexplicable case of I80-1, which shows very high extreme values. The Lankershim datasets present the greatest extreme biases while the US101 present the lowest.
- *Mean bias* and *percentage of vehicle pairs with mean bias in the range [−1, 1]*. This value is expected to be near to zero, since almost all the vehicles analysed appeared both as leaders and followers. Values are generally around zero, apart from the case of Lankershim and I80-1 datasets, where a bias is also revealed at the level of individual pairs.
- *RMSE* and *RMSPE*. These statistics substantially confirm the above results, in particular concerning the shortcomings of Lankershim and I80-1 datasets.

Finally, it can be said that the datasets that exhibited the highest number of unphysical situations, i.e. I-80, were not those with the highest biases, i.e. Lankershim. This apparent contradiction may be again explained by the higher congestion of the I-80 datasets (entailing more stop-and-go situations but lower average biases).

4.4. Internal consistency analysis

As already stated, internal consistency analysis aims to verify whether or not the differentiation of the vehicle trajectory function yields consistent speeds and accelerations, Eqs. (8) being respected. Such analysis is performed on individual vehicles rather than on vehicle pairs as the platoon consistency analysis.

As regards the application of Eqs. (8) to the NGSIM datasets, $\hat{s}(k)$ and $\hat{s}(0)$ were taken directly from the "Local Y" field, and $\hat{v}(k)$ and $\hat{v}(0)$ from the "Vehicle Velocity" field while the integrals at the right hand sides of Eqs. (8) were calculated through numerical integration assuming, respectively, a trapezoidal and a rectangular rule.

First in Table 7 some general features of the datasets are reported, in particular:

- *the total number of vehicles*, around 1300 per dataset;
- *the mean observation period of vehicles*. This values are obviously higher than their counterpart for the vehicle pairs and around one minute for all datasets;
- *the longest observation period for a vehicle:* not longer than 4 min (Peach-2).

The outcomes of internal consistency analysis are subdivided into Table 8 and Table 9, respectively for consistency between spaces and speeds (Eq. (8a)), and for consistency between speeds and accelerations (Eq. (8b)).

As regards the former (Table 8), internal consistency is largely unsatisfied for all datasets from arterial streets, in which very high errors are revealed. It is even more significant that there is a considerable number of vehicles with coarse errors: for arterial datasets (i.e. Lank. and Peach.), more than 75% vehicles reveal a systematic error (mean error outside the interval $[-1, 1]$) and more than 35% of vehicles have an RMSPE greater than 10%. Conversely, the outcomes for freeways are much better: in particular, US101 datasets verify internal consistency with a good level of reliability (mean error outside the interval $[-1, 1]$ ranging from 0.2% to 0.4%). The poor results were expected for the Lankershim datasets, given the results of platoon consistency, but not for Peachtree. This could be related to the fact that vehicles with the highest errors were

presumably not included in platoon consistency analysis, in which only vehicle trajectories with inter-vehicle spacing lower than 50 m were considered (besides, Peachtree datasets are those with the smallest percentage of analysed vehicle pairs).

The second group of results (speed/acceleration, Table 9) reveals significant inconsistency throughout all the datasets. In particular, the percentage of vehicles with RMSPE higher than 10% ranges between 35% and 96%. Again, US101 seems the case with the lowest inconsistency (albeit, very high) and Lankershim that with the highest. Importantly, these results confirm the unrealistic values of the estimated accelerations, already pointed out in jerk analysis.

4.5. Spectral analysis

In Figs. 4–6 results from the analysis of four out of ten NGSIM databases are reported. The graphs give the continuum amplitude spectra of speeds, accelerations and jerking (i.e. the amplitude of frequency components against frequency). Analysis is limited to frequencies up to 5 Hz according to the Nyquist–Shannon sampling theorem.

Amplitude density plots show similar patterns for all databases: the magnification of the high-frequency components of the signal moving from speeds to jerking is self-evident. Interestingly, comparing the plots of speeds and jerking, it may be noted that in the latter the noise components almost cover the original signal for all the datasets.

