

City Bus Mass Estimation based on GPS and CAN-Collected Tracking Data

Branimir Škugor

University of Zagreb

Faculty of Mechanical Engineering and Naval Architecture
Zagreb, Croatia
branimir.skugor@fsb.hr

Joško Deur

University of Zagreb

Faculty of Mechanical Engineering and Naval Architecture
Zagreb, Croatia
josko.deur@fsb.hr

Abstract — This paper deals with estimation of city bus mass based on regularly, on-line collected engine CAN data and vehicle tracking GPS data. Since mass estimation based on instantaneous recording is not usable due to imprecise inputs and, correspondingly, large variance of estimates, an alternative approach is employed based on using multiple consecutive data recordings for the interval between two subsequent bus stations. The estimation methodology is based on bus longitudinal dynamics model and least squares method. The key characteristic, which needs to be properly reconstructed for accurate mass estimation, is related to dependence of total wheel power on measured engine power load. The proposed estimation method is parameterized by using the data recorded on a real bus route in the city of Dubrovnik, and validated by using passenger mass measurement by means of passenger counting.

Index Terms — GPS vehicle tracking data, vehicle mass estimation, city buses, longitudinal dynamics model, least squares method

I. INTRODUCTION

The vehicle mass significantly impacts profiles of vehicle velocity and acceleration, and consequently energy/fuel consumption and vehicle safety features [1]. Knowing vehicle mass could be valuable for adaptation of powertrain control strategy [2], especially in the case of battery and hybrid electric vehicles, as well as for energy demand prediction and vehicle monitoring purposes, e.g. in the case of delivery vehicle fleets, and economic vehicle cruise control [3]. This is particularly emphasized in the case of delivery vehicles or busses, whose total mass can substantially vary. Since costly sensors are needed for direct vehicle mass measurement, different approaches to estimation of vehicle mass based on regular on-board measurements have been proposed in literature, such as e.g. Kalman filtering [4]-[6], recursive least squares [1, 3], particle filtering [7], etc. Vehicle mass estimation methods used in literature mostly rely on longitudinal vehicle dynamics model [1, 2, 3, 6, 8, 9] and to smaller extent on vertical vehicle dynamics model [10]. Mass estimation methods can be divided into three distinctive groups regarding a road slope [8]: methods which neglect a road slope, methods which estimate mass and road slope simultaneously [2, 4, 6], and methods that use an off-line calculated road slope [8]. However, in order to achieve a

It is gratefully acknowledged that this work has been supported by the Croatian Science Foundation under the project No. IP-2018-01-8323 (Project Acronym: ACHIEVE, web site: <http://achieve.fsb.hr/>).

highly accurate mass estimation, a road slope information is needed [11].

This paper deals with estimation of city bus mass from engine-CAN and GPS data recordings collected by using a regular low-cost vehicle tracking device with the sampling time of 1 second. The estimation is based on the longitudinal vehicle dynamics model and the least squares method. Due to the presence of noise and other inaccuracies in the recorded data, the mass estimation based on instantaneous recordings cannot be reliable. In order to overcome this difficulty, the multiple consecutive data are buffered and consolidated, and used for vehicle mass estimation between two subsequent bus station, where the vehicle mass does not change. It is assumed that the road slope is known in advance, i.e. stored in the on-board memory, based on off-line road slope recording or its reconstruction from navigation maps.

The main hypothesis meant to be tested in this paper can be formulated as follows: regularly recorded city bus GPS and CAN tracking data contain rich enough information for effective on-board vehicle mass estimation, provided that the road slope is a-priori known. The proposed vehicle mass estimation algorithm constitutes the main contribution of the paper. The method is also applicable to delivery trucks and other similar GPS tracked vehicles.

II. DATA COLLECTION AND ANALYSIS

A. Data Collection

For the purpose of city bus transport system characterization in the city of Dubrovnik [12], ten 12-meter buses covering most of the city bus routes have been equipped with GPS/GPRS-based telemetry equipment. Data collection have been performed over one-year period, 24 hours per day, with the sampling time of one second. The recorded data relate to standard vehicle tracking-available data, and they include: vehicle ID, GPS coordinates (longitude, latitude, and altitude), vehicle speed, engine status, cumulative fuel consumption, engine rotational speed, engine load, mileage, acceleration pedal position, ambient temperature, and current selected gear. It is assumed that the road slope is known in advance; in the particular case, it has been reconstructed off-line based on vertical and horizontal vehicle velocity acquired

by using a separate, accurate GPS receiver on the selected city bus route [13, 14]. It should be mentioned that the on-line recorded tracking data are acquired on a cloud server and used off-line for the purpose of on-line mass estimation. In the real application, however, the estimation algorithm should be run on-line (on the vehicle tracking electronic unit), because it is not practical to regularly broadcast the tracking data on the sampling base of 1 second. The road slope information would be connected to GPS coordinates and stored in the tracking unit's memory.

