Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 52 (2021) 195–202

Transportation Research Procedia

www.elsevier.com/locate/procedia

23rd EURO Working Group on Transportation Meeting, EWGT 2020, 16-18 September 2020, Paphos, Cyprus

# Modeling Car Following with Feed-Forward and Long-Short Term Memory Neural Networks

Chiara Colombaroni[a], Gaetano Fusco[a], Natalia Isaenko[a]*

[a]*Department of Civil, Environmental and Constructional Engineering, Sapienza University of Rome, Via Eudossiana 18, Rome 00184, Italy*

**Abstract**

The paper investigates the capability of modeling the car following behavior by training shallow and deep recurrent neural networks to reproduce observed driving profiles, collected in several experiments with pairs of GPS-equipped vehicles running in typical urban traffic conditions. The input variables are relative speed, spacing, and vehicle speed. In the model, we assume that the reaction is not instantaneous. However, it may occur during a time interval of the order of a few tenth of seconds because of both the psychophysical driver's reaction process and the mechanical activation of braking or dispensing the traction power to the wheels. Experimental results confirm the reliability of this assumption and highlight that the deep recurrent neural network outperforms the simpler feed-forward neural network.



*Keywords: Car following; Driver behavior; Traffic models; Artificial neural networks; Machine learning.*

## 1. Introduction

Studies on car-following models started more than 60 years ago, with the first experiments conducted almost simultaneously in the late 50s at General Motors in the USA (Chandler et al., 1958) and Japan (Kometani and Sasaki, 1961). Twenty years later and more, the use in simulation software stimulated the introduction of complex models that are harder to study theoretically but are effective in a simulation framework (Gipps, 1981; Wiedemann, 1975). More recently, artificial intelligence was recognized as an effective paradigm to reproduce complex driver behavior. A recent systematic literature review on the different applications of modeling driver-vehicle-environment systems has proved the capability of machine learning methods to accurately reproduce the driving behavior (Elamrani Abou Elassad et al., 2020).

In an earlier paper (Colombaroni and Fusco, 2014), the authors developed a car-following model based on a traditional feed-forward Neural Network trained by a swarm algorithm over a set of driving data collected during a specifically settled experiment. That model introduced a static relationship between the acceleration of the follower

* Corresponding author. Tel.: +39-06-44585-145.
*E-mail address:* natalia.isaenko @uniroma1.it.


10.1016/j.trpro.2021.01.022

and the relative speed and spacing from the leader, after a time lag. Big computational advances and new technologies made it possible to collect big amounts of traffic data useful for purposes of traffic state monitoring and modeling (Felice et al., 2014; Isaenko et al., 2017). After the earliest applications of simple neural networks for modeling car-following behavior (Fusco and Gori, 1995), deep neural networks are becoming popular to relate variables of a highly non-linear nature. Transferring information by multiple layers enables modeling of complex data relations and capturing different patterns. Several deep learning models have been introduced in the literature to derive the car-following mechanism. In particular, NGSIM data were used to model speed profile of the following vehicle by a recurrent neural network (Wang et al., 2018); a long-short term memory network (Huang et al., 2018), and a combination of autoencoder and long-short term memory networks (Fan et al., 2019).

In this paper, we introduce several conceptual enhancements concerning the static model. We assume that the driver's perception is a continuous process that undergoes the speed variation of the leading vehicle. So, the driver's reaction starts after a reaction time and, after that, develops depending on the stimuli perceived during a time interval of the order of a few tenth of seconds. In addition to the psychophysical driver's reaction process, the duration of the stimulus-reaction process is due to the mechanical activation of braking or dispensing the traction power to the wheels.

In addition to the typical car-following model, formulated in terms of the follower's acceleration, we investigate the features of an alternative modeling formulation based on the follower's speed, which is frequently used in machine-learning models. Unlike other studies in literature, we perform a comparison of different models formulated in terms of speed or acceleration on the same dataset.

Finally, we perform an analysis of different deep learning structures to compare performance for car-following modeling. The main goal is to assess the capability of deep learning models to generalize different behavioral patterns to analyze the impact of the input variables on the model performance.