However, it is worth noting that the results of this analysis confirm that Lankershim data are definitely the worst. This is evident across all the spectra, i.e. on speeds, accelerations and jerking (jerk values are truncated in Fig. 6 to allow them to be represented at the same scale as the other datasets).

4.6. Examples of noisy trajectories

In this section four examples of noisy trajectories from NGSIM databases are shown. They were selected applying the platoon consistency criterion (9) and allowed us to identify the major sources of data inaccuracy, as detailed in the following.

The top diagrams in Figs. 7–10 show two speed profiles for the same vehicle: one is drawn using values in the field "Velocity" while the other is obtained differentiating space data in the field "Local Y". Their difference represents an error in internal consistency.

The bottom chart of each figure, instead, show the inter-vehicle spacing of two following vehicles, one of which is the same vehicle shown in the top chart. The two profiles depicted correspond to the right and left hand sides of Eq. (9), respectively: the former (continuous line) is calculated integrating the values in the fields "Velocity" of the two following vehicles, while the latter (dashed line) is the difference of the vehicles' "Local Y" values at the same instant (i.e. $\Delta S_{np}^{obs}(k) = LocalY_n(k) - LocalY_p(k)$). The difference in inter-vehicle spacing profiles represents an error in platoon consistency (due to errors in internal consistency in the single vehicle).

At this point it is worth remembering again that, being $\Delta S_{np}^{obs}(k)$ a memory-less measure, a local noise in the data of a vehicle, like that revealed by the spike in the speed diagram of Fig. 7, for instance, does not affect the spacing elsewhere. In other words, except for the noisy instants, we may rely in the spacing calculated from the "Local Y" field as an unbiased measure.

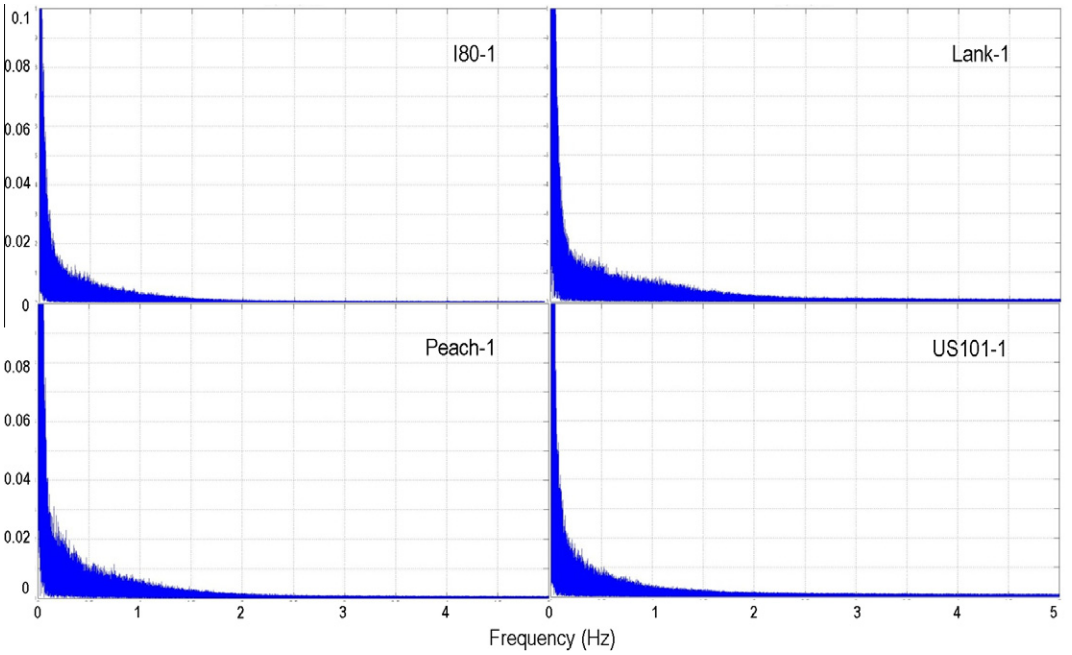

Fig. 4. Error statistics of NGSIM trajectories – spectral analysis. Amplitude spectra of vehicles *speed* for 4 out of 10 datasets.