In order to provide ground truth for parameterization and testing of vehicle mass estimation algorithm, counting of passengers was performed during two multi-hour single-day intervals, both in the morning and afternoon hours, and for the same route for which the tracking data and the road slope information were available (the route Babin kuk-Pile-Babin kuk, Fig. 1). Fig. 2 shows the recorded time profile of number of passengers, which indicates that passenger number varies significantly, i.e. from zero to almost maximum number of 80 passengers. The geographic locations corresponding to bus stations are depicted in Fig. 1.

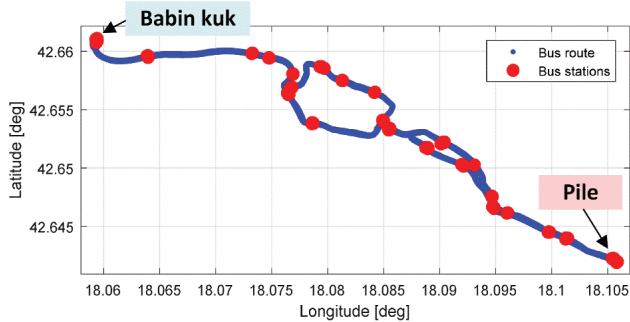


Figure 1. Circular route Babin kuk-Pile-Babin kuk for which city bus tracking data have been collected.

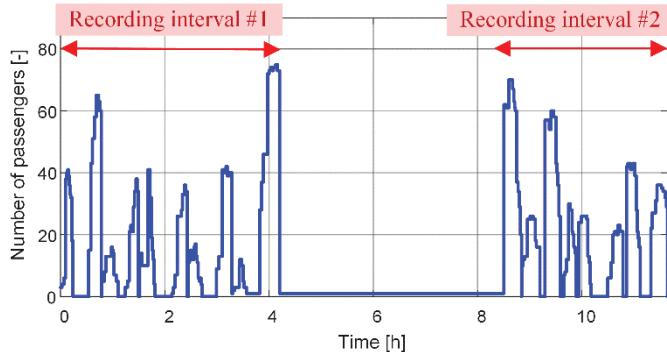


Figure 2. Recorded time profile of number of passengers in bus operating on circular route shown in Fig. 1, which is obtained by means of counting in two single-day intervals.

B. Data Processing and Analysis

The plot in Fig. 3 shows that the histogram of number of passengers is quite reach, i.e. well-distributed. Next, Fig. 4 shows the histogram of difference in number of passengers between two consecutive bus stations, which indicates that the mass change is mostly within ± 20 passengers. Finally, Fig. 5 shows the histogram of time intervals between two consecutive bus stations, which reveals the amount of data

which can be used for mass estimation (around 100 samples in average, for the mean interval time of around 100 s and the sampling time of 1 s).

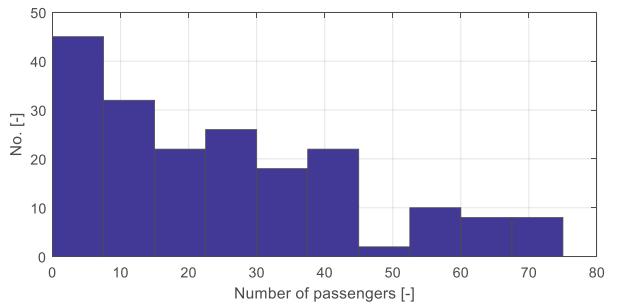


Figure 3. Distribution of number of passengers (note that maximum number of passengers for particular bus is 80).

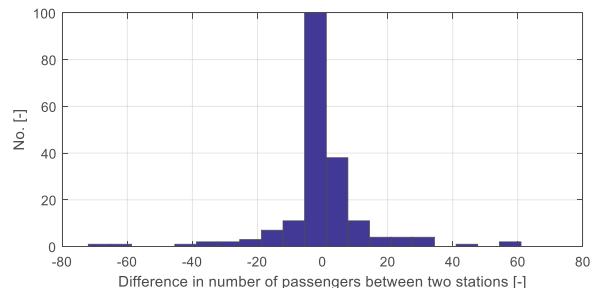


Figure 4. Distribution of difference in number of passengers between two consecutive bus stations.