## 2. Model Formulation

Car-following behavior is formalized as a data-driven model considering different kinematic features of the couple of vehicles, the leader vehicle *n,* and the follower vehicle *n*+1, to describe the response of the follower.

The reaction of the following vehicle occurs after a certain number of reaction lags from the perceived stimulus. Traditionally, the explanatory variables used in the data-driven car following models include spacing and relative speed. In this study, also the speed of the following vehicle is included in the model, according to the most general form of the GHR model (Gazis et al., 1961). Also, instead of considering only one constant lag between a reaction and a stimulus, as in most papers in literature, we assume that the follower's reaction starts after a lag $\eta$, and develops during a time interval $\tau$.

Two alternative formulations are introduced for the car-following model.

$$a_{n+1}(t) = f\left(\mathbf{u}(t-\eta-\tau), \mathbf{u}(t-\eta-\tau+1), \ldots, \mathbf{u}(t-\eta)\right) \quad (1)$$

$$v_{n+1}(t) = f\left(\mathbf{u}(t-\eta-\tau), \mathbf{u}(t-\eta-\tau+1), \ldots, \mathbf{u}(t-\eta)\right) \quad (2)$$

where: $\mathbf{u}(t) = \left\{v_{n+1}(t), \Delta v_{n,n+1}(t), \Delta x_{n,n+1}(t)\right\}$

$v_{n+1}(t)$: Speed of the follower at time interval *t*;

$a_{n+1}(t)$: Acceleration of the follower at time interval *t*;

$\Delta v_{n,n+1}(t)$: Relative speed between the leader and the follower at time interval *t*;

$\Delta x_{n,n+1}(t)$: speed between the leader and the follower at time interval *t*.

$\eta$: Reaction lag, i.e., the time interval during which the stimulus is not yet perceived;

$\tau$: Reaction duration; i.e., the time interval during which the stimulus is perceived.

Model (1) corresponds to a behavioral representation of stimulus-response mechanism and models the possible actions the driver can do: acting either on the accelerator of the brake pedal depending on its own speed and the relative speed and the spacing from the leading vehicle.

Model (2) corresponds to a phenomenological representation of the vehicle speed variation process depending on the kinematic variables of the leader and the follower vehicles, expressed as in Model (1). In this case, the car-following problem can be seen as a time series modeling, since the speed itself is an input element of the model, with

exogenous variables defined by the relative speed and the relative spacing. This model specification corresponds to the concept that the vehicle speed is the realization of a stochastic process affected by external variables that represent the way the driver interacts with the external environment.

The time diagram structure for the data of the proposed models is illustrated in Fig. 1. Instead of considering only one lagged value for the stimulus perceived, a time sequence is introduced for the three input variables $v_{n+1}$, $\Delta x_{n,n+1}$, $\Delta v_{n,n+1}$ during an interval $\tau$. Coherently to the driving behavior, the response at the beginning of the stimulus occurs after a time interval $\eta$, depending on the driver's reaction time, during which no stimulus is perceived.

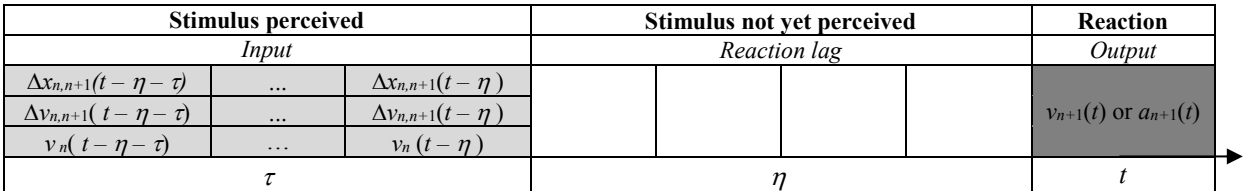


Fig. 1. Structure of input data for the proposed models

## 3. Model specification

*Feed-Forward Neural Network (FF)*

Feed-Forward Neural Network (FF) is a static non-linear vector multivariate function that derives the value of the forecasted variable as a complex non-linear combination of independent input variables. The feed-forward structure implies that the inputs of each neuron propagate in one direction, are taken from a previous layer, and do not depend on the internal state of the neuron. The structure of the FF model applied for the car following is reported in Fig. 2. Each of the components of the vector **u** is modeled as an independent input of the model. Thus, the total number of input neurons is 3×number of the time intervals used for the stimulus perception.