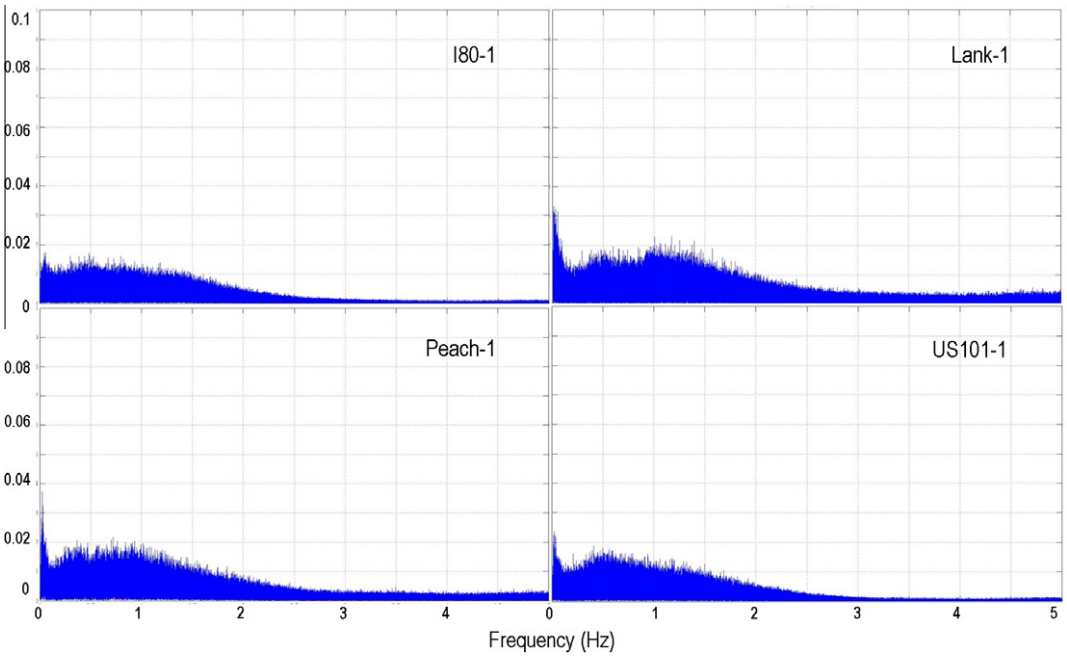

Fig. 5. Error statistics of NGSIM trajectories – spectral analysis. Amplitude spectra of vehicles *acceleration* for 4 out of 10 datasets.