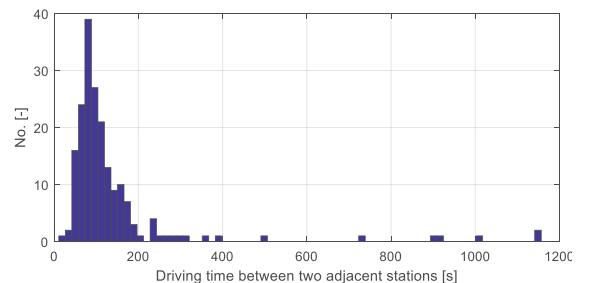


Figure 5. Distribution of driving time between two consecutive bus stations.

The raw data of road slope reconstructed from the GPS vertical and horizontal velocity (Subsection II.A) are shown by green dots in Fig. 6. The Gaussian processes-based regression method is employed to reconstruct the final, smooth road slope (denoted by red line in Fig. 6) [13, 14].

To improve the accuracy of ground truth mass reconstruction based on counting the number of passengers, the varying fuel mass is taken into consideration (see Subsection II.C). The fuel mass is determined from the available time profile of fuel level in the tank expressed in percentage in Fig. 7, after approximating it by a low-order polynomial in order to remove the noise/outliers, and the time profile of cumulative fuel consumption.

For the sake of bus mass estimation, the information on internal combustion engine power P_e is used. This information is calculated by multiplying the engine speed and the engine load, where the latter is expressed in percentage and multiplied with the declared maximum engine torque.

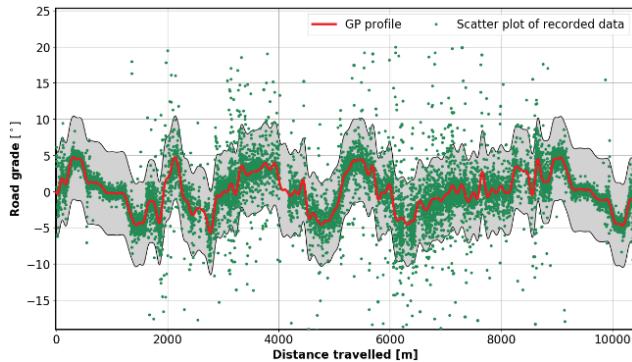


Figure 6. Road slope values calculated by using GPS vertical and horizontal velocity recordings (green dots) and road slope profile reconstructed by using Gaussian processes-based regression method (red line) including 95% confidence interval represented by grey area.

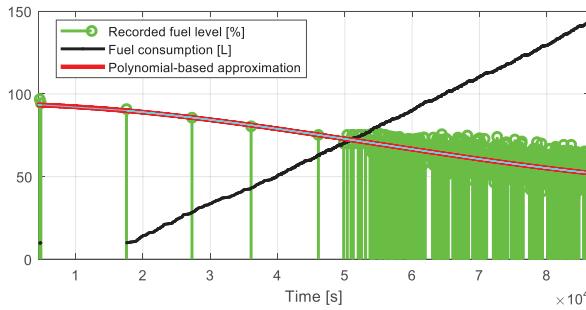


Figure 7. Recorded profile of fuel level in tank along with its polynomial approximation used for reconstructing time-varying fuel mass.

The distribution of the recorded engine operating points is shown in Fig. 8, along with declared maximum values of available torque and rotational speed. This plot indicates that most of the points are located within the declared limits, thus confirming the validity of recorded points.

Since the total vehicle mass, which includes the mass of passengers, is constant while driving between two consecutive bus stations, the recorded data are divided into driving cycles, where j^{th} driving cycle begins at the departure from the j^{th} bus station and ends at the arrival at the $(j+1)^{\text{th}}$ bus station. The vehicle mass is then estimated for each individual driving cycle. Fig. 9 shows a segment of the recorded data, which illustrates the organization of data into driving cycles. The

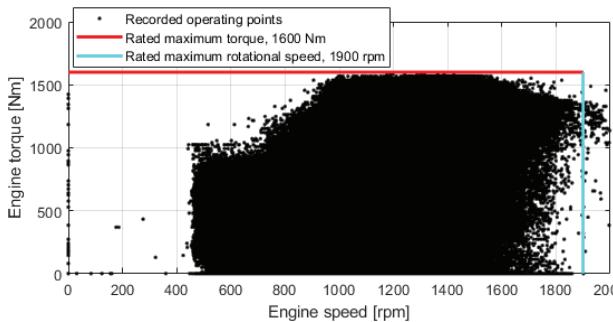


Figure 8. Distribution of recorded engine operating points along with declared torque and speed limits.

total number of recorded driving cycles with known number of passengers is equal to 191.