The complexity of the non-linear relationship is typically ensured by the number of hidden units and the hidden layers, which change accordingly with the number of independent variables. In this study, one hidden layer was adopted; the number of hidden units was taken as twice the number of input variables.

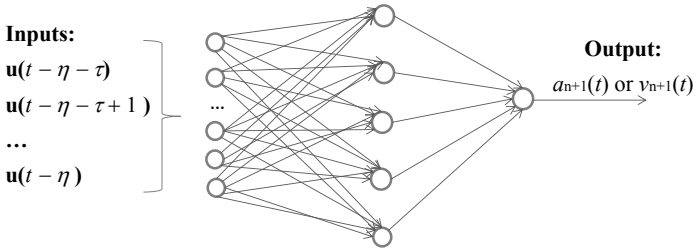


Fig. 2. Feed-Forward neural Network (FF) structure

*Long Short Term Memory Network (LSTM)*

Long Short Term Memory Network (LSTM) is a particular kind of recurrent neural network, especially suitable for working with sequences of data. Thanks to a particular memory mechanism, LSTM works as a feedback network. The output of the model depends not only on the inputs in each time interval but as well on the induced internal state of the network, represented by a memory cell. The structure of the LSTM model used for the car-following is reported in Fig. 3, where solid lines indicate interactions with the current state of the gates while dashed lines are used in case the state from the previous time interval is used.

The model includes four elements: an input gate, a recurrent connection, actuated by the memory cell, a forget gate, and an output gate. The gates regulate interactions between the input and the final output. Firstly, the model takes into account both the input of the current time step and the output of the model from the previous time step. Based on these

data, the input gate, the forget gate, performs some transformations of the data and compute values of the relative functions. Based on these values, the new state of the cell is computed: it can either stay unchanged with respect to the previous time interval or assume a new value computed based on the input gate value. Once the new state of the memory cell has been computed, the new value of the output gate and the output itself can be obtained. Thus, while FF is a memoryless system, whose current output does not depend on the output in the previous time interval, LSTM model propagates the state of the system and the output values obtained in the previous time intervals over time.

In this study, after a preliminary analysis of model performance, the structure including two sequential LSTM, with 16 hidden neurons each, was adopted.

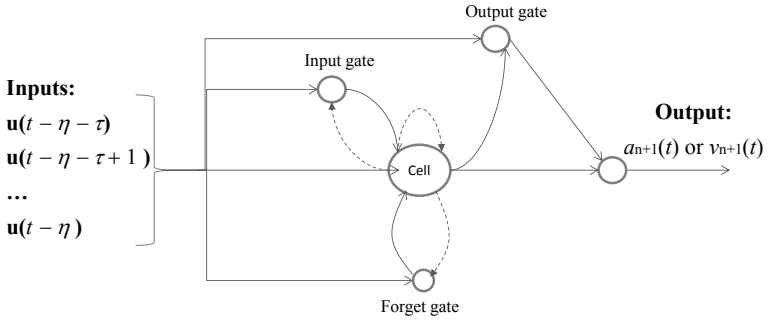


Fig. 3. Long Short Term Memory Network (LSTM) structure

## 4. Experiments

### *4.1. Dataset*

The dataset is composed of measures of car positions collected by high precision GPS devices mounted on different cars running in usual traffic conditions on different urban roads in the town of Rome. Position data were collected at 10 Hz frequency and were post-processed to eliminate systematic errors by using the GPS reference station at Sapienza University of Rome. After this process, the average errors resulted in 2 mm in stillness and ranged from 0.2 to 17 cm in motion. Speed and accelerations were computed through numerical derivation. Several experiments were conducted under typical traffic conditions, including stopping and starting at signals, as well as approaching and moving on roundabouts. As a whole, the experiments covered about 8 km and lasted 16 minutes.