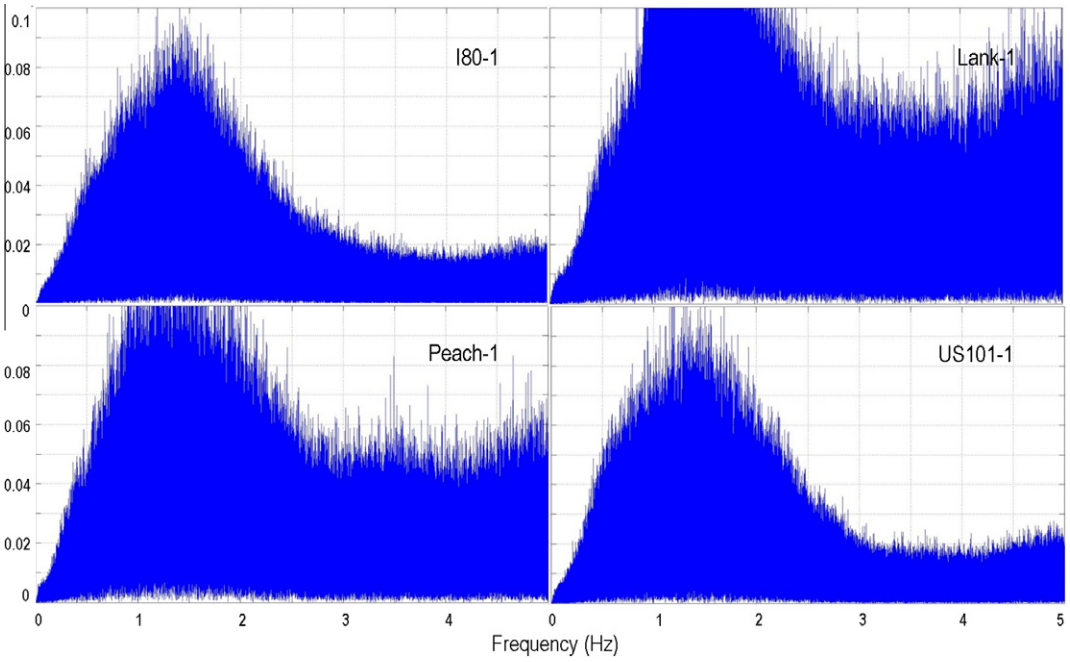

Fig. 6. Error statistics of NGSIM trajectories – spectral analysis. Amplitude spectra of vehicles *jerk* for 4 out of 10 datasets.

Therefore, it would be expected that consistent estimated speeds, once integrated, gave an inter-vehicle spacing almost overlapped to that obtained from the "Local Y" field. This is not the case of the examples presented in Figs. 7–10, which reveal significant inconsistencies in the estimated speeds (i.e. in "Vehicle Velocity").

In the top graph of Fig. 7, around the fourth second, it can be observed that the values of "Local Y" result in negative speeds i.e. in a backward motion of the vehicle. In the database field "Velocity" we find the same values of speed simply

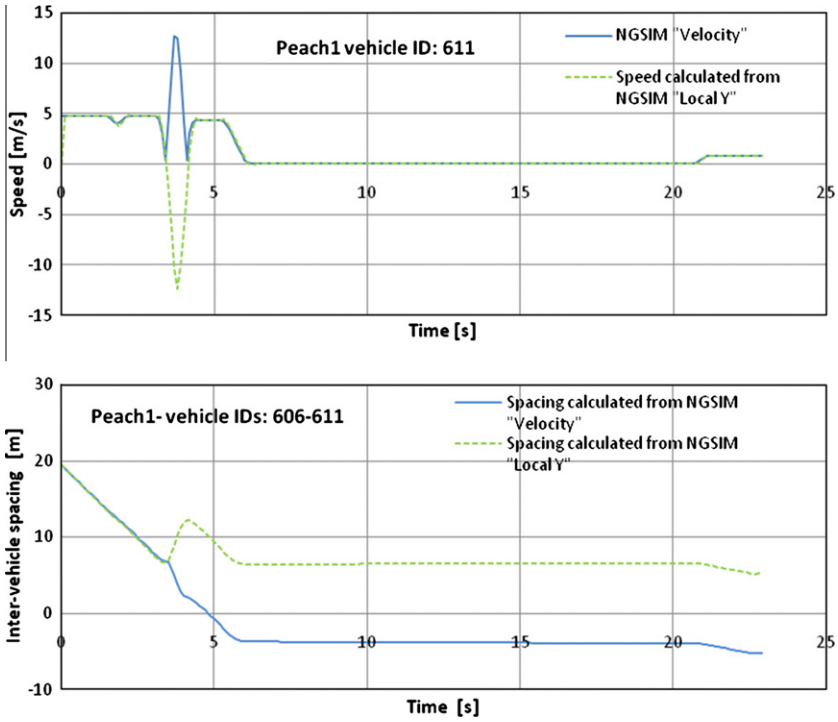

Fig. 7. Internal consistency (top) and platoon consistency (down) [Peach1 dataset].