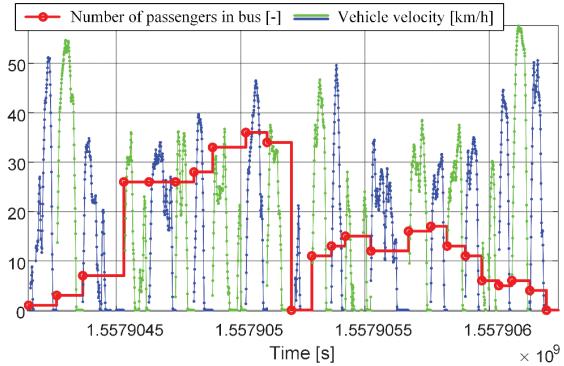


Figure 9. Illustration of vehicle velocity profile separation into driving cycles, i.e. portions of velocity profile between consecutive bus stations, between which number of passengers changes (even and odd driving cycles are shown in different colors for better recognition).

C. Establishing Ground Truth for Buss Mass

The real vehicle mass $m_{v,\text{real}}$, aimed to be used as a benchmark for vehicle mass estimation, is calculated based on the following expression:

$$m_{v,\text{real}} = m_{v,\text{empty}} + \frac{75 \cdot (n_{\text{pass}} + 4)}{\text{Mass of passengers}} + \frac{0.8508 \cdot \text{fuel}_{\text{level}} \cdot 250}{\frac{100}{\text{Fuel mass}}} \quad (1)$$

where $m_{v,\text{empty}} = 12031$ kg represents the empty bus mass, n_{pass} is the number of passengers in the bus, $\text{fuel}_{\text{level}}$ is the fuel level in the tank in percent (where, 100% means full tank), and 250 and 0.8508 stand for the fuel tank capacity in liters (L) and the diesel fuel density in kg/L, respectively, while the average mass of passengers is assumed to be 75 kg, where the number of passengers n_{pass} is increased by 4 persons, involving the bus driver and three examiners that were performing passengers counting.

III. VEHICLE MASS ESTIMATION METHOD

A. Longitudinal Vehicle Dynamics Model

The vehicle mass estimation algorithm is based on the longitudinal vehicle dynamics model [15]. The wheel power at k^{th} discrete time instant $P_{\text{wheel},\text{est}}(k)$ is described as:

$$P_{\text{wheel},\text{est}}(k) = \left[\underbrace{m_v a_v(k)}_{\text{Acceleration}} + \underbrace{m_v g \sin(\alpha(k))}_{\text{Hill climbing}} + \underbrace{m_v R_o \cos(\alpha(k))}_{\text{Rolling resistance}} + \underbrace{0.5 \rho_{\text{air}} C_d A_f v_v(k)^2}_{\text{Aerodynamical drag}} \right] v_v(k), \quad (2)$$

where m_v represents the vehicle mass, a_v is the vehicle acceleration, v_v is the vehicle velocity, α is the road slope, g is the gravity acceleration, R_o is the coefficient of rolling resistance, ρ_{air} is the air density, C_d is the coefficient of

aerodynamic drag, and A_f is the vehicle frontal surface. As stated in Section II, the vehicle mass m_v is considered to be constant between two consecutive bus stations.

The vehicle velocity $v_v(k)$ is directly available from the set of GPS measurements. The acceleration $a_v(k)$ is then calculated as:

$$a_v(k) = \frac{v(k+1) - v(k-1)}{t(k+1) - t(k-1)}, \quad (3)$$

where t is the precise absolute time, which is also available within the GPS tracking data. The time profile of road slope $\alpha(k)$ is reconstructed by associating the recorded/actual geographical coordinates in the time step k with the referent ones for which the road slope is known (as shown in Fig. 6). Equation (2) can be rearranged as

$$\begin{aligned} & \underbrace{P_{wheel,est}(k) - 0.5\rho_{air}C_dA_fv_v(k)^3}_{y(k)} \\ &= \underbrace{v_v(k)[a_v(k) + g \sin(\alpha(k)) + R_o \cos(\alpha(k))] m_v}_{x(k)}. \end{aligned} \quad (4)$$

The following linear regression can be established based on (4):

$$\begin{aligned} & \underbrace{[y(1) \ y(2) \ \dots \ y(n)]^T}_y \\ &= \underbrace{[x(1) \ x(2) \ \dots \ x(n)]^T m_v}_x, \end{aligned} \quad (5)$$

where each vector element corresponds to one discrete time step (see (4)), and n represents the total number of time steps within a single driving cycle. Under assumption that the wheel power can be effectively calculated/estimated from the engine power, it is possible to determine the vehicle mass by using the least squares method as [16]:

$$\hat{m}_v = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}. \quad (6)$$

Two variants of estimating the wheel power $P_{wheel,est}$ from the engine power P_e are considered and analyzed: (i) $P_{wheel,est} = P_e$ where the drivetrain losses are disregarded, and (ii) $P_{wheel,est} = f(P_e)$ where f is a nonlinear fitting function that accounts for the drivetrain losses. The engine power is calculated from engine speed and load, as described in Section II.