It is worth mentioning that the only directly measured variable is the position of the vehicles. Thus, both the speed and the acceleration of the vehicles are obtained by numeric derivation of the trajectories. Model (2) directly considers the speed of the vehicle, computed as the first derivative of the measured trajectory. Speed values were directly used in the model without applying any filter; acceleration data were processed by a median filter with 21 windows size, meaning that each value is substituted with a median of 21 neighboring values. After that, the first 60% of the data collected is used in the training phase; the remaining 40% of the data is used for validation.

### *4.2. Performance analysis*

In order to identify the most suitable structure of the model inputs, a test for evaluating the model performances was carried out.

Specifically, the reaction duration $\tau$ was changed along with the reaction lag $\eta$, which represents a shift window for the input data. The results of the test for the speed model are reported in Fig. 4. The performances of the models are compared in terms of root mean square error (RMSE). The graphs referring to the acceleration model were very similar and are not reported here for the limited space of the paper. Three different values for $\tau$ are tested: 0.1 s, 0.4 s, 0.8 s. By assuming a discrete formulation of the model with a time step of 0.1 s, the extension of the perceived stimulus from 0.4 s to 0.8 s implies an increase of the input size from a 4×3 to an 8×3 vector. The reaction lag $\eta$, during which the stimulus is not yet perceived, is varied from 0.6 s to 1 s, coherently with other studies on the car following behavior.

In the literature, indeed, even values slightly longer than 1 s are considered; these values are not in contrast with our model that assumes the reaction starts after a reaction lag varying from a minimum of $\eta = 0.6$ s to a maximum of $\eta = 1.0$ s and ends after a time interval varying from a minimum of 0.7 s ($\eta = 0.6$ s; $\tau = 0.1$ s) and a maximum of 1.8 s ($\eta = 1.0$ s; $\tau = 0.8$ s). The case $\tau = 0.1$ s corresponds to the usual car-following model with a lagged and instantaneous reaction.

The results obtained for the FF and the LSTM models are represented in Fig. 4-a and Fig. 4-b, respectively. For the FF model, a gain in accuracy by using more than one time lag as input is achieved for all the values of the reaction lags, except for $\eta = 0.9$ s, more frequently for $\tau = 0.4$ s. For the LSTM model, the relative gain is lower. For what concerns the selection between the other two structures, both models provide the best performances for a reaction duration of 0.4 s; that is, a four-interval structure, although the gain is more evident for the LSTM model.

Fig. 5 reports a scatterplot of the results obtained for the NN and LSTM acceleration models with a reaction lag of 0.6 s and different values of reaction duration. Fig. 5–a relates to the simplest case of the model, with a stimulus duration $\tau = 0.1$ s. The FF model generally underestimates the values of acceleration while the LSTM model provides a better fit to the data but underestimates the highest values of both acceleration and deceleration. The introduction of an extended duration of the stimulus (Fig. 5-b and Fig. 5-c) increases the accuracy of the model, reduces the general underestimation of the acceleration values, and fits the estimated values closer to the bisector line of the plot in the whole interval of the observed acceleration. Both models report an unavoidable dispersion of the results, almost uniform around the bisect line, with a limited systematic underestimation of deceleration values lower than –1 $ms^{-2}$. This result can likely be attributed to an abrupter process than the regular deceleration corresponding to the car-following model.