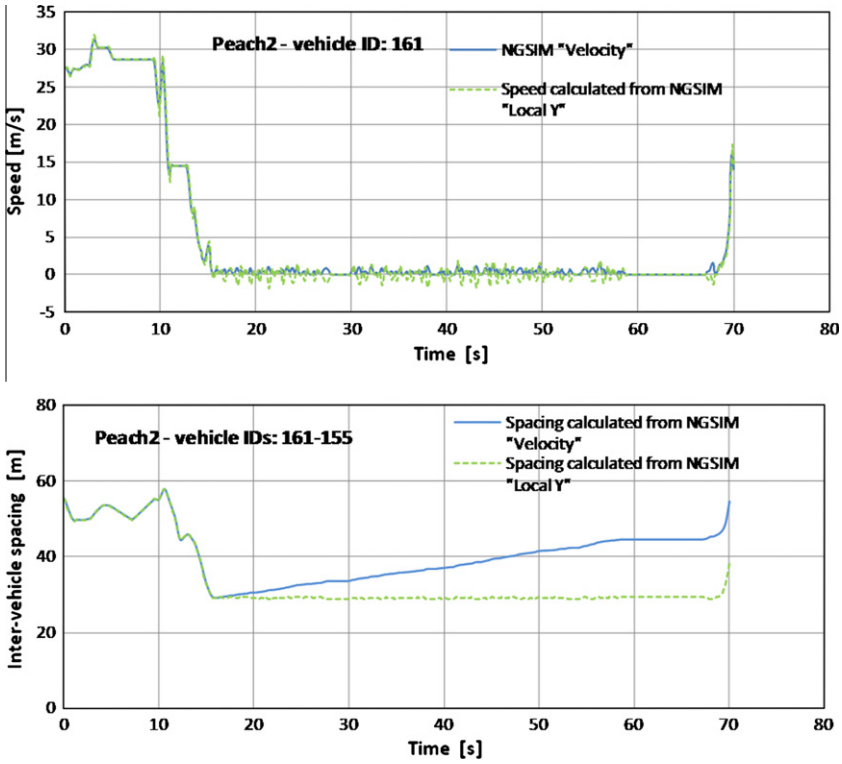

Fig. 8. Internal consistency (top) and platoon consistency (down) [Peach2 dataset].

reversed in sign. The effect on platoon dynamics of this wrong estimation of vehicles "Velocity" clearly appears observing the inter-vehicle spacing obtained by integrating those values of "Velocity", as such spacing becomes even negative.

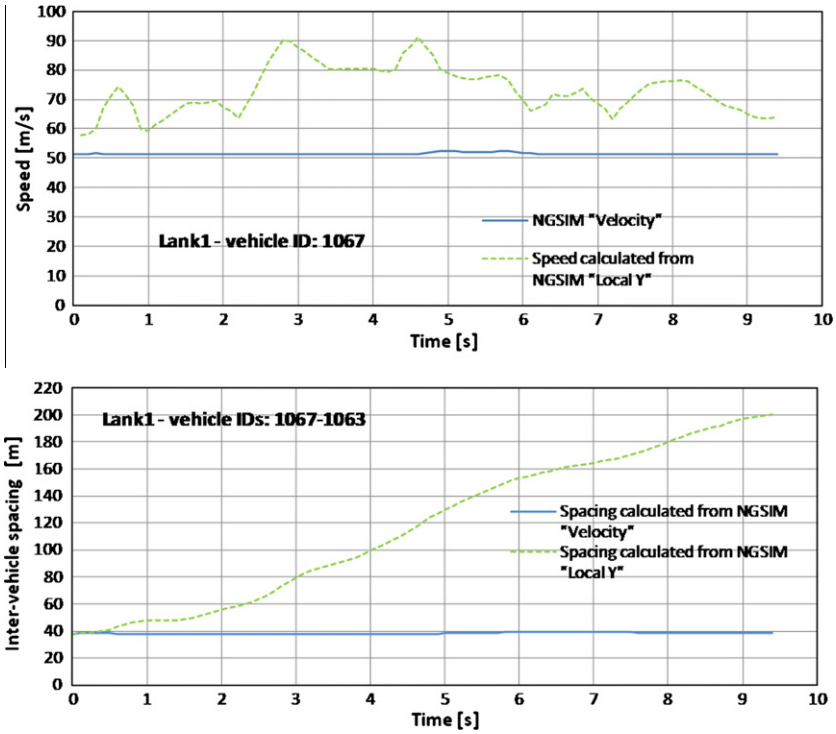

Fig. 9. Internal consistency (top) and platoon consistency (down) [Lank1 dataset].