The functional dependence $P_{wheel,est} = f(P_e)$ is established by using the driving cycles in which the vehicle mass m_v is measured (by passenger counting, Fig. 9), in which case the wheel power $P_{wheel,est}$ can be readily reconstructed from (2). Fig. 10 shows the obtained plot $P_{wheel,est}$ vs. P_e , along with the fitting function $f(P_e)$ based on a 2nd order polynomial

$$P_{wheel,est} = f(P_e) = a_2 P_e^2 + a_1 P_e + a_0, \quad (7)$$

whose parameters a_0 , a_1 , and a_2 are obtained by using Matlab function *polyfit()*. Note that very high correlation between $P_{wheel,est}$ and P_e points (red points in Fig. 10), which equals 0.94 (i.e. close to ideal value of 1; calculated by using Matlab function *corrcoef()*), confirms the quality of recorded data. Fig. 10 also shows the straight line that corresponds to the crude approximation $P_{wheel,est} = P_e$. Evidently, most of the measurement-based points (red dots) fall below the straight-line approximation, thus confirming the occurrence of

(positive) drivetrain losses ($P_{wheel,est} < P_e$). The intersection of polynomial approximation function with x-axis can be interpreted as idle power losses (here equal to $P_e = 10.4$ kW).

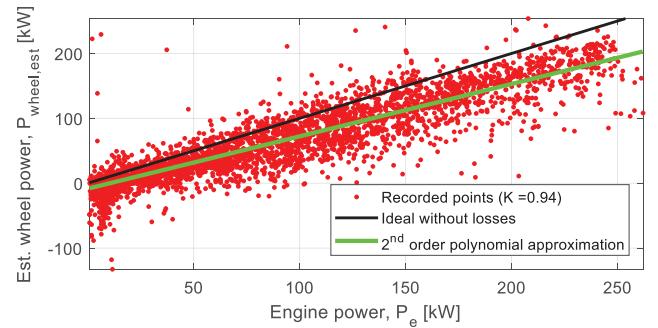


Figure 10. Wheel power reconstructed from longitudinal vehicle dynamics model vs. recorded engine power, approximated by 2nd order polynomial fitting function.

Since the braking torque/power measurement is not available, the mass estimation is limited to non-braking occasions where $P_e > 0$ and *accelerator_pedal* > 0 conditions hold, i.e. only the driving conditions-related samples are included in the output vector \mathbf{y} in (5) and (6).

IV. RESULTS

Fig. 11a illustrates a driving cycle as a station-station segment (red line) of the overall route (blue line). Figs. 11c and 11d indicate a very good correlation between the wheel power estimated based on (7) and the “actual” one, i.e. the one reconstructed from (2) by using the vehicle mass estimate. This is reflected in quite accurate vehicle mass estimation (i.e. the mass estimation error is equal to -5.3% according to the data included in the title of Fig. 11).

Fig. 12 shows estimated vehicle mass with respect to real vehicle mass (i.e. the one reconstructed from the number of passengers) where the following cases are considered: (i) zero road slope ($\alpha = 0$) and no power losses in transmission ($P_{wheel,est} = P_e$), (ii) zero road slope and existent power losses in transmission (7), (iii) actual/reconstructed road slope and no power losses, and (iv) actual/reconstructed road slope and existent power losses. These results indicate that when neglected the road slope, the vehicle mass estimates have very large variance and significant dc error/bias; consequently, correlation is very low in both cases (less than 0.1, see the correlation factors included in the figure legend). On the other hand, when using the actual/reconstructed road slope, the variance of mass estimates is significantly reduced (i.e. correlation index is equal to around 0.55 in both cases). It can be observed that both approaches have certain estimation offset, but this offset is significantly reduced in absolute value when using the more accurate, polynomial wheel power model (1144.1 kg vs. 4327.6 kg).