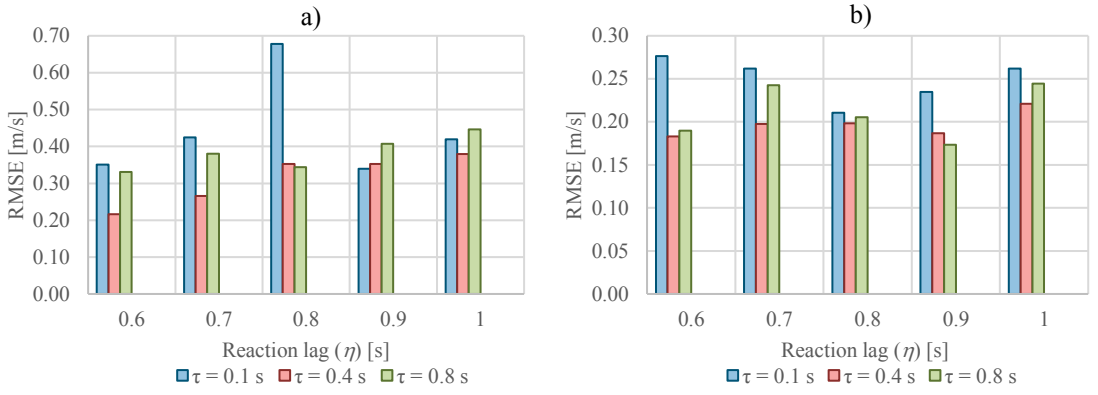


Fig. 4. Analysis of model performances based on the duration of reaction lag ($\eta$) and stimulus duration ($\tau$): a) FF model, b) LSTM model.

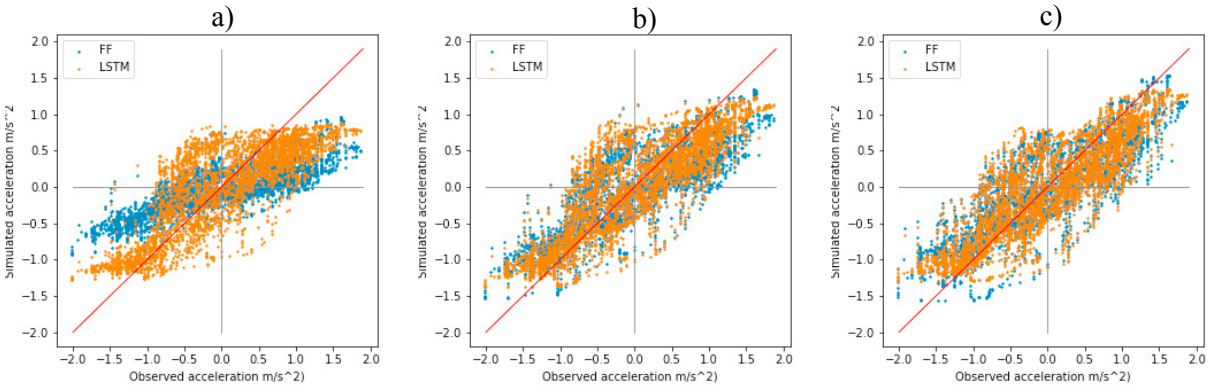


Fig. 5. Examples of observed acceleration values and simulated acceleration values provided by FF model (blue) and LSTM model (orange) with reaction lag $\eta = 0.6$ s and different values of stimulus duration $\tau$: a) $\tau = 0.1$ s; b) $\tau = 0.4$ s; c) $\tau = 0.8$ s.

### 4.3. Validation

In the validation phase, the models are applied to a set of data that were not used during the calibration phase. The performance analysis has permitted to individuate the most suitable structure for both models, corresponding to a structure based on the stimulus perceived during a time interval of 0.4 s. While this number is maintained constant, a further test was made on the reaction lags, which was varied from 0.6 s to 1 s with a step of 0.1 s. The determined structures were then tested in both speed and acceleration car-following models. The corresponding results are summarized by the main error indicators: Mean Absolute Error (MAE), Mean Absolute Relative Error (MARE), Root Mean Square Error (RMSE), Root Mean Square Error Relative (RMSER).

Table 1 and Table 2 report the results obtained by the application of FF and LSTM for the car-following model based on speed and acceleration, respectively. In both forms, LSTM model outperformed FF model for all the values of reaction lags. For the speed model, the gain of the accuracy of LSTM compared to FF, computed for different values of the reaction lags, is, on average, around 20% in terms of MAE and around 40% in terms of RMSE. For the acceleration model, the accuracy gain is around 24% in terms of MAE and around 40% in terms of RMSE.

In Fig. 6, the speed profiles provided by the LSTM and FF speed car-following models. Both models show a good approximation of the observed speed values. However, the LSTM model exhibits a better capability to capture the peaks of the data. The simulated acceleration profiles are reported in Fig. 7. The obtained curves do not fit the data as precisely as the profiles obtained by the application of the models on the speed data.