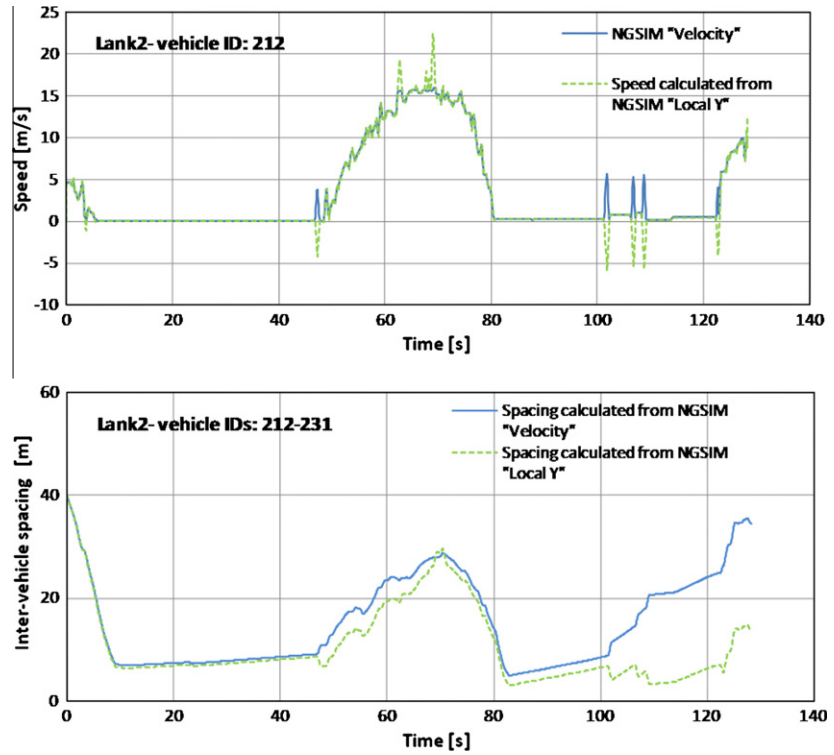

Fig. 10. Internal consistency (top) and platoon consistency (down) [Lank2 dataset].

A similar problem arises in Fig. 8 where the estimation technique applied fails to recognize that the vehicle is stopped and gives not null though very small values of "Velocity".

Fig. 9, instead, shows a case where the estimated vehicle "Velocity" is sensibly different from that obtained from the positions stored in the field "Local Y".

Finally, Fig. 10 shows the bias in "Velocity" yielded by noisy "Local Y" data in both the phases of vehicle cruising and stopping.

5. Conclusions

This paper focused on the problem of estimating vehicle trajectories from observed positions. As observations are obviously always affected by measurement errors, accuracy of estimated data should be verified, if such data have to be of any use. It was noted that in the literature few efforts have been devoted to this problem, perhaps due to the underestimation of the bias arising in application results when using data not properly processed.

Therefore in order to draw up a quantitative method to inspect the quality of trajectory data, the structure of the error on point measurements and its propagation on the space travelled have been first investigated. Analytical evidence of the bias propagated in the vehicle trajectory functions has been given for independent measurement errors. This yielded a criterion to verify whether trajectories of following vehicles are unbiased, referred to as the platoon consistency criterion. A review of the literature on estimation/filtering techniques followed accordingly.

A method to inspect trajectory data accuracy, based on jerks' analysis, consistency analysis and spectral analysis, has been presented and successively applied to the overall set of the most recent datasets available on-line within the NGSIM program. While the application results support the effectiveness of the method to quantify trajectory data accuracy, they also confirm the need to perform such an analysis before carrying out any traffic study on this type of data.

Application to NGSIM data allowed us to verify the types and sources of the errors in the trajectory data and to quantify them. Moreover, the different accuracy of the datasets was captured, also highlighting the impact of traffic conditions on the type and the magnitude of errors arising in the trajectory estimation process. Above all, the Lankershim datasets exhibited the highest errors among those analysed, for all the measurements.