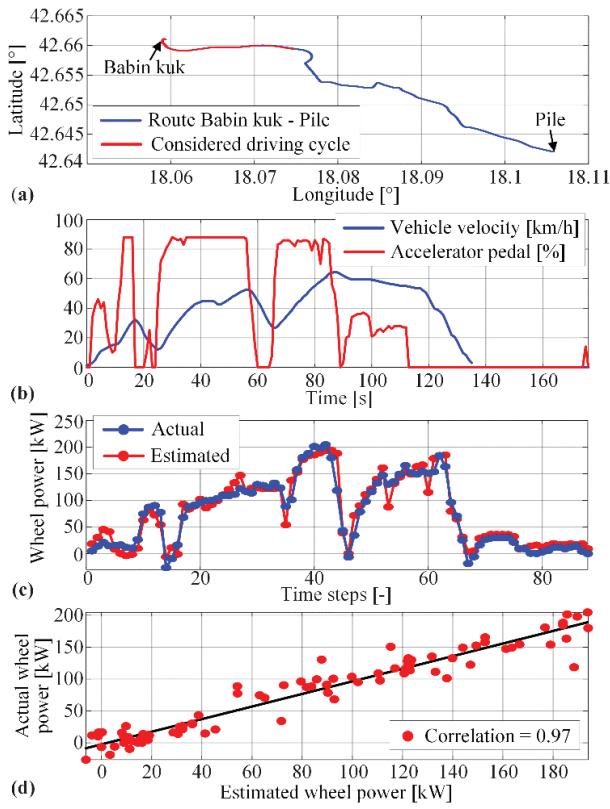


Figure 11. Example of driving cycle for which mass estimation is performed (number of valid points for estimation is 88; real/target bus mass is equal to 12513 kg, and estimated bus mass to 11847 kg, which is -5.3% with respect to real mass).

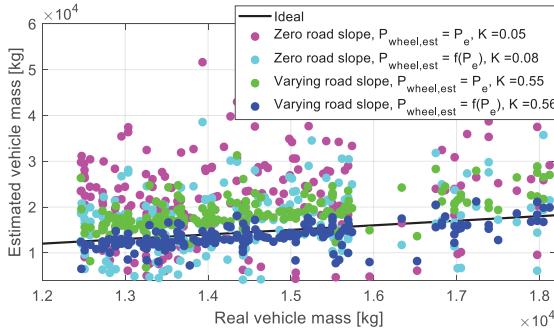


Figure 12. Estimated vs. actual vehicle mass for different cases of wheel power estimation and road slope used.

The residuals of mass estimates are analyzed with respect to (a) number of samples/points present in the driving cycle and used for mass estimation and (b) accuracy of wheel power estimation (for this purpose actual/). The results are shown in Fig. 13, and they point out that large mass estimation residuals mostly occur in the case of low number of driving cycle points or low estimated wheel power vs. actual/reconstructed wheel power correlation coefficient. Therefore, those driving cycles which satisfy the following straightforward condition are excluded from the further analysis: the number of driving cycle samples/points is lower than 20 or the correlation index for estimated vs. actual wheel power dependence is lower than 0.7. After applying this filtering rule 105 driving cycles out of 191 (55%) remained. It should be noted that the driving cycles with low values of estimated vs. actual wheel power

correlation index (red points in Fig. 13) are largely connected with locally unreliable road slope reconstruction (e.g. due to absent or weak GPS signal; see Fig. 6).

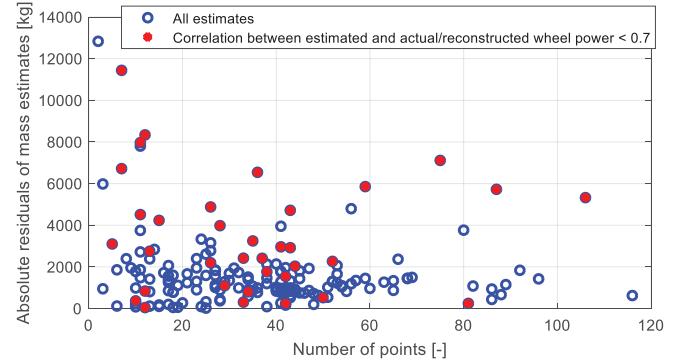


Figure 13. Dependence of absolute mass estimation residuals (i.e. estimated mass – real mass) on number of driving cycle samples/points used in estimation.

The remaining/filtered estimated mass points are shown with respect to actual vehicle mass in Figs. 14 and 15. The slight offset previously noticed in Fig. 12 can again be observed. Otherwise, the estimated vehicle mass faithfully follows the actual vehicle mass, which is reflected in the relatively high correlation index value of 0.81 (Fig. 15). Finally, Fig. 16 shows distribution of mass estimation residuals. The offset (i.e. mean value of residuals) is equal to -863.6 kg, while the standard deviation of residuals is equal to 1244.5 kg, which is less than 10% of total vehicle mass. The distribution of relative mass estimates residuals is given in Fig. 17, and it shows that most of relative residuals fall in the interval between -15% and -5%.