For both models, the accuracy generally decreases as the reaction lag increases. This trend is more evident for the speed model, and this result can likely be attributed to the autoregressive feature of this model, where traditionally, the prediction for longer horizons becomes a more challenging task. The speed-based LSTM model with $\eta = 0.6$ s and $\tau = 0.4$s provides the best accuracy (MARE = 0.05 and RMSER = 0.01).

In the case of the acceleration model, unlike the speed-based model, the dependent variable is formulated as a function of its primitive instead of its previous values. Thus, the acceleration-based LSTM model does not exploit the simulated variable as an input variable, and the simulation is not corrected every time based on the most recent data. However, also in this application, the LSTM model approximates data peaks better than FF for both acceleration and deceleration phases. Higher errors are experienced in this case (MARE = 0.39 and RMSER = 0.23).

It is worth noting that by applying a numerical derivation of the speed profile to estimate the acceleration, we obtain a noisier profile than that of the acceleration-based LSTM and FF models with higher error indicators.

Table 1. Error indicators (Mean Absolute Error, Mean Absolute Percentage Error, Root Mean Square Error, Root Mean Square Error Percentage) for the feed-forward neural network (FF) and the long short-term memory (LSTM) neural network, and different values of the initial instant of reaction ($\eta$) and the duration of the stimulus-reaction ($\tau$) of the speed car-following model.

| Reaction Lag ($\eta$) | Stimulus-Reaction duration ($\tau = 4$) | LSTM | | | | FF | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAE (m/s) | MARE adim. | RMSE (m/s) | RMSER adim. | MAE (m/s) | MARE adim. | RMSE (m/s) | RMSER adim. |
| (0.0s ÷ 0.6s) | (0.7s ÷ 1.1s) | 0.37 | 0.05 | 0.22 | 0.01 | 0.53 | 0.07 | 0.46 | 0.02 |
| (0.0s ÷ 0.7s) | (0.8s ÷ 1.2s) | 0.42 | 0.06 | 0.27 | 0.01 | 0.52 | 0.07 | 0.43 | 0.02 |
| (0.0s ÷ 0.8s) | (0.9s ÷ 1.3s) | 0.48 | 0.07 | 0.35 | 0.02 | 0.61 | 0.08 | 0.59 | 0.03 |
| (0.0s ÷ 0.9s) | (1s ÷ 1.4s) | 0.48 | 0.06 | 0.35 | 0.01 | 0.58 | 0.08 | 0.51 | 0.02 |
| (0.0s ÷ 1.0s) | (1.1s ÷ 1.5s) | 0.50 | 0.07 | 0.38 | 0.02 | 0.62 | 0.08 | 0.62 | 0.02 |

Table 2. Error indicators (Mean Absolute Error, Mean Absolute Percentage Error, Root Mean Square Error, Root Mean Square Error Percentage) for the feedforward neural network (FF) and the long short-term memory (LSTM) neural network, and for different values of the initial instant of reaction ($\eta$) and the duration of the stimulus-reaction ($\tau$) of the acceleration car-following model.

| Reaction Lags ($\eta$) | Stimulus-Reaction duration ($\tau = 4$) | LSTM | | | | FF | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAE ($m/s^2$) | MARE adim. | RMSE ($m/s^2$) | RMSER adim. | MAE ($m/s^2$) | MARE adim. | RMSE ($m/s^2$) | RMSER adim. |
| (0.0s ÷ 0.6s) | (0.7s ÷ 1.1s) | 0.36 | 0.39 | 0.18 | 0.23 | 0.49 | 0.53 | 0.32 | 0.37 |
| (0.0s ÷ 0.7s) | (0.8s ÷ 1.2s) | 0.37 | 0.41 | 0.20 | 0.25 | 0.54 | 0.59 | 0.36 | 0.43 |
| (0.0s ÷ 0.8s) | (0.9s ÷ 1.3s) | 0.37 | 0.41 | 0.20 | 0.25 | 0.45 | 0.50 | 0.28 | 0.34 |
| (0.0s ÷ 0.9s) | (1.0s ÷ 1.4s) | 0.36 | 0.40 | 0.19 | 0.25 | 0.51 | 0.56 | 0.34 | 0.41 |
| (0.0s ÷ 1.0s) | (1.1s ÷ 1.5s) | 0.40 | 0.43 | 0.22 | 0.27 | 0.48 | 0.52 | 0.31 | 0.37 |