As regards errors discovered, they are mostly present in vehicle stopped conditions and in the initial and final parts of trajectories. Those at low speeds in the field "Velocity", in particular, are due to the apparent practice of reversing the sign of negative speeds in the process of estimating vehicle "Velocity" from the "Local Y". These considerations suggest to carefully handle values in "Velocity" and "Acceleration" fields of NGSIM datasets and, whenever possible, to directly estimate speeds and accelerations from "Local Y". In light of the reasoning made throughout the paper, it seems that the most effective way to do this would be that of smoothing the "Local Y" and then to differentiate to speeds (when smoothing the "Local Y" a particular care should be devoted to cut-off measurement errors arising at the vehicle stoppages; see Fig. 8). Then, in order to dampen down the residual noise, it is advisable to smooth speeds applying whatever filter that does not introduce a bias on the space travelled, before differentiating to accelerations and smoothing again. This process would guarantee smooth speeds and accelerations while preserving the internal and platoon consistency of trajectories.

In conclusion the methodology introduced may represent a first step in the analysis of the impact of trajectory data accuracy on results of traffic studies. The methodology for error analysis indeed allows the data accuracy resulting from different estimation techniques or collection technologies to be captured.

From now on NGSIM data may also be used as a benchmark for trajectory data quality.

Acknowledgements

The authors are most grateful to the NGSIM Program for the availability of data and to the anonymous referees for their precious and constructive comments. Research contained within this paper benefited from participation in EU COST Action TU0903 MULTITUDE – Methods and tools for supporting the Use caLibration and validaTIon of Traffic simUlation moDEls.

References

Brackstone, M., McDonald, M., 1999. Car-following: a historical review. Transportation Research Part F 2, 181–196.

Brockfeld, E., Kuhne, R.D., Wagner, P., 2004. Calibration and validation of microscopic traffic flow models. Transportation Research Record: Journal of the Transportation Research Board 1876, 62–70.

Chandler, R.E., Herman, R., Montroll, E.W., 1958. Traffic dynamics: studies in car following. Operations Research 6, 165–184.

Coifman, B., Beymer, D., McLauchlan, P., Malik, J., 1998. A real-time computer vision system for vehicle tracking and traffic surveillance. Transportation Research Part C: Emerging Technologies 6 (4), 271–288.

Daganzo, C., 1997. Fundamentals of Transportation and Traffic Operations. Pergamon, Elsevier Science Ltd..

Duret, A., Buisson, C., Chiabaut, N., 2008. Estimating individual speed-spacing relationship and assessing the Newell's car-following model ability to reproduce trajectories. Transportation Research Record: Journal of the Transportation Research Board 2088, 188–197.

Ervin, R.D., MacAdam, C.C., Gilbert, K., Tchoryk, Jr., P., 1991. Quantitative characterization of the vehicle motion environment (VME). In: Proceedings of 2nd Vehicle Navigation and Information Systems Conference, 1991, pp. 1011–1029.

Gurusinghe, G.S., Nakatsuji, T., Azuta, Y., Ranjitkar, P., Tanaboriboon, Y., 2002. Multiple car-following data using real time kinematic global positioning system. Transportation Research Record: Journal of the Transportation Research Board 1802, 166–180.

Hamdar, S.H., Mahmassani, H.S., 2008. Driver car following behavior: from a discrete event process to a continuous set of episodes. In: TRB 87th Annual Meeting Compendium of Papers, Transportation Research Board of the National Academies, Washington, DC, 13–17 January 2008.