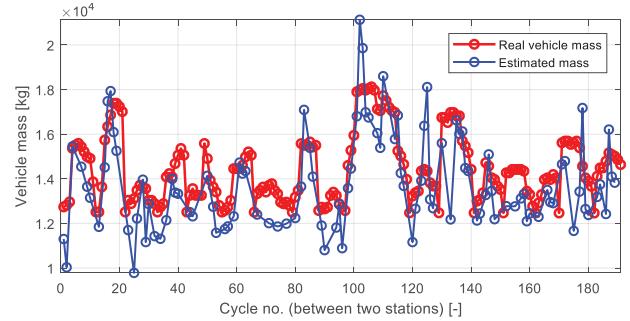


Figure 14. Estimated vs. real vehicle mass for driving cycles which satisfy condition: number of driving cycle samples/points ≥ 20 and correlation between estimated and actual wheel power ≥ 0.7 .

It should be noted that mass estimation errors are partly induced by the errors of passenger mass reconstruction due to variation of passengers' mass (while the reconstruction relies on the assumption of mean mass of 75 kg, see (1)). Also, the observed offset in mass estimates may be caused by errors in vehicle longitudinal dynamics model parameters. It could be potentially removed by adding absolute value of mean residuals to new mass estimates. Alternatively, it could be systematically removed by identifying the longitudinal dynamics model parameters from recorded driving data.

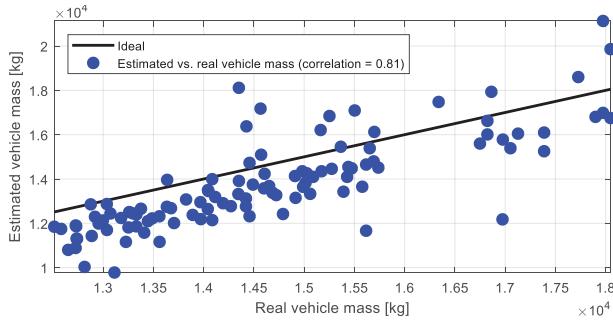


Figure 15. Estimated vs. real vehicle mass (replotted from Fig. 14).

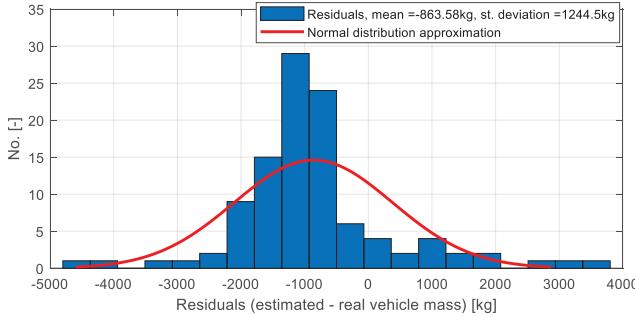


Figure 16. Distribution of mass estimate residuals.

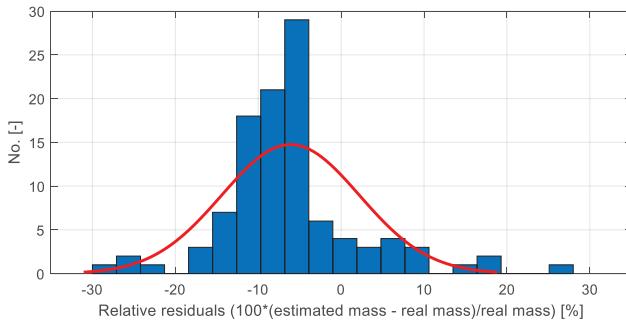


Figure 17. Distribution of relative mass estimate residuals.

V. CONCLUSION

This paper has presented a method of vehicle mass estimation based on on-board GPS and CAN data regularly acquired through tracking devices for city buses, delivery trucks and similar vehicles. First, the significant impact of road slope on vehicle mass estimation accuracy has been demonstrated, by assuming zero and actual road slope profiles and comparing the results. Next, two methods of estimating the wheel power, as an input of the vehicle longitudinal dynamics model, have been applied in vehicle mass estimation, where the impact of neglecting drivetrain power losses on overestimating the vehicle mass has been revealed. Finally, the residuals of mass estimates are analyzed with respect to number of driving cycles samples and accuracy of wheel power estimate. The straightforward rule has been applied to isolate 'irregular data' driving cycles. It has been pointed out that the final mass estimation method with properly selected driving cycle data provides relatively high correlation of mass estimates with respect to actual masses (the correlation index is equal to 0.81 on the scale of 1). The mean value of residuals of mass estimates is equal to -863.6

kg, while the standard deviation is equal to 1244.5 kg (less than 10% of the total vehicle mass).