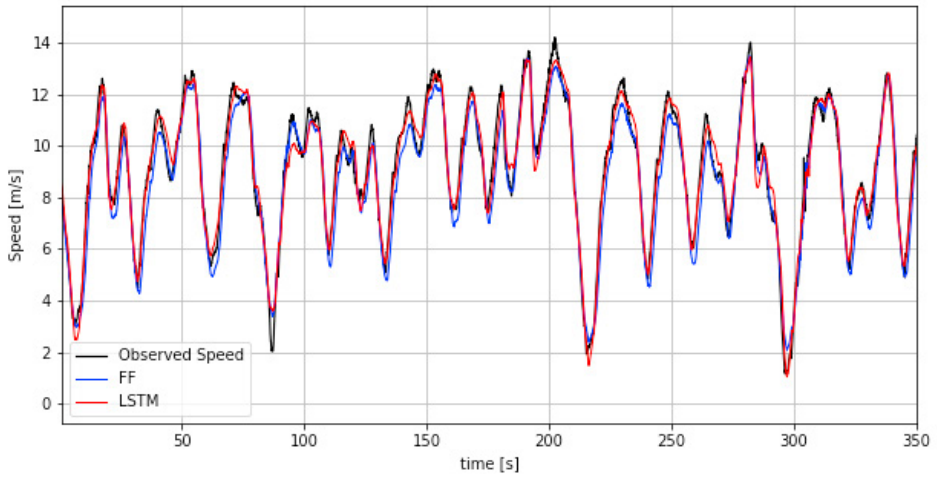


Fig. 6. Car-following models performances obtained by the application of FF and LSTM models for speed modeling

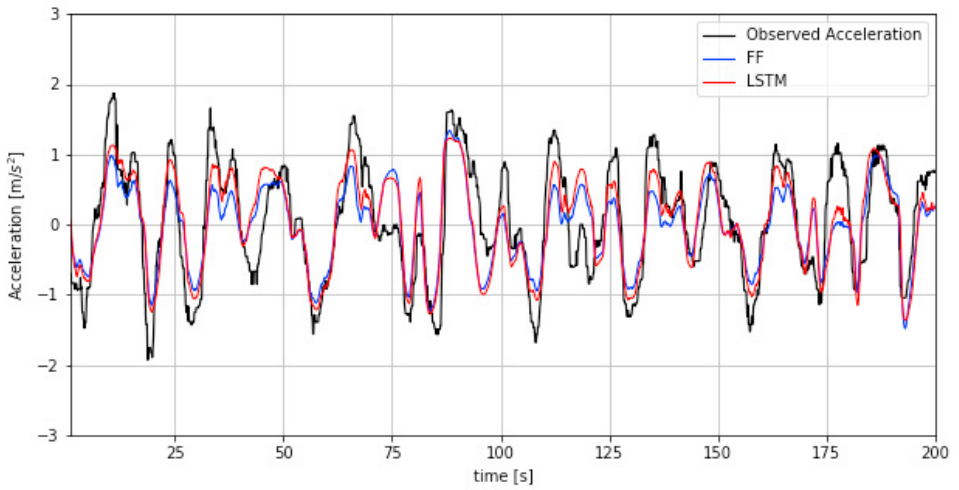


Fig. 7. Car-following models performances obtained by the application of FF and LSTM for acceleration models.

## 5. Conclusions

This study dealt with the modeling of car-following behavior by the application of machine learning methods. Two alternative formulations of the car-following model were proposed: the traditional formulation concerning the follower's acceleration and an alternative formulation based directly on the follower's speed. Two models, namely a traditional feed-forward neural network and a more recent long short-term memory neural network, were then applied to reproduce the behavior of the following vehicle. Both models were based on kinematic features of the movement of two vehicles: that are the speed of the vehicles itself, and the relative speed and the spacing from the leading vehicle. Different tests on the performances of the models analyzed different extensions for the stimulus perception and different values of reaction time for the driver.

The models were trained and tested on a real dataset of highly precise GPS data collected with 10 Hz frequency during 16 minutes on an 8 km path.

The obtained results demonstrated that both models were suitable for car-following modeling. Specific analysis of the models' sensitivity to the number of stimulus intervals included in the model showed that including additional intervals significantly improved the predictions, especially for the feed-forward network. Long short term memory model outperformed the traditional feed-forward network for both formulations: the accuracy increased by 20% and by 24% in terms of mean absolute error, respectively, for the speed and the acceleration model.

The car following modeling in terms of the speed variable showed, in general, better performance respect to the modeling of the acceleration variable. Indeed, in case of the speed model, the simulated variable is formulated as a function of its previous values, and can, therefore, be seen as a time series modeling, while in case of the acceleration model is formulated as a function of its primitive, a more difficult task for the prediction. Nevertheless, modeling of the acceleration of the following vehicle must be seen as an essential task, since it represents the directly executed changes of the behavior, made by the driver. A further analysis based on the derivation of the acceleration profile from the speed profile suggested a possible improvement of the model by the implementation of data filtering techniques in order to smooth the obtained profile.

The undergoing research is, therefore, concentrated on the analysis of the models applied to multiple couples of the vehicles and the derivation of the acceleration profile based on the speed model.

## References

Chandler, R.E., Herman, R., Montroll, E.W., 1958. Traffic Dynamics: Studies in Car Following. Oper. Res. 6, 165–184.

Colombaroni, C., Fusco, G., 2014. Artificial neural network models for car following: Experimental analysis and calibration issues. J. Intell. Transp. Syst. Technol. Planning, Oper. 18, 5–16. https://doi.org/10.1080/15472450.2013.801717

Elamrani Abou Elassad, Z., Mousannif, H., Al Moatassime, H., Karkouch, A., 2020. The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. Eng. Appl. Artif. Intell. 87, 103312. https://doi.org/10.1016/j.engappai.2019.103312

Fan, P., Guo, J., Zhao, H., Wijnands, J.S., Wang, Y., 2019. Car-following modeling incorporating driving memory based on autoencoder and long short-term memory neural networks. Sustain. 11. https://doi.org/10.3390/su11236755

Felice, M. De, Baiocchi, A., Cuomo, F., Fusco, G., Colombaroni, C., 2014. Traffic monitoring and incident detection through VANETs. pp. 122–129. https://doi.org/10.1109/WONS.2014.6814732

Fusco, G., Gori, S., 1995. The Use of Artificial Neural Networks in Advanced Traveler Information and Traffic Management Systems, in: 4th International Conference on Applications of Advanced Technologies in Transportation Engineering. ASCE, New York, NY, United States, Capri, Italy. pp. 341–345.

Gazis, D.C., Herman, R., Rothery, R.W., 1961. Nonlinear Follow-the-Leader Models of Traffic Flow. Oper. Res. 9, 545–567. https://doi.org/10.1287/opre.9.4.545

Gipps, P.G., 1981. A behavioural car-following model for computer simulation. Transp. Res. Part B 15, 105–111. https://doi.org/10.1016/0191-2615(81)90037-0

Huang, X., Sun, J., Sun, J., 2018. A car-following model considering asymmetric driving behavior based on long short-term memory neural networks. Transp. Res. Part C Emerg. Technol. 95, 346–362. https://doi.org/10.1016/j.trc.2018.07.022

Isaenko, N., Colombaroni, C., Fusco, G., 2017. Traffic dynamics estimation by using raw floating car data, in: 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings. https://doi.org/10.1109/MTITS.2017.8005604

Kometani, E., Sasaki, T., 1961. Dynamic Behavior of Traffic with a Nonlinear Spacing-Speed Relationship. Theory Traffic Flow 105–119.

Wang, X., Jiang, R., Li, L., Lin, Y., Zheng, X., Wang, F.Y., 2018. Capturing Car-Following Behaviors by Deep Learning. IEEE Trans. Intell. Transp. Syst. 19, 910–920. https://doi.org/10.1109/TITS.2017.2706963

Wiedemann, R., 1975. Simulation des Strassenverkehrsflusses, Schriftenreihe des Instituts für Verkehrswesen der Universität Karlsruhe, Band 8, Karlsruhe, Germany.