Herman, R., Potts, R.B., 1959. Single lane traffic theory and experiment. In: Proceedings of the Symposium on Theory of Traffic Flow, Research Labs, General Motors, New York, Elsevier, pp. 147–157.
Herrera, J.C., Bayen, A.M., 2008. Traffic flow reconstruction using mobile sensors and loop detector data. In: TRB 87th Annual Meeting Compendium of Papers, Transportation Research Board of the National Academies, Washington DC, 13–17 January 2008.
Hoogendoorn, S.P., Van Zuylen, S.P., Schreuder, M., Gorte, B., Vosselman, G., 2004. Microscopic traffic data collection by remote sensing. Transportation Research Record: Journal of the Transportation Research Board 1855, 121–128.
Jagacinski, R.J., Flach, J.M., 2003. Control theory for humans – quantitative approaches to modeling performance. Lawrence Erlbaum Associates, Mahwah, New Jersey.
Kometani, E., Sasaki, T., 1959. Dynamic behaviour of traffic with a nonlinear spacing-speed relationship. In: Proceedings of the Symposium on Theory of Traffic Flow, Research Laboratories, General Motors, New York, Elsevier, pp. 105–119.
Lu, X., Skabardonis, A., 2007. Freeway traffic shockwave analysis: exploring the NGSIM trajectory data. In: TRB 86th Annual Meeting Compendium of Papers, Transportation Research Board of the National Academies, Washington, DC, 21–25 January 2007.
Ma, X., Andreasson, I., 2006. Statistical analysis of driver behavior data in different regimes of the car-following stage. Transportation Research Record: Journal of the Transportation Research Board 2018, 87–96.
Martinez, J.-J., Canudas-de-Wit, C., 2007. A safe longitudinal control for adaptive cruise control and stop-and-go scenarios. IEEE Transactions on Control Systems Technology 15 (2), 246–258.
Oehlert, G.W., 1992. A note on the Delta method. The American Statistician 46 (1), 27–29.
Ossen, S., Hoogendhoorn, S.P., 2008. Validity of trajectory-based calibration approach of car-following models in the presence of measurement errors. Transportation Research Record: Journal of the Transportation Research Board 2088, 117–125.
Ozaki, H., 1993. Reaction and anticipation in the car following behaviour. In: Proceedings of the 13th International Symposium on Traffic and Transportation Theory, Lyon, France, 24–26 July 1993, pp. 349–366.
Punzo, V., Formisano, D.J., Torrieri, V., 2005. Nonstationary Kalman filter for estimation of accurate and consistent car-following data. Transportation Research Record: Journal of the Transportation Research Board 1934, 3–12.
Punzo, V., Simonelli, F., 2005. Analysis and comparison of car-following models using real traffic microscopic data. Transportation Research Record: Journal of the Transportation Research Board 1934, 53–63.
Rayner, J.N., 1971. An introduction to Spectral Analysis. London, Pion.
Schultz, G.G., Rilett, R.L., 2005. Calibration of distributions of commercial motor vehicles in corsim. Transportation Research Record: Journal of the Transportation Research Board 1934, 246–255.
Thiemann, C., Treiber, M., Kesting, A., 2008. Estimating acceleration and lane-changing dynamics based on NGSIM trajectory data. Transportation Research Record: Journal of the Transportation Research Board 2088, 90–101.
Toledo, T., Koutsopoulos, H.N., Ahmed, K.I., 2007. Estimation of vehicle trajectories with locally weighted regression. Transportation Research Record: Journal of the Transportation Research Board 1999, 161–169.
Toledo, T., Koutsopoulos, H.N., Ben-Akiva, M., 2009. Estimation of an integrated driving behavior model. Transportation Research Part C: Emerging Technologies 17 (4), 365–380.
Treiterer, J., Myers, J.A., 1974. The hysteresis phenomenon in traffic flow. In: Proceedings of the Sixth International Symposium on Transportation and Traffic Theory, Sydney, pp. 13–38.
US Department of Transportation – FHWA, 2008a. NGSIM – Next Generation SIMulation. <http://www.ngsim.fhwa.dot.gov/>. (accessed 31.10.08).
US Department of Transportation – FHWA, 2008b. Interstate 80 Freeway Dataset. <http://www.tfhrc.gov/about/06137.htm>. (accessed 10.04.09).
Wu, J., Brackstone, M., McDonald, M., 2003. The validation of a microscopic simulation model: a methodological case study. Transportation Research Part C: Emerging Technologies 11 (6), 463–479.
Xin, W., Hourdos, J., Michalopoulos, P., 2008. A vehicle trajectory collection and processing methodology and its implementation to crash data. In: TRB 87th Annual Meeting Compendium of Papers, Transportation Research Board of the National Academies, Washington DC, 13–17 January 2008.
Xing, J., 1995. A parameter identification of a car following model. In: Proceedings of the Second World Congress on ATT. Yokohama, November, 1739–1745.