The future work will be directed towards identification of the vehicle longitudinal dynamics model parameters based on recorded driving data, as well as investigating application of data-driven regression methods for mass estimation.

APPENDIX

Parameters of longitudinal vehicle dynamics model: $m_{v,empty} = 12031$ [kg], $C_d = 0.7$ [-], $R_o = 0.012$ [-], $A_f = 7.4625 \text{ m}^2$. Other constants used: $g = 9.81 \text{ m/s}^2$, $\rho_{air} = 1.224 \text{ kg/m}^3$.

REFERENCES

- [1] D. Kim, S. B. Choi, and M. Choi, "Integrated vehicle mass estimation for vehicle safety control using the recursive least-squares method and adaptation laws," Proceedings of the Institution of Mechanical Engineers, Part D, Journal of Automobile Engineering, vol. 229, no. 1, pp. 14-24, 2014.
- [2] S. Jiang, C. Wang, C. Zhang, H. Bai, L. Xu, "Adaptive estimation of road slope and vehicle mass of fuel cell vehicle," eTransportation, vol. 2, 2019.
- [3] N. Lin, C. Zong, S. Shi, "The Method of Mass Estimation Considering System Error in Vehicle Longitudinal Dynamics", Energies, vol. 12, no. 1:52, 2019.
- [4] P. Lingman, B. Schmidtbauer, "Road Slope and Vehicle Mass Estimation Using Kalman Filtering," Vehicle System Dynamics, vol. 37, 2002.
- [5] B. L. Boada, M. J. L. Boada, and H. Zhang, "Sensor Fusion Based on a Dual Kalman Filter for Estimation of Road Irregularities and Vehicle Mass Under Static and Dynamic Conditions," in IEEE/ASME Transactions on Mechatronics, vol. 24, no. 3, pp. 1075-1086, 2019.
- [6] S. Hao, P. Luo, and J. Xi, "Estimation of vehicle mass and road slope based on steady-state Kalman filter," IEEE International Conference on Unmanned Systems (ICUS), Beijing, pp. 582-587, 2017.
- [7] S. Sun, N. Zhang, P. Walker, and C. Lin, "Intelligent estimation for electric vehicle mass with unknown uncertainties based on particle filter," in IET Intelligent Transport Systems, vol. 14, no. 5, pp. 463-467, 2020.
- [8] H. Jonhed, "Mass esstimation using mapped road grade data in heavy duty vehicles", Master of Science Thesis, Stockholm, Sweden, 2013.
- [9] J. Ghosh, S. Foulard, and R. Fietzek, "Vehicle Mass Estimation from CAN Data and Drivetrain Torque Observer", SAE Technical Paper 2017-01-1590, 2017.
- [10] J. Jordan, N. Hirsenkorn, F. Klanner, and M. Kleinstuber, "Vehicle mass estimation based on vehicle vertical dynamics using a multi-model filter," 17th International IEEE Conference on Intelligent Transportation Systems (ITSC), Qingdao, pp. 2041-2046, 2014.
- [11] A. Erikson, "Implementation and Evaluation of a Mass Estimation Algorithm", Masters' Degree Project, Stockholm, Sweden, 2009.
- [12] J. Topić, J. Soldo, F. Maletić, B. Škugor, and J. Deur, "Virtual Simulation of Electric Bus Fleets for City Bus Transport Electrification Planning", Energies, vol. 13, no. 13:3410, 2020.
- [13] B. Škugor and J. Deur, "GPS measurement-based road grade reconstruction with application to electric vehicle simulation and analysis," 10th Conference on Sustainable Development of Energy, Water and Environment Systems (10th SDEWES), Dubrovnik, Croatia, 2015.
- [14] J. Topić, B. Škugor, and J. Deur, "Analysis of City Bus Driving Cycle Features for the Purpose of Multidimensional Driving Cycle Synthesis," SAE Technical Paper No. 2020-01-1288, 2020 SAE World Congress, Detroit, MI, 2020.
- [15] L. Guzzella, A. Sciarretta, Vehicle Propulsion Systems. 2nd ed., Springer Verlag, Berlin, 2007.
- [16] I. Goodfellow, J. Bengio, A. Courville, Deep Learning. MIT Press, 